USAID
FROM THE AMERICAN PEOPLE

SECOND FOLLOW-UP REPORT

IMPACT EVALUATION OF WOMEN'S LEADERSHIP IN SMALL AND MEDIUM ENTERPRISES IN THE KYRGYZ REPUBLIC

FEBRUARY 13, 2017

This publication was produced for review by the United States Agency for International Development. It was prepared by Management Systems International, A Tetra Tech Company, for the E3 Analytics and Evaluation Project.

COVER PHOTO

Caption: WLSME participants learning from each other during a site visit of other women-owned small businesses in the garment industry.

Credit: Gulmira Asanbaeva, WLSME Activity

SECOND FOLLOW-UP REPORT

IMPACT EVALUATION OF WOMEN'S LEADERSHIP IN SMALL AND MEDIUM ENTERPRISES IN THE KYRGYZ REPUBLIC

Contracted under AID-OAA-M-13-00017

E3 Analytics and Evaluation Project

Prepared by:
Alberto Chong, Principal Investigator, MSI
Irene Velez, MSI

CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

The evaluation team would like to thank the Kyrgyz women who shared their time to participate in this study, M-Vector Research and Consulting for conducting the surveys, and ACDI/VOCA for its assistance and guidance throughout the study. They would also like to thank Angelo Cozzubo for his statistical support and Altyn Kapalova for providing quality oversight during data collection. The evaluation team also benefited from the shared thoughts and experiences of the Women's Leadership in Small and Medium Enterprises (WLSME) Community of Practice, including the WLSME Peru team (GRADE/Sector 3 Social Venture Group) and WLSME India team (CARE), as well as technical reviews from Gary Glass, Jacob Patterson-Stein, and Jeremy Gans at MSI. Finally, the evaluation team would like to thank Anastasia de Santos and Wade Channell of USAID for their continued support of this study, and FHI 360 for its work with the original evaluation design and baseline phase.

ACRONYMS

ANCOVA	Analysis of Covariance
APR	Annual Percentage Rate
BMT	Business Management Trainings
BT Fund	Bai-Tushum Innovations Fund
DID	Difference-in-Differences
E3	Bureau for Economic Growth, Education, and Environment (USAID)
GenDev	Office of Gender Equality and Women's Empowerment (USAID/E3)
LWA	Leader with Associates
ML	Market Linkages
MSI	Management Systems International
RCT	Randomized Controlled Trial
SME	Small and Medium Enterprise
TS/AF	Technical Skills/Access to Finance
USAID	United States Agency for International Development
USD	United States Dollars
WLSME	Women's Leadership in Small and Medium Enterprises

EXECUTIVE SUMMARY

This report is the second follow-up to the impact evaluation of the Kyrgyzstan Women's Leadership in Small and Medium Enterprises (WLSME) activity commissioned by the Office of Gender Equality and Women's Empowerment in the United States Agency for International Development's Bureau for Economic Growth, Education, and Environment (USAID/E3/GenDev). The report summarizes the data collection and analysis methods and provides findings and conclusions at 12 months after the end of the Kyrgyzstan WLSME activity. It also includes findings from the first follow-up round, conducted at the end of the activity in 2015. The impact evaluation plans to conduct one more follow-up round, at 24 months' post-intervention, to examine longer-term impacts of the activity.

The medium-term findings of this second follow-up report are somewhat encouraging yet still inconclusive. A year after the WLSME activity ended, there is:

- An increased number of paid household members;
- Increased perception that there are no gender inequality problems in the women's community;
- Decreased likelihood of being an active member of any social group; and
- Increased recording of an owner's salary in a registry or computer.

WLSME Kyrgyzstan Activity Description

USAID's WLSME initiative aims to address women's relative absence in the small and medium enterprise (SME) sector. The WLSME Kyrgyzstan activity directly addresses two critical barriers: (1) agency constraints, which impede adequate accumulation of human capital and managerial capital, and (2) relationship constraints, which limit women's access to information and reduce opportunities to build and draw on social capital.

ACDI/VOCA and its partner organization, Bai-Tushum Innovations Fund, implemented the Kyrgyzstan WLSME activity nationwide between September 2013 and September 2015. The activity targeted 960 women in the garment, tourism, and agro-processing sectors who were identified as potential high-growth entrepreneurs. The activity consisted of three components: (1) Business Management Trainings (BMT), which covered topics such as negotiating skills, business planning, marketing, financial planning, productivity, and human resource management; (2) Market Linkages (ML), including stakeholder meetings, trade fairs, workshops on value chains and sub-sectors, semi-annual value chain stakeholder meetings, web page resources, and an annual business plan competition; and (3) Technical Skills/Access to Finance (TS/AF), where only finalists and semi-finalists from the business plan competitions received customized assistance, including technical training, targeted technical assistance, mentorship, and exchange visits. They were also eligible to apply for small grants and a loan product.

Evaluation Design

The WLSME Kyrgyzstan activity is based on the development hypothesis that if women business owners have greater human capital, social capital, and access to market information, then they will be more likely to grow their businesses and become entrepreneurial leaders. This impact evaluation aims to test this hypothesis with a randomized controlled trial, where eligible applicants to the WLSME activity are randomly assigned to a treatment and control group. Only women in the treatment group were given access to the WLSME activity, but once enrolled, participation in the components of the activity was demand-driven. The key evaluation questions to be answered are:

1. Primary Question (combined impact T vs. C): Compared to participants in the control group, do participants who are randomly assigned to receive the program have higher mean values on the following, post-intervention outcomes: entrepreneurial leadership, business growth, business knowledge/practices, and social/business networks?

2. Secondary Questions (separate estimates across T arms):

    - Compared to participants who only receive BMT, do participants also exposed to ML have higher mean values on the same set of outcomes listed under the Primary Objective?
    - Compared to participants who only receive BMT, do participants also exposed to TS/AF have higher mean values on the same set of outcomes listed under the Primary Objective?

The baseline sample consisted of 843 women, of whom 568 were randomly assigned to the treatment group and 275 women were randomly assigned to the control group. This excludes the agro-processing sector participants who were not randomized into treatment. Out of the 568 participants enrolled in the activity, 461 actually participated, 378 completed the BMT component, 251 participated in the ML component, and 126 participated in the TS/AF component. The evaluation team conducted the first follow-up survey at the end of the activity, between August and October 2015, with an 82 percent response rate. The team conducted the second follow-up survey between August and October 2016, with an 84 percent response rate. The second follow-up sample consisted of 704 women, of whom 472 are in the treatment group and 232 in the control group.

Balance tests were originally conducted at baseline but are reported again in this report to show that the remaining treatment group, after excluding the agro-processing participants, and the control group shared similar social and economic conditions before the WLSME activity started. Analysis of survey non-responses showed that non-response was random, as there is no correlation with treatment assignment and survey non-response does not depend on observable characteristics.

To answer the primary questions, the evaluation team used two empirical models to estimate the Kyrgyzstan WLSME activity's overall impact on the key set of outcome variables: difference-in-differences (DID) and analysis of covariance (ANCOVA). The ANCOVA findings for both follow-up rounds are reported below; the DID results can be found in Annex D. To answer the secondary questions, the evaluation team used the DID model. However, given that selection into the activity components was not randomized, it is not possible to look at a causal treatment effect within the activity, but only at correlations. This is because participants who self-selected into participating in the ML or TS/AF components are different in ways that are correlated to the outcomes. For example, women who attended the additional WLSME components may have been the most motivated or had a stronger support network to be able to dedicate more time away from home compared to women who only participated in the BMT, so they would have done better with their business and leadership even in absence of the WLSME activity.

Findings

The majority of outcomes do not yield a statistically significant links to the overall WLSME activity at the medium-term, but there are positive causal links for a small number of outcome variables. The short-term effects that had been detected in the first follow-up round are no longer detected in the second follow-up round. The third and final follow-up round, at 24 months' post-intervention, could shed more insights into the overall impact of the activity. The findings reported below demonstrate the impact of the overall Kyrgyzstan WLSME activity (Primary Question).

Business Growth

- Sales and profits have moved in a positive direction when comparing treatment and control. This means that both sales and profits of the treatment group tend to show an increase with respect to the control group. While this evidence is not statistically significant, the positive shift and economically sizable magnitude in the coefficients are encouraging.
- While the short-term capital investment effect is no longer significant in the second follow-up round, there is now a small employment effect, with an increase of 0.32 number of paid household workers. While total number of employees has not increased, women in the treatment group are remunerating more of the help they are receiving from their household. It is uncertain whether this is due to greater value placed on human capital resources, perception of growth potential, or leading to future increased sales and profits.

Entrepreneurial Leadership

- The evaluation team finds mixed, yet not statistically significant, results for entrepreneurial leadership outcomes. Some of the coefficients have corresponding signs that are not consistent with the expected behavioral change of increased leadership. In particular, for some outcomes in decision-making and independence, women in the treatment group were less likely to participate in different aspects of the business and more likely to ask for advice in general business planning, deciding what type of work to do, and staffing of business.
- While there is some agreement with questions that point to gender roles being at play and hard to change, there is a statistically significant increase of six percentage points in the perception that there are no gender inequalities in the community.

Networks

- The short-term effects of increased implementation of professional advice and the number of people to ask for business advice are no longer detected in the second follow-up round. In the medium term, the only statistically significant effect is on the likelihood of being an active member of any social group, which has a counterintuitive effect of a 10 percentage point reduction.

Business Knowledge and Practices

- Only one outcome variable shows a statistically significant effect. There is an 11.2 percentage point increase in recording the owner's salary in a registry or computer for the treatment group compared to the control group.
- The effects on the other outcome variables are not statistically significant and are also very small in magnitude.

The evaluation team's analysis of the Secondary Questions shows that the different treatment components (ML and TS/AF) may have incremental effects over the BMT component with respect to a few outcomes measures. However, unlike the results presented above, those results cannot be interpreted as causally valid, but only as relevant correlations.

An important caveat to these findings is that, while a significant number of outcome variables show a change in the coefficient in the expected direction, this does not occur for all outcome variables surveyed. Furthermore, the vast majority of outcome variables do not show any statistical significance, regardless of the empirical method employed. The third and final follow-up round may help clarify the extent to which the WLSME activity has had a "true" positive effect.

Conclusions

The short-term and medium-term impacts of the WLSME activity are somewhat encouraging yet still inconclusive. The limited statistical significance in the outcome variables may still be related to the temporal aspect of changing behavior, as it takes more time to generate change in tasks that require more assertiveness or effort than in others. However, the lack of consistency between the first and second follow-ups may indicate either statistical power issues or difficulty in measuring some of these outcomes. Nonetheless, given the slight improvement in response rate for the second follow-up round, the evaluation team expects to conduct the third follow-up round as originally planned in September 2017.

Given the non-random selection of participants into the activity's components, it is not possible to conclude whether agency or relationship constraints are greater barriers in the Kyrgyz context and whether activities to address these constraints have different levels of effectiveness. Understanding more about how these constraints can be addressed programmatically is particularly important given the prevalence of business training activities around the world.

INTRODUCTION

This report is the second follow-up to the impact evaluation of the Kyrgyzstan Women's Leadership in Small and Medium Enterprises (WLSME) activity commissioned by the Office of Gender Equality and Women's Empowerment in the United States Agency for International Development's Bureau for Economic Growth, Education, and Environment (USAID/E3/GenDev). The E3 Analytics and Evaluation Project provided post-baseline support for the evaluation, including ongoing implementation monitoring and follow-up data collection.[1] The impact evaluation uses a randomized controlled trial (RCT) design to test how the WLSME interventions affect women and their businesses across four primary sets of indicators – business growth, entrepreneurial leadership, business knowledge/practices, and social/business networks.

This second follow-up report summarizes the data collection and analysis methods and provides findings and conclusions at 12 months after the end of the Kyrgyzstan WLSME activity. It also includes findings from the first follow-up round, which the evaluation team conducted at the end of the activity in 2015. The evaluation team expects to conduct one more follow-up round at 24 months' post-intervention, to examine longer-term impacts.

WLSME INITIATIVE DESCRIPTION

USAID's WLSME initiative aims to address women's relative absence in the small and medium enterprise (SME) sector by implementing specific measures to reduce critical barriers, so that women may benefit from labor market participation both in the short- and medium-term. These barriers include: (1) agency constraints, which impede adequate accumulation of human capital and managerial capital, and thus limit women's knowledge and business practices; (2) relationship constraints, which limit women's access to information and, as a consequence, reduce the opportunities for women entrepreneurs to build and draw on social capital; and (3) external constraints, which place gender-specific barriers that limit the presence and success of women entrepreneurs. The Kyrgyzstan WLSME activity directly addresses only the first two constraints.

In September 2012, USAID awarded three WLSME activities in Kyrgyzstan, India, and Peru through a competitive process, each with a performance period of three years and a budget of around $1.5 to $2 million. USAID also initiated impact evaluations for each of these three activities under the Financial Integration, Economic Leveraging, Broad-Based Dissemination (FIELD-Support) Leader with Associates Cooperative Agreement with FHI 360 that concluded on September 30, 2014, following the completion of evaluation design and baseline data collection and analysis activities. Subsequently, USAID transferred implementation of two of the impact evaluations (for Kyrgyzstan and India) to the E3 Analytics and Evaluation Project, while the Peru evaluation is being separately funded and completed through the Multilateral Investment Fund of the Inter-American Development Bank.

Kyrgyzstan WLSME Activity Description

ACDI/VOCA, in collaboration with its partner organization Bai-Tushum Innovations Fund (BT Fund), implemented the WLSME activity in Kyrgyzstan. The activity operated nationwide from September 2013 to September 2015 and targeted 960 women who operated enterprises in priority sectors (garment, tourism, and agro-processing), met minimum employee and loan size requirements, and were identified

[1] Team lead Management Systems International (MSI, implements USAID's E3 Analytics and Evaluation Project in collaboration with partners Development and Training Services (dTS) and NORC at the University of Chicago.

as potential high-growth entrepreneurs. Activity components were sequenced, with standard services and courses offered to all participants. As components became more tailored and specialized, they progressively focused on fewer women. The most intensive mentoring and skills development components were reserved for the most promising women entrepreneurs. The components of this activity were nested within each other and consisted of subsets of participants:

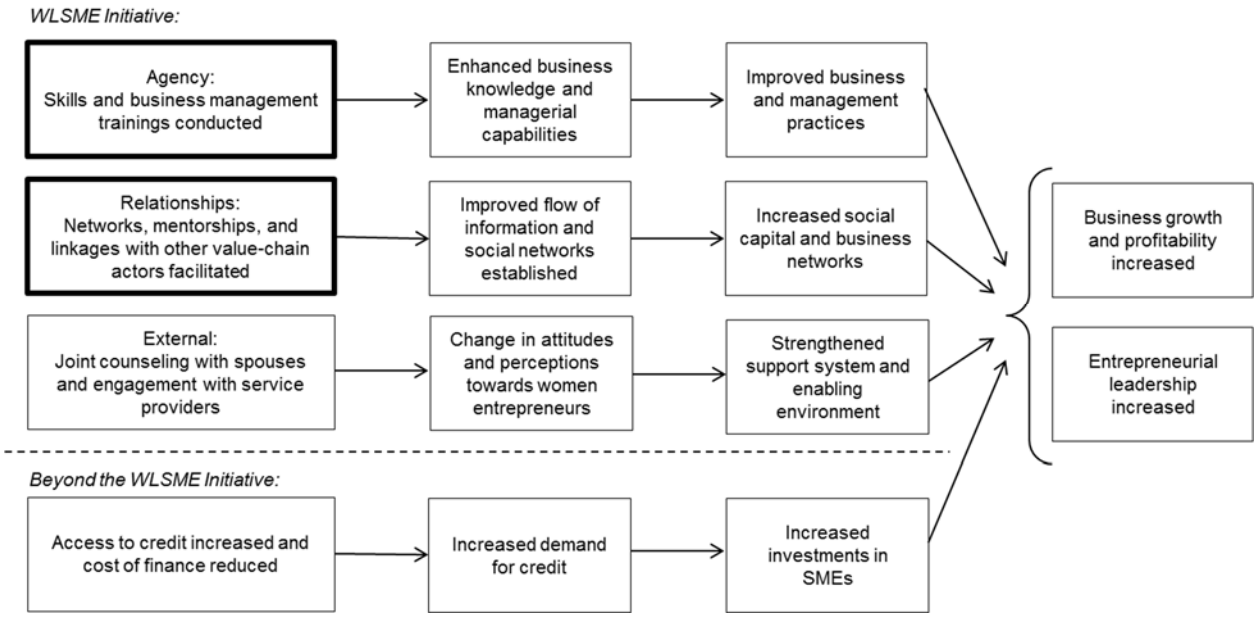
- **Component 1 – Human Capital Gap (Agency):** Business Management Trainings (BMT) covered topics such as negotiating skills, business planning, marketing, financial planning, productivity, and human resource management. The BMT were demand driven and market oriented, meaning that women had a choice to some extent on which topics and when they attended, based on their interest. However, they had to attend a minimum of 24 hours of business training (4 days of 6-hour training sessions) to complete this component, plus a 3-day Business Planning Seminar. Women could continue with more training, up to a maximum of 72 hours per participant.
- **Component 2 – Information and Social Capital Gap (Relations):** This Market Linkages (ML) component included stakeholder meetings, trade fairs, workshops on value chains and sub-sectors, semi-annual value chain stakeholder meetings, web page resources, and an annual business plan competition. Originally, only participants that completed their BMT requirements from Component 1 were invited to the activities under Component 2. However, starting in January 2015, this requirement was dropped to increase the take-up rate of the overall activity.
- **Component 3 – Technical Skills/Access to Finance (TS/AF):** Finalists and semifinalists from the business plan competitions in Component 2 received this customized assistance. Assistance included technical training, targeted technical assistance, mentorship, and exchange visits. Participants were also eligible to apply for small grants of between USD $200 and $2,000. Instead of the initially planned loan guarantee fund, BT&P Bank offered a new loan product with a lower interest rate.[2]

DEVELOPMENT HYPOTHESES

USAID's development hypotheses for the WLSME initiative are displayed graphically in Figure 1, highlighting each of the intended results of the components and the presumed causal linkages (arrows). While the diagram focuses on the three parallel constraints that are hypothesized to impede business growth and entrepreneurial leadership, which the WLSME initiative aim to address, the Kyrgyzstan activity directly addresses only the first two – agency and relationship constraints – shown in bolded boxes. The third constraint (external constraints), addressed through joint counseling with spouses and engagement with service providers, was not part of the Kyrgyzstan activity. The Kyrgyzstan activity also includes the possibility of increased investments in SMEs that may occur as a result of increased access and reduced cost of finance from the BT Fund partnership in Kyrgyzstan. However, this final path of interest is beyond the WLSME initiative and is not being delivered exclusively to activity beneficiaries, so it is depicted below the dotted line.

[2] Regulatory changes made the planned loan guarantee fund no longer feasible. BT&P Bank's loan product had a 19 Annual Percentage Rate (APR), lower than the 28 APR market rate but higher than the state bank loans with a 10 APR. All activity participants, including the control group, had general access to loans from the BT&P Bank.

FIGURE 1: THEORY OF CHANGE



EVALUATION QUESTIONS

Little empirical research exists that provides convincing evidence about which interventions have the greatest chance of success in terms of creating female-led SMEs and helping female business owners grow their businesses (see Annex A: Literature Review). The purpose of this evaluation is to provide a learning, accountability, and decision-making platform by clarifying the most important constraints to women's business growth and leadership, and thereby the most effective means to unleash the potential of women's entrepreneurship in the SME sector in Kyrgyzstan. This evidence is expected to be useful to USAID staff, other donors, host governments, and stakeholders to improve future programming in order to better address the barriers to women's entrepreneurship at the SME level.

USAID's evaluation questions included here are taken directly from the Evaluation Protocol designed by FHI 360.

1. Primary Question (combined impact T vs. C): Compared to participants in the control group, **do participants who are randomly assigned to receive the program have higher mean values on the following, post-intervention outcomes: entrepreneurial leadership, business growth, business knowledge/practices, and social/business networks?**

2. Secondary Questions (separate estimates across T arms):

    - Compared to participants who only receive Business Management Trainings, do participants also exposed to Market Linkages have higher mean values on the same set of outcomes listed under the Primary Objective?
    - Compared to participants who only receive Business Management Trainings, **do participants also exposed to Technical Skills/Access to Finance have higher mean values on the same set of outcomes listed under the Primary Objective?**

Outcome Measures

A number of outcomes (dependent variables) linked to the WLSME theory of change were defined to measure whether and how much change the activity caused for women entrepreneurs. These include:

- **Business Growth:** This includes, but is not limited to: measures of sales, profits, number of employees, number and type of paid employees, hours worked, investments, and formality.
- **Entrepreneurial Leadership:** This includes measures on decision-making in the business, entrepreneurial vocation, level of independence, and women's empowerment.
- **Business Knowledge and Practices:** This includes measures on marketing, inventory management, costing and record keeping, and financial planning.
- **Social/Business Networks:** This includes measures regarding participants' involvement in professional networks, such as the number of other business owners with whom the woman discusses business matters, as well as commercial networks.

Gender Aspects of the Questions

USAID evaluation guidance calls upon Agency staff and evaluation teams to examine evaluation questions from a gender perspective and to incorporate gender issues into study designs. This WLSME activity is targeted at women only, and the evaluation does not intend to collect data from male stakeholders (either spouses or male value chain actors). Thus, it will not be possible to disaggregate data collected in this evaluation by gender or to look at the differential gender effects of the activity components. Nonetheless, the main objective of the WLSME initiative is to close the multiple existing gaps between women and men in SMEs.

EVALUATION DESIGN

This impact evaluation is based on an RCT design. Eligible applicants to the Kyrgyzstan WLSME activity were interviewed for the baseline and then randomly assigned to participate in the activity on a rolling basis, starting in July 2013. Randomization into the treatment and control groups was done at the batch level, on average 32 women, in a 2:1 ratio. The baseline sample consisted of 843 women, of whom 568 were randomly assigned to the treatment group and 275 women were randomly assigned to the control group. This excludes the 354 agro-processing sector participants who were not randomized into treatment.[3] All numbers in the rest of this report exclude the agro-processing sector participants.

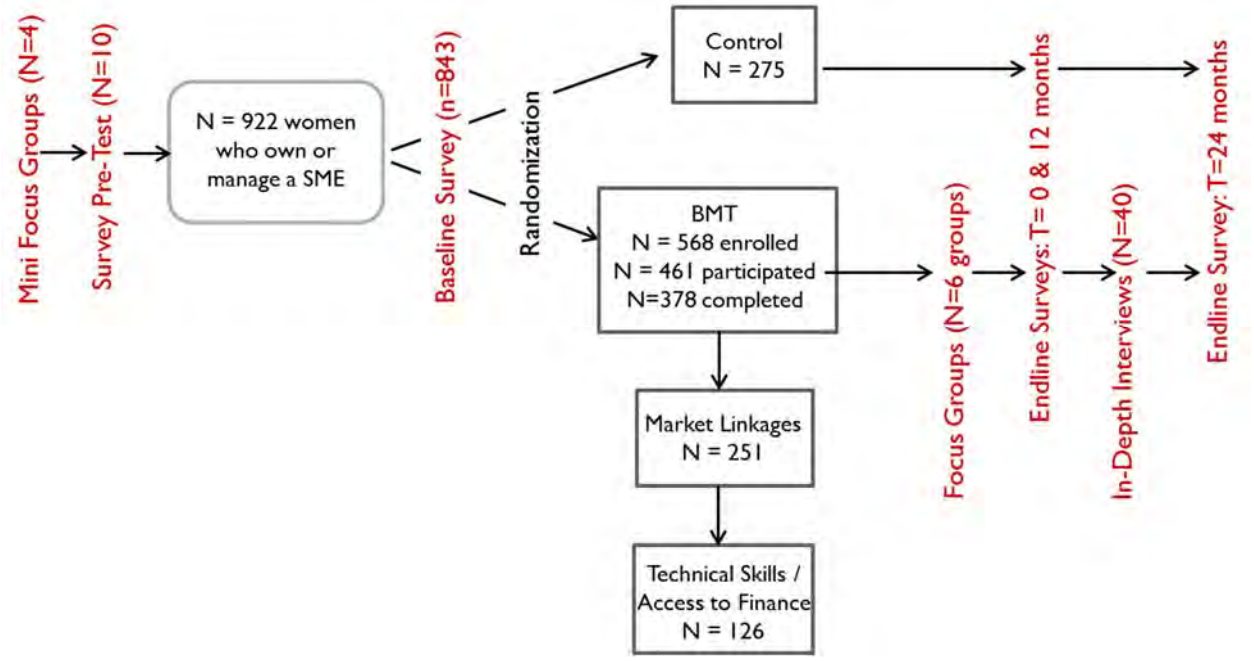
The WLSME activity offered BMTs to the treatment group only, of which 461 participants (81 percent take-up) attended at least one BMT session and 378 participants (66 percent take-up) completed the BMT component. This is consistent with the average take-up rate across different studies for individuals offered business training, as reported by McKenzie and Woodruff (2014). For most of the duration of the activity, participants had to complete their BMT to access the ML activities. However, the activity changed this in January 2015 to increase take-up of the overall activity components. A total of 251 women participated in the ML component. The activity offered the TS/AF component to only 126 participants who were finalists and semifinalists of the Business Plan competition.

The evaluation team conducted the first follow-up survey at the end of the WLSME activity, between August and October 2015, with an 82 percent response rate. The team conducted the second follow-up survey 12 months after the end of the activity, between August and October 2016, with an 84 percent

[3] ACDI/VOCA, USAID, and the FHI 360 evaluation team made the decision to exclude the agro-processing sector participants from the evaluation sample. See Briefing Note submitted to USAID on February 11, 2015.

response rate. The next section further analyzes the non-response rate, and Annex C summarizes the evaluation's data collection and quality assurance process. The overall evaluation design, timeline, and participation numbers are shown in Figure 2.

FIGURE 2: KYRGYZSTAN EVALUATION DESIGN AND PROCESS



Empirical Methods

To answer the Primary Question about the overall impact of the activity, the RCT methodology provides the greatest potential for causal inference and minimizes potential bias from unmeasured confounding factors since access to the activity is not based on any external measure or individual characteristics that may correlate with the outcomes or participation in the activity. Since eligible participants were randomly assigned to the treatment and comparison group, both groups have, on average, the same characteristics and thus would be statistically identical in the absence of the activity. If both groups are identical, differing only in exposure to the activity, then any difference in outcomes at the end of the evaluation can be attributed to the activity. Similar to the first follow-up report, the evaluation team estimates the direct effect of the overall WLSME activity on several outcome variables, using analysis of covariance (ANCOVA) and difference-in-difference (DID) models. Since ANCOVA is the preferred estimation method and this report includes findings from the first and second follow-up rounds, the ANCOVA results are reported in the findings section and the DID results can be found in Annex D. The two rigorous methods are summarized below, and additional details can be found in Annex B.

- **Analysis of Covariance:** ANCOVA is a statistical method that takes advantage of the low autocorrelation of certain outcome variables in this study, such as business profits and sales, to improve power beyond what a DID approach can attain with the same sample size. Baseline data for these outcome measures have little predictive power for future outcomes, so it is inefficient to fully correct for baseline imbalances between treatment and control groups using DID. Instead, an ANCOVA model can adjust the degree of correction for baseline difference in means according to the degree of correlation between past and future outcomes actually observed in the data (McKenzie 2012). The ANCOVA specification used for estimations in this evaluation used information from the first and second follow-up surveys, as well as the baseline:

$$Y_{i,t} = \beta_0 + \beta_1 D_i + \beta_2 Y_{i,t-1} + \beta_3 Y_{i,t-2} + \delta X_i + \varepsilon_{i,t}$$

In this case, $Y_{i,t-2}$ is the first follow-up value of the outcome variable, $Y_{i,t-1}$ denotes the baseline value of the outcome variable, and β_1 is the ANCOVA treatment effect. X_i is a matrix of relevant covariates for identification to increase the efficiency of β_1. X_i contains the following variables: age, marital status, higher education, business ownership, number of full-time workers from the household and also non-family workers, participation in previous training or seminars, and number of children under 18. The evaluation team also included sector and region fixed effects.

- **Difference-in-Differences:** The DID estimator is one of the most popular methodologies for applied research in economics. To answer the hypotheses, DID estimates causal relationships among variables by comparing the difference in outcomes before and after an intervention between groups of beneficiaries and nonparticipants (Bertrand et al. 2004). The first "difference" in this method is the difference before the intervention (baseline) and after the intervention (endline). The second "difference" is between the beneficiary group (treatment) and nonparticipant group (control). Thus, two rounds of data are required. Within this framework and in order to estimate the impacts of the WLSME activity, for each outcome of interest the evaluation team employed the following DID specification:

  $$\Delta Y_i = \beta_o + \beta_1 D_i + \delta X_i + \varepsilon_i$$

  Here, ΔY_i is the first difference in the outcome variable between the baseline and the second round, D_i is the treatment status dummy that remains unchanged between periods, and β_1 is the treatment effect. The X_i covariates matrix is evaluated in the baseline period and contains the same variable that was used in the ANCOVA model. This specification proposed by Bernal and Peña (2014) has the advantage of applying an analogous approach to a RCT, as it focuses on the changes of one key variable of interest between periods. This is not the case for standard DID approaches, where period variations of the other explanatory variables also impact the outcome variable.

  The evaluation team employed a DID panel model specified as:

  $$Y_{i,t} = \lambda_t + c_i + \beta_1 D_{i,t} + \delta X_{i,t} + \varepsilon_{i,t}\,, \qquad t = 1, \ldots, T$$

  Where the same individuals are compared at baseline and each follow-up round, λ_t denotes the time fixed effects, c_i are the individual fixed effects, and X_i is a matrix of relevant covariates for each period and individual. β_1 is the treatment effect. In other words, this technique uses both follow-up surverys as well as the baseline information. The results for both DID specifications are very similar.

The Secondary Questions cannot be objectively answered from an attribution perspective, since access to the activity components was not randomized but participants instead either self-selected or the "most promising entrepreneurs" were selected. The evaluation team could argue that more motivated women, more ambitious entrepreneurs, or those who would have succeeded even without the activity are more likely to have participated in the activity components. Therefore, comparing women who participated across the sub-groups would systematically miscalculate the impact estimates of each activity component since changes in the outcome variables would not only reflect their participation in the components, but also the set of characteristics (e.g., ambition, motivation) that led the group to participate in the activity components in the first place. Given this non-randomized selection into the activity components, the DID approach is most useful to account for all observable baseline differences between the groups. Thus, the results presented in the Secondary Questions findings section only correspond to the DID model, not ANCOVA. However, these estimates cannot be concluded as impact or causality, but rather are merely a sign of association between the activity components and the outcome variables.

BALANCE AMONG TREATMENT AND CONTROL GROUPS

Balance across key demographic and outcome variables is necessary to show that the treatment and control groups are the same prior to the start of the intervention. The baseline report showed balance across the two groups except in two instances: number of years the business operated and whether the respondent borrowed credit for her business (from any source) in the past 12 months. Given changes to the sample due to exclusion of the agro-processing sector and non-responses to the follow-up survey, the evaluation team conducted balance tests again on the baseline sample. The team applied the student t test – also known as a t-test – for two independent samples with unequal variances. This test provided previewing evidence about the differences between the control and treatment groups before the intervention started. The evaluation team reported the tables below in the first follow-up report and is providing them again here.

The evaluation team chose 10 characteristics or dimensions to establish whether there were significant differences between treatment and control groups at the baseline round. If the p-value associated with the t-test is small (p-value <0.05), there is evidence to suggest that the average is different for both groups. Namely, the mean difference is significantly different from zero. On the contrary, when the p-value associated with the test is not small, it can be concluded that the means of both groups are not different.

Table 1 shows the tests of balance between the treatment group and control at the baseline. It is expected that participants in these two groups have similar characteristics at the stage previous to the intervention, since the treatment group was randomly selected at an individual level, but this provides evidence that the remaining sample (after the exclusion of the agro sector) was still balanced at baseline. As shown, the sample proves to be balanced for all variables except borrowed credit in the last 12 months. At baseline, the control group was more likely to have had a loan in the previous 12 months. No other variable shows statistically significant differences in the presented model. Thus, the data show that the control and treatment groups shared similar social and economic conditions before the WLSME activity took place, and therefore may be compared validly by experimental methods.

TABLE 1: BASELINE BALANCE TEST

	Control	Treatment	Difference	(p-value)
Age	44.27	44.79	-0.52	0.53
Married (=1)	0.79	0.77	0.01	0.64
Higher education (=1)	0.48	0.51	-0.03	0.42
Owner (=1)	0.53	0.51	0.02	0.65
Years of operation	7.27	7.00	0.27	0.55
Full-time workers from the household	0.98	0.91	0.07	0.47
Full-time non-family workers	3.68	3.89	-0.21	0.73
Previous trainings/seminars (=1)	0.34	0.37	-0.03	0.39
Number of children under 18	1.69	1.82	-0.13	0.23
Sells in external markets	0.22	0.24	-0.01	0.72
Borrowed credit for business in past 12 months	0.59	0.51	0.08	0.03**
Number of observations	568	275	-	-

Notes: Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1

ANALYSIS OF NON-RESPONSE AT FOLLOW-UP

Given the sample distribution between the baseline and two follow-up survey rounds, as shown in Table 2, the evaluation team examined whether non-response is random and not correlated with treatment assignment.

TABLE 2: SAMPLE DISTRIBUTION BY SURVEY ROUND

	Baseline	First Follow-up		Second Follow-up		All Rounds	
		N=	% of sample	N=	% of sample	N=	% of sample
Treatment Group	568	459	80.8%	472	83.1%	432	76.1%
Control Group	275	228	82.9%	232	84.4%	216	78.5%
Total	**843**	**687**	**81.5%**	**704**	**83.5%**	**648**	**76.9%**

There is no statistically significant difference (p-value=0.46, 0.64, and 0.42) in non-response between the treatment and control groups across each round, as shown in Table 3.

TABLE 3: NON-RESPONSE RATES WITH RESPECT TO BASELINE

	First Follow-up	Second Follow-up	All Rounds
Treatment Group	19.2%	16.9%	23.9%
Control Group	17.1%	15.6%	21.5%
Total	**18.5%**	**16.5%**	**23.1%**
p-value	*0.46*	*0.64*	*0.42*

The evaluation team also examined whether non-response depends on observable characteristics. The team regressed follow-up survey completion on the same set of baseline variables, treatment status, and interaction terms of those baseline variables with the treatment variable across each round. Looking at the F-test on the interaction variable coefficients for the first follow-up round (p-value=0.17), second follow-up round (p-value=0.38), and both rounds combined (p-value=0.60), the evaluation team does not find differences in the observable composition of the treatment versus control groups despite differential non-response rates. Thus, although the non-response rate for all rounds is 23 percent, it does not affect the validity of the findings presented below.

PRIMARY QUESTION FINDINGS – OVERALL IMPACT

This section presents the short-term (first follow-up at the end of the activity) and medium-term (second follow-up at 12 months after the end of the activity) estimates, using the ANCOVA model, of the overall Kyrgyzstan WLSME activity on the key outcome measures.[4] Findings from the DID models can be found in Annex D. The estimations below can be interpreted as intent to treat, which represents the average effect of having access to the activity (i.e., all of the women assigned to the treatment group remain part of the treatment group, regardless of actual participation).

[4] In order to test for robustness, the evaluation team also computed Bonferroni-type corrections in all empirical results. The number of hypothesis (m) considered for this correction depends on the number of tests applied for each outcome variable within any specific category. In particular, the number of hypothesis considered are as follows: business growth outcomes (m=10), entrepreneurial leadership category (m=27), networks (m=7), and business knowledge and practices outcomes (m=22). The corrected p-values at 90 percent are 0.0010, 0.004, 0.014, and 0.005, respectively. The corrected p-values at 95 percent are 0.005, 0.002, 0.007, and 0.002, respectively. While the key findings in this report do not change, some of the outcome variables lose statistical significance in most categories. Given the fact that Bonferroni tends to give false negatives and requires high power – a structural weakness of the sample – the evaluation team would not want to put excessive emphasis on these additional results at this point.

Business Growth

Business growth outcomes includes a set of variables: related sales, profit, business cycle, time spent working in the business, number and type of paid employees, investments, and loans. As shown in Table 4, most of the outcome variables are not statistically significant. However, most of the signs of the coefficient are consistent with an improvement.

Self-reported average sales and average profit show an improvement, regardless of month. Coefficients that had inconsistent signs in the short term, including number of good months, number of bad months, and sales in the last 12 months, have been corrected; however, none of them are statistically significant. The lack of statistical significance can be due to a combination of low statistical power and difficulty in measurement in the sales variable; the latter is frequently reported in related literature (McKenzie 2012).

In the short term, the evaluation found an 11 percentage point increase in the likelihood of investment in capital inputs in the last 12 months. However, the evaluation no longer detected this capital effect in the medium term. In the second follow-up round, the evaluation found an employment effect that had not been captured in the first round. There is a statistically significant increase of 0.32 paid household workers. This shows an increase in the labor demand and a better contracting situation, since all other indicators of the number of workers reveal a positive relationship (although not significant). In contrast, the estimation reveals a negative impact in the likelihood of applying for a loan that can be counterintuitive in relation to the activity's objectives or could be interpreted as a decreased need for liquidity due to business improvement

These results seem to indicate that capital investments were done in the short term as an immediate way to grow the business, and that employment gains only begin in the medium term. However, it is uncertain whether this increased remuneration of household members is due to greater value placed on human capital resources, increased perception of growth potential, or leading to future increased sales and profits.

TABLE 4: BUSINESS GROWTH OUTCOMES

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Average sales on a good month (in ranges)+	0.040	0.168
		(0.137)	(0.097)
2	Average sales on a bad month (in ranges)+	-0.069	0.060
		(0.110)	(0.129)
3	Average sales on an average month (in ranges)+	-0.039	0.154
		(0.117)	(0.111)
4	Average profit on a good month (in ranges)+	0.370	0.101
		(0.228)	(0.142)
5	Average profit on a bad month (in ranges)+	0.044	0.068
		(0.095)	(0.090)
6	Average profit on an average month (in ranges)+	0.081	0.052
		(0.156)	(0.122)
7	Number of good months in the last year	-0.079	0.074
		(0.204)	(0.103)
8	Number of bad months in the last year	0.191	-0.060
		(0.119)	(0.082)
9	Sales in the last 12 months (Soms)	-6,919	13,583

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
		(16,731)	(23,713)
10	Likelihood of having a bank account for business only	-0.038	-0.028
		(0.032)	(0.036)
11	Number of household members who have worked in business in the last 12 months	0.0342	0.010
		(0.105)	(0.040)
12	Number of paid household workers	0.0921	0.319**
		(0.165)	(0.122)
13	Number of non-household people who have worked in business in the last 12 months	-0.297	0.006
		(0.586)	(0.153)
14	Number of paid non-household workers	-0.102	-0.109
		(0.588)	(0.257)
15	Number of months per year spent working in the business owned or managed	0.033	-0.282
		(0.215)	(0.233)
16	Number of days per week spent working in the business owned or managed	0.179	-0.104
		(0.111)	(0.092)
17	Number of hours per day spent working in the business owned or managed	0.148	0.135
		(0.217)	(0.290)
18	Likelihood of purchasing raw materials, goods, or equipment for business with a loan in the last 12 months	0.110**	-0.031
		(0.050)	(0.049)
19	Likelihood of applying for a loan from a financial institution in the last 12 months	0.014	-0.026
		(0.039)	(0.040)
20	Likelihood of loan approval	-0.026	-0.095
		(0.087)	(0.092)

Note: Coefficients were obtained by ANCOVA regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.

Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

[+] These outcomes are ordinal variables with the following categories: 1) None; 2) Less than 5,000 Soms; 3) 5,001 – 10,000; 4) 10,001 – 20,000; 5) 20,001 – 40,000; 6) 40,001 – 60,001; 7) 60,000 – 80,000; 8) 80,001 – 100,000; 9) 100,001 – 150,000; 10) 150,001 – 200,000; 11) 200,001 – 500,000; 12) More than 500,000. The coefficient is a measure of increase towards the next higher category. While interpretation of this coefficient (as an ordinal variable) is not straightforward, the lack of statistical significance does not change if a multinomial logistic regression is used.

Entrepreneurial Leadership

The second key outcome category, entrepreneurial leadership, includes variables related to decision-making in the business, entrepreneurial vocation, level of independence, and women's empowerment, as shown in Table 5. Similar to the set of business growth outcomes, most variables are not statistically significant at conventional levels.

Unlike the previous set, not all corresponding signs of the coefficients are consistent with the expected behavioral change. In particular, the direction of the effect for the role in decision-making (rows 3 to 10) are negative for half of these outcomes. In other words, women in the treatment group are less likely to participate in decision-making of different aspects of the business. However, these changes are not statistically significant.

The outcomes related to decisions taken without consulting anyone else (rows 11 to 18) also show mixed results, as some of the coefficients show negative signs. In other words, women in the treatment group are more likely to ask for advice in making general business planning, deciding what type of work to do, and staffing of business. However, none of these changes are statistically significant. Moreover, the short-term effect of the increased likelihood of managing sales and client relationships independently is no longer statistically significant in the medium term.

The empowerment outcomes (rows 19 to 27) also show mixed results. First, the short-term statistically significant effects have been lost. Moreover, the unexpected signs in some of the coefficients, such as less agreement with the following statements: "the role of women is to earn money and take care of her family," "a mother who works can establish a relationship as warm and solid with her children as a mother who does not work," and "father's and mother's dedication is equally important for the learning and effective development of children," depict a context where external constraints and gender roles may be at play and hard to change. However, this is contradicted by the statistically significant increase of 6 percentage points in the agreement that there are no gender inequalities in the community.

The lack of consistent and statistically significant findings for this set of outcomes may be due to a lack of statistical power or to the inelasticity of decision-making power and gender dynamics.

TABLE 5: ENTREPRENEURIAL LEADERSHIP OUTCOMES

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that prefers to work as an employee in a business instead of managing/owning one	-0.017 (0.0273)	0.006 (0.037)
2	Likelihood that people ask me for business advice (very) often	-0.043 (0.039)	-0.007 (0.019)
3	Likelihood that I (with my partner/spouse or another household member) am in charge of general business planning decisions	-0.007 (0.032)	-0.006 (0.047)
4	Likelihood that I (with my partner/spouse or another household member) decide what inputs to buy for production	0.027 (0.036)	0.007 (0.049)
5	Likelihood that I (with my partner/spouse or another household member) am in charge of sales and client relations	-0.024 (0.033)	0.018 (0.037)
6	Likelihood that I (with my partner/spouse or another household member) decide if I should apply for a loan	-0.019 (0.038)	-0.035 (0.037)
7	Likelihood that I (with my partner/spouse or another household member) decide my own singular wage	0.005 (0.034)	0.050 (0.036)
8	Likelihood that I (with my partner/spouse or another household member) decide what type of work I will do	0.021 (0.032)	0.006 (0.035)
9	Likelihood that I (with my partner/spouse or another household member) am in charge of marketing and advertising decisions	0.031 (0.037)	-0.039 (0.060)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
10	Likelihood that I (with my partner/spouse or another household member) am in charge of staffing of business decisions	0.045 (0.036)	-0.040 (0.033)
11	Likelihood that often (or always) makes general business planning decisions without consulting anyone else	0.026 (0.038)	-0.046 (0.041)
12	Likelihood that often (or always) decides what inputs to buy for production without consulting anyone else	0.033 (0.039)	0.026 (0.076)
13	Likelihood that often (or always) manages sales and client relations without consulting anyone else	0.072* (0.039)	0.015 (0.030)
14	Likelihood that often (or always) decides whether to apply for a loan without consulting anyone else	-0.011 (0.044)	0.051 (0.031)
15	Likelihood that often (or always) decides own singular wage without consulting anyone else	0.026 (0.043)	0.042 (0.032)
16	Likelihood that often (or always) decides what type of work to do without consulting anyone else	0.052 (0.039)	-0.029 (0.033)
17	Likelihood that often (or always) makes marketing and advertising decisions without consulting anyone else	0.054 (0.044)	0.014 (0.051)
18	Likelihood that often (or always) makes staffing of business decisions without consulting anyone else	0.036 (0.041)	-0.019 (0.050)
19	Likelihood that (strongly) agrees that "women should do what men say"	-0.028 (0.036)	-0.032 (0.047)
20	Likelihood that (strongly) agrees that "women must share their income with their husbands"	0.014 (0.030)	-0.008 (0.024)
21	Likelihood that (strongly) agrees that "it is OK if men abandon women if they wish to"	-0.037 (0.030)	0.016 (0.027)
22	Likelihood that (strongly) agrees that "it is OK if men chide women because they went out without any permission"	-0.065* (0.036)	-0.017 (0.028)
23	Likelihood that (strongly) agrees that "it is OK if men chide women if they do not take care of children"	0.038 (0.028)	-0.015 (0.024)
24	Likelihood that (strongly) agrees that "the role of women is to earn money and take care of her family"	-0.065* (0.038)	-0.019 (0.034)
25	Likelihood that (strongly) agrees that "a mother who works can establish a relationship as warm and solid with her children as a mother who does not work"	0.017 (0.018)	-0.019 (0.024)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
26	Likelihood that (strongly) agrees that "father's and mother's dedication is equally important for the learning and effective development of children"	0.001 (0.014)	-0.013 (0.016)
27	Likelihood that (strongly) agrees that there are no gender inequality problems in my community	-0.014 (0.027)	0.059** (0.020)

Note: Coefficients were obtained by ANCOVA regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

Networks

For the networks outcomes, the short-term effects of increased implementation of professional advice and number of people to ask for business advice are no longer detected in the medium term. In the medium term, there are no statistically significant effects, except for a 10 percentage point reduction in the likelihood of being an active member of any social group, which is an unexpected effect with respect to expanding networks. There are also counterintuitive, but not statistically significant, negative effects in confidence negotiating with suppliers and buyers.

TABLE 6: NETWORKS OUTCOMES

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that there are social groups in my community	0.026 (0.037)	0.045 (0.032)
2	Likelihood of being an active member of any social group	0.031 (0.053)	-0.102* (0.045)
3	Likelihood of participation in Trade Shows or Fairs	0.032 (0.037)	-0.054 (0.034)
4	Likelihood of implementation of professional advice (business adviser, lawyer or accountant) during the past two years	0.106*** (0.034)	-0.003 (0.041)
5	Number of people I can go to ask business advice	0.980** (0.393)	0.604 (0.349)
6	Likelihood that feels sometimes (very) confident negotiating lower prices with suppliers	0.027 (0.033)	-0.035 (0.053)
7	Likelihood that feels sometimes (very) confident negotiating higher prices with buyers	-0.030 (0.030)	-0.017 (0.035)

Note: Coefficients were obtained by ANCOVA regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

Business Knowledge and Practices

Table 7 shows the evaluation team's findings with respect to business knowledge and practices outcomes, which include implementation of marketing, operations, and accounting practices; negotiation skills; recording of budget; and future expectations, among other messages provided by the training delivered.

In the second follow-up, only one outcome variable is statistically significant. There is an 11.2 percentage point increase in recording the owner's salary in a registry or computer. The estimates on the other outcome variables are not statistically significant and also very small in magnitude. It is difficult to explain the lack of statistical significance in these variables, but women may not be fully convinced of the usefulness of some of the practices taught or investing in them may require more assertiveness and effort.

TABLE 7: BUSINESS KNOWLEDGE AND PRACTICES OUTCOMES

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that no marketing activities implemented in the last 3 years	-0.019 (0.020)	0.019 (0.038)
2	Likelihood that does not use internet for marketing purposes or to sell products/services	-0.050 (0.033)	0.014 (0.030)
3	Likelihood that does not formally keep track of business' products and materials	-0.007 (0.024)	0.012 (0.018)
4	Likelihood that does not perform a physical validation of inventory levels	-0.003 (0.031)	-0.035 (0.036)
5	Likelihood that business runs out of inventory at least one time a month	-0.004 (0.039)	-0.034 (0.022)
6	Likelihood that tried to negotiate a lower price with suppliers during the last three months	-0.023 (0.036)	0.002 (0.047)
7	Likelihood that does compared price and quality of inputs with other suppliers' products during the last three months	-0.050* (0.029)	0.017 (0.043)
8	Likelihood that fixed salary for the owner	0.012 (0.035)	0.045 (0.026)
9	Likelihood that records salary of the owner in a notebook, registry or computer	0.018 (0.056)	0.112* (0.052)
10	Likelihood that does not keep track of business purchases and sales	-0.014 (0.024)	0.003 (0.020)
11	Likelihood that has a written expense budget	0.013 (0.035)	0.050 (0.031)
12	Likelihood that has no written goals for next 12 months	-0.054** (0.026)	0.034 (0.025)
13	Likelihood that has no accountancy documents prepared annually	0.006 (0.037)	-0.042 (0.031)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
14	Likelihood that has no changes planned over the next 12 months	-0.005 (0.021)	0.036 (0.030)
15	Likelihood that (strongly) agrees that "my workspace is well organized"	0.034 (0.031)	0.030 (0.025)
16	Likelihood that (strongly) agrees that "I often communicate clear objectives to my colleagues and employees"	-0.008 (0.012)	0.009 (0.014)
17	Likelihood that (strongly) agrees that "I develop work plans at regular intervals"	0.018 (0.025)	0.032 (0.036)
18	Likelihood that (strongly) agrees that "I sometimes miss deadlines"	-0.013 (0.040)	0.014 (0.050)
19	Likelihood that (strongly) agrees that "I believe employees should be treated like family"	-0.009 (0.026)	-0.006 (0.031)
20	Likelihood that (strongly) agrees that "I am sometimes late for appointments or meetings"	-0.016 (0.038)	-0.028 (0.050)
21	Likelihood that (strongly) agrees that "I often attempt to anticipate future circumstances and plan how I/my company will deal with them"	-0.009 (0.017)	-0.021 (0.012)
22	Likelihood that (strongly) agrees that "I am constantly collecting information about the market in which my company operates"	0.002 (0.019)	-0.018 (0.017)

Note: Coefficients were obtained by ANCOVA regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

SECONDARY QUESTIONS FINDINGS – COMPONENT ANALYSIS

In the previous sections, this report addressed the overall impact from participation in the WLSME activity. Next, it presents an analogous analysis that compares the correlation of two activity components (ML and TS/AF) relative to the other (BMT). The group of analysis consists of women who engaged in BMT only, which the evaluation team compared to Component (A), BMT + ML, and Component (B), BMT + ML + TS/AF. The number of participants in these three groups is listed in Table 8.

TABLE 8: TREATMENT SAMPLE DISTRIBUTION BY ACTIVITY COMPONENT

	Component Comparison Group		Component A		Component B	
	Only BMT		BMT + ML		BMT + ML+TS/AF	
	N	%	N	%	N	%
Participant	348	75.8	211	60.6	117	33.6
Non-participant	111	24.2	137	39.4	231	66.4
Total	459	100.0	348	100.0	348	100.0

Note: The evaluation sample for this analysis excludes the following individuals: (1) activity participants in the agriculture sector, and (2) activity participants who were assigned to receive a specific component but rejected to participate in the second follow-up survey, in spite of being registered in the baseline survey.

Since selection into the activity components was not randomized, it is not possible to look at a *causal* treatment effect within the intervention, but only at correlations. This is because participants who self-selected into participating in ML or TS/AF are different in ways that are correlated to the outcomes. Given the non-randomized selection into the activity components, the DID approach is most useful to account for all observable baseline differences between the groups. Thus, the results presented below to estimate the links of the activity's components on various outcomes only correspond to the DID model, not ANCOVA. An additional statistical method proposed in the Evaluation Design Proposal, propensity score matching, was not included in this report due to the small sample size.

In the tables below, the estimates for Component A correspond to the added value of the ML component relative to the BMT only. The estimates for Component B corresponds to the added value of the ML+TS/AF components relative to the BMT only.

Summary of Findings

The evaluation team presents key findings for the Secondary Questions below, followed by detailed explanations and tables.

- For **business growth** outcomes, both Components A and B increase the likelihood of having a bank account for the business and reduce the number of days per week spent working in the own business. Component B increases the number of bad months, increases the number of non-household people who worked in the business, and increases the number of paid non-household workers.
- For **entrepreneurial leadership** outcomes, Components A and B are consistent. Both components are associated with an increase in the likelihood of being in charge of the marketing decisions. Component A reveals a counterintuitive association of preferring to work as an employee instead of managing a business and a decrease in agreeing with the affirmation "a mother who works can establish a relationship as warm with her children as a mother who does not work." There is also a reduction in the likelihood of agreeing that there are no gender inequality problems in the community. On the other hand, Component B reveals a positive association on the likelihood of people asking participants for business advice, while the probability of being in charge of sales, deciding their own wage, deciding what type of work to do, and managing sales and client relations are negatively affected.
- For **networks** outcomes, only Component B shows statistically significant associations, corresponding to a 17 percentage point increase in implementing professional advice and an increase of 1.4 people available for asking business advice. Although not statistically significant, Component A also had a positive association with both of these variables.

- For **business knowledge and practices** outcomes, both Components A and B are associated with a higher likelihood of preparing accountancy documents annually and agreeing on missing some deadlines. Component A shows a reduction in the likelihood of not performing physical validation of inventories, while Component B shows an increased likelihood of keeping a written expense budget and a reduction in the likelihood of believing employees should be treated like family.

Business Growth

The evaluation team finds that most of the business growth outcome variables are not statistically significant, as shown in Table 9. Nonetheless, in most cases the signs of the coefficients point in the direction consistent with treatment. Interestingly, five variables do yield statistically significant associations. For both Components A and B, the estimation showed an increase in the likelihood of having a bank account for business and a decrease in the number of days per week spent working in their own business. Component B increased the number of bad months, increased the number of non-household members that worked in the business, and increased the number of paid non-household workers. The labor gains from Component B show that these women are formally growing their workforce beyond their household help.

TABLE 9: BUSINESS GROWTH – BY ACTIVITY COMPONENT

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
1	Average sales on a good month (in ranges)+	0.187 (0.294)	-0.100 (0.351)	-0.095 (0.280)	-0.039 (0.169)
2	Average sales on a bad month (in ranges)+	0.130 (0.294)	0.001 (0.351)	0.105 (0.198)	0.085 (0.276)
3	Average sales on an average month (in ranges)+	0.187 (0.247)	0.093 (0.294)	0.096 (0.241)	0.195 (0.181)
4	Average profit on a good month (in ranges)+	0.303 (0.254)	0.157 (0.316)	0.070 (0.124)	0.114 (0.146)
5	Average profit on a bad month (in ranges)+	0.252 (0.179)	0.185 (0.225)	0.202 (0.136)	0.143 (0.181)
6	Average profit on an average month (in ranges)+	0.230 (0.205)	0.119 (0.250)	0.084 (0.145)	0.099 (0.131)
7	Number of good months in the last year	0.181 (0.365)	-0.223 (0.378)	-0.047 (0.220)	-0.327 (0.244)
8	Number of bad months in the last year	-0.147 (0.311)	0.495 (0.317)	-0.089 (0.512)	0.836*** (0.223)
9	Sales in the last 12 months (Soms)	-37,192 (59,839)	-31,837 (89,127)	-1,384 (45,440)	-3,608 (51,940)
10	Likelihood of having a bank account for business purposes only	-0.159** (0.063)	-0.280*** (0.072)	0.104* (0.054)	0.197** (0.067)

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
11	Number of household members that have worked in business in past 12 months	0.168 (0.186)	-0.062 (0.189)	0.168 (0.091)	0.101 (0.097)
12	Number of paid household members	-0.058 (0.259)	0.151 (0.252)	-0.039 (0.290)	0.421 (0.282)
13	Number of non-household people that have worked in business in past 12 months	1.044 (0.721)	1.322 (0.858)	0.642 (0.563)	1.632*** (0.354)
14	Number of paid non-household workers	0.770 (0.826)	1.058 (0.948)	0.383 (0.475)	1.422** (0.408)
15	Number of months per year spent on the business owned or managed	-0.386 (0.495)	0.326 (0.498)	-0.644 (0.457)	0.052 (0.464)
16	Number of days per week spent on the business owned or managed	0.024 (0.206)	0.010 (0.195)	-0.415** (0.146)	-0.290* (0.125)
17	Number of hours per day spent on the business owned or managed	0.668 (0.456)	0.474 (0.488)	0.254 (0.506)	0.169 (0.427)
18	Likelihood of purchasing raw materials, goods, or equipment for business with a loan in the last 12 months	-0.008 (0.099)	-0.051 (0.098)	-0.096 (0.063)	0.090 (0.105)
19	Likelihood of applying for a loan from a financial institution in the last 12 months	-0.021 (0.084)	-0.101 (0.087)	-0.005 (0.099)	0.092 (0.066)
20	Likelihood of loan approval	0.155* (0.093)	0.025 (0.086)	-0.096 (0.072)	0.007 (0.074)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.
[+] These outcomes are ordinal variables with the following categories: 1) None; 2) Less than 5,000 Soms; 3) 5,001 – 10,000; 4) 10,001 – 20,000; 5) 20,001 – 40,000; 6) 40,001 – 60,001; 7) 60,000 – 80,000; 8) 80,001 – 100,000; 9) 100,001 – 150,000; 10) 150,001 – 200,000; 11) 200,001 – 500,000; 12) More than 500,000. The coefficient is a measure of increase towards the next higher category. While interpretation of this coefficient (as an ordinal variable) is not straightforward, the lack of statistical significance does not change if a multinomial logistic regression is used.

Entrepreneurial Leadership

Table 10 shows findings for entrepreneurial leadership outcomes by activity component. The findings in this module are inconsistent with some outcome variables showing statistically significant associations in the unexpected direction. For instance, both components are associated with a preference to work as an employee in a business instead of managing/owning one; this result is statistically significant for Component A. Both components are associated with an increase in the likelihood of being in charge of the marketing decisions.

Component A shows a statistically significant reduction in agreeing that there are no gender inequality problems in the community. This component also shows a counterintuitive reduction in agreeing with

the affirmation "a mother who works can establish a relationship as warm with her children as a mother who does not work." On the other hand, Component B shows a positive association with the likelihood of being asked for business advice, while the probability of being in charge of sales, deciding one's own wage, deciding what type of work to do, and managing sales and client relations are negatively affected.

TABLE 10: ENTREPRENEURIAL LEADERSHIP – BY ACTIVITY COMPONENT

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
1	Likelihood that prefers to work as an employee in a business instead of managing/owning one	0.272*** (0.066)	0.129** (0.064)	0.273** (0.082)	0.120 (0.096)
2	Likelihood that people ask me for business advice (very) often	-0.075 (0.084)	0.173** (0.087)	-0.065 (0.068)	0.181** (0.059)
3	Likelihood that I (with my partner/spouse or another household member) am in charge of general business planning decisions	0.031 (0.073)	0.003 (0.075)	0.002 (0.044)	0.010 (0.052)
4	Likelihood that I (with my partner/spouse or another household member) decide what inputs to buy for production	0.010 (0.079)	-0.114 (0.077)	0.000 (0.038)	-0.071 (0.058)
5	Likelihood that I (with my partner/spouse or another household member) am in charge of sales and client relations	-0.018 (0.075)	-0.123* (0.074)	-0.037 (0.054)	-0.094* (0.041)
6	Likelihood that I (with my partner/spouse or another household member) decide if I should apply for a loan	-0.021 (0.080)	-0.100 (0.081)	-0.075 (0.061)	-0.048 (0.059)
7	Likelihood that I (with my partner/spouse or another household member) decide my own singular wage	0.153 (0.341)	0.097 (0.232)	-0.100 (0.057)	-0.164** (0.062)
8	Likelihood that I (with my partner/spouse or another household member) decide what type of work I will do	0.027 (0.073)	-0.085 (0.071)	0.015 (0.040)	-0.095* (0.045)
9	Likelihood that I (with my partner/spouse or another household member) am in charge of marketing and advertising decisions	0.093 (0.077)	0.049 (0.076)	0.123** (0.045)	0.150** (0.056)
10	Likelihood that I (with my partner/spouse or another household member) am in charge of staffing of business decisions	0.065 (0.074)	-0.047 (0.074)	0.028 (0.046)	0.001 (0.036)
11	Likelihood that often (or always) makes general business planning decisions without consulting anyone else	-0.040 (0.081)	0.016 (0.081)	-0.047 (0.075)	-0.045 (0.095)

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
12	Likelihood that often (or always) decides what inputs to buy for production without consulting anyone else	0.024 (0.082)	-0.064 (0.083)	0.047 (0.056)	-0.079 (0.071)
13	Likelihood that often (or always) manages sales and client relations without consulting anyone else	-0.098 (0.079)	-0.176** (0.080)	-0.078 (0.098)	-0.177** (0.053)
14	Likelihood that often (or always) decides whether to apply for a loan without consulting anyone else	-0.118 (0.083)	-0.143* (0.085)	-0.069 (0.078)	-0.126 (0.132)
15	Likelihood that often (or always) decides own singular wage without consulting anyone else	0.134 (0.090)	0.043 (0.092)	0.139 (0.109)	0.031 (0.095)
16	Likelihood that often (or always) decides what type of work will do without consulting anyone else	-0.124 (0.084)	-0.061 (0.085)	-0.112 (0.088)	-0.080 (0.046)
17	Likelihood that often (or always) makes marketing and advertising decisions without consulting anyone else	-0.002 (0.089)	-0.064 (0.088)	0.060 (0.117)	-0.066 (0.137)
18	Likelihood that often (or always) makes staffing of business decisions without consulting anyone else	0.024 (0.084)	-0.046 (0.082)	0.054 (0.068)	-0.034 (0.117)
19	Likelihood that (strongly) agrees that "women should do what men say"	-0.054 (0.071)	0.047 (0.076)	-0.067 (0.092)	0.025 (0.082)
20	Likelihood that (strongly) agrees that "Women must share their income with their husbands"	-0.037 (0.062)	-0.103 (0.066)	0.007 (0.046)	-0.017 (0.109)
21	Likelihood that (strongly) agrees that "it is OK if men abandon women if they wish to"	0.130** (0.060)	-0.017 (0.061)	0.113 (0.060)	0.006 (0.058)
22	Likelihood that (strongly) agrees that "it is OK if men chide women because they went out without any permission"	-0.028 (0.078)	0.080 (0.081)	0.039 (0.063)	0.057 (0.059)
23	Likelihood that (strongly) agrees that "it is OK if men chide women if they do not take care of children"	-0.108** (0.054)	-0.092 (0.062)	-0.060 (0.041)	-0.064 (0.054)
24	Likelihood that (strongly) agrees that "the role of women is to earn money and take care of her family"	-0.051 (0.080)	0.038 (0.084)	-0.059 (0.067)	0.067 (0.060)
25	Likelihood that (strongly) agrees that "a mother who works can establish a relationship as warm and solid with her children as a mother who	-0.075** (0.036)	-0.041 (0.037)	-0.057* (0.026)	-0.045 (0.056)

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
	does not work"				
26	Likelihood that (strongly) agrees that "father's and mother's dedication is equally important for the learning and effective development of children"	-0.045* (0.025)	-0.035 (0.029)	-0.019 (0.027)	-0.004 (0.027)
27	Likelihood that (strongly) agrees that there are no gender inequality problems in my community	-0.045* (0.025)	-0.035 (0.029)	-0.178** (0.060)	-0.102 (0.092)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

Networks

Table 11 shows the network outcomes by activity components. The results in this module are more encouraging than in the case of entrepreneurial leadership and business growth outcomes for Component B. This may be because the behavioral effort to produce change may be relatively less taxing than in the other two cases. Overall, this is also consistent with the evaluation's more general results. The evaluation team finds few statistically significant associations present for Component B only.

The statistically significant associations correspond to the likelihood of implementation of professional advice, which increased by 17 percentage points, and to the number of people available to ask for business advice, which increased by approximately 1.4 people. Although not significant, Component A also had the same positive association with both variables. However, there is a counterintuitive association, albeit not statistically significant, for both components regarding a decrease in the feeling of confidence while negotiating with suppliers and buyers.

TABLE 11: NETWORKS – BY ACTIVITY COMPONENT

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
1	Likelihood that there are social groups in my community	-0.097 (0.083)	-0.042 (0.085)	-0.126 (0.112)	-0.067 (0.134)
2	Likelihood of being an active member of any social group	-0.090 (0.059)	-0.028 (0.067)	-0.008 (0.095)	0.019 (0.069)
3	Likelihood of participation in Trade Shows or Fairs	0.300*** (0.084)	0.395*** (0.086)	-0.017 (0.069)	0.040 (0.056)
4	Likelihood of implementation of professional advice (business adviser, lawyer or accountant) during the past two years	0.109 (0.072)	0.179** (0.074)	0.051 (0.070)	0.170* (0.075)

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A	Comp. B	Comp. A	Comp. B
		(std. error)	(std. error)	(std. error)	(std. error)
5	Number of people I can go to ask business advice	0.273 (0.791)	0.382 (1.007)	1.115 (1.039)	1.394** (0.552)
6	Likelihood that feels sometimes (very) confident negotiating lower prices with suppliers	-0.076 (0.074)	0.006 (0.079)	-0.060 (0.063)	-0.004 (0.113)
7	Likelihood that feels sometimes (very) confident negotiating higher prices with buyers	-0.122* (0.071)	-0.003 (0.076)	-0.093 (0.080)	-0.011 (0.082)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

Business Knowledge and Practices

Table 12 presents business knowledge and practices outcomes by activity component. Both components have a sustained reduction, between first and second follow-up rounds, in the likelihood that no accountancy documents are prepared annually. While both components also show an increase in the likelihood of sometimes missing deadlines, the magnitude of this association has decreased substantially from the first round, especially for Component B. Component A continues to show a reduction in the likelihood of not performing physical validation of inventories. Component B shows a large increase in the likelihood of keeping a written expense budget and a reduction in the likelihood of believing that employees should be treated like family.

TABLE 12: BUSINESS KNOWLEDGE AND PRACTICES – BY ACTIVITY COMPONENT

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A	Comp. B	Comp. A	Comp. B
		(std. error)	(std. error)	(std. error)	(std. error)
1	Likelihood that no marketing activities implemented in the last 3 years	0.016 (0.021)	0.019 (0.032	0.009 (0.023)	-0.033 (0.020)
2	Likelihood that does not use internet for marketing purposes or to sell products/services	0.039 (0.053)	-0.066 (0.049)	0.046 (0.052)	-0.040 (0.046)
3	Likelihood that does not formally keep track of business' products and materials	0.052 (0.042)	-0.057 (0.058)	0.055 (0.053)	-0.022 (0.060)
4	Likelihood that does not perform a physical validation of inventory levels	-0.122** (0.049)	-0.116 (0.090)	-0.093* (0.041)	-0.149 (0.091)
5	Likelihood that business runs out of inventory at least one time a month	0.006 (0.074)	0.039 (0.115)	0.029 (0.064)	0.055 (0.092)
6	Likelihood that tries to negotiate a lower price with suppliers during the last three months	0.018 (0.112)	0.039 (0.080)	0.083 (0.084)	0.127 (0.088)

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
7	Likelihood that compares price and quality of inputs with other suppliers' products during the last three months	0.015 (0.083)	0.004 (0.041)	0.043 (0.077)	0.064 (0.046)
8	Likelihood that has fixed salary for the owner	-0.047 (0.098)	0.046 (0.116)	-0.035 (0.075)	0.085 (0.093)
9	Likelihood that records salary of the owner in a notebook, registry or computer	-0.062 (0.131)	-0.057 (0.102)	-0.043 (0.102)	-0.013 (0.107)
10	Likelihood that does not keep track of business purchases and sales	-0.031 (0.030)	-0.078 (0.044)	-0.026 (0.044)	-0.045 (0.059)
11	Likelihood that has a written expense budget	-0.047 (0.050)	-0.047 (0.050)	-0.044 (0.050)	0.219*** (0.039)
12	Likelihood that has no written goals for next 12 months	0.004 (0.030)	0.004 (0.049)	-0.004 (0.033)	-0.049 (0.043)
13	Likelihood that has no accountancy documents prepared annually	-0.228** (0.088)	-0.284** (0.098)	-0.218* (0.103)	-0.250** (0.090)
14	Likelihood that has no changes planned over the next 12 months	-0.014 (0.025)	-0.074* (0.032)	-0.038 (0.044)	-0.104 (0.055)
15	Likelihood that (strongly) agrees that "my workspace is well organized"	0.055 (0.050)	-0.002 (0.049)	0.064 (0.049)	0.032 (0.059)
16	Likelihood that (strongly) agrees that "I often communicate clear objectives to my colleagues and employees"	0.000 (0.027)	-0.007 (0.022)	-0.016 (0.033)	-0.020 (0.044)
17	Likelihood that (strongly) agrees that "I develop work plans at regular intervals"	0.029 (0.053)	-0.052 (0.051)	0.030 (0.040)	0.001 (0.034)
18	Likelihood that (strongly) agrees that "I sometimes miss deadlines"	0.158* (0.071)	0.202** (0.069)	0.110* (0.051)	0.049* (0.025)
19	Likelihood that (strongly) agrees that "I believe employees should be treated like family"	-0.068 (0.040)	-0.070** (0.023)	-0.060 (0.056)	-0.059* (0.026)
20	Likelihood that (strongly) agrees that "I am sometimes late for appointments or meetings"	0.129 (0.113)	0.128 (0.177)	0.092 (0.081)	0.056 (0.120)
21	Likelihood that (strongly) agrees that "I often attempt to anticipate future circumstances and plan how I/my company will deal with them"	0.011 (0.047)	-0.018 (0.056)	0.004 (0.028)	0.001 (0.018)

Row	Outcome Variable	Follow-up 1		Follow-up 2	
		Comp. A (std. error)	Comp. B (std. error)	Comp. A (std. error)	Comp. B (std. error)
22	Likelihood that (strongly) agrees that "I am constantly collecting information about the market in which my company operates"	0.008 (0.036)	0.020 (0.039)	0.000 (0.046)	0.035 (0.022)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Robust standard errors in parentheses. Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

CONCLUSIONS

This report presents findings regarding the short- and medium-term impacts of the Kyrgyzstan WLSME activity. The impact evaluation expects to collect one more follow-up round of data at 24 months' post-intervention, to examine longer-term effects of the activity. Whereas the majority of outcomes considered do not yield a statistically significant link to the activity, there are rather promising causal links for a small but encouraging number of outcome variables. In particular, the activity appears to have short-term capital investment effects and medium-term employment effects. While self-reported sales and profits are not statistically significant, they are all increasing in magnitude. The evaluation detected the following medium-term effects:

- Increased number of paid household members;
- Increased perception that there are no gender inequality problems in the women's community;
- Decreased likelihood of being an active member of any social group; and
- Increased recording of owner's salary in a registry or computer.

When comparing the second follow-up findings with those from the first follow-up, there is little consistency with the outcome variables that had shown significant impact when the activity had just ended. The evaluation found that the following short-term effects are not sustained into the medium term:

- Investment of capital inputs;
- Likelihood of managing sales and client relationships independently;
- Likelihood of having written business goals;
- Disagreement that it is okay for men to chide women when they go out without permission; and
- Business support networks, with respect to the number of people participants can ask for advice and participants' likelihood of implementing professional advice.

As in the case of the first follow-up report, the somewhat limited statistical significance in several of the outcome variables may still be related to the temporal aspect of changing behavior, as it takes more time to generate change in tasks that require more assertiveness or effort than in others. However, the fact that the evaluation did not find full consistency between the first and second follow-ups may indicate statistical power issues or difficulty in measuring some of these outcomes. Nonetheless, these quantitative findings are still consistent with the qualitative findings obtained from focus groups that the evaluation team conducted last year (reported in the first follow-up report). The third follow-up round can shed more insight into the inconsistency in the results.

Given the non-random selection of participants into the activity's components, it is not possible to conclude whether agency or relationship constraints are greater barriers in the Kyrgyz context and whether activities to address these constraints have different levels of effectiveness. Understanding more

about how these constraints can be addressed programmatically is particularly important given the prevalence of business training activities around the world.

ANNEX A: LITERATURE REVIEW

Existing approaches to supporting growth-oriented women entrepreneurs have been heterogeneous in their design and delivery although they have provided some suggestive evidence on the key issues that should be taken into account (Cirera and Qasim 2014). Recent efforts recognize this and have attempted to provide a more unified effort on how to best insert women in the productive process and, at the same time, help maximize their contributions to the well-being of societies (Buvinic et al. 2013). In the context of the current empirical research on the barriers faced by women entrepreneurs and based on the existing literature, the WLSME initiative identified (i) agency limitations, (ii) external constraints, and (iii) lack of relationships as critical issues that should receive support in order to remove crippling limitations to women's productive advancement and contribution to the economy.

Agency Constraints

Whereas most academic and development policy discussions about female entrepreneurs focus on credit constraints, many studies and discussions assume that entrepreneurs manage their businesses optimally. In fact, human capital is treated as fixed with a focus on the process of infusion of financial capital into micro-enterprises, not human or managerial capital, but assuming that entrepreneurs have the latter in optimal amounts (Karlan and Valdivia 2012). Clearly, this is not necessarily the case as the relatively poor among the self-employed rarely have any formal training in business skills. In particular, it has been argued that one must develop "managerial capital" in order to help entrepreneurs affect their firm's business practices, including improving strategic and operational decisions and productivity of factors of production by helping to use them more efficiently (Bruhn et al. 2012). Managerial capital appears to be a fundamental constraint for microenterprise development as business training may enable entrepreneurs to better identify profitable business opportunities, leading to changes in business practices and ultimately to higher sales, profits and happiness (Berge et al. 2012).

The WLSME Kyrgyzstan activity tries to reduce agency constraints by improving human capital of female entrepreneurs with particular emphasis on their managerial capabilities. The key question asked is thus, the following: is lack of managerial capital a first order impediment to firm results, profitability, and growth? In fact, it has been shown in other studies that small-firm entrepreneurs are constrained in the acquisition of these skills, in particular if they require formal training (Caselli and Gennaioli 2005). In particular, the design of the treatment arms in this activity follow a systematic pattern that tries to condense the approaches taken by a growing number of microfinance organizations attempting to build the human and managerial capital of micro-entrepreneur activities, which previously have been vastly idiosyncratic and heterogeneous, and as a consequence, have provided limited external validity. This is perhaps the reason why the current literature on human and managerial capital shows a mixed record. For instance, Karlan and Valdivia (2012) and Cole et al (2011) show that basic microenterprise training seems to affect the command of accounting practices for microenterprises, but has limited to no effects on actual firm outcomes and performance, including profits and sales. Similarly, Bruhn and Zia (2013) and Giné and Mansuri (2014) find that training in managerial capital leads to improvements in business practices but has only limited effects on business performance and sales. On the other hand, Drexler et al, (2012) show that training activities increase in impact if they are targeted to the owner's level as training has significant impact on real outcomes for micro entrepreneurs who have low educational attainment and poor business practices prior to the intervention. Along the same lines, Field et al (2010) find positive treatment effects on upper- caste Hindus, but no such effects on either lower-caste Hindus or Muslims.

External Constraints

In spite of the importance of human capital and managerial capital, a consistent finding in recent academic research is that business training is vastly more effective for male entrepreneurs than for female entrepreneurs (Berge et al. 2012; de Mel et al. 2014; Gine and Mansuri 2014). The differences are striking. Even though female entrepreneurs benefit from training in terms of business knowledge, researchers are unable to find a positive effect on their business-related outcomes. In fact, whereas it has been found that there are no differences in business knowledge between males and females, the former report better business practices, lower business failures, higher investment, and even more household expenditures (Gine and Mansuri 2014). These findings point to the need for more comprehensive measures to promote the businesses of female entrepreneurs as any positive effect of the business training is contingent on gender.

Furthermore, it appears that deeper factors than lack of business knowledge seem to constrain the development of female owned microenterprises. In fact, female and male entrepreneurs fundamentally differ in terms of mind-set and household constraints, which may indicate that more comprehensive measures are necessary in order to promote development among female entrepreneurs, paying greater attention to their motivation for being involved in business activities and to external constraints that may limit their opportunities (Berge et al. 2012). In the context above, the WLSME activities aim at removing external constraints by promoting an enabling environment and a more positive attitude toward women entrepreneurs, with particular emphasis on social norms, which is the most plausible explanation for the gender differences about the role of women in the workplace.

Relationship Constraints

Despite the fact that informal social mechanisms, such as word of mouth, may help reduce external constraints, there is a limit to them. For instance, it has been shown that a significant share of women say that their (male) spouses are responsible for most of their business decisions, suggesting that female businesses show no improvement because women have little decision-making control. In fact, female entrepreneurs are less willing to share income information with their spouse than male entrepreneurs, which may suggest that female entrepreneurs are taxed by their husbands and thus may have less to gain from expanding their businesses (Berge et al. 2012). In this context it is important to develop specific formal direct channels in which women entrepreneurs are able to interact with all the actors involved in the productive process. The fact that women are also less willing than males to compete, suggests that they to a lesser extent have an entrepreneurial mind-set focused on business competition and growth (Berge et al. 2012).

The WSLME initiative aims to reduce information and social gaps in the productive process of women entrepreneurs by facilitating effective relationships between women and the value chain actors and in particular, by increasing cohesion in the productive process. In addition, it is expected that the components included in the activity can help increase the sense of empowerment in women entrepreneurs, in particular, those with specific leadership skills. This, by taking advantage of a combination of women's increased economic activity and control over income resulting from access to a larger network (Mayoux 2001; Kulkani 2011). In particular, these activities can help enhance the status of women entrepreneurs within the community, which are reinforced by the formation of the networks that are part of this activity. This is consistent with an empowerment paradigm that advocates for explicit strategies that support women's ability to protect their individual and collective gender interests (Mayoux 2001).

ANNEX B: EMPIRICAL METHODS

ANCOVA

The evaluation team's preferred method, analysis of covariance (ANCOVA), is a statistical method based on variance, multiple regression, and correlation analysis. It is used to increase the precision of comparison between groups and reduce the probability of Type II errors, i.e., when a false null hypothesis is not rejected (Miller and Chapman 2001; Huck 2012). ANCOVA is thought to improve statistical power and control as long as the relationship between the dependent variable and the covariate within each group is linear and parallel, the covariate is unaffected by other independent variables, and if data are collected under a completely randomized design and before any treatment is applied (Schwarz 2015; Huck 2012). When complying with these assumptions, ANCOVA can have a higher explanatory power than DID *only* if autocorrelation is low. In the context of this evaluation, ANCOVA takes advantage of the low autocorrelation of certain outcome variables in this study, such as business profits and sales, to improve power beyond what a DID approach can attain with the same sample size. Baseline data for these outcome measures have little predictive power for future outcomes, so it is inefficient to fully correct for baseline imbalances between treatment and control groups using DID. Instead, an ANCOVA model can adjust the degree of correction for baseline difference in means according to the degree of correlation between past and future outcomes actually observed in the data (McKenzie 2012).

The ANCOVA specification used for estimations in this evaluation used information from the first and second follow-up survey, as well as the baseline:

$$Y_{i,t} = \beta_0 + \beta_1 D_i + \beta_2 Y_{i,t-1} + \beta_3 Y_{i,t-2} + \delta X_i + \varepsilon_{i,t}$$

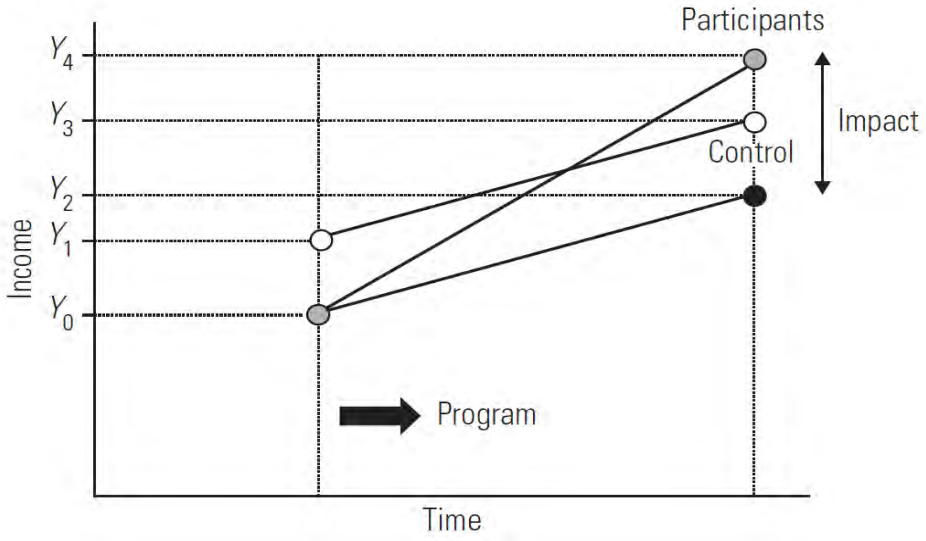
In this case, $Y_{i,t-2}$ is the first follow-up value of the outcome variable, $Y_{i,t-1}$ denotes the baseline value of the outcome variable, and β_1 is the ANCOVA treatment effect. X_i is a matrix of relevant covariates for identification to increase the efficiency of β_1. Specifically, X_i contains the following variables: age, marital status, higher education, business ownership, number of full-time workers from the household and also non-family workers, participation in previous training or seminars, and number of children under 18. The evaluation team also included sector and region-fixed effects were also included.

Difference-in-Differences

The difference-in-differences (DID) estimator is one of the most popular methodologies for applied research in economics. To answer the hypotheses, DID estimates causal relationships among variables by comparing the difference in outcomes before and after an intervention between groups of beneficiaries and nonparticipants (Bertrand et al. 2004). The first "difference" in this method is the difference before the intervention (baseline) and after the intervention (endline). The second "difference" is between the beneficiary group (treatment) and nonparticipant group (control). Thus, two rounds of data are required.

The main advantage of this approach is that in takes into account both observed and unobserved factors which reduces endogeneity problems and provides a tractable way to incorporate both types of variables in the analysis of the effects program interventions have over beneficiaries (Bertrand et al. 2004; Khander et al. 2010). However, this method only remains unbiased as long as interventions are random, and the difference between treatment and control groups' outcomes are time-invariant. This means that to avoid any over- or underestimation of a program's effects, it is crucial to ensure that both treatment and control groups are similar (Ravallion et al. 2005; Khander et al. 2010). For this evaluation, the estimate of the overall activity benefits from the randomized assignment of the intervention; however, the estimates of the components do not. Unlike single differences-in-means, the DID method can be generalized to consider various periods in time.

FIGURE 3: DIFFERENCE-IN-DIFFERENCES METHOD



SOURCE:

KHANDKER, KOOLWAL & SAMAD (2010)

Within this framework and in order to estimate the impacts of the WLSME activity, for each outcome of interest the evaluation team employed the DID specification proposed by Bernal and Peña (2014)[5] as follows:

$$\Delta Y_i = \beta_o + \beta_1 D_i + \delta X_i + \varepsilon_i$$

Here, ΔY_i is the first difference in the outcome variable between the baseline and the second round, D_i is the treatment status dummy that remains unchanged between periods, and β_1 is the treatment effect. The X_i covariates matrix are evaluated in the baseline period and contains the same variable as those used in the ANCOVA model. This specification has the advantage of applying an analogous approach to a randomized controlled trial, as it focuses on the changes of one key variable of interest between periods, which is not the case of standard DID approaches, where period variations of the other explanatory variables also impact the outcome variable.

The evaluation team also employed a DID panel model, following Imbens and Wooldridge(2007), to account for the longitudinal characteristic of the data and generalize the DID specification for several time periods. With many time periods, the DID panel estimation is specified as:

$$Y_{i,t} = \lambda_t + c_i + \beta_1 D_{i,t} + \delta X_{i,t} + \varepsilon_{i,t}\,, \qquad t = 1, \ldots, T$$

Where the same individuals are compared at baseline and each follow-up round, λ_t denotes the time fixed effects, c_i are the individual fixed effects, and X_i is a matrix of relevant covariates for each period and individual. β_1 is the treatment effect. In other words, this technique uses both follow-up surverys as well as the baseline information. The results for both DID specifications are very similar.

[5] The Bernal and Peña (2014) approach allows the evaluation team to only use the baseline and the second survey round.

ANNEX C: DATA COLLECTION AND QUALITY ASSURANCE

Data Collection Process

A local survey partner subcontracted by MSI, M-Vector Research and Consulting, collected quantitative data for both follow-up rounds, with close collaboration, supervision, and quality oversight provided by the evaluation team. The surveys were administered face to face, at the participants' places of business, home, or other location that was convenient for them. The survey took 40 to 60 minutes to complete. The enumerators obtained oral informed consent from each participant, prior to the start of the survey, to confirm that she was willing to participate. M-Vector provided a mobile phone card with 200 Soms (equivalent of USD $2.70) to respondents after the survey was completed as a token of appreciation. Prior to the start of data collection, the survey was piloted and enumerators were trained over the course of three days. Data quality assurance processes were put in place internally by M-Vector, and also independently by the evaluation team.

Baseline data collection conducted by FHI 360 started in July 2013 on a rolling basis over a year as participants applied and were randomly enrolled in batches into the activity. The first follow-up survey conducted by the E3 Analytics and Evaluation Project took place at the end of the activity, between August and October 2015, with an 82 percent response rate. Non-response rates for this round were evenly distributed between treatment and control groups; however, within the treatment group, those who had dropped out of the activity (i.e., did not start or complete the BMT component) were more likely to not respond to the follow-up survey. Among those who did not participate in the follow-up survey, the main reasons given included moved or changed telephone number so could not be reached (9 percent), refusal to participate (6 percent), unavailable after multiple rescheduled appointments (2 percent), passed away (1 percent), and moved abroad (1 percent). The second follow-up survey, also conducted by the E3 Analytics and Evaluation Project, took place 12 months after the end of the WLSME activity, between August and October 2016, with an 84 percent response rate. Similarly to the previous round, non-response rates were evenly distributed between treatment and control groups; however, within the treatment group, those who had dropped out of the activity (i.e., did not start or complete the BMT component) were more likely to not respond to the second follow-up survey. Among those who did not participate in the second follow-up survey, the main reasons given included moved or changed telephone number so could not be reached (9 percent), refusal to participate (4 percent), unavailable after multiple rescheduled appointments (2 percent), and moved abroad (1 percent).

Data Quality Assurance

M-Vector followed standard operating procedures for data collection, including verification procedures conducted both at the site and at headquarters in Bishkek, double entry of survey data, and data query. The survey firm employed the following set of quality control procedures during each data collection round:

- The field manager and supervisors constantly managed the workflow to ensure all enumerators followed the agreed timeline and procedures. Field managers were in contact with the evaluation team to find proper solutions to any unexpected challenges.
- Each supervisor reviewed all completed questionnaires on site, including reading through all questions and answers in the questionnaire to ensure that there were no blanks, skip mistakes, logical inconsistencies, etc. If the supervisor noticed missed questions, skipped questions, or unclear writing, questionnaires were marked and returned to interviewers.
- Supervisors accompanied at least 10 percent of the interviews conducted.

- Completed surveys were sent to the survey firm's headquarters on a weekly basis, where an inspector reviewed each survey for completeness and adequacy prior to data entry.
- Double data entry was performed by two individuals, and the second data entry was done without knowledge or cross reference to the first data entry. Any discrepancies between the two entries of data were resolved by a third person.
- Datasets and progress reports were submitted to the evaluation team on a weekly basis. The progress report included number of contact attempts for pending surveys and reasons for pending status.

The evaluation team provided additional oversight and monitoring during each data collection round:

- Accompaniments of enumerators during interviews at regular intervals. The local coordinator on the evaluation team observed the enumerators' familiarity with and comprehension of the questionnaire and clarity in asking questions.
- The local coordinator also conducted back-checks on 10 percent of completed surveys for each round. Surveys to be back checked were selected randomly and stratified by enumerator to ensure each one was checked on an equal basis. During the back-check call, several validation questions were asked, including interview location, age, household size, type of business, receipt of mobile phone card, and friendliness of the enumerator. Only a few minor discrepancies were found. No back check resulted in significant variance from the reported data.
- From the remaining surveys that were not back checked, a random sample of the scanned paper surveys were compared with the database (five percent). This audit showed no meaningful data entry errors; minor discrepancies were fixed.
- Each week, the evaluation team conducted additional checks to compare each enumerator's average performance to the total sample averages in terms of interview length, number of completed codes, number of "do not knows," scale usage, section skips, and ranges of numerical values. No significant outliers were found.

ANNEX D: DIFFERENCE-IN-DIFFERENCES RESULTS

Difference-in-Difference Results

TABLE 13: BUSINESS GROWTH OUTCOMES – DID

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Average sales on a good month (in ranges)+	0.120 (0.165)	0.198 (0.138)
2	Average sales on a bad month (in ranges)+	0.038 (0.124)	0.114 (0.138)
3	Average sales on an average month (in ranges)+	0.035 (0.113)	0.126 (0.101)
4	Average profit on a good month (in ranges)+	0.420 (0.327)	0.138 (0.170)
5	Average profit on a bad month (in ranges)+	0.181 (0.164)	0.214* (0.092)
6	Average profit on an average month (in ranges)+	0.089 (0.156)	0.024 (0.154)
7	Number of good months in the last year	0.247 (0.158)	0.231 (0.131)
8	Number of bad months in the last year	0.202 (0.123)	-0.331** (0.171)
9	Sales in the last 12 months (Soms)	-168,078 (216,152)	-8,368 (40,964)
10	Likelihood of having a bank account for business only	-0.017 (0.042)	-0.027 (0.046)
11	Number of household members who have worked in business in the last 12 months	0.194 (0.119)	0.205 (0.124)
12	Number of paid household workers	0.068 (0.188)	0.406** (0.381)
13	Number of non-household people who have worked in business in the last 12 months	-0.097 (0.664)	0.119 (0.150)
14	Number of paid non-household workers	-0.369 (0.849)	0.266 (0.529)
15	Number of months per year spent working in the business owned or managed	0.013 (0.337)	-0.089 (0.396)
16	Number of days per week spent working in the business owned or managed	0.117** (0.043)	-0.096 (0.125)
17	Number of hours per day spent working in the business owned or managed	0.064 (0.242)	0.040 (0.486)
18	Likelihood of purchasing raw materials, goods, or equipment for business with a loan in the last 12 months	0.0342 (0.029)	0.050 (0.103)
19	Likelihood of applying for a loan from a financial institution in the last 12 months	0.112** (0.042)	-0.080* (0.040)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
20	Likelihood of loan approval	0.031 (0.082)	-0.079 (0.077)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.
Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.
[+] These outcomes are ordinal variables with the following categories: 1) None; 2) Less than 5,000 Soms; 3) 5,001 – 10,000; 4) 10,001 – 20,000; 5) 20,001 – 40,000; 6) 40,001 – 60,001; 7) 60,000 – 80,000; 8) 80,001 – 100,000; 9) 100,001 – 150,000; 10) 150,001 – 200,000; 11) 200,001 – 500,000; 12) More than 500,000. The coefficient is a measure of increase towards the next higher category. While interpretation of this coefficient (as an ordinal variable) is not straightforward, the lack of statistical significance does not change if a multinomial logistic regression is used.

TABLE 14: ENTREPRENEURIAL LEADERSHIP OUTCOMES – DID

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that prefers to work as an employee in a business instead of managing/owning one	-0.005 (0.028)	0.027 (0.046)
2	Likelihood that people ask me for business advice (very) often	-0.039 (0.075)	0.007 (0.033)
3	Likelihood that I (with my partner/spouse or another household member) am in charge of general business planning decisions	-0.014 (0.049)	-0.027 (0.044)
4	Likelihood that I (with my partner/spouse or another household member) decide what inputs to buy for production	-0.012 (0.049)	-0.031 (0.060)
5	Likelihood that I (with my partner/spouse or another household member) am in charge of sales and client relations	-0.068 (0.052)	-0.030 (0.073)
6	Likelihood that I (with my partner/spouse or another household member) decide if I should apply for a loan	-0.033 (0.046)	-0.036 (0.060)
7	Likelihood that I (with my partner/spouse or another household member) decide my own singular wage	-0.120 (0.086)	0.047 (0.063)
8	Likelihood that I (with my partner/spouse or another household member) decide what type of work I will do	0.002 (0.031)	-0.002 (0.029)
9	Likelihood that I (with my partner/spouse or another household member) am in charge of marketing and advertising decisions	-0.006 (0.056)	-0.060 (0.048)
10	Likelihood that I (with my partner/spouse or another household member) am in charge of staffing of business decisions	-0.018 (0.040)	-0.089 (0.057)
11	Likelihood that often (or always) makes general business planning decisions without consulting anyone else	-0.011 (0.018)	-0.055 (0.054)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
12	Likelihood that often (or always) decides what inputs to buy for production without consulting anyone else	-0.026 (0.027)	-0.019 (0.062)
13	Likelihood that often (or always) manages sales and client relations without consulting anyone else	0.060 (0.046)	0.029 (0.061)
14	Likelihood that often (or always) decides whether to apply for a loan without consulting anyone else	-0.039 (0.050)	0.051 (0.053)
15	Likelihood that often (or always) decides own singular wage without consulting anyone else	-0.023 (0.054)	0.050 (0.051)
16	Likelihood that often (or always) decides what type of work to do without consulting anyone else	-0.019 (0.073)	-0.056 (0.067)
17	Likelihood that often (or always) makes marketing and advertising decisions without consulting anyone else	0.071 (0.086)	0.021 (0.062)
18	Likelihood that often (or always) makes staffing of business decisions without consulting anyone else	0.004 (0.057)	-0.004 (0.058)
19	Likelihood that (strongly) agrees that "women should do what men say"	-0.029 (0.038)	-0.058 (0.056)
20	Likelihood that (strongly) agrees that "women must share their income with their husbands"	0.060* (0.030)	0.024 (0.029)
21	Likelihood that (strongly) agrees that "it is OK if men abandon women if they wish to"	-0.009 (0.050)	0.047 (0.035)
22	Likelihood that (strongly) agrees that "it is OK if men chide women because they went out without any permission"	0.026 (0.081)	0.036 (0.059)
23	Likelihood that (strongly) agrees that "it is OK if men chide women if they do not take care of children"	0.031 (0.036)	-0.003 (0.012)
24	Likelihood that (strongly) agrees that "the role of women is to earn money and take care of her family"	-0.083 (0.057)	-0.067* (0.034)
25	Likelihood that (strongly) agrees that "a mother who works can establish a relationship as warm and solid with her children as a mother who does not work"	0.047** (0.020)	0.024 (0.026)
26	Likelihood that (strongly) agrees that "father's and mother's dedication is equally important for the learning and effective development of children"	0.010 (0.018)	0.010 (0.027)
27	Likelihood that (strongly) agrees that there are no gender inequality problems in my community	-0.021 (0.037)	0.049 (0.039)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)

Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

TABLE 15: NEWORKS OUTCOMES – DID

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that there are social groups in my community	0.047 (0.049)	0.064 (0.045)
2	Likelihood of being an active member of any social group	0.006 (0.054)	-0.113 (0.060)
3	Likelihood of participation in Trade Shows of Fairs	0.033 (0.082)	-0.037 (0.045)
4	Likelihood of implementation of professional advice (business adviser, lawyer or accountant) during the past two years	0.094 (0.068)	-0.001 (0.064)
5	Number of people I can go to ask business advice	0.858 (0.687)	0.516 (0.446)
6	Likelihood that feels sometimes (very) confident negotiating lower prices with suppliers	0.046 (0.045)	0.016 (0.052)
7	Likelihood that feels sometimes (very) confident negotiating higher prices with buyers	-0.054 (0.032)	-0.013 (0.036)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.
Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

TABLE 16: BUSINESS KNOWLEDGE AND PRACTICES OUTCOMES – DID

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that no marketing activities implemented in the last 3 years	0.013 (0.032)	0.025 (0.055)
2	Likelihood that does not use internet for marketing purposes or to sell products/services	-0.028 (0.021)	0.012 (0.020)
3	Likelihood that does not formally keep track of business' products and materials	0.060 (0.038)	0.091*** (0.023)
4	Likelihood that does not perform a physical validation of inventory levels	0.043 (0.050)	0.009 (0.053)
5	Likelihood that business runs out of inventory at least	-0.028	-0.064

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
	one time a month	(0.038)	(0.060)
6	Likelihood that tried to negotiate a lower price with suppliers during the last three months	0.022 (0.062)	0.063 (0.056)
7	Likelihood that does compared price and quality of inputs with other suppliers' products during the last three months	-0.058 (0.051)	-0.007 (0.047)
8	Likelihood that fixed salary for the owner	0.002 (0.044)	0.012 (0.030)
9	Likelihood that records salary of the owner in a notebook, registry or computer	-0.074 (0.070)	-0.064 (0.122)
10	Likelihood that does not keep track of business purchases and sales	0.036 (0.032)	0.057 (0.036)
11	Likelihood that has a written expense budget	-0.046 (0.050)	-0.037 (0.036)
12	Likelihood that has no written goals for next 12 months	-0.067 (0.041)	0.006 (0.022)
13	Likelihood that has no accountancy documents prepared annually	0.012 (0.042)	-0.081 (0.056)
14	Likelihood that has no changes planned over the next 12 months	0.008 (0.018)	0.045** (0.017)
15	Likelihood that (strongly) agrees that "my workspace is well organized"	0.030 (0.028)	0.035 (0.022)
16	Likelihood that (strongly) agrees that "I often communicate clear objectives to my colleagues and employees"	0.010 (0.008)	0.020 (0.013)
17	Likelihood that (strongly) agrees that "I develop work plans at regular intervals"	0.011 (0.027)	0.018 (0.032)
18	Likelihood that (strongly) agrees that "I sometimes miss deadlines"	-0.047 (0.032)	0.005 (0.044)
19	Likelihood that (strongly) agrees that "I believe employees should be treated like family"	-0.015 (0.027)	-0.023 (0.028)
20	Likelihood that (strongly) agrees that "I am sometimes late for appointments or meetings"	-0.057 (0.064)	-0.065 (0.066)
21	Likelihood that (strongly) agrees that "I often attempt to anticipate future circumstances and plan how I/my company will deal with them"	-0.006 (0.023)	-0.012 (0.015)
22	Likelihood that (strongly) agrees that "I am constantly	0.003	-0.003

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
	collecting information about the market in which my company operates"	(0.016)	(0.014)

Note: Coefficients were obtained by DID regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.
Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

DID Panel Results

TABLE 17: BUSINESS GROWTH OUTCOMES – DID PANEL

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Average sales on a good month (in ranges)+	0.045 (0.182)	0.110 (0.138)
2	Average sales on a bad month (in ranges)+	-0.038 (0.143)	0.033 (0.167)
3	Average sales on an average month (in ranges)+	-0.042 (0.147)	0.034 (0.108)
4	Average profit on a good month (in ranges)+	0.379 (0.253)	0.266 (0.270)
5	Average profit on a bad month (in ranges)+	0.144 (0.127)	0.111 (0.119)
6	Average profit on an average month (in ranges)+	0.080 (0.132)	0.003 (0.114)
7	Number of good months in the last year	0.131 (0.268)	0.265*** (0.059)
8	Number of bad months in the last year	0.280 (0.214)	0.087 (0.110)
9	Sales in the last 12 months (Soms)	-6,652 (20,286)	-6,490 (26,886)
10	Likelihood of having a bank account for business only	-0.012 (0.042)	-0.018 (0.042)
11	Number of household members who have worked in business in the last 12 months	0.186 (0.127)	0.208** (0.073)
12	Number of paid household workers	0.191 (0.215)	0.222 (0.179)
13	Number of non-household people who have worked in business in the last 12 months	-0.287 (0.589)	-0.178 (0.331)
14	Number of paid non-household workers	-0.081 (0.588)	-0.148 (0.577)
15	Number of months per year spent working in the business owned or managed	0.093 (0.285)	-0.113 (0.338)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
16	Number of days per week spent working in the business owned or managed	0.130 (0.141)	0.003 (0.077)
17	Number of hours per day spent working in the business owned or managed	0.119 (0.232)	0.153 (0.341)
18	Likelihood of purchasing raw materials, goods, or equipment for business with a loan in the last 12 months	0.109** (0.050)	-0.021 (0.071)
19	Likelihood of applying for a loan from a financial institution in the last 12 months	0.098 (0.060)	-0.106** (0.041)
20	Likelihood of loan approval	-0.006 (0.097)	-0.089** (0.030)

Note: Coefficients were obtained by DID Panel regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.
Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.
[+] These outcomes are ordinal variables with the following categories: 1) None; 2) Less than 5,000 Soms; 3) 5,001 – 10,000; 4) 10,001 – 20,000; 5) 20,001 – 40,000; 6) 40,001 – 60,001; 7) 60,000 – 80,000; 8) 80,001 – 100,000; 9) 100,001 – 150,000; 10) 150,001 – 200,000; 11) 200,001 – 500,000; 12) More than 500,000. The coefficient is a measure of increase towards the next higher category. While interpretation of this coefficient (as an ordinal variable) is not straightforward, the lack of statistical significance does not change if a multinomial logistic regression is used.

TABLE 18: ENTREPRENEURIAL LEADERSHIP OUTCOMES – DID PANEL

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that prefers to work as an employee in a business instead of managing/owning one	-0.012 (0.042)	-0.003 (0.029)
2	Likelihood that people ask me for business advice (very) often	-0.038 (0.055)	-0.022 (0.045)
3	Likelihood that I (with my partner/spouse or another household member) am in charge of general business planning decisions	-0.009 (0.047)	0.016 (0.042)
4	Likelihood that I (with my partner/spouse or another household member) decide what inputs to buy for production	-0.004 (0.050)	0.004 (0.045)
5	Likelihood that I (with my partner/spouse or another household member) am in charge of sales and client relations	-0.070 (0.048)	-0.031 (0.067)
6	Likelihood that I (with my partner/spouse or another household member) decide if I should apply for a loan	-0.028 (0.053)	-0.006 (0.039)
7	Likelihood that I (with my partner/spouse or another household member) decide my own singular wage	-0.116 (0.100)	0.018 (0.048)
8	Likelihood that I (with my partner/spouse or another household member) decide what type of work I will do	-0.003 (0.047)	-0.000 (0.026)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
9	Likelihood that I (with my partner/spouse or another household member) am in charge of marketing and advertising decisions	0.002 (0.052)	-0.020 (0.042)
10	Likelihood that I (with my partner/spouse or another household member) am in charge of staffing of business decisions	-0.004 (0.052)	-0.042 (0.048)
11	Likelihood that often (or always) makes general business planning decisions without consulting anyone else	0.010 (0.052)	-0.060* (0.025)
12	Likelihood that often (or always) decides what inputs to buy for production without consulting anyone else	0.001 (0.053)	-0.044 (0.032)
13	Likelihood that often (or always) manages sales and client relations without consulting anyone else	0.081 (0.053)	0.045 (0.051)
14	Likelihood that often (or always) decides whether to apply for a loan without consulting anyone else	-0.000 (0.058)	0.028 (0.044)
15	Likelihood that often (or always) decides own singular wage without consulting anyone else	0.003 (0.058)	0.026 (0.060)
16	Likelihood that often (or always) decides what type of work to do without consulting anyone else	-0.037 (0.058)	-0.064 (0.057)
17	Likelihood that often (or always) makes marketing and advertising decisions without consulting anyone else	0.120** (0.056)	0.046 (0.070)
18	Likelihood that often (or always) makes staffing of business decisions without consulting anyone else	0.033 (0.052)	-0.016 (0.051)
19	Likelihood that (strongly) agrees that "women should do what men say"	-0.043 (0.046)	-0.021 (0.046)
20	Likelihood that (strongly) agrees that "women must share their income with their husbands"	0.062 (0.039)	0.032 (0.025)
21	Likelihood that (strongly) agrees that "it is OK if men abandon women if they wish to"	-0.016 (0.043)	0.013 (0.042)
22	Likelihood that (strongly) agrees that "it is OK if men chide women because they went out without any permission"	0.020 (0.051)	0.052 (0.076)
23	Likelihood that (strongly) agrees that "it is OK if men chide women if they do not take care of children"	0.040 (0.040)	0.011 (0.024)
24	Likelihood that (strongly) agrees that "the role of women is to earn money and take care of her family"	-0.074 (0.054)	-0.077* (0.040)
25	Likelihood that (strongly) agrees that "a mother who works can establish a relationship as warm and solid with her children as a mother who does not work"	0.044* (0.024)	0.027 (0.020)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
26	Likelihood that (strongly) agrees that "father's and mother's dedication is equally important for the learning and effective development of children"	0.010 (0.021)	0.003 (0.018)
27	Likelihood that (strongly) agrees that there are no gender inequality problems in my community	0.007 (0.038)	0.020 (0.032)

Note: Coefficients were obtained by DID Panel regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.
Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

TABLE 19: NEWORKS OUTCOMES – DID PANEL

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that there are social groups in my community	0.054 (0.052)	0.058 (0.037)
2	Likelihood of being an active member of any social group	0.027 (0.060)	-0.066 (0.071)
3	Likelihood of participation in Trade Shows of Fairs	0.041 (0.057)	-0.015 (0.041)
4	Likelihood of implementation of professional advice (business adviser, lawyer or accountant) during the past two years	0.082* (0.046)	0.033 (0.065)
5	Number of people I can go to ask business advice	1.076*** (0.401)	0.848*** (0.204)
6	Likelihood that feels sometimes (very) confident negotiating lower prices with suppliers	0.066 (0.048)	0.001 (0.054)
7	Likelihood that feels sometimes (very) confident negotiating higher prices with buyers	-0.034 (0.047)	-0.039 (0.029)

Note: Coefficients were obtained by DID Panel regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.
Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

TABLE 20: BUSINESS KNOWLEDGE AND PRACTICES OUTCOMES – DID PANEL

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
1	Likelihood that no marketing activities implemented in the last 3 years	0.008 (0.032)	0.028 (0.038)

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
2	Likelihood that does not use internet for marketing purposes or to sell products/services	-0.026 (0.037)	-0.005 (0.013)
3	Likelihood that does not formally keep track of business' products and materials	0.056* (0.033)	0.079** (0.032)
4	Likelihood that does not perform a physical validation of inventory levels	0.038 (0.044)	0.040 (0.054)
5	Likelihood that business runs out of inventory at least one time a month	-0.032 (0.055)	-0.070 (0.050)
6	Likelihood that tried to negotiate a lower price with suppliers during the last three months	0.050 (0.051)	0.029 (0.055)
7	Likelihood that does compared price and quality of inputs with other suppliers' products during the last three months	-0.056 (0.042)	-0.030 (0.029)
8	Likelihood that fixed salary for the owner	-0.004 (0.051)	0.020 (0.037)
9	Likelihood that records salary of the owner in a notebook, registry or computer	-0.004 (0.051)	-0.101 (0.080)
10	Likelihood that does not keep track of business purchases and sales	0.037 (0.035)	0.063 (0.037)
11	Likelihood that has a written expense budget	-0.048 (0.046)	-0.056 (0.045)
12	Likelihood that has no written goals for next 12 months	-0.081** (0.038)	-0.019 (0.026)
13	Likelihood that has no accountancy documents prepared annually	0.006 (0.046)	0.014 (0.047)
14	Likelihood that has no changes planned over the next 12 months	0.006 (0.027)	0.029 (0.020)
15	Likelihood that (strongly) agrees that "my workspace is well organized"	0.036 (0.033)	0.031 (0.028)
16	Likelihood that (strongly) agrees that "I often communicate clear objectives to my colleagues and employees"	-0.003 (0.021)	0.020** (0.009)
17	Likelihood that (strongly) agrees that "I develop work plans at regular intervals"	0.017 (0.027)	0.008 (0.021)
18	Likelihood that (strongly) agrees that "I sometimes miss deadlines"	-0.051 (0.046)	-0.027 (0.021)
19	Likelihood that (strongly) agrees that "I believe	-0.051	-0.025

Row	Outcome Variable	Follow-up 1 Treatment Effect (std. error)	Follow-up 2 Treatment Effect (std. error)
	employees should be treated like family"	(0.046)	(0.025)
20	Likelihood that (strongly) agrees that "I am sometimes late for appointments or meetings"	-0.065 (0.050)	-0.057 (0.070)
21	Likelihood that (strongly) agrees that "I often attempt to anticipate future circumstances and plan how I/my company will deal with them"	-0.004 (0.020)	-0.016 (0.019)
22	Likelihood that (strongly) agrees that "I am constantly collecting information about the market in which my company operates"	0.006 (0.020)	-0.001 (0.008)

Note: Coefficients were obtained by DID Panel regressions with Region and Sector fixed effects. Age, marital status, education level, business ownership, number of workers, participation in previous trainings, number of children under 18 were included as control variables. Robust standard errors in parentheses.
Statistical significance is denoted by the following system: *** p<0.01, ** p<0.05, * p<0.1.
Outcomes variables stated as "likelihood" can be interpreted as percentage point change by multiplying the coefficient (treatment effect) by 100.

ANNEX E: REFERENCES

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010) "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." Journal of the American Statistical Association, 105(490).

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, Kartika Sari Juniwaty, and Bertil Tungodden (2012) "Business Training in Tanzania: From Research Driven Experiment to Local Implementation", Journal of African Economics, forthcoming.

Bernal, Raquel and Eduardo Peña (2014) "Diferencias en Diferencias: Notas", Manuscrito, Universidad de Los Andes.

Bruhn, Miriam, Dean Karlan and Antoinette Schoar (2012) "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico", Yale Economics Department Working Paper 100.

Bruhn, Miriam and Bilal Zia (2013) "Stimulating managerial capital in emerging markets: the impact of business training for young entrepreneurs" Journal of Development Effectiveness, Taylor & Francis Journals, vol. 5(2), pages 232-266, June.

Buvinic, Mayra, Rebecca Furst-Nichols, and Emily Courey Pryor (2013) "A Roadmap for Promoting Women's Economic Empowerment", United Nations Foundation, New York.

Caselli, Francesco, and Nicola Gennaioli (2005) "Credit Constraints, Competition, and Meritocracy." Journal of the European Economic Association 3(2-3): 679-689.

Cirera, Xavier and Qursum Qasim (2014) "Innovation, Technology, and Entrepreneurship, Policy Note", World Bank 5, September, Washington, DC.

Cole, Shawn A., Thomas Sampson, and Bilal Zia (2011) "Prices or Knowledge? What Drives Demand for Financial Services in Emerging Markets?" Journal of Finance 66, no. 6: 1933–1967.

De Mel, Suresh, David McKenzie and Christopher Woodruff (2014) "Business Training and Female Enterprise Start-up, Growth, and Dynamics: Experimental evidence from Sri Lanka", Journal of Development Economics, 106, 199-210.

_____ (2009) "Measuring Microenterprise Profits: Must We Ask How the Sausage is Made?" Journal of Development Economics, 88(1), 19-31.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar (2012) "Keeping it Simple: Financial Literacy and Rules of Thumb", Manuscript, London School of Economics.

Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*, *4*, 3895-3962.

Field, Erica, Seema Jayachandran and Rohini Pande (2010) "Do traditional institutions constrain female entrepreneurship? A field experiment on business training in India", American Economic Review Papers and Proceedings 100(2): 125-29.

Gathani, Sachin, and Dimitri Stoelinga (2013) "Export Similarity Networks and Proximity Control Methods for Comparative Case Studies." Journal of Globalization and Development, 3(2): 1-39.

Gine, Xavier and Ghazala Mansuri (2014) "Money or Ideas: A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan". Policy Research Paper, World Bank 6959, Washington, DC.

Hejazi, Seyed Amir Mousavi, Mohamad Mehdi Maleki, and M. J. Naeji (2012) "Designing a Scale for Measuring Entrepreneurial Leadership in SMEs." International Conference on Economics Marketing and Management, 28: 71-77. Singapore: IACSIT Press.

Imbens G. and Wooldridge J. (2007) "Differences-in-Differences Estimation", Lecture Notes 10, What's New in Econometrics?, NBER, Summer 2007.

Karlan, Dean and Martin Valdivia (2012) 'Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions", The Review of Economics and Statistics, 93, 2: 510-527.

Kulkani, Vami (2011) "Women's Empowerment and Microfinance: An Asian Perspective Study", Occasional Paper 13, International Fund for Agricultural Development, Washington, DC

Mayoux, Linda (2001) "Microfinance and the Empowerment of Women: A Review of the Issues" International Labor Organization, Geneva.

McKenzie, David (2012) "Beyond Baseline and Follow-up: The Case for More T in Experiments." Journal of Development Economics, 99(2): 210-221.

McKenzie, David and Christopher Woodruff (2014) "What Are We Learning From Business Training and Entrepreneurship Evaluations Around the Developing World?" World Bank Research Observer, Forthcoming.

Waldinger, Fabian (2014) "Differences-in-Differences", Lecture Notes, University of Warwick, United Kingdom.