USAID
FROM THE AMERICAN PEOPLE

# FIRST FOLLOW-UP REPORT

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

**JULY 1, 2016**

This publication was produced for review by the United States Agency for International Development. It was prepared for the E3 Analytics and Evaluation Project by Management Systems International, a Tetra Tech Company; Development and Training Services; and NORC at the University of Chicago.

## COVER PHOTO

**Caption:** Garment production owner and WLSME activity participant with her sewing staff in Bishkek.

**Credit:** Irene Velez, MSI

# FIRST 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 Joan Jennifer Martinez for her statistical support, Isaac Morrison for his qualitative analytical support, and Zura Mendikulova for her support in facilitating the focus groups and providing quality oversight during data collection. The evaluation team also benefited from the shared thoughts and experiences of the 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 Molly Hageboeck (MSI), Benjamin Linkow (NORC), Jeremy Gans (MSI), Joseph Schumacher (MSI), Jacob Patterson-Stein (MSI), and Katherine Jackstadt (dTS). Finally, the evaluation team would like to thank Anastasia de Santos (USAID/E3/TRR) and Wade Channell (USAID/E3/GenDev) for their continued support of this study, and FHI 360 for its work with the original evaluation design and baseline data.

# ACRONYMS

| | |
|---|---|
| ANCOVA | Analysis of Covariance |
| APR | Annual Percentage Rate |
| BMT | Business Management Trainings |
| BT Fund | Bai-Tushum Innovations Fund |
| DID | Difference-in-Differences |
| dTS | Development and Training Services |
| E3 | Bureau for Economic Growth, Education, and Environment (USAID) |
| FIELD-Support | Financial Integration, Economic Leveraging, Broad-Based Dissemination and Support |
| FGD | Focus Group Discussion |
| 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 |
| WLSME | Women's Leadership in Small and Medium Enterprises |

# EXECUTIVE SUMMARY

This report is the first 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 (GenDev) in the United States Agency for International Development's Bureau for Economic Growth, Education and Environment (USAID/E3). The report summarizes the data collection and analysis methods and provides initial findings and conclusions based on baseline and post-intervention data collection and analysis of the Kyrgyzstan WLSME activity. The impact evaluation plans to collect two more follow-up rounds at 12 months and 24 months post-intervention; thus, this report contains only initial findings that could be observed at the end of the activity.

The short-term findings of this first follow-up report are somewhat encouraging yet still inconclusive. The overall WLSME activity has had a positive impact on increasing:

- 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.

It is possible that the limited statistical significance in the findings may be related to the temporal aspect of changing behavior, and that it takes more time to generate change in these outcome measures. This suggests that positive results may consolidate in the future follow-up rounds.

## 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: (i) agency constraints, which impede adequate accumulation of human capital and managerial capital, and (ii) relationship constraints, which limit women's access to information and reduce opportunities to build and draw on social capital.

The activity was implemented nationwide between September 2013 and September 2015 by ACDI/VOCA and its partner organization, Bai-Tushum Innovations Fund, and targeted 960 women in the garment, tourism, and agro-processing sectors who were identified as potential high-growth entrepreneurs. It was composed 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 had greater human capital, social capital, and access to market information, then they would be more likely to grow their businesses and become entrepreneurial leaders. This impact evaluation aims to test this hypothesis with a randomized controlled trial (RCT), where eligible applicants to the WLSME activity are

randomly assigned to a treatment (T) and control (C) 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?

Sample size at baseline consisted of 568 participants in the treatment group and 275 women in 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 first follow-up survey took place at the end of the activity, between August and October 2015, with an 81 percent response rate. The final follow-up sample consisted of 459 participants in the treatment group and 228 women in the control group.

Balance tests were conducted again, showing that the remaining treatment and control groups shared similar social and economic conditions before the WLSME activity started. Moreover, an analysis of the survey non-responses showed that non-response was random, so it is not correlated with treatment assignment and does not depend on observable characteristics.

To answer the primary questions, two empirical models were used to estimate the overall impact of the WLSME on the key set of outcome variables, Difference-in-Differences (DID) and Analysis of Covariance (ANCOVA). The findings using both approaches are reported below. To answer the secondary questions, the DID model was used. 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.

In addition, 6 focus groups with a total of 70 WLSME participants were conducted to explore specific opinions and experiences with the activity in greater depth, as well as produce narratives that address the continuity of personal experiences over time.

## Findings

The majority of outcomes do not yield a statistically significant link to the overall WLSME activity, but there are rather promising causal links for a small number of outcome variables. More importantly, the evaluation team finds indications that future follow-ups may see improved results. First of all, in a large number of cases, the signs of the effect are consistent with the expected direction of change. This suggests that the reason why the evaluation team is unable to find statistically significant results so far may have to do with a lack of statistical power. Moreover, the ANCOVA model, which uses data more efficiently to improve power beyond what DID can attain with the same sample size, tends to show stronger causal links than the DID model. Finally, longer post-intervention time and additional follow-up rounds of data can provide additional gains in statistical power, which may result in more conclusive statistical findings. The findings reported below correspond to the impact of the overall WLSME activity (Primary Question).

### Business Growth

- Sales and profits have moved in the "right" direction when comparing treatment and control. This means that both sales and profits of treated groups tend to show an increase with respect to the control group. While this evidence is not statistically significant, the fact that there is a positive shift in the coefficient and that in some cases the coefficient appears to be economically sizable is encouraging.
- There is a statistically significant positive impact on investments in capital inputs of about 11 percentage points with the ANCOVA model, which might lead to increased sales and profits down the road.

### Entrepreneurial Leadership

- The evaluation team finds some statistically significant results only with the ANCOVA model. In particular, there is a 7.2 percentage point increase in the likelihood of managing sales and client relations independently and a 6.5 percentage point decrease in agreement that "it is okay if men chide women because they went out without any permission."
- However, there is a negative effect of 6.45 percentage points on the agreement with the statement that the role of women is to earn money and take care of their family.

### Networks

- The ANCOVA model yields statistically significant positive results in some of the outcome variables. In particular, there is a 10.6 percentage point increase in the implementation of professional advice and an increase of one additional person women business owners can ask for advice.

### Business Knowledge and Practices

- The direction of the effects is not consistent under the two different models. While the ANCOVA model corrects most of this inconsistency, the evaluation team still finds mixed results.
- The only two outcome variables that yield statistically significant coefficients point in different directions. For instance, treated women entrepreneurs are 5 percentage points less likely to have compared price and quality of inputs with other suppliers' products during the last three months. However, they are 5 percentage points more likely to have written goals for their business for the next 12 months. This last activity was directly encouraged by the WLSME activity, so it is reassuring to see a positive significant impact there.

With respect to the Secondary Questions, the component analysis 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 the findings above is that, whereas 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 statistically significant sign, regardless of the empirical method employed. Future follow-ups will help clarify the extent to which the WLSME activity has had a "true" positive effect. A comment from a WLSME participant during one of the focus groups fully reflects that limited statistical significance in the findings may be related to the temporal aspect of changing behavior, and that it takes more time to generate change in tasks that require more assertiveness or effort than in others:

"It is hard to tell. We have been participating in the project for one year. We absorb all the information, but we cannot apply all of it at once. It is not possible to succeed right after the training. It does not work this way. We have been learning for one year, and now apply our skills and knowledge step by step, and make changes. Our thinking somewhat changed. We apply it now but will see the results in the future. I would not say our financial situation has improved over the year. I think we will see the results in one year."

## Conclusions

The short-term impact of the WLSME activity is somewhat encouraging yet still inconclusive. While the evaluation team expects that future follow-ups will build on these results, given the higher than expected non-response rate, it will be important to assess the viability of the third follow-up round (24 months post-intervention) after the upcoming 12-month follow-up round is completed later this year.

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 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 first 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 (GenDev) in the United States Agency for International Development's Bureau for Economic Growth, Education and Environment (USAID/E3). Post-baseline support for the evaluation, including ongoing implementation monitoring and follow-up data collection has been provided by the Bureau's E3 Analytics and Evaluation Project.[1] The Kyrgyzstan impact evaluation consists of a randomized controlled trial (RCT) designed 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 first follow-up report summarizes the data collection and analysis methods and provides initial findings and conclusions based on baseline and post-intervention data analysis of the Kyrgyzstan WLSME activity. The impact evaluation team plans to collect two more follow-up rounds at 12 months and 24 months post-intervention; thus, this report contains only initial findings that could be observed at the end of the activity.

# WLSME INITIATIVE DESCRIPTION

USAID's WLSME initiative aims to address women's relative absence in the 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: (i) agency constraints, which impede adequate accumulation of human capital and managerial capital, and thus limit women's knowledge and business practices; (ii) 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 (iii) external constraints, which place gender-specific barriers that limit the presence and success of women entrepreneurs. (The Kyrgyzstan activity directly addresses only the first two constraints. The third was not part of the Kyrgyzstan activity.)

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. Impact evaluations for each of these three activities were also initiated under the FIELD-Support LWA Cooperative Agreement with FHI 360 that concluded on September 30, 2014, following the evaluation design and baseline data collection and analysis. Subsequently, implementation of two of the impact evaluations (for Kyrgyzstan and India) were transferred 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 between September 2013 and September 2015 and targeted 960 women who operate enterprises in priority sectors (garment, tourism, and agro-processing), meet minimum employee and loan size requirements, and are identified as potential high-growth entrepreneurs. Activity components were sequenced, with standard services and

[1] The E3 Analytics and Evaluation Project team consists of a team lead, Management Systems International (MSI), and team partners Development and Training Services (dTS) and NORC at the University of Chicago.

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:

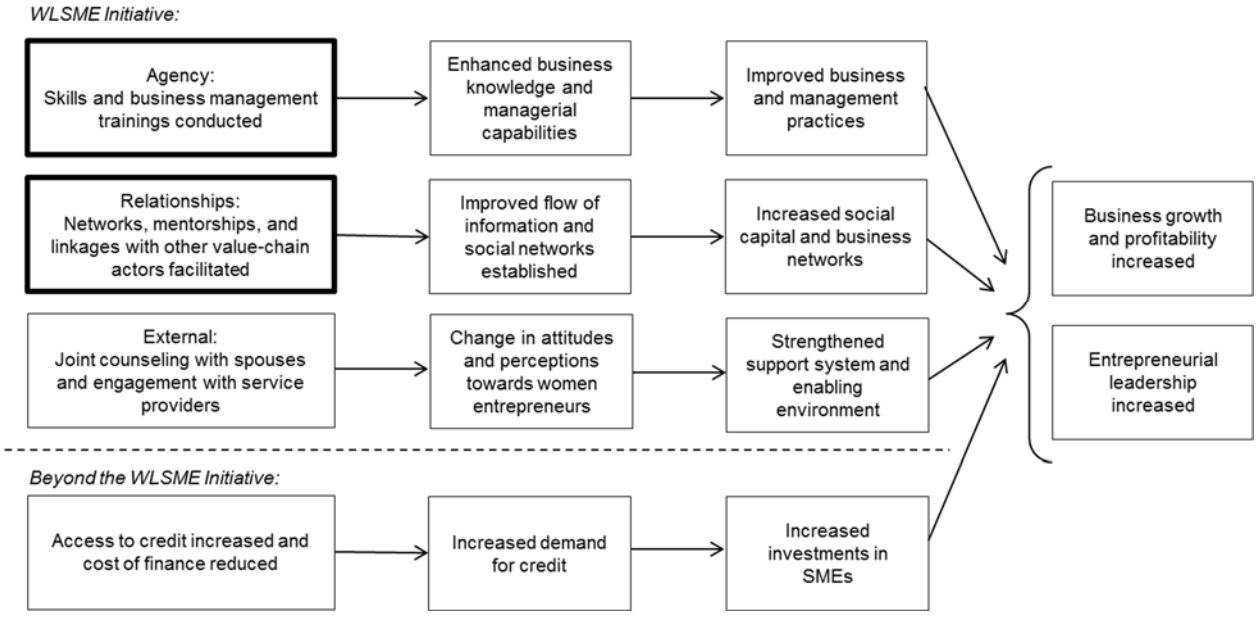
- **Component 1 – Human Capital Gap (Agency):** Business Management Trainings (BMT) covered topics such as negotiating skills, business planning, marketing, financial planning, productivity, and human resource management. The BMT were demand driven and market oriented, meaning that women had a choice 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:** Finalists and semi- finalists from the business plan competitions in Component 2 received this customized assistance (TS/AF). 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 impact 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.

Evaluation questions for an impact evaluation are structured around the development hypotheses being tested. The evaluation questions included here are taken directly from the Evaluation Protocol designed by FHI 360. As such, they include references to the evaluation design that are directly addressed in subsequent sections of this Evaluation Design Proposal. The actual evaluation questions are highlighted in bold.

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 theory of change were defined to measure whether and how much change WLSME activities caused for women entrepreneurs. These include:

- **Business Growth:** Business growth measures include, but are 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.

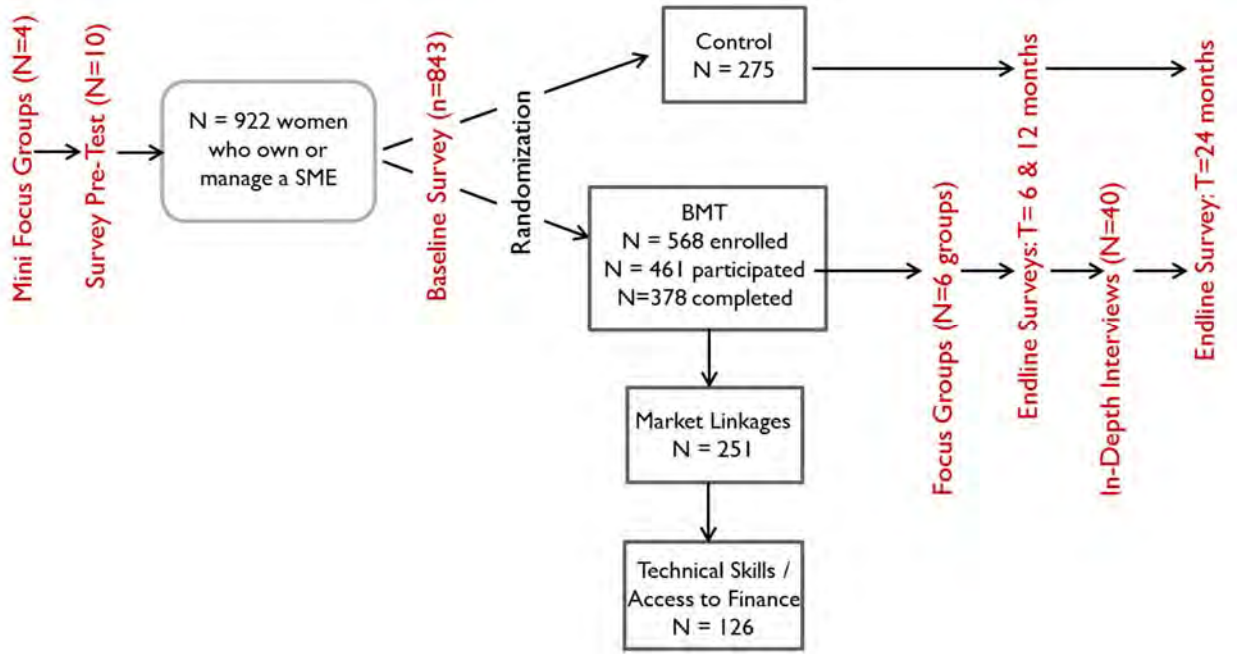
# EVALUATION DESIGN

This impact evaluation consists of an RCT design. Eligible applicants to the 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. Sample size at baseline consisted of 568 participants in the treatment group and 275 women in the control group, excluding the 354 agro-processing sector participants who were not randomized into treatment.[3] Note that all numbers in the rest of this report exclude the agro-processing sector participants.

BMTs were offered 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, this was changed in January 2015 to increase take-up of the overall activity components. A total of 251 women participated in the ML component. Only 126 participants, who were finalists and semi-finalists of the Business Plan competition, were offered the TS/AF component. The first follow-up survey took place at the end of the activity, between August and October 2015, with 81 percent response rate (see Annex C for Data Collection and Quality Assurance Process). The overall evaluation design, timeline, and participation are shown in Figure 2.

[3] ACDI/VOCA, USAID, and the 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.

**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. The direct effect of the overall WLSME activity is therefore estimated on several outcome variables, using rigorous methods drawn from the impact evaluation literature and summarized below (additional details can be found in Annex B).

- **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. 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 that follows:

$$Y_i = \beta_o + \beta_1 D_i + \beta_2 I_t + \beta_3 D_i I_t + \delta X_i + \varepsilon_i$$

Here, $D_i$ is the treatment status dummy, $I_t$ is the follow-up period dummy, $D_i I_t$ is an interaction term of treatment status and follow-up period, and $X_i$ is a matrix of relevant covariates for identification to increase the efficiency of $\beta_3$. 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. Sector and region fixed effects were also included. In the DID specification, $\beta_3$ is the treatment effect. The evaluation team also employs a DID panel model, where the same individuals are compared at baseline and endline. Given that these results are very similar to the cross-sectional DID estimates, the DID panel results are found in Annex D.

- **Analysis of Covariance:** ANCOVA, the evaluation team's preferred method, 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 is the following:

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

  In this case, $Y_{i,t-1}$ is the baseline value of the outcome variable and $\beta_1$ is the ANCOVA treatment effect. Covariates are the same as those used in the DID model, and region and fixed effects are included as well.

The Secondary Questions cannot be objectively answered from an attribution perspective since access to the activity components was not randomized; instead, participants either self-selected or the "most promising entrepreneurs" were selected. The evaluation team could argue that the 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 (namely, ambition and motivation, among others) that led the group to participate in the activity components in the first place. While the same DID method was also used for the analysis across activity components, the estimates cannot be concluded as impact or causality, rather merely a sign of association between the activity components and the outcome variables.

Finally, the evaluation also includes a qualitative component designed to help interpret and better understand the quantitative analyses. The qualitative component included focus group discussions (FGDs) immediately after the activity ended with women who completed the different activity components. The FGDs explored specific opinions and experiences of the activity in greater depth and produced narratives that address the continuity of personal experiences over time. These discussions provide insights into the social and cultural dynamics by which the effects happen and help explain *why* and *how* the WLSME activity worked. Qualitative analysis of the focus group data was conducted using MaxQDA software. Focus group transcripts, translated to English, were uploaded into the software and then reviewed in their entirety with special attention given to participant opinions and feedback related to WSLME activity effectiveness, utility, and outcomes, as well as perceptions on networks and personal empowerment. Responses were sorted categorically and assigned descriptive "codes" to facilitate frequency and demographic cross-comparison. Common trends and themes were tracked across all six FGDs, as were divergences and outliers from those trends and themes. Relevant informant quotations were also extracted from the interview transcripts to serve as examples of participant opinion and sentiment on specific topics.

Qualitative analysis was done parallel to the quantitative analysis and could not be combined.[4] The key findings from this analysis are presented in the next section along with the empirical results. A summary analysis of the qualitative data is included as Annex E.

# 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 or not 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 balance tests were conducted again on the baseline sample. The evaluation 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.

In particular, 10 characteristics or dimensions were chosen to establish whether there were significant differences between treatment and control groups at the baseline and follow-up 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, then 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 exclusion of 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 shows that the control and the treatment groups shared similar social and economic conditions before the WLSME activity took place, and therefore may be compared validly by experimental methods.

[4] The evaluation team chose not to video record FGDs to avoid inhibiting participants in voicing their opinions. Since only audio recording was used, it was not possible to tag each voice to a specific participant and then link the qualitative data to the survey data. Therefore, analysis for the two data sources was done in parallel and could not be combined.

### 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** |
| Total | 568 | 275 | - | - |

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

# ANALYSIS OF NON-RESPONSE IN THE FOLLOW-UP

Given the non-response between the baseline and first follow-up survey round, as shown in Table 2 below, the evaluation team examined whether non-response is random and not correlated with treatment assignment. They found no statistically significant difference (p-value=0.46) in non-response between the treatment and control groups.

### TABLE 2: SAMPLE DISTRIBUTION BY SURVEY ROUND

| | Sample Size | | |
|---|---|---|---|
| | Baseline | First Follow-up | % of Sample |
| Treatment Group | 568 | 459 | 80.95% |
| Control Group | 275 | 228 | 82.91% |
| Total | 842 | 687 | 81.59% |

The evaluation team also examined whether non-response depends on observable characteristics. They 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. Looking at the F-test on the interaction variable coefficients (p-value=0.33), the evaluation team does not find differences in the observable composition of the treatment versus control groups, based on characteristics observed at the baseline. Thus, although the non-response rate is almost 20 percent, it does not affect the validity of the findings presented below.

# PRIMARY QUESTION FINDINGS – OVERALL IMPACT

The following section reports the findings using the DID and ANCOVA approaches.[5] Panel DID findings are very similar to those shown under the DID model below and 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; that is, all of the women assigned to the treatment group remain part of the treatment group regardless of actual participation.

## Business Growth

Business Growth 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. According to Table 3, most of the outcome variables are not statistically significant; however, most of the signs of the coefficient are consistent with an improvement. The DID model shows statistically significant effects (at the 5 percent level) for number of days per week spent working in the business and whether the individual applied to a loan. These results seem to indicate that WLSME participants are spending slightly more days per week in their businesses rather than getting help from household members or growing their businesses by hiring non-household employees. This increased time spent in their businesses, however, is not yet reflected in an increase in sales. Furthermore, the increased probability of applying for a loan is expected, as this was directly encouraged by the WLSME activity. However, this does not yet translate into being more likely to be approved for a loan. In the ANCOVA model, these two variables are not statistically significant, but there is a statistically significant (at the 5 percent level) increase in capital inputs purchased on a loan. Specifically, there is an 11 percentage point increase in the likelihood of purchasing capital inputs on a loan.

In addition, it is important to emphasize that average sales and average profit show an improvement, regardless of the type of month, even though the variable capturing sales yields a negative sign. The reasons for these seemingly paradoxical results may be a combination of lack of statistical power and difficulty in measurement in the sales variable, the latter frequently reported in the related literature (McKenzie 2012). It might also be that changes to business growth take longer to materialize, so future follow-up rounds will be beneficial in analyzing the trends in these outcomes.

[5] 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; whereas 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, except in the case of Networks (
Table 5). 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.

## TABLE 3: BUSINESS GROWTH OUTCOMES

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| 1 | Average sales on a good month (in ranges)+ | 0.12 | 0.04 |
| | | (0.165) | (0.137) |
| 2 | Average sales on a bad month (in ranges)+ | 0.0378 | -0.0697 |
| | | (0.124) | (0.11) |
| 3 | Average sales on an average month (in ranges)+ | 0.0354 | -0.0398 |
| | | (0.113) | (0.117) |
| 4 | Average profit on a good month (in ranges)+ | 0.42 | 0.37 |
| | | (0.327) | (0.228) |
| 5 | Average profit on a bad month (in ranges)+ | 0.181 | 0.044 |
| | | (0.164) | (0.0948) |
| 6 | Average profit on an average month (in ranges)+ | 0.089 | 0.081 |
| | | (0.156) | (0.101) |
| 7 | Number of good months in the last year | 0.247 | -0.079 |
| | | (0.158) | (0.204) |
| 8 | Number of bad months in the last year | 0.202 | 0.191 |
| | | (0.123) | (0.119) |
| 9 | Sales in the last 12 months (Soms) | -168,078 | -6,919 |
| | | (216152) | (16731) |
| 10 | Likelihood of having a bank account for business purposes only | -0.0174 | -0.0377 |
| | | (0.0424) | (0.0320) |
| 11 | Number of people from household who have worked in business in the last 12 months | 0.194 | 0.0342 |
| | | (0.119) | (0.105) |
| 12 | Number of household people who worked in business and were remunerated with cash | 0.0675 | 0.0921 |
| | | (0.188) | (0.165) |
| 13 | Number of people from outside household who have worked in business in the last 12 months | -0.0971 | -0.297 |
| | | (0.664) | (0.586) |
| 14 | Number of non-household people who worked in business and were remunerated with cash | -0.369 | -0.102 |
| | | (0.849) | (0.588) |
| 15 | Number of months per year spent working in the business owned or managed | 0.0127 | 0.0332 |
| | | (0.337) | (0.215) |
| 16 | Number of days per week spent working in the business owned or managed | 0.117** | 0.179 |
| | | (0.0429) | (0.111) |
| 17 | Number of hours per day spent working in the business owned or managed (hours) | 0.0635 | 0.148 |
| | | (0.242) | (0.217) |
| 18 | Likelihood of purchasing raw materials, goods, or equipment for business with a loan in the last 12 months | 0.0342 | 0.110** |
| | | (0.0286) | (0.0498) |
| 19 | Likelihood of applying for a loan from a financial institution in the last 12 months | 0.112** | 0.0139 |

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| | | (0.0423) | (0.0386) |
| 20 | Likelihood of loan approval | 0.0305 | -0.0264 |
| | | (0.0820) | (0.0875) |

Note: Coefficients were obtained by DID and ANCOVA regressions with Region and Sector FE. 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 category, Entrepreneurial Leadership, includes variables related to decision-making in the business, entrepreneurial vocation, level of independence, and women's empowerment, among others. This is shown in Table 4. Similar to the set of Business Growth outcomes, most variables are not statistically significant at conventional levels. Unlike the previous set, not all the corresponding signs of the coefficients are consistent with behavioral change. In particular, the direction of the effect for role in decision-making (rows 3 to 10) are mainly negative under the DID model. In other words, treated women are less likely to participate in decision-making of different aspects of the business. However, these changes are not statistically significant. Moreover, the negative sign disappears for most of these variables under the ANCOVA model, but results are still not statistically significant.

The outcomes related to decisions taken without consulting anyone else (rows 11 to 18) show mixed results under the DID model, with most of the coefficients showing negative signs. In other words, treated women are more likely to ask for advice in decision-making of different aspects of the business. However, none of these changes are statistically significant. The direction of the effects is again corrected under the ANCOVA model, and one outcome is statistically significant (at the 10 percent level). There is a 7.2 percentage point increase in the likelihood of always/often managing sales and client relations without consulting anyone else.

The empowerment outcomes (rows 19 to 27) are more consistent with the expected sign direction, but results are mixed. Under the DID model, two outcome variables yield statistically significant coefficients. The first one is the treatment effect on the proportion of individuals who (strongly) agree that "women must share their income with their husband," which shows that treated individuals are six percentage points more likely to agree with that statement. Unfortunately, this result is bad news for women's economic empowerment. However, the second statistically significant outcome variable in Table 4 shows that treated individuals are 4.6 percentage points more likely to agree that "a mother who works can establish a relationship as warm and solid with her children as a mother who does not work". Under the ANCOVA model, however, these two variables are no longer statistically significant. However, there is an increased rejection of the statement "it is OK if men chide women because they went out without permission." Treated women are 6.5 percentage points less likely to agree with this statement, significant at the 10 percent level.

As in the case of Business Growth, these overall findings may be interpreted as showing a problem related to lack of statistical power, but they may also be signaling a transition towards some behavioral change, although at this point it is quite premature to affirm that this is the case given the ingrained nature of empowerment issues of women in the Kyrgyz Republic. It is expected that the next follow-up will provide additional insights on the trends with respect to changes in this set of outcome variables.

**TABLE 4: ENTREPRENEURIAL LEADERSHIP OUTCOMES**

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| 1 | Likelihood that prefers to work as an employee in a business instead of managing/owning one | -0.0048 | -0.0171 |
| | | (0.028) | (0.0273) |
| 2 | Likelihood that people ask me for business advice (very) often | -0.0393 | -0.0426 |
| | | (0.0747) | (0.0393) |
| 3 | Likelihood that I (with my partner/spouse or another household member) am in charge of general business planning decisions | -0.0141 | -0.00703 |
| | | (0.0493) | (0.0322) |
| 4 | Likelihood that I (with my partner/spouse or another household member) decide what inputs to buy for production | -0.0118 | 0.0269 |
| | | (0.0491) | (0.0355) |
| 5 | Likelihood that I (with my partner/spouse or another household member) am in charge of sales and client relations | -0.068 | -0.0237 |
| | | (0.0516) | (0.0332) |
| 6 | Likelihood that I (with my partner/spouse or another household member) decide if I should apply for a loan | -0.0325 | -0.0186 |
| | | (0.0463) | (0.0375) |
| 7 | Likelihood that I (with my partner/spouse or another household member) decide my own singular wage | -0.12 | 0.00456 |
| | | (0.0863) | (0.0342) |
| 8 | Likelihood that I (with my partner/spouse or another household member) decide what type of work I will do | 0.00241 | 0.0201 |
| | | (0.0308) | (0.0322) |
| 9 | Likelihood that I (with my partner/spouse or another household member) am in charge of marketing and advertising decisions | -0.00643 | 0.0306 |
| | | (0.0558) | (0.0373) |
| 10 | Likelihood that I (with my partner/spouse or another household member) am in charge of staffing of business decisions | -0.0181 | 0.045 |
| | | (0.0398) | (0.0357) |
| 11 | Likelihood that often (or always) makes general business planning decisions without consulting anyone else | -0.0112 | 0.0257 |
| | | (0.0178) | (0.0375) |
| 12 | Likelihood that often (or always) decides what inputs to buy for production without consulting anyone else | -0.0257 | 0.0328 |
| | | (0.027) | (0.0385) |

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| 13 | Likelihood that often (or always) manages sales and client relations without consulting anyone else | 0.0603 | 0.0720* |
| | | (0.0455) | (0.039) |
| 14 | Likelihood that often (or always) decides whether to apply for a loan without consulting anyone else | -0.0391 | -0.0105 |
| | | (0.0501) | (0.0443) |
| 15 | Likelihood that often (or always) decides own singular wage without consulting anyone else | -0.0226 | 0.026 |
| | | (0.0543) | (0.0432) |
| 16 | Likelihood that often (or always) decides what type of work to do without consulting anyone else | -0.019 | 0.0517 |
| | | (0.0731) | (0.0393) |
| 17 | Likelihood that often (or always) makes marketing and advertising decisions without consulting anyone else | 0.0706 | 0.0543 |
| | | (0.0856) | (0.0444) |
| 18 | Likelihood that often (or always) makes staffing of business decisions without consulting anyone else | 0.00352 | 0.0355 |
| | | (0.0572) | (0.041) |
| 19 | Likelihood that (strongly) agrees that "women should do what men say" | -0.0285 | -0.0279 |
| | | (0.0382) | (0.0355) |
| 20 | Likelihood that (strongly) agrees that "women must share their income with their husbands" | 0.0604* | 0.0136 |
| | | (0.0299) | (0.0295) |
| 21 | Likelihood that (strongly) agrees that "it is OK if men abandon women if they wish to" | -0.00895 | -0.0372 |
| | | (0.0503) | (0.0295) |
| 22 | Likelihood that (strongly) agrees that "it is OK if men chide women because they went out without any permission" | 0.0257 | -0.0653* |
| | | (0.081) | (0.0364) |
| 23 | Likelihood that (strongly) agrees that "it is OK if men chide women if they do not take care of children" | 0.031 | 0.0379 |
| | | (0.0363) | (0.028) |
| 24 | Likelihood that (strongly) agrees that "the role of women is to earn money and take care of her family" | -0.083 | -0.0645* |
| | | (0.0573) | (0.0375) |
| 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" | 0.0467** | 0.0166 |
| | | (0.0195) | (0.0177) |
| 26 | Likelihood that (strongly) agrees that "father's and mother's dedication is equally important for the learning and effective development of children" | 0.0097 | 0.00131 |
| | | (0.018) | (0.0141) |
| 27 | Likelihood that (strongly) agrees that there are no gender inequality problems in my community | -0.021 | -0.0144 |
| | | (0.0369) | (0.0271) |

Note: Coefficients were obtained by DID and ANCOVA regressions with Region and Sector FE. 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

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|

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[6], findings are consistent with the ones from Business Growth and Entrepreneurship. Under the DID model, there are no statistically significant effects on the network-related outcomes. However, the findings are more encouraging under the ANCOVA model, where there is statistical significance for two outcomes. Treated women have approximately one more person to ask for business advice in comparison to the control group and they are 10.6 percentage points more likely to implement professional advice. As in the previous cases, lack of statistical power and difficulty in observing changes in behavior in the short run may be the culprits of not finding additional impacts on networks. While there is no direct implication that positive impact might be detectable in the future, the fact that most of the outcome variables point in the "correct" direction is somewhat reassuring.

**TABLE 5: NETWORKS OUTCOMES**

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| 1 | Likelihood that there are social groups in my community | 0.0465 | 0.0262 |
| | | (0.0492) | (0.037) |
| 2 | Likelihood of being an active member of any social group | 0.00583 | 0.0309 |
| | | (0.0537) | (0.0531) |
| 3 | Likelihood of participation in Trade Shows or Fairs | 0.0329 | 0.0316 |
| | | (0.0822) | (0.0372) |
| 4 | Likelihood of implementation of professional advice (business adviser, lawyer or accountant) during the past two years | 0.0938 | 0.106*** |
| | | (0.0677) | (0.0338) |
| 5 | Number of people I can go to ask business advice | 0.848 | 0.980** |
| | | (0.687) | (0.393) |
| 6 | Likelihood that feels sometimes (very) confident negotiating lower prices with suppliers | 0.0464 | 0.0269 |
| | | (0.0453) | (0.0327) |
| 7 | Likelihood that feels sometimes (very) confident negotiating higher prices with buyers | -0.0539 | -0.0295 |
| | | (0.0316) | (0.0297) |

Note: Coefficients were obtained by DID and ANCOVA regressions with Region and Sector FE. 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.

[6] The outcome variables that are statistically significant in Table 5 remain so even after including Bonferroni corrections, in spite of the fact that this type of procedure is particularly taxing in terms of power requirements. Also, see Footnote 7.

## Business Knowledge and Practices

Table 6 shows the evaluation team's findings with respect to Business Knowledge and Practices, which include implementation of marketing, operations, and accounting practices; negotiation skills; recording of budget; and future expectations, among others. Under the DID model, the treatment effect is not statistically significant for any of the outcome variables, and there is no consistency with the direction of the effect expected for several statements related to business practices. While the coefficients reflecting agreement with statements related to better organization, communication, and basic planning are consistent with the treatment, the sign of the coefficients does not hold for outcome variables related to activities that require somewhat more proactivity and effort. This is the case of recording and writing of plans, expenses and budget, implementation of marketing activities, keeping track of products and materials, and organized salary recording.

Under the ANCOVA model, the inconsistency in the sign of these outcomes is mostly corrected. However, the evaluation team finds mixed results as the only two outcome variables that yield statistically significant coefficients point in different directions. For instance, treated women entrepreneurs are 5 percentage points less likely to have compared price and quality of inputs with other suppliers' products during the last three months (statistically significant at the 10 percent level). However, they are 5 percentage points more likely to have written goals for their business for the next 12 months (statistically significant at the 5 percent level). This last activity was directly encouraged by the WLSME activity, so it is reassuring to see a positive significant impact there.

At this point it is premature to try to explain this divergence in the direction of outcomes. This may be occurring for the same reasons mentioned above, namely lack of statistical power and inherent difficulty in impacting behavior in the short run. For instance, 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, especially given that they are already spending more time on the business overall. In this context, the next follow-up may shed some additional light to the findings in this section.

**TABLE 6: BUSINESS KNOWLEDGE AND PRACTICES OUTCOMES**

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| 1 | Likelihood that no marketing activities implemented during the last three years | 0.0128 | -0.0189 |
| | | (0.0315) | (0.0196) |
| 2 | Likelihood that made special offers during the last three months | 0.0210 | 0.0277 |
| | | (0.0399) | (0.0386) |
| 3 | Likelihood that does not use internet for marketing purposes or to sell products/services | -0.0284 | -0.0495 |
| | | (0.0208) | (0.0328) |
| 4 | Likelihood that does not formally keep track of business' products and materials | 0.0599 | -0.00718 |
| | | (0.0382) | (0.0243) |
| 5 | Likelihood that does not perform a physical validation of inventory levels | 0.0427 | -0.00259 |
| | | (0.0502) | (0.0309) |
| 6 | Likelihood that business runs out of inventory at least one time a month | -0.0282 | -0.00398 |
| | | (0.0381) | (0.0389) |
| 7 | Likelihood that tried to negotiate a lower price with | 0.0215 | -0.0229 |

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| | suppliers during the last three months | | |
| | | (0.0615) | (0.0357) |
| 8 | Likelihood that does compared price and quality of inputs with other suppliers' products during the last three months | -0.0584 | -0.0502* |
| | | (0.0506) | (0.0288) |
| 9 | Likelihood that fixed salary for the owner | 0.00188 | 0.0117 |
| | | (0.044) | (0.0354) |
| 10 | Likelihood that records salary of the owner in a notebook, registry or computer | -0.074 | 0.0182 |
| | | (0.0696) | (0.056) |
| 11 | Likelihood that does not keep track of business purchases and sales | 0.0357 | -0.0141 |
| | | (0.0318) | (0.0242) |
| 12 | Likelihood that has a written expense budget | -0.0458 | 0.0132 |
| | | (0.0501) | (0.0348) |
| 13 | Likelihood that has no written goals for next 12 months | -0.067 | -0.0543** |
| | | (0.0412) | (0.0257) |
| 14 | Likelihood that has no accountancy documents prepared annually | 0.0124 | 0.00599 |
| | | (0.0422) | (0.0369) |
| 15 | Likelihood that has no changes planned over the next 12 months | 0.00777 | -0.00523 |
| | | (0.018) | (0.0213) |
| 16 | Likelihood that (strongly) agrees that "my workspace is well organized" | 0.0295 | 0.0343 |
| | | (0.0282) | (0.0313) |
| 17 | Likelihood that (strongly) agrees that "I often communicate clear objectives to my colleagues and employees" | 0.0102 | -0.00847 |
| | | (0.00826) | (0.0198) |
| 18 | Likelihood that (strongly) agrees that "I develop work plans at regular intervals" | 0.0111 | 0.0176 |
| | | (0.0265) | (0.0251) |
| 19 | Likelihood that (strongly) agrees that "I sometimes miss deadlines" | -0.0475 | -0.0134 |
| | | (0.0316) | (0.0403) |
| 20 | Likelihood that (strongly) agrees that "I believe employees should be treated like family" | -0.0149 | -0.0089 |
| | | (0.0265) | (0.0256) |
| 21 | Likelihood that (strongly) agrees that "I am sometimes late for appointments or meetings" | -0.0568 | -0.0155 |
| | | (0.0637) | (0.0376) |
| 22 | Likelihood that (strongly) agrees that "I often attempt to anticipate future circumstances and plan how I/my company will deal with them" | -0.00568 | -0.00912 |
| | | (0.0227) | (0.0174) |
| 23 | Likelihood that (strongly) agrees that "I am constantly collecting information about the market in which my | 0.00278 | 0.00217 |

| Row | Outcome Variable | DID Treatment Effect (std. error) | ANCOVA Treatment Effect (std. error) |
|---|---|---|---|
| | company operates" | | |
| | | (0.0156) | (0.0185) |

Note: Coefficients were obtained by DID and ANCOVA regressions with Region and Sector FE. 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 covered 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 7.

### TABLE 7: 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 | 59.0 | 211 | 60.6 | 117 | 33.6 |
| Non-participant | 111 | 41.0 | 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) Participants who were assigned to receive a specific activity component but rejected to participate in the follow-up survey in spite of being registered in the baseline survey.

It is important to mention that, given that 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, Propensity Score Matching, was not included in this report due to small sample size.

In the tables below, the treatment effect of Component A corresponds to the value-add of the ML component relative to the BMT only. The treatment effect of Component B corresponds to the value-add of the ML+TS/AF components relative to the BMT only.

## Summary of Findings

A summary of the key findings for the Secondary Questions is below. Detailed explanations and tables follow.

- For the **Business Growth** outcomes, both Components A and B decrease the association of having a bank account for business purposes. The ML component increases the association of loan approval.
- For the **Entrepreneurial Leadership** outcomes, findings are generally in the opposite direction of what is expected. In particular, both Components A and B are associated with a preference to work as an employee instead of managing/owning a business. The TS/AF component is associated with a decrease in decision-making role and independence in making decisions with respect to the business. The ML component, on the other hand, is associated with negative attitudes towards gender equity and empowerment. The single positive association is that the TS/AF participants are more likely to be asked for business advice very often.
- For the **Networks** outcomes, both Components A and B have increased participation in trade shows or fairs, since this was directly encouraged by the WLSME activity. Moreover, the TS/AF component is associated with increased implementation of professional advice. Given the one-on-one mentoring support provided under this component, this might imply that women entrepreneurs are more likely to implement professional advice when it is given directly to them rather than to a larger training group.
- For the **Business Knowledge and Practices** outcomes, both Components A and B are associated with accountancy documents prepared annually, while the Market Linkages component is also associated with performing physical validation of inventory levels.

## Business Growth

The evaluation team finds that most of the Business Growth outcome variables are not statistically significant at conventional levels, as shown in Table 8. Nonetheless, in most cases the signs of the coefficients point in the direction consistent with treatment. Interestingly, two variables do yield signs that are statistically significant. The first outcome is whether the bank account is used for business purposes only, which unfortunately yields a negative sign in both Components. In particular, the ML component is associated with a decrease in having a bank account for business purposes by 15.9 percentage points (significant at the 5 percent level), while the TS/AF component is associated with a decrease of 28 percentage points (significant at the 1 percent level). The second outcome is whether the individual was approved a loan, which while positive for both Components, is statistically significant for Component A, only. The ML component is associated with an increased loan approval of 15.5 percentage points (significant at the 10 percent level).

**TABLE 8: BUSINESS GROWTH – BY ACTIVITY COMPONENT**

| Row | Outcome Variable | Component A Treatment Effect | Component B Treatment Effect |
|---|---|---|---|
| | | (std. error) | (std. error) |
| 1 | Average sales on a good month (in ranges)+ | 0.187 | -0.0997 |
| | | (0.294) | (0.351) |
| 2 | Average sales on a bad month (in ranges)+ | 0.130 | 0.00107 |
| | | (0.294) | (0.351) |

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| 3 | Average sales on an average month (in ranges)+ | 0.187 | 0.0934 |
| | | (0.247) | (0.294) |
| 4 | Average profit on a good month (in ranges)+ | 0.303 | 0.157 |
| | | (0.254) | (0.316) |
| 5 | Average profit on a bad month (in ranges)+ | 0.252 | 0.185 |
| | | (0.179) | (0.225) |
| 6 | Average profit on an average month (in ranges)+ | 0.230 | 0.119 |
| | | (0.205) | (0.250) |
| 7 | Number of good months in the last year | 0.181 | -0.223 |
| | | (0.365) | (0.378) |
| 8 | Number of bad months in the last year | -0.147 | 0.495 |
| | | (0.311) | (0.317) |
| 10 | Sales in the last 12 months (Soms) | -37,192 | -31,837 |
| | | (59,839) | (89,127) |
| 11 | Likelihood of having a bank account for business purposes only | -0.159** | -0.280*** |
| | | (0.0626) | (0.0724) |
| 12 | Number of people from household that have worked in business in past 12 months | 0.168 | -0.0620 |
| | | (0.186) | (0.189) |
| 13 | Number of household people that worked in business and were remunerated with cash | -0.0578 | 0.151 |
| | | (0.259) | (0.252) |
| 14 | Number of people from outside household that have worked in business in past 12 months | 1.044 | 1.322 |
| | | (0.721) | (0.858) |
| 15 | Number of non-household people that worked in business and were remunerated with cash | 0.770 | 1.058 |
| | | (0.826) | (0.948) |
| 16 | Number of months per year spent on the business owned or managed | -0.386 | 0.326 |
| | | (0.495) | (0.498) |
| 17 | Number of days per week spent on the business owned or managed | 0.0243 | 0.00975 |
| | | (0.206) | (0.195) |
| 18 | Number of hours per day spent on the business owned or managed | 0.668 | 0.474 |
| | | (0.456) | (0.488) |
| 19 | Likelihood of purchasing raw materials, goods, or equipment for business with a loan in the last 12 months | -0.00835 | -0.0507 |
| | | (0.0988) | (0.0980) |
| 20 | Likelihood of applying for a loan from a financial institution in the last 12 months | -0.0213 | -0.101 |
| | | (0.0835) | (0.0866) |
| 21 | Likelihood of loan approval | 0.155* | 0.0251 |
| | | (0.0925) | (0.0857) |

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|

Note: Coefficients were obtained by Diff-in-Diff regressions with Region and Sector FE. 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 9 shows findings for Entrepreneurial Leadership by activity component. For the most part, the findings in this module are rather inconsistent as some outcome variables yield the "wrong" sign and in some cases are even statistically significant. For instance, both Components are associated with a preference to work as an employee in a business instead of managing/owning one, statistically significant at conventional levels.

We find additional negative and statistically significant associations for the TS/AF component with respect to decision-making roles and independence in running the business, as seen in the 12.3 percentage point decrease in participating in the decision of sales and client relations and 14.3 percentage point decrease in deciding whether to apply for a loan without consulting anyone else. The ML component is associated with negative attitudes towards gender equity and empowerment. The participants from this component recognize the gender inequality problems in their community, as shown by the statistically significant association of 4.5 percentage points, which explains why their answers to four gender empowerment questions are negative and statistically significant. The results above are clearly troubling and may be the result of small and biased samples, given the fact that the activity components were not randomized.

The single positive association is that TS/AF participants are more likely to be asked for business advice very often, as shown by a 17.3 percentage point increase in that outcome variable.

### TABLE 9: ENTREPRENEURIAL LEADERSHIP – BY ACTIVITY COMPONENT

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| 1 | Likelihood that prefers to work as an employee in a business instead of managing/owning one | 0.272*** | 0.129** |
| | | (0.0663) | (0.0641) |
| 2 | Likelihood that people ask me for business advice (very) often | -0.0746 | 0.173** |
| | | (0.0835) | (0.0866) |
| 3 | Likelihood that I (with my partner/spouse or another household member) am in charge of general business planning decisions | 0.0310 | 0.00255 |
| | | (0.0734) | (0.0749) |
| 4 | Likelihood that I (with my partner/spouse or another household member) decide what inputs to buy for production | 0.0102 | -0.114 |

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| | | (0.0787) | (0.0771) |
| 5 | Likelihood that I (with my partner/spouse or another household member) am in charge of sales and client relations | -0.0176 | -0.123* |
| | | (0.0753) | (0.0741) |
| 6 | Likelihood that I (with my partner/spouse or another household member) decide if I should apply for a loan | -0.0213 | -0.100 |
| | | (0.0798) | (0.0812) |
| 7 | Likelihood that I (with my partner/spouse or another household member) decide my own singular wage | 0.153 | 0.0967 |
| | | (0.341) | (0.232) |
| 8 | Likelihood that I (with my partner/spouse or another household member) decide what type of work I will do | 0.0270 | -0.0853 |
| | | (0.0732) | (0.0711) |
| 9 | Likelihood that I (with my partner/spouse or another household member) am in charge of marketing and advertising decisions | 0.0931 | 0.0488 |
| | | (0.0767) | (0.0764) |
| 10 | Likelihood that I (with my partner/spouse or another household member) am in charge of staffing of business decisions | 0.0649 | -0.0473 |
| | | (0.0744) | (0.0741) |
| 11 | Likelihood that often (or always) makes general business planning decisions without consulting anyone else | -0.0397 | 0.0160 |
| | | (0.0810) | (0.0807) |
| 12 | Likelihood that often (or always) decides what inputs to buy for production without consulting anyone else | 0.0240 | -0.0640 |
| | | (0.0819) | (0.0827) |
| 13 | Likelihood that often (or always) manages sales and client relations without consulting anyone else | -0.0979 | -0.176** |
| | | (0.0790) | (0.0804) |
| 14 | Likelihood that often (or always) decides whether to apply for a loan without consulting anyone else | -0.118 | -0.143* |
| | | (0.0827) | (0.0845) |
| 15 | Likelihood that often (or always) decides own singular wage without consulting anyone else | 0.134 | 0.0425 |
| | | (0.0895) | (0.0919) |
| 16 | Likelihood that often (or always) decides what type of work will do without consulting anyone else | -0.124 | -0.0607 |
| | | (0.0839) | (0.0847) |
| 17 | Likelihood that often (or always) makes marketing and advertising decisions without consulting anyone else | -0.00205 | -0.0641 |
| | | (0.0889) | (0.0880) |
| 18 | Likelihood that often (or always) makes staffing of business decisions without consulting anyone else | 0.0235 | -0.0462 |
| | | (0.0838) | (0.0822) |
| 19 | Likelihood that (strongly) agrees that "women should do what men say" | -0.0537 | 0.0467 |

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| | | (0.0714) | (0.0760) |
| 20 | Likelihood that (strongly) agrees that "Women must share their income with their husbands" | -0.0372 | -0.103 |
| | | (0.0615) | (0.0661) |
| 21 | Likelihood that (strongly) agrees that "it is OK if men abandon women if they wish to" | 0.130** | -0.0168 |
| | | (0.0597) | (0.0612) |
| 22 | Likelihood that (strongly) agrees that "it is OK if men chide women because they went out without any permission" | -0.0278 | 0.0797 |
| | | (0.0781) | (0.0807) |
| 23 | Likelihood that (strongly) agrees that "it is OK if men chide women if they do not take care of children" | -0.108** | -0.0920 |
| | | (0.0536) | (0.0619) |
| 24 | Likelihood that (strongly) agrees that "the role of women is to earn money and take care of her family" | -0.0510 | 0.0375 |
| | | (0.0798) | (0.0838) |
| 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" | -0.0753** | -0.0412 |
| | | (0.0357) | (0.0374) |
| 26 | Likelihood that (strongly) agrees that "father's and mother's dedication is equally important for the learning and effective development of children" | -0.0454* | -0.0349 |
| | | (0.0253) | (0.0292) |
| 27 | Likelihood that (strongly) agrees that there are no gender inequality problems in my community | -0.0454* | -0.0349 |
| | | (0.0253) | (0.0292) |

Note: Coefficients were obtained by DID regressions with Region and Sector FE. 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 10 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. 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 a positive and statistically significant association with several outcome variables in one or both Components, such as whether women participate in trade shows of fairs, and whether they implement professional advice during the past two years. Trade shows and fairs were part of the ML component, so it is not surprising to see such large increases in participation relative to those who completed the only BMT component. Given that the TS/AF component provided one-on-one mentoring support to participants, this might imply that women are more likely to implement professional advice when it is given directly to them rather than to a larger training group.

### TABLE 10: NETWORKS – BY ACTIVITY COMPONENT

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| 1 | Likelihood that there are social groups in my community | -0.0967 | -0.0415 |
| | | (0.0826) | (0.0851) |
| 2 | Likelihood of being an active member of any social group | -0.0898 | -0.0279 |
| | | (0.0581) | (0.0666) |
| 3 | Likelihood of participation in Trade Shows or Fairs | 0.300*** | 0.395*** |
| | | (0.0838) | (0.0863) |
| 4 | Likelihood of implementation of professional advice (business adviser, lawyer or accountant) during the past two years | 0.109 | 0.179** |
| | | (0.0719) | (0.0735) |
| 5 | Number of people I can go to ask business advice | 0.273 | 0.382 |
| | | (0.791) | (1.007) |
| 6 | Likelihood that feels sometimes (very) confident negotiating lower prices with suppliers | -0.0758 | 0.00556 |
| | | (0.0737) | (0.0785) |
| 7 | Likelihood that feels sometimes (very) confident negotiating higher prices with buyers | -0.122* | -0.00254 |
| | | (0.0710) | (0.0764) |

Note: Coefficients were obtained by DID regressions with Region and Sector FE. 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 11 presents Business Knowledge and Practices by activity component. The findings in this module are similar to the ones found under the Primary Question. On one hand, some outcome variables provide somewhat encouraging results, but other variables show contradictory findings that are not consistent with the treatment. On a positive note, some specific variables that are statistically significant and yield the expected signs are whether women "do not perform a physical validation of inventory levels," which yields a negative coefficient and is statistically significant for the ML component, and "no accountancy documents prepared annually," which is also negative and statistically significant for both activity Components.

### TABLE 11: BUSINESS KNOWLEDGE AND PRACTICES – BY ACTIVITY COMPONENT

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| 1 | Likelihood that no marketing activities implemented during the last three years | 0.0155 | 0.0190 |
| | | (0.0205) | (0.0320) |
| 2 | Likelihood that made special offers during the last three months | - | - |

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| 3 | Likelihood that does not use internet for marketing purposes or to sell products/services | 0.0389 | -0.0664 |
| | | (0.0534) | (0.0490) |
| 4 | Likelihood that does not formally keep track of business' products and materials | 0.0520 | -0.0568 |
| | | (0.0423) | (0.0577) |
| 5 | Likelihood that does not perform a physical validation of inventory levels | -0.122** | -0.116 |
| | | (0.0495) | (0.0897) |
| 6 | Likelihood that business runs out of inventory at least one time a month | 0.00576 | 0.0389 |
| | | (0.0739) | (0.115) |
| 7 | Likelihood that tries to negotiate a lower price with suppliers during the last three months | 0.0180 | 0.0390 |
| | | (0.112) | (0.0796) |
| 8 | Likelihood that compares price and quality of inputs with other suppliers' products during the last three months | 0.0147 | 0.00350 |
| | | (0.0827) | (0.0410) |
| 9 | Likelihood that has fixed salary for the owner | -0.0466 | 0.0456 |
| | | (0.0977) | (0.116) |
| 10 | Likelihood that records salary of the owner in a notebook, registry or computer | -0.0622 | -0.0574 |
| | | (0.131) | (0.102) |
| 11 | Likelihood that does not keep track of business purchases and sales | -0.0310 | -0.0778 |
| | | (0.0299) | (0.0442) |
| 12 | Likelihood that has a written expense budget | -0.0474 | -0.0474 |
| | | (0.0496) | (0.0496) |
| 13 | Likelihood that has no written goals for next 12 months | 0.00427 | 0.00376 |
| | | (0.0302) | (0.0488) |
| 14 | Likelihood that has no accountancy documents prepared annually | -0.228** | -0.284** |
| | | (0.0876) | (0.0984) |
| 15 | Likelihood that has no changes planned over the next 12 months | -0.0139 | -0.0740* |
| | | (0.0251) | (0.0319) |
| 16 | Likelihood that (strongly) agrees that "my workspace is well organized" | 0.0554 | -0.00170 |
| | | (0.0501) | (0.0492) |
| 17 | Likelihood that (strongly) agrees that "I often communicate clear objectives to my colleagues and employees" | -2.40e-05 | -0.00718 |
| | | (0.0271) | (0.0223) |
| 18 | Likelihood that (strongly) agrees that "I develop work plans at regular intervals" | 0.0289 | -0.0524 |
| | | (0.0531) | (0.0519) |

| Row | Outcome Variable | Component A Treatment Effect (std. error) | Component B Treatment Effect (std. error) |
|---|---|---|---|
| 19 | Likelihood that (strongly) agrees that "I sometimes miss deadlines" | 0.158* | 0.202** |
| | | (0.0712) | (0.0689) |
| 20 | Likelihood that (strongly) agrees that "I believe employees should be treated like family" | -0.0675 | -0.0698** |
| | | (0.0404) | (0.0232) |
| 21 | Likelihood that (strongly) agrees that "I am sometimes late for appointments or meetings" | 0.129 | 0.128 |
| | | (0.113) | (0.177) |
| 22 | Likelihood that (strongly) agrees that "I often attempt to anticipate future circumstances and plan how I/my company will deal with them" | 0.0107 | -0.0183 |
| | | (0.0465) | (0.0557) |
| 23 | Likelihood that (strongly) agrees that "I am constantly collecting information about the market in which my company operates" | 0.00759 | 0.0201 |
| | | (0.0358) | (0.0394) |

Note: Coefficients were obtained by DID regressions with Region and Sector FE. 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.

# QUALITATIVE FINDINGS

In conjunction with the survey activity that was conducted during this evaluation's research phase, six FGDs were also held with WLSME participants as the activity ended. The FGDs were held across five regions and separated by sector, with three focused on participants from the tourism sector and three from the garment sector. The FGDs varied in size from 9 to 14, and comprised of beneficiaries who had participated in the different WLSME activity components (see Table 12 below). The heterogeneity of exposure to the different activity components provided insights on the value of partial activity attendance relative to full activity attendance. All the groups discussed a consistent but open-ended series of questions related to the results of the WLSME activity. These lines of questioning generally corresponded to the four sets of outcome variables; however, the format and venue of the FGDs was intended to provide more opportunities for follow-up questioning and open-ended participant response.

**TABLE 12: FOCUS GROUP SAMPLE DISTRIBUTION**

| Region | Sector | BMT Only | BMT + Market Linkages | BMT+ML+ TS/AF | Total |
|---|---|---|---|---|---|
| Issyk-Kul 1 | Tourism | 0 | 4 | 5 | 9 |
| Issyk-Kul 2 | Tourism | 0 | 8 | 2 | 10 |
| Naryn | Tourism | 3 | 7 | 4 | 14 |
| Bishkek | Garment | 3 | 0 | 9 | 12 |
| Jalal-Abad | Garment | 2 | 3 | 4 | 12 |
| Osh | Garment | 4 | 5 | 9 | 13 |
| Total | | 12 | 27 | 31 | 70 |

The FGDs were recorded as audio files, and the facilitators took notes throughout the meetings, but in deference to the cultural context, the FGDs were not videotaped. As a result, although participant identity was tracked to the best of the facilitator's ability, in some cases it was not possible to ascertain the identity of the individual speakers at certain points in the audio recordings. As a result, precise numbers for individual respondents on specific topics are not always possible, and nonverbal affirmations cannot be identified. In the cases where the transcript is not clear as to whether a single respondent is voicing an opinion twice, or two separate respondents are voicing similar sentiments concurrently, the more conservative numeric estimate is the one cited (e.g., "one respondent voiced an opinion, and at least two other members of the focus group agreed.").

## General Perception

Attitudes towards the WLSME activity were strongly positive across all of the FGDs, which is consistent with the general direction of the quantitative findings. However, some qualifiers were expressed in all three of the tourism FGDs. In particular, one informant felt that the seminars were disorganized, disagreed with elements of the seminar instruction, and thought she saw signs of family favoritism in the selection of participants[7]. No other FGD attendees in this or any other group mentioned this problem. Overall, the most representative comment, one that fully matches with the findings in the quantitative section, is reflected below:

> "It is hard to tell. We have been participating in the project for one year. We absorb all the information, but we cannot apply all of it at once. It is not possible to succeed right after the training. It does not work this way. We have been learning for one year, and now apply our skills and knowledge step by step, and make changes. Our thinking somewhat changed. We apply it now but will see the results in the future. I would not say our financial situation has improved over the year. I think we will see the results in one year."

In fact, this comment fully reflects that while the direction of the quantitative findings is positive, as expected, the limited statistical significance in the findings may be related to the temporal aspect of changing behavior, and that it takes more time to generate change in tasks that require more assertiveness or effort than in others. This may be the reason why, relative to Business Practices, Networks show more consistency and have more outcome variables that are statistically significant. This is also a somewhat encouraging result as positive results may consolidate in the future.

[7] Issyk-Kul Focus Group.

## Specifics: Business Growth

Although all of the aspects were spoken of highly, the aspect of the training that was most commonly cited as being important was systematic planning, particularly through the development of a work plan. Consistently, training participants described work plan development as something they had adopted, but had not previously been engaged in. Some representative comments are the following:

> "I would not say that financial situation got better. It is the same, stable. But we changed our attitude to personnel and finance. We became more serious. We learnt to increase our expenses. We never counted them before. Now everything is clear. We started counting all expenses and income. The attitude to business changed after the trainings."

> "A consultation was a great help to me. A consultant from the U.S. worked with me. She truly helped me and I liked it. First, she made me angry as she provided too much detail. But later I understood that it was useful. The exchange of experience also helped. You start looking at old things and practices from a different perspective."

> "I still have difficulties, as I did not win a grant. I did not receive a certificate and did not write a project proposal. Now I am wonder whether it is possible to participate again and get trained. I look at them and get interested. It turns out that I have missed out many things."

> "We never wrote a business plan before, just talked about it. At the training, we learnt to write a business plan and do our business in general. I only attended two trainings, but a grant would have been very useful, too. It turned out that we did some marketing, but it was haphazard, now we know how to do it on paper."

> "All our business is described in a business plan – how and when we should work. Even if there is no work in autumn, we will use this time to make a research. This is very useful to us. Some trainings were delivered twice, as knowledge is more important than a grant."

## Specifics: Leadership

The impressions of the participants were very good. In particular, the increased confidence and the ensuing respect of friends and family were both cited as clear outcomes from both participation in the activity and the success that followed. Discussions with the various FGD participants revealed a very close relationship between the self-confidence that they gained from participation in the WLSME activity and their work-related success and accomplishments. Participants in all six of the FGDs explicitly stated that their participation in the activity improved their standing in the eyes of their immediate family members, earning greater respect in their day-to-day lives and increased levels of support and encouragement for their sustained and expanded business activities. Similarly, they noticed an increase in their status within their communities. Participants attributed the knowledge and skills they gained from the activity as a key factor in their increased confidence, which in turn helped them to scale up or expanded their business activities.

Some representative comments are the following:

> "When I attended training, my husband complained all the time. Later, when we were preparing documents and after we received a grant, he finally understood that it was a good program and started helping me. Today, he told me I should go to this meeting and say thank you. I got some connections in network marketing. We talked on the phone. They told me I was a smart and intelligent woman. They could tell I was trained. They promised to call me back."

"Family members are more trustful and respectful now. They ask for my advice all the time. My neighbors and friends say they want to participate, too, and become as successful as me."

"My husband let me go anywhere I want, I have become more powerful, a business lady. I tell him that I learnt a lot."

"If [a] mother can achieve success, it means she is still good for something. It is not like when you get old and you are done. Women should aspire to success even if aged. I won a grant. My children respect me.

"I showed the catalog we produced to the Mayor's Office. We have more weight now. They listen to us. For example, if I ask for a land plot, I think they will consider my request."

"People have changed their attitude and now treat me as a leader. They even came and made a video of how I do business. I feel more self-confident now. They also came from Washington. It had effect on my family's attitude to me. They now see me differently."

## Specifics: Networks

FGD attendees provided several helpful insights into how the WLSME activity affected their access to and use of client and peer networks. This was seen in four areas: Knowledge exchange with peers, client referrals from tourism competitors, cultivation of client contacts, and the use of online social-network resources. Interestingly, knowledge exchange with peers, particularly other WLSME participants, was specifically referenced by participants in the focus groups. WLSME activity participants working in the tourism sector appeared to benefit strongly from networking, something that was much less evident in the garment/sewing sector attendees. All three of the tourism sector FGDs specifically cited client-sharing and client referrals with other actors in the tourism business, and increasing their connections regionally. This was part of their increased efforts to connect more broadly within that sector. However, more deliberate cultivation and development of client contacts was common across all of the FGDs. Online social network usage was limited, and was specifically mentioned by only three of the FGD attendees, but more general use of the internet for promotion and networking was mentioned in all of the sessions. Several of the participants mentioned having their success publicly recognized and praised by government institutions. For example, participants of the sewing group had invitations to participate in fashion shows to display their work. In fact, one area that was particularly helpful to activity participants was the opportunity to travel to other institutions to observe and learn from how other operators in their respective sectors conducted their work and serve their customers. This expanded cross-cultural understanding, which was cited as useful in both the tourism and garment sectors.

Some representative comments follow:

"Whatever knowledge we cannot get from the program, we get from one another. We share opinions, information and meet new people."

"We exchange telephone numbers. Gulnara calls me from her guesthouse. Once she told me tourists wanted to see the southern shore and waterfalls. She sent tourists to my place and greet them here. The project taught us how to approach tourists. I accommodate them and cook meals. We know each other's prices. She tells me what tourists like, and what they dislike. This is a good thing. Then I call her and tell that everything is all right. I got several references from her."

"I am a member of the Community-Based Tourism network. I met tourist operators in Supara this year. We exchanged contacts and kept in touch via email. They refer tourists directly to us with no third parties. We got more tourists this year."

"I would also like to say that we met many people from Issyk-Kul oblast at that fair held within the framework of the program. We exchanged our business cards. After the fair, they called us and referred tourists to us. We give them our business cards." "I met many people. I was advised to start a guestbook. Guests from Russia left their feedback. I recorded all the detail about tourists' stay – how many days they stayed, how much money they paid. I also have their contact information. I will call them in April and ask whether they will come again."

"I used to have partners only in Karakol only, now I even have connections in Alai." "I met many people through the website and in WhatsApp. We now distribute our business cards. We exchange experience with cooks and private persons from Bishkek."

"I added tourist operators on social networks."

# CONCLUSIONS

This evaluation report is based on the baseline and first follow-up round for the impact evaluation of the USAID-funded WLSME activity in Kyrgyzstan. The impact evaluation plans to collect two more follow-up rounds of data at 12 months and 24 months post-intervention; thus, this report contains only initial findings that could be observed at the end of the activity.

The short-term findings of this first follow-up are somewhat encouraging yet still inconclusive. Whereas the majority of outcomes considered do not yield a statistically significant link to the WLSME activity, there are rather promising causal links for a small but encouraging number of outcome variables, in particular when using the ANCOVA model. In particular, the WLSME activity has had a positive impact on increasing:

- 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.

Nonetheless, the limited statistical significance in the findings may 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. These quantitative findings are very consistent with the qualitative findings obtained from focus groups.

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

The evaluation team expects that future follow-ups will build on these results given that in a significant number of cases, the signs of the coefficients are consistent with the expected direction of change, which suggests that with more statistical power and/or time of observation, the evaluation team may be able to find improved results. However, given the higher than expected non-response rate, it will be important to assess the viability of the third follow-up round (24 months post-intervention) after the upcoming 12-month follow-up round is completed later this year.

# 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 has 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).

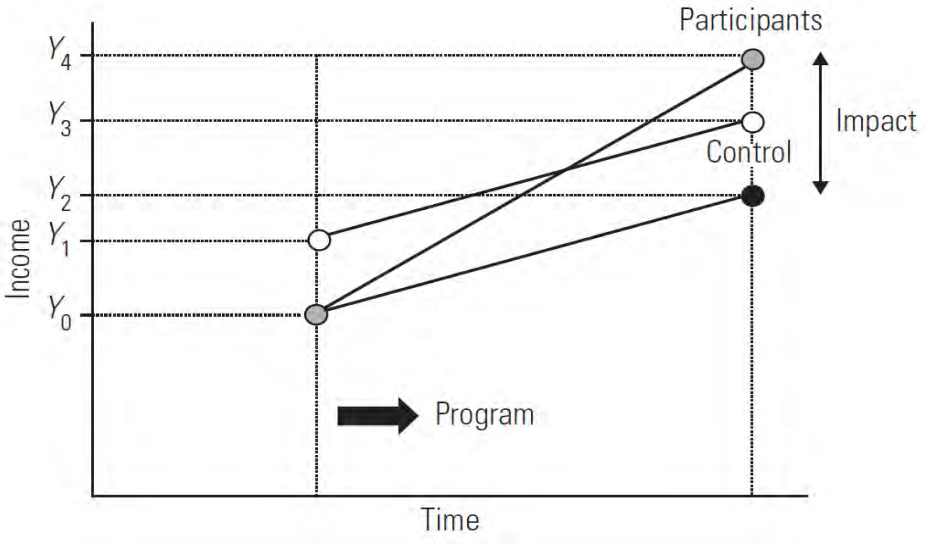
# ANNEX B: EMPIRICAL METHODS

## 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 program 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. In following reports, the evaluation team expects to exploit follow-up data to control time-invariant and unobserved heterogeneity characteristics.

**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 that follows:

$$Y_i = \beta_o + \beta_1 D_i + \beta_2 I_t + \beta_3 D_i I_t + \delta X_i + \varepsilon_i$$

Here, $D_i$ is the treatment status dummy, $I_t$ is the follow-up period dummy, $D_i I_t$ is an interaction term of treatment status and follow-up period, and $X_i$ is a matrix of relevant covariates for identification to increase the efficiency of $\beta_3$. 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. Sector and region fixed effects were also included. In the DID specification, $\beta_3$ is the treatment effect. Two alternative error estimation methods were implemented for the calculation of standard errors: robust errors and clustered errors by region —for the latter region fixed effects were not included. The evaluation team also employs a difference-in-difference panel model, where the same individuals are compared at baseline and endline. Specifically, the DID Panel specification is the following:

$$\Delta Y_i = \beta_0 + \beta_1 D_i + \delta X_i + \varepsilon_i$$

In this case, $\beta_1$ is an unbiased estimator of treatment effect if $E(\varepsilon_i|D_i) = 0$ holds; that is, there is no correlation between the treatment and the error term. Covariates are the same as the first specification. Region and fixed effects are included in all regressions. Given that these results are very similar to the cross-sectional DID estimates, the DID Panel results are found in Annex D.

## ANCOVA

ANCOVA, the evaluation team's preferred method, refers to the Analysis of Covariance which is a statistical method based on variance, multiple regression and correlation analysis 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 is 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 is the following:

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

In this case, $Y_{i,t-1}$ is the baseline value of the outcome variable and $\beta_1$ is the ANCOVA treatment effect. Covariates are the same as those used in the DID model, and region and fixed effects are included as well.

# ANNEX C: DATA COLLECTION AND QUALITY ASSURANCE

## Data Collection Process

Quantitative data collection for the first follow-up endline was carried out by a local survey partner subcontracted by MSI, M-Vector Research and Consulting, 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 the participant. The survey took between 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. A mobile phone card with 200 Soms (equivalent of $2.70 USD) was provided to respondents after the survey was completed as a token of appreciation. Prior to the start of data collection, the survey was pre-tested 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 Analytis and Evaluation Project took place at the end of the activity, between August and October 2015, with 81 percent response rate. 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 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 reach (9 percent), refusal to participate (6 percent), unavailable after multiple rescheduled appointments (2 percent), passed away (1 percent), and moved abroad (1 percent). The evaluation team addressed balance among treatment groups and the non-response rate in the report.

Qualitative data collection was led by the local qualitative researcher from the E3 Analytics and Evaluation Project with logistical support from M-Vector. A subsample of WLSME participants from each selected region and sector were randomly selected to participate in the FGDs. Recruitment accounted for a no-show rate of 50 percent and variability of participation across the activity components. The FGDs were held in a convenient and central location within each region. Participants' transportation expenses were covered and light refreshments were provided. Prior to the start of each FGD, each participant received and signed an informed consent form ensuring confidentiality and voluntary participation. FGDs lasted 90 minutes and consisted of open-ended guided questions. Each FGD was audio recorded and transcribed verbatim into Russian, and then translated into English. Translations were audited independently by another member of the evaluation team.

## Data Quality Assurance

Standard operating procedures for data collection were followed by the survey firm, including verification procedures conducted both at the site and at headquarters in Bishkek, double entry of survey data, and data query. Specifically, the survey firm employed the following set of quality control procedures:

- 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 of the quality of data collected:

- 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 80 completed surveys (11 percent). 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 (6 percent). This audit showed no meaning 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 PANEL RESULTS

### TABLE 13: BUSINESS GROWTH OUTCOMES – DID PANEL

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| 1 | Average sales on a good month (in ranges)+ | 0.0453 |
| | | (0.182) |
| 2 | Average sales on a bad month (in ranges)+ | -0.0383 |
| | | (0.143) |
| 3 | Average sales on an average month (in ranges)+ | -0.0415 |
| | | (0.147) |
| 4 | Average profit on a good month (in ranges)+ | 0.379 |
| | | (0.253) |
| 5 | Average profit on a bad month (in ranges)+ | 0.144 |
| | | (0.127) |
| 6 | Average profit on an average month (in ranges)+ | 0.0802 |
| | | (0.132) |
| 7 | Number of good months in the last year | 0.131 |
| | | (0.268) |
| 8 | Number of bad months in the last year | 0.28 |
| | | (0.214) |
| 10 | Sales in the last 12 months (soms) | -6,652 |
| | | (20286) |
| 11 | Likelihood that has a bank account for business purposes only | -0.0119 |
| | | (0.0417) |
| 12 | Number of people from household that worked in the business in the last 12 months | 0.186 |
| | | (0.127) |
| 13 | Number of household people that worked in the business and were remunerated with cash | 0.191 |
| | | (0.215) |
| 14 | Number of people from outside household that worked in the business in the last 12 months | -0.287 |
| | | (0.589) |
| 15 | Number of non-household people that worked in the business and were remunerated with cash | -0.0811 |
| | | (0.588) |
| 16 | Number of months per year spent working in the business you own or manage | 0.0933 |
| | | (0.285) |
| 17 | Number of days per week spent working in the business you own or manage | 0.130 |
| | | (0.141) |
| 18 | Number of hours per day spent working in the business you own or manage | 0.119 |
| | | (0.232) |
| 19 | Likelihood of purchasing raw materials, goods, or equipment for business with a loan, in the last 12 months | 0.109** |
| | | (0.0496) |
| 20 | Likelihood of applying for a loan from a financial institution in the | 0.0978 |

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| | last 12 months | |
| | | (0.0599) |
| 21 | Likelihood of loan approval | -0.00567 |
| | | (0.0967) |

Note: Coefficients were obtained by DID Panel regressions with Region and Sector FE. 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 PANEL

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| 1 | Likelihood that prefers to work as an employee in a business instead of managing/owning one | -0.0119 |
| | | (0.0417) |
| 2 | Likelihood that people ask me for business advice (very) often | -0.0376 |
| | | (0.0547) |
| 3 | Likelihood that I (with my partner/spouse or another household member) am in charge of general business planning decisions | -0.00894 |
| | | (0.0466) |
| 4 | Likelihood that I (with my partner/spouse or another household member) decide what inputs to buy for production | -0.00374 |
| | | (0.0497) |
| 5 | Likelihood that I (with my partner/spouse or another household member) am in charge of sales and client relations | -0.0703 |
| | | (0.0479) |
| 6 | Likelihood that I (with my partner/spouse or another household member) decide if I should apply for a loan | -0.0283 |
| | | (0.0529) |
| 7 | Likelihood that I (with my partner/spouse or another household member) decide my own singular wage | -0.116 |
| | | (0.1) |
| 8 | Likelihood that I (with my partner/spouse or another household member) decide what type of work I will do | -0.00255 |
| | | (0.0468) |
| 9 | Likelihood that I (with my partner/spouse or another household member) am in charge of marketing and advertising decisions | 0.00171 |
| | | (0.0512) |
| 10 | Likelihood that I (with my partner/spouse or another household member) am in charge of staffing of business decisions | -0.00393 |
| | | (0.0511) |
| 11 | Likelihood that often (or always) makes general business planning decisions without consulting anyone else | 0.0103 |

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| | | (0.0521) |
| 12 | Likelihood that often (or always) decides what inputs to buy for production without consulting anyone else | 0.0012 |
| | | (0.0526) |
| 13 | Likelihood that often (or always) manages sales and client relations without consulting anyone else | 0.0806 |
| | | (0.0531) |
| 14 | Likelihood that often (or always) decides whether to apply for a loan without consulting anyone else | -0.000285 |
| | | (0.0579) |
| 15 | Likelihood that often (or always) decides own singular wage without consulting anyone else | 0.00318 |
| | | (0.0576) |
| 16 | Likelihood that often (or always) decides what type of work will do without consulting anyone else | -0.0361 |
| | | (0.0575) |
| 17 | Likelihood that often (or always) makes marketing and advertising decisions without consulting anyone else | 0.120** |
| | | (0.0558) |
| 18 | Likelihood that often (or always) makes staffing of business decisions without consulting anyone else | 0.0326 |
| | | (0.0515) |
| 19 | Likelihood that (strongly) agrees that "women should do what men say" | -0.0433 |
| | | (0.0464) |
| 20 | Likelihood that (strongly) agrees that "Women must share their income with their husbands" | 0.0623 |
| | | (0.0392) |
| 21 | Likelihood that (strongly) agrees that "it is OK if men abandon women if they wish to" | -0.0158 |
| | | (0.0434) |
| 22 | Likelihood that (strongly) agrees that "it is OK if men chide women because they went out without any permission" | 0.0199 |
| | | (0.0506) |
| 23 | Likelihood that (strongly) agrees that "it is OK if men chide women if they do not take care of children" | 0.0404 |
| | | (0.0395) |
| 24 | Likelihood that (strongly) agrees that "the role of women is to earn money and take care of her family" | -0.0735 |
| | | (0.0539) |
| 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" | 0.0443* |
| | | (0.0236) |
| 26 | Likelihood that (strongly) agrees that "father's and mother's dedication is equally important for the learning and effective development of children" | 0.0104 |
| | | (0.0206) |
| 27 | Likelihood that (strongly) agrees that there are no gender inequality problems in my community | 0.00684 |

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| | | (0.0384) |

Note: Coefficients were obtained by DID Panel regressions with Region and Sector FE. 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 15: NEWORKS OUTCOMES – DID PANEL

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| 1 | Likelihood that there are social groups in my community | 0.0543 |
| | | (0.052) |
| 2 | Likelihood of being an active member of any social group | 0.0273 |
| | | (0.0597) |
| 3 | Likelihood of participation in Trade Shows of Fairs | 0.0407 |
| | | (0.0572) |
| 4 | Likelihood of implementation of professional advice (business adviser, lawyer or accountant) during the past two years | 0.0822* |
| | | (0.0457) |
| 5 | Number of people I can go to ask business advice | 1.076*** |
| | | (0.401) |
| 6 | Likelihood that feels sometimes (very) confident negotiating lower prices with suppliers | 0.0664 |
| | | (0.048) |
| 7 | Likelihood that feels sometimes (very) confident negotiating higher prices with buyers | -0.0337 |
| | | (0.0468) |

Note: Coefficients were obtained by DID Panel regressions with Region and Sector FE. 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 PANEL

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| 1 | Likelihood that no marketing activities implemented during the last three years | 0.0077 |
| | | (0.0315) |
| 2 | Likelihood that made special offers during the last three months | 0.0311 |
| | | (0.0509) |
| 3 | Likelihood that does not use internet for marketing purposes or to sell products/services | -0.0256 |
| | | (0.037) |
| 4 | Likelihood that does not formally keep track of business' | 0.0596* |

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| | products and materials | |
| | | (0.0332) |
| 5 | Likelihood that does not perform a physical validation of inventory levels | 0.0375 |
| | | (0.0439) |
| 6 | Likelihood that business runs out of inventory, at least one time a month | -0.0324 |
| | | (0.0546) |
| 7 | Likelihood that tried to negotiate a lower price with suppliers during the last three months | 0.0499 |
| | | (0.0513) |
| 8 | Likelihood that compared price and quality of inputs with other suppliers' products during the last three months | -0.0557 |
| | | (0.0416) |
| 9 | Likelihood that fixed salary for the owner | -0.00361 |
| | | (0.0512) |
| 10 | Likelihood that records salary of the owner in a notebook, registry or computer | -0.00361 |
| | | (0.0512) |
| 11 | Likelihood that does not keep track of business purchases and sales | 0.0369 |
| | | (0.0347) |
| 12 | Likelihood that has a written expense budget | -0.0476 |
| | | (0.0458) |
| 13 | Likelihood that has no written goals for next 12 months | -0.0806** |
| | | (0.0379) |
| 14 | Likelihood that has no accountancy documents prepared annually | 0.00639 |
| | | (0.0462) |
| 15 | Likelihood that has no changes planned over the next 12 months | 0.00581 |
| | | (0.0265) |
| 16 | Likelihood that (strongly) agrees that "my workspace is well organized" | 0.0356 |
| | | (0.0328) |
| 17 | Likelihood that (strongly) agrees that "I often communicate clear objectives to my colleagues and employees" | -0.00278 |
| | | (0.0208) |
| 18 | Likelihood that (strongly) agrees that "I develop work plans at regular intervals" | 0.0168 |
| | | (0.0265) |
| 19 | Likelihood that (strongly) agrees that "I sometimes miss deadlines" | -0.0509 |
| | | (0.0457) |
| 20 | Likelihood that (strongly) agrees that "I believe employees should be treated like family" | -0.0509 |
| | | (0.0457) |
| 21 | Likelihood that (strongly) agrees that "I am sometimes late for appointments or meetings" | -0.0646 |
| | | (0.05) |
| 22 | Likelihood that (strongly) agrees that "I often attempt to anticipate future circumstances and plan how I/my company will | -0.00362 |

| Row | Outcome Variable | Treatment Effect (std. error) |
|---|---|---|
| | deal with them" | |
| | | (0.0195) |
| 23 | Likelihood that (strongly) agrees that "I am constantly collecting information about the market in which my company operates" | 0.00612 |
| | | (0.0204) |

Note: Coefficients were obtained by DID Panel regressions with Region and Sector FE. 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: SUMMARY OF FOCUS GROUPS

| | Bishkek - Garment | Issyk-Kul 2 - Tourism | Naryn - Tourism | Issyk-Kul 1 - Tourism | Jalal-Abad - Garment | Osh - Garment |
|---|---|---|---|---|---|---|
| General Comments | All positive. | All positive except for three, all of which were unhappy about not having their project proposals approved, but in particular were frustrated at not getting clear feedback on why those proposals were not accepted. One of those also felt that the seminars were disorganized, disagreed with seminar instruction, and showed signs of favoritism towards family, though no other participants identified this problem. | All positive, but one did not understand why her grant proposal was not accepted | All positive, but one did not understand why her grant proposal was not accepted | All positive | All positive |

| | Bishkek - Garment | Issyk-Kul 2 - Tourism | Naryn - Tourism | Issyk-Kul 1 - Tourism | Jalal-Abad - Garment | Osh - Garment |
|---|---|---|---|---|---|---|
| Most useful aspect of the WLSME activity | • learning to write a business plan and sharing experience<br>• sharing experience<br>• consultation<br>• seminars and grant support<br>• consultation<br>• seminars on financial management | • practical training<br>• tract training<br>• marketing and experience exchange<br>• technical knowledge and increased self-confidence<br>• networking | • exchange of ideas<br>• exchange of ideas<br>• exchange of ideas<br>• the spread out pace of the activity, and the exchange of ideas<br>• networking | • exchange of ideas and Internet exposure<br>• exchange of ideas<br>• exchange of ideas and Internet exposure<br>• individual consultation and Internet exposure<br>• consultation and exchange of ideas<br>• exchange of ideas<br>• exchange of ideas | • practical training<br>• practical training in exchange of ideas<br>• private consultation<br>• grant<br>• grant<br>• grant<br>• grant and training<br>• grant<br>• exchange of experience<br>• business planning | • marketing training<br>• greater awareness of demand<br>• technical training<br>• grant<br>• business planning and marketing<br>• business planning<br>• increased technical knowledge<br>• technical training<br>• marketing<br>• technical training<br>• exchange of ideas<br>• marketing training |

| | Bishkek - Garment | Issyk-Kul 2 - Tourism | Naryn - Tourism | Issyk-Kul 1 - Tourism | Jalal-Abad - Garment | Osh - Garment |
|---|---|---|---|---|---|---|
| Business improvement since the WLSME activity | • Optimistic but no change<br>• optimistic but no change<br>• smart economic decisions (liquidating stock)<br>• made smarter economic decisions (liquidating stock)<br>• smart economic decisions (monitoring expenditure and adapting production methods) seeing positive changes<br>• nothing has changed – do not feel any improvements<br>• too soon to tell, but optimistic<br>• making smarter economic decisions – not better, but surviving negative economic situation. | • Training and marketing helped improve the financial situation<br>• Won a grant and improved economic management<br>• learned how to use the Internet to attract customers<br>• improved economic management<br>• Won a grant | • improved faster than would have otherwise and better record keeping<br>• better record keeping and economic management<br>• better record keeping and economic management<br>• better record keeping and improved compliance | • small improvement due to the activity<br>• improved economic decision-making and increase business from referrals<br>• increased business from referrals and advertising<br>• improved economic decision-making<br>• increased income thanks to expanded activity inspired by the activity<br>• improved economic decision-making | • improved thanks to the activity<br>• business has expanded thanks to the activity<br>• no longer reliant on loans and planning expansion thanks to increased sales and marketing plan<br>• no improvement due to not applying gained knowledge yet<br>• some improvements through increased knowledge<br>• minor improvements, but unable to afford expansion at this time<br>• improved due to expansion thanks to activity support and planning<br>• expanded using grant funds<br>• expanded using grant funds | • improved thanks to increase productivity<br>• improved economic decision-making<br>• expansion of business<br>• improved economic decision-making |

# ANNEX F: 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.

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.

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.