



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,
Education and Research EAER
State Secretariat for Economic Affairs SECO



100
1919-2019
SOCIAL JUSTICE
DECENT WORK

DOES BUNDLING
CREDIT AND
ENTERPRISE
DEVELOPMENT
SERVICES PROMOTE
GROWTH OF
MICRO AND SMALL
ENTERPRISES?



**Evidence from a randomized
control trial in Indonesia**

DOES BUNDLING CREDIT AND
ENTERPRISE DEVELOPMENT SERVICES
PROMOTE GROWTH OF MICRO AND
SMALL ENTERPRISES?

**Evidence from a randomized
control trial in Indonesia**

This impact evaluation was by conducted by Alexandra Avdeenko¹, Markus Frölich^{1,2,3,4}, Simona Helmsmüller^{1,2} with research support by Ulugbek Aminjonov¹, Adelina Garamow⁵



Copyright © International Labour Organization 2019

First published 2019

Publications of the International Labour Office enjoy copyright under Protocol 2 of the Universal Copyright Convention. Nevertheless, short excerpts from them may be reproduced without authorization, on condition that the source is indicated. For rights of reproduction or translation, application should be made to ILO Publications (Rights and Permissions), International Labour Office, CH-1211 Geneva 22, Switzerland, or by email: pubdroit@ilo.org. The International Labour Office welcomes such applications.

Libraries, institutions and other users registered with reproduction rights organizations may make copies in accordance with the licences issued to them for this purpose. Visit www.ifro.org to find the reproduction rights organization in your country.

Does bundling credit and enterprise development services promote growth of micro and small enterprises? Evidence from a randomized control trial in Indonesia, Kantor Jakarta; ILO, 2019

viii; 38 p.

ISBN: 978-92-2-031586-6 (print)
978-92-2-031587-3 (web pdf)

ILO Cataloguing in Publication Data

The designations employed in ILO publications, which are in conformity with United Nations practice, and the presentation of material therein do not imply the expression of any opinion whatsoever on the part of the International Labour Office concerning the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its frontiers.

The responsibility for opinions expressed in signed articles, studies and other contributions rests solely with their authors, and publication does not constitute an endorsement by the International Labour Office of the opinions expressed in them.

Reference to names of firms and commercial products and processes does not imply their endorsement by the International Labour Office, and any failure to mention a particular firm, commercial product or process is not a sign of disapproval.

ILO publications and electronic products can be obtained through major booksellers or ILO local offices in many countries, or direct from ILO Publications, International Labour Office, CH-1211 Geneva 22, Switzerland. Catalogues or lists of new publications are available free of charge from the above address, or by email: pubvente@ilo.org

Visit our website: www.ilo.org/publns

This report was prepared by Owais Parray, Chief Technical Advisor (Economist), ILO Jakarta. The ILO would like to thank the University of Padjajaran (UNPAD) for carrying out the survey.

Printed in Indonesia

Preface



Acknowledgements

We would like to thank the University of Padjajaran (UNPAD) and Institut Teknologi Sepuluh Nopember (ITS) for conducting the baseline and endline surveys. This work would not have been possible without the key contributions of the ILO team, in particular Owais Parray, Muce Mochtar, Yanis Saputra, Agustinus Simon Petrus Siregar, and Yousra Hamed. The ILO team has been supportive in every step of the impact evaluation, thoroughly following up with the implementing institutions, providing us with access to the monitoring data, quick updates from the field, and most importantly reviewing the questionnaires and providing us with further explanations for the results and ideas for further analysis.

In Mannheim, excellent research assistance and intellectual input was provided by Peter Derheim, Lisa Jahn, and Asmus Zoch. A special thanks goes to Asmus Zoch, who participated in the inception workshop of this impact evaluation and provided excellent input to the baseline survey. Finally, to Tatjana Kulp, an excellent research manager, who led the endline data collection and supported the monitoring of implementation.

Executive Summary

The fact that only few entrepreneurs make the transition from micro to small and from small to medium enterprises suggests that there may be a whole range of factors inhibiting growth of small businesses. Access to finance is one of the key constraints faced by small businesses. Besides finance, weak managerial skills and business practices also affect productivity of small businesses.

In Indonesia, among more than 59 million businesses, 99.8% are classified as micro and small enterprises (MSEs). Despite a stable macroeconomic environment, light taxation, large and segmented financial sector, many MSEs operate at low productivity levels.

Typically, it is assumed that once there is adequate management capacity, small businesses can access credit enabling them to reinvest and move towards a higher growth path. However, there is limited evidence to show that this is indeed the case. An important question, therefore, is whether after becoming “bankable” do small businesses continue to grow?

This question can also be framed from a supply side: can FSPs, at least those that tend to have a development mandate, do more than just lend to small businesses to create greater social and economic value?

Working with 13 financial service providers (FSPs) in Indonesia, the International Labour Organization (ILO) tested this hypothesis using a randomized control trial (RCT) in a pilot project that helped FSPs design “bundled” loan and enterprise development services.

In the RCT, the bundled services were tested with over 5,000 clients in two provinces of Indonesia. As part of the intervention, loan officers first went through a training of trainers (TOTs) and training of counselors (TOCs) training courses. More than 170 loan officers were trained to become trainers and/or counselors. Trained loan officers then delivered a two-day classroom training and/ or carried out five individual counseling sessions with the clients. It was expected that training, counseling, or a combination of the two would improve the performance of the businesses.

Results were compared between clients who received bundled services (treatment) and those who only had a loan from the FSP (control). The theory of change for the intervention articulated that enterprise development would increase knowledge, which would then lead to a behavioural change i.e. improved business practices. Over time, it was expected that gains in knowledge and improved practices would increase revenue and profit.

FSPs that participated in the testing were savings and credit cooperatives (KSPs), rural banks (BPRs), and regional development banks (BPDs). The client base of the FSPs varied substantially from 1,000 to over 400,000 active borrowers. The proportion of female borrowers also varied from 30% to 100%. In all the FSPs, the number of active depositors exceeded the number of borrowers. In general, all the selected FSPs have a dual mission, i.e. while pursuing profitability they also aim to contribute towards social welfare of their clients.

The impact evaluation included collecting detailed information on client background, practices, and challenges before the pilot started. On average, two out of three interviewed clients were female and the average household comprised of three adults, of whom two earned an income. The majority of businesses were small enterprises without written contracts for their workers, operating from the same location where the family of the owner lives, and mainly selling their products locally.

More than 70% of clients reported having some marketing-related strategies in place, but less than 40% prepared a business plan or kept records of all the transactions. The most common loan type was a



business/individual loan and the average client usually took one loan per year from the respective FSP. On average, every fifth client reported being late making loan repayments. More than 70% of the clients planned to borrow more for their businesses.

In the endline survey, the research team collected self-reported data from the clients. Almost two thirds of the respondents reported participating in training or counseling. The most common topics discussed during the training or counseling were marketing and financial management. This is in line with the objectives of the program and confirmed that the services provided were consistent with the design of the intervention.

Client satisfaction with the pilot amongst participants was high: approximately 87% of the respondents considered training helpful, while 80% reported the same for counseling sessions. The self-reported data on participation was validated through monitoring data collected by the loan officers during the testing phase. The loan officers were requested to complete a short questionnaire after each training/counseling session. This information was available for half of the clients who were assigned to participate in the pilot. On average 16 clients participated in each training session. Around four other persons were present at the training sessions, who were neither clients nor FSP staff.

Each counseling session lasted (on average) 32 minutes. The most common reason for not participating in the pilot was that the client could not leave his/her businesses unattended, followed by client being busy with household chores and childcare. In 59% of the cases, the loan officer and the client agreed on a business plan during the counseling sessions. In every fourth counseling session, family members of the client were present.

Clients regarded the technical assistance provided by the FSPs as useful. The loan officers were asked to rate the interest of the group (training) or individual (counseling) on a scale of 1 to 10: with 1 showing very low interest and 10 showing very high interest. The average score was 8 and 7 for training and counseling respectively.

While the evidence reported by the clients and loan officers suggested that the intervention was very well received, the impact evaluation went a step further to compare changes in knowledge, practices, and business performance between the clients in the treatment and the control group. Two thirds of the clients were assigned to the treatment group while one third of the total were assigned to the control. The latter did not receive any additional technical assistance.

The randomization of the clients into groups ensured that the clients assigned to the treatment and the control group were, on average, comparable across a wide range of pre-pilot characteristics. Thus, any differences after the intervention could be attributed to the pilot. Additionally, in some FSPs we assigned clients to one of the three treatments arms: two-day classroom training, individual counseling, or both. The purpose was to see which of the three treatments yielded better results. The endline survey included panel data collected from 4,292 clients, i.e. the same clients who were interviewed at the baseline and then again at the endline.

Before turning to the results, it is important to note that the timeframe for the implementation of the pilot and the endline survey was short. For several FSPs, the endline survey was conducted immediately after the last training and counseling session. As such, changes in the final outcomes (such as revenue and profit) were not expected. Additionally, implementation of the pilot varied by FSPs. Several FSPs were not able to deliver the minimum number of training and counselling sessions.

Furthermore, it is likely that there was some degree of spillover to the control group i.e. some clients in the control group may have been targeted or happened to learn from those clients in their neighbourhood who participated in (this or similar) training or counselling sessions. Moreover, some clients in the control group may have benefited from trained loan officers who served clients in both groups. Naturally, these factors may have affected the estimation when comparing clients in the treatment and the control group.

Still, at the level of intermediate outcomes, there is sizable increase in the share of clients who, after the intervention, include a cash flow analysis in their business plans. Given the importance of loan repayments and cashflow for a business, this result is meaningful and can mutually benefit clients and the FSPs. The increase in use of cashflow is higher for clients who received counseling. Providing counseling is a more cost-effective intervention compared to classroom training. Renting a classroom often exceeds the loan officer's basic monthly wage.

An important finding of our study is that results varied across FSPs. In terms of aggregate changes in the final (impact) outcomes we do not detect any significant changes yet. In other words, when data from all the FSPs are combined together, there is little or no evidence of changes in key indicators. However, we can see several positive effects - both at the intermediate and final outcomes - for different sub-groups.

We report selected results only for those outcomes that have a statistically significant coefficient for at least one subgroup of clients. For example, we observe that clients from the trading sector had a higher likelihood (5.9 percentage points to others) to know their payment type (difference significant at the 0.1% significance level).

Clients served by one sub-group of FSPs had a higher likelihood (9.6 percentage points) of preparing business or financial plan. In one of the FSPs, the likelihood of preparing business or financial plan was even much higher (18.6 percentage points). A larger number of clients now include cash flow, including for clients of FSPs with above average amount of loans disbursed in the last 12 months (difference of 4.7 percentage points to other FSPs), FSPs with above average number of active borrowers (by 5.4 percentage points), and clients aged between 35 and 50 (by 6.5 percentage points).

Results under the final (impact) outcomes are significant mostly at the 1% significance level. We observed that the revenue of clients served by one of the FSPs increased by Indonesian Rupiah (IDR) 22.331 million (USD 1,595), profit generated in the last 30 days was higher by IDR 1.667 million (USD 119) for treated clients who initially reported tough competition as the main business barrier.

Furthermore, the same subgroup also experienced higher profit in the worst month by IDR 1.157 million (USD 82) and higher profit in the best month by IDR 2.540 (USD 181). As mentioned, while overall (aggregated by 13 FSPs) there are no effects on the likelihood that profits increased, we find that clients from one of the FSPs were likely to benefit more by 16.5 percentage points. Also, clients who (at baseline) cited managing business as the main business barrier also were likely to benefit more by 16.6 percentage points.

On a scale from 1-10, life satisfaction improved by up to 0.3 for clients with primary education. Also, life satisfaction improved by 0.775 for treated clients having no education, and by 0.34 for clients from BPDs. A number of clients now show a smaller likelihood to be currently in loan default or behind with the repayment of any loans. In particular, clients who at baseline reported that business brings high income show a reduction in loan default rate by 7.2 percentage points), clients from the manufacturing sector show a reduction of 9 percentage points, and clients with more personnel support show a reduction of 11.4 percentage points. Lastly, clients who already prepared business plans at the time of baseline were 5.3 percentage points less likely to report late repayment.

JEL Classification: O17, L26, O12, D22

Keywords: entrepreneurship training, self-employment skills, microfinance, micro and small enterprises, training, program evaluation, randomized control trial



Table of Contents

Foreword	iii
Acknowledgement	iv
Executive Summary	v
1 Introduction	1
2 The Intervention and Intended Effects	5
2.1 Relevance and Challenges of Micro and small enterprises in Indonesia	5
2.2 The Intervention	5
2.3 Intended Effects	6
3 Experimental Design and Treatment Implementation	9
3.1 Participating FSPs and clients' baseline characteristics	9
3.2 Randomization	13
3.3 Implementation fidelity	14
4 Data and Estimation Method	23
4.1 Survey data	23
4.2 Econometric approach	25
5 Main Results	27
5.1 Intermediate and final outcomes	27
5.2 Comparison across treatment arms: Training vs. counseling	39
6 Heterogeneous Effects	45
7 Robustness Checks	61
8 Cost Benefit Analysis	67
8.1 Financial cost benefit analysis	68
8.2 Social cost benefit analysis	70
9 Conclusion	73
References	76

- A Additional Information on the FSPs and the Program
 - A.1 Course outline
 - A.2 Descriptive statistics from online monitoring data
- B Additional Information on the Impact Evaluation
 - B.1 Pre-analysis plan: Hypotheses
 - B.2 Variables
 - B.3 Additional information on the randomization
 - B.4 Additional information on the survey data collection and data cleaning
- C Additional Background Statistics and Results
 - C.1 Balance Statistics
 - C.2 Summary Statistics for the Clients Reporting Participation in the Program
 - C.3 Local Average Treatment Effects
 - C.4 Additional Results
 - C.5 Results for Selected FSPs (FSPs 1, 2, 3, and 5)
 - C.6 Additional Charts on Program Participation (Disaggregated)
 - C.7 Correlation of Actual Participation and Outcomes



List of Figures

1	Timeline of Baselines and Randomization	19
2	Assignment of Clients to Different Treatment Arms per Batch	23
3	Reported Implementation Fidelity by FSP: Monitoring and Survey Data Evidence	25
4	Distribution of Endline Survey Respondents over FSPs	27
5	Program Participation by Treatment Arms: Evidence from the Endline Survey	28
6	Reported Reasons for Non-Participation in the Program	29
7	Topics Discussed During Training or Counseling Sessions	31
8	Respondent Opinion on the Helpfulness of Classroom Training and Counseling Sessions	32
9	Respondent Recommendation of Classroom Training and Counseling Sessions	33
10	Usefulness of Classroom Training and Counseling Sessions	34
11	Location of Clients	36
12	Answers to Knowledge Questions in Treatment and Control Group	42
13	Business Practices in Treatment and Control Group	44
14	Profits, Revenues, Costs and Number of Workers in Treatment and Control Group	45
15	Increased Profits and Coverage of Business Expenses in Treatment and Control Group	46
16	Entrepreneurial Attitudes and Business Start-ups in Treatment and Control Group	47
17	Household Outcomes in Treatment and Control Group	48
18	Loan Behavior Treatment and Control Group	49
A.1	Program Participation by Treatment Arms and Gender: Evidence from the Endline Survey	173
A.2	Program Participation by Treatment Arms and Age: Evidence from the Endline Survey	174
A.3	Program Participation by Treatment Arms and Education: Evidence from the Endline Survey	175
A.4	Program Participation by Treatment Arms, Old and New Businesses: Evidence from the Endline Survey	176
A.5	Program Participation by Treatment Arms and Business Size: Evidence from the Endline Survey	177
A.6	Reported Reasons for Non-Participation in the Program by Gender	178
A.7	Reported Reasons for Non-Participation in the Program by Age Groups	179
A.8	Reported Reasons for Non-Participation in the Program by Education	180
A.9	Reported Reasons for Non-Participation in the Program by Old and New Businesses	181
A.10	Reported Reasons for Non-Participation in the Program by Business Size	182
A.11	Usefulness of Classroom Training and Counseling Sessions by Gender	
A.12	Usefulness of Classroom Training and Counseling Sessions by Age Groups	184
A.13	Usefulness of Classroom Training and Counseling Sessions by Education	185
A.14	Usefulness of Classroom Training and Counseling Sessions by Old and New Businesses	186
A.15	Usefulness of Classroom Training and Counseling Sessions by Business Size	187

List of Tables

1	Baseline Characteristics of Estimation Sample (All FSPs)	20
2	Intermediate Outcomes	40
3	Final Outcomes	41
4	Intermediate Outcomes by Treatment Arm (Batch I)	51
5	Final Outcomes by Treatment Arm (Batch I)	52
6	Effects by Entrepreneur and Business Characteristics (All FSPs)	54
7	Effects by FSP Characteristics (All FSPs)	55
8	Selected Heterogenous Effects - Intermediate Outcomes	59
9	Selected Heterogenous Effects - Final Outcomes	62
10	Robustness Checks: All FSPs	67
11	Financial Costs and Benefits	71
12	Social Costs and Benefits	74
A.1	Key Characteristics of FSPs - Batch I	82
A.2	Key Characteristics of FSPs - Batch II and Batch III	84
A.3	Summary Statistics on Program Implementation (Online Monitoring Data).....	87
A.4	Overview of Hypotheses on Average Treatment Effects	91
A.5	Variables	100
A.6	Balance Tests	134
A.7	Characteristics of the Clients who Report Having Participated in the Program.....	139
A.8	All Outcomes - Local Average Treatment Effects (All FSPs)	140
A.9	Effects by Entrepreneur and Business Characteristics (All FSPs)	142
A.10	Effects by FSP Characteristics (All FSPs)	143
A.11	Intermediate Outcomes by Treatment Arms - Robustness	144
A.12	Final Outcomes by Treatment Arms - Robustness	145
A.13	Effects by Entrepreneur and Business Characteristics (FSP 8)	146
A.14	Intermediate Outcomes by FSPs	147
A.15	Final Outcomes by FSPs	148
A.16	Household Income and Expenditure	149
A.17	Effects by Entrepreneur and Business Characteristics I - Intermediate Outcomes (All FSPs)..	150
A.18	Effects by Entrepreneur and Business Characteristics I - Final Outcomes (All FSPs)	151
A.19	Effects by Entrepreneur and Business Characteristics II - Intermediate Outcomes (All FSPs)	152
A.20	Effects by Entrepreneur and Business Characteristics II - Final Outcomes (All FSPs)	153
A.21	Effects by Entrepreneur and Business Characteristics III - Intermediate Outcomes (All FSPs)	154
A.22	Effects by Entrepreneur and Business Characteristics III - Final Outcomes (All FSPs)	155
A.23	Effects by Reported Barriers to Business I - Intermediate Outcomes (All FSPs)	156
A.24	Effects by Reported Barriers to Business I - Final Outcomes (All FSPs)	157
A.25	Effects by Reported Barriers to Business II - Intermediate Outcomes (All FSPs)	158
A.26	Effects by Reported Barriers to Business II - Final Outcomes (All FSPs)	159



A.27	Intermediate Outcomes (FSPs 1, 2, 3 and 5)	161
A.28	Final Outcomes (FSPs 1, 2, 3 and 5)	162
A.29	Intermediate Outcomes by Treatment Arm (FSPs 1, 2, 3 and 5)	163
A.30	Final Outcomes by Treatment Arm (FSPs 1, 2, 3 and 5)	164
A.31	Intermediate Outcomes - Local Average Treatment Effects (FSPs 1, 2, 3 and 5)	165
A.32	Final Outcomes - Local Average Treatment Effects (FSPs 1, 2, 3 and 5)	166
A.33	Effects by Entrepreneur and Business Characteristics I - Intermediate Outcomes (FSPs 1, 2,3 and 5)	167
A.34	Effects by Entrepreneur and Business Characteristics I - Final Outcomes (FSPs 1, 2, 3 and 5)	168
A.35	Effects by Entrepreneur and Business Characteristics II - Intermediate Outcomes (FSPs 1, 2, 3 and 5)	169
A.36	Effects by Entrepreneur and Business Characteristics II - Final Outcomes (FSPs 1, 2, 3 and 5)	170
A.37	Effects by Entrepreneur and Business Characteristics III - Intermediate Outcomes (FSPs 1, 2, 3 and 5)	171
A.38	Effects by Entrepreneur and Business Characteristics III - Final Outcomes (FSPs 1, 2, 3 and 5)	172
A.39	Intermediate Outcomes: Naive Estimation	188
A.40	Final Outcomes: Naive Estimation	189

List of Abbreviations

AEA	American Economic Association
C	consulting treatment arm
Ctrl	control group
CAPI	computer assisted personal interviews DK do not know
FSP	micro-Financial Service Provider
GDP	gross domestic product
HH	household
IDR	Indonesian Rupiah
ILO	International Labour Organization
ITT	intention-to-treat
LATE	local average treatment effect MDE minimum detectable effect
MIS	Management Information System
MSEs	Micro and Small Enterprises
PROMISE-IMPACT	Promoting Micro and Small Enterprises through Improved Entrepreneurs' Access to Financial Services
PAP	Pre-Analysis Plan
PAR	portfolio at risk
RA	refused to answer
RCT	randomized control trial
s.d.	standard deviation
s.e.	standard error
SECO	Swiss State Secretariat for Economic Affairs
SIYB	Start and Improve Your Business
T	training treatment arm
TC	training and consulting treatment arm ToC training-of-counselors
ToT	training-of-trainers



Glossary: Key Terms

This box provides explanations of general terminology and concepts used in this report - for better readability, understanding, and interpretation.

Survey information and outcomes

Baseline and Endline Data. We conducted two waves of data collection: One prior to the program implementation (the baseline) and one after the program implementation (the endline). With the term baseline and endline data, we refer to these two waves of data collection. That is, we differentiate in time, not in the type or category of variables. For example, if we write about comparison of baseline and endline outcomes, we refer to changes over time for given outcomes.

Intermediate and Final Outcomes. We have deliberately categorized all outcome variables into these two groups before collecting endline data. They follow a theory of change developed prior to the data collection. The core idea is: What do we expect to change first? What is more likely to follow in the chain of changes?

Intermediate outcomes are those variables that are expected to be directly affected by the program, mainly changes in knowledge and business practices. We have derived these intermediate outcomes after carefully reviewing the manuals for training including material used for training of loan officers. Thus, they capture the outputs of the program as reported by the clients and observed by the enumerators.

Final outcomes are those variables which presumably change through the intermediate outcomes and which are more directly related to the client's living conditions. Hence, these outcomes are at the core of the program's goals. We expect that the changes are more difficult to achieve, especially in a short period of time. Examples of final outcomes are business profits, household spending and life satisfaction.

Interpreting Results in Outcome Tables

Control Mean. The control mean, usually reported in column (1), reports average values of each outcome variable for the control group. Since the treatment was randomized, this average can be interpreted as counterfactual for the treatment group, i.e., without the treatment, the treatment group would have reported the same average outcomes. The control mean helps to get a sense of the relative size of the effect.

Standard Deviation of the Control Mean. The standard deviation (SD or s.d.), usually in column (2), of the control mean is a measure of how much an outcome varies in the control group.

ITT. To estimate the intention-to-treat effect, we compare the outcomes of clients that were assigned the treatment with the outcomes of clients that were not assigned the treatment. The difference in outcomes is called the ITT. The original assignment to treatment and control group took place prior to the program start and was made by the authors of this study in a pool of eligible clients and using a randomization algorithm. Among the clients that were assigned to the treatment, some did not receive the treatment, either because they were not offered the treatment or because they were not interested in participating. We nevertheless consider them as part of the treatment group. The rationale is that we want to estimate the effects of allocating the program to clients, not the effects on the subgroup of clients who participate. For FSPs and program stakeholders this will provide more valuable information for future roll-outs as not all clients will be offered or will choose to participate in the program. In addition, while we asked loan officers to submit information on conducted training and counseling activities, this information was available to all clients, especially those clients who received counseling and therefore did not report themselves as treated in the endline survey. Thus, calculating treatment effects only on those clients who we identify as treated would lead to very small sample sizes for some FSPs.

Statistical Significance. In our study we assess the program effects based on a sample of individuals, not the entire population. Is the effect we calculate true for the whole population? Or did we simply capture it by chance? In other words, what is the likelihood our results are caused by chance (and not the program)? To answer these questions many factors need to be considered, such as the sample size (which might be too small to generalize effects to the whole population with much confidence), the size of the observed effect, etc. It is common practice to display different levels of confidence, marking the results according to the p-values. A p-value is the probability that random chance could explain our results; thus, the lower the probability, the better. We mark the p-values with stars: the more stars you observe, the more confident we are about the results. While we display all results (significant and not significant ones), for prudence; we interpret and discuss only those out-comes that are significant (marked with additional stars). In our case, we focus on the interpretation of outcomes that are significant at the 1% (three stars) or 0.1% significance level (four stars), since these effects are more robust. They tell us that with a chance of only 1% or 0.1% what we discuss is caused by random chance and not the program itself. If an effect is below this significance level, we are more careful with its interpretation given that the observed effect could be due to chance and not actually caused by the program.



1.

INTRODUCTION



ACCOUNTING for a majority of businesses, Micro and Small Enterprises (MSEs) are the backbone of many developing and emerging economies (Ayyagari, Beck, and Demirguc-Kunt 2007; Berge, Bjorvatn, and Tungodden 2015). Often MSEs are the only source of employment for individuals who lack access to formal labor market. While low growth and productivity of MSEs has long focused on financial constraints (e.g., Banerjee, Duflo, Glennerster, and Kinnan 2015), more recently, the discourse has shifted to business and managerial skills. To this end, donors and practitioners have advocated for a vast range of training courses, be it as stand-alone entrepreneurship program or as an addition to financial services.

Our study evaluates pilot projects that the International Labour Organization (ILO) in Indonesia implemented with 13 financial service providers (FSPs). These pilots were carried out as part of the broader Promoting Micro and Small Enterprises through Improved Entrepreneurs' Access to Financial Services (PROMISE-IMPACT) project. PROMISE IMPACT aims to promote inclusive finance in Indonesia through value-added financial services for small businesses, and improve the enabling policy environment. FSPs that participated in the pilots include savings and credit cooperatives, rural banks, and development banks in East Java and West Java.

The pilots were intended to test bundling of loans with enterprise development services. The services included classroom training and individual counseling to MSEs that were already clients of the FSPs. Training and business counseling focused on financial management and marketing practices. More than 170 loan officers of the FSPs were trained to become trainers and/ or counselors in two separate types of courses: (1) Start and Improve Your Business (SIYB) and (2) a training-of-counselors (ToC) course based on the SIYB methodology. SIYB is one of the largest business training courses in the world. It has been used by over 65,000 trainers in around 100 countries. Loan officers who successfully completed the ToT course then conducted two-day classroom training for clients of their respective FSP. For this, the ILO developed a training-of-counselors (ToC) course based on the SIYB methodology for a ToC to enable loan officers to provide individual counseling services to their clients in a semi-structured fashion.

Our research design was a randomized control trial (RCT) in which a total number of 5,131 clients of the participating FSPs were randomly assigned to one of the three treatment groups or a control group. The clients three treatment arms were: (1) only classroom training, (2) only individual counseling, or (3) classroom training combined with individual counseling in follow-up visits. In some FSPs all the three treatments were offered, while in others the random assignment was between control and classroom training combined with individual counseling in follow-up visits. We reviewed the training material



and a-priori interviewed local experts. Subsequently and before the endline data was collected, we uploaded a Pre-Analysis Plan (PAP) on the AEA website (RCT ID: AEARCTR-0003625), which forms the basis of our analysis.¹

We build on a growing body of literature evaluating enterprise development interventions, which have given reason for both optimism and caution. An earlier string of literature evaluates classroom training aimed at improving the livelihoods of many microentrepreneurs simultaneously. The evaluated interventions differ substantially and hence it is not surprising that previous studies measure varying effects, both in terms of behavioral changes, business practices, and changes in financial business outcomes. Whereas in an early work, Karlan and Valdivia (2011) find no overall impact of training on profits or employment, more recent research sheds light on the importance of training content and delivery mode. For example, Drexler, Fischer, and Schoar (2010) measure a significant impact of a rule-of-thumb training on accounting principles, but cannot detect any such impact for the standard fundamentals-based training.

More recently, Lafortune, Riutort, and Tessada (2017) find that a visit of a role model in a classroom training setting can increase the training impact relative to only classroom training, suggesting that motivational aspects in training might play an important part in the impact channel. McKenzie and Woodruff (2013) provide a comprehensive overview of similar evaluations and the general caveats in providing classroom training. However, the aim of supporting a large number of microentrepreneurs simultaneously is not immune to criticism. After all, micro businesses are seldom registered, hence do not pay taxes, often providing employment to family members only, and not adhering to standards such as occupational safety. Their productivity is typically low, and some subgroups have been found to operate at a loss (de Mel, McKenzie, and Woodruff (2008)). This casts doubt on the efficiency of artificially boosting a potentially unprofitable sector of the economy.

Another approach to MSE development puts more emphasis on the Schumpeterian process of creative destruction: Here, the aim is to specifically support those entrepreneurs who have the largest potential of growing their business to a more formalized level and therewith greater capacity to generate job opportunities for others. More recent studies consequently also aim at determining the right target group for an efficient allocation of business support. Much of the literature here has focused on the gender divide. In this regard, Berge et al. (2015) and Gine and Mansuri (2014) both find that male participants are more likely to benefit from training in terms of business outcomes, with neither study finding significant changes in income or assets for women. In contrast, Field, Jayachandran, and Pande (2010) find training increased business income for upper-caste Hindu women in India, but not for lower-castes or Muslim women. The increased analysis of heterogeneous treatment effects is accompanied by the emergence of a new branch of literature focusing on more tailor-made and more intensive individual consulting approaches (Bloom, Eifert, Mahajan, McKenzie, and Robert 2013; Bruhn, Karlan, and Schoar 2013; Karlan, Knight, and Udry 2012, 2015; Lafortune et al. 2017). Results are again mixed with heterogeneous effects across population subgroups and possibly sustained only in the short run.²

Our research contributes to both strings of literature by tackling the following research questions: First, we investigate whether MSEs in Indonesia on average benefit from SIYB in general, and from classroom training and/ or individual counseling provided by their loan officers in particular; second, we investigate who benefits the most from this business support.

1 <https://www.socialscisceregistry.org/trials/3625>

2 While the channels through which training take effect are generally not identified, a theoretical model developed by Karlan et al. (2012) provides some intuition. Here, entrepreneurs learn their type by experimenting with business innovations and return to their prior practices if they identify themselves to be of the low type.

The ToT-ToC approach in the ILO intervention has the advantage of being scalable and comparably cost-effective. On the downside, classroom trainings are short and provided by loan officers (as opposed to full time and experienced trainers), so the average effects are expected to remain small. As McKenzie and Woodruff (2013) point out, most existing studies face serious challenges regarding their power. With a sample size, typically ranging between 300 and 500 individuals in a treatment group and these being quite heterogeneous in baseline outcomes, the authors estimate the power to detect a 25% increase in revenues to be as low as 7% and seldom above 25%. In our study, we are able to pool the clients of 13 different FSPs, hence achieving a relatively large sample size of 2,830 clients in the treatment group. This allows us to detect comparably small average treatment effects. Nevertheless, we do not find any significant effects on final outcomes such as business profits, household spending or life satisfaction. This is in line with our inability to detect large behavioral changes or even an increase in theoretical knowledge. The notable exception is an increase in the share of clients whose business plan includes a cash flow analysis. This result is robust across model specifications.

Additionally, we make use of our relatively large sample size that allows for heterogeneity analyses. As discussed above, previous literature has shown that the need for different forms of technical assistance is divided along the lines of gender, but also along educational background, age of the business, as well as degree of knowledge and motivation at baseline. However, we cannot confirm these findings in our study. We measure some negative effects for more educated clients and only find few positive effects for registered businesses or businesses outside the trading sector.³

In addition, we investigate how the effects differ by different implementing partners. More importantly, our results vary greatly across the different FSPs. Clients from FSPs that predominantly lend to male clients are less likely to default on their loans due to the intervention. The fact that this result holds only when considering FSP heterogeneity, and not when comparing male and female clients individually, suggests that unobserved factor on FSP level influences the results.

We organize the rest of this report as follows: In Section 2 we describe the set-up and intended effects of the intervention we evaluate. In Section 3, we detail the eligible sample and the randomization procedure, and address treatment compliance. In Section 4, we present our data set and explain our identification strategy. Our results are discussed in Section 5 for average treatment effects and Section 6 for heterogeneity analysis. We demonstrate the robustness of our results in Section 7. We provide our conclusion in Section 9.

3 Among those clients who did not switch business sector since baseline, 59% of businesses were engaged in trading sector, 21% in manufacturing, 12% in services, and 6% in agriculture.



2.



THE INTERVENTION AND INTENDED EFFECTS

2.1 Relevance and Challenges of Micro and small enterprises in Indonesia

CONSIDERED an emerging middle-income country, Indonesia is a member of the G20 with robust growth rates over the past years. Whereas it has managed to cut its poverty rate substantially from 19% in 2000 to below 10% in 2018 (Statistics Indonesia 2018), poverty and inequality remain real development challenges. In 2010, of the over 50 million firms, 99.9% were MSEs (Aldianto, Rudito, Mirzanti, Situmorang, and Larso 2010). Their relatively small contribution to the gross domestic product (GDP) (43%) illustrates the MSEs' low productivity (International Labour Organization 2018). The ILO also estimated that labor productivity grew at only 1.1% between 2005 and 2013 in manufacturing MSEs, while the growth rate in medium and large manufacturing firms lay at 4.3%.

Indonesia boasts a large and innovative microfinance sector, which has been recognized as a global leader for decades (Rosengard and Prasetyantoko 2011).⁴ However, the fact that only few entrepreneur transition from micro to small businesses suggests that factors other than financing also inhibit growth. This is especially true given that some of the microenterprises are potentially necessity entrepreneurs who start their own business because they lack formal employment opportunities. These individuals are less likely to identify promising investment opportunities or to manage business growth effectively (Atmadja, Su, and Sharma 2016; Aldianto et al. 2010). The productivity of microentrepreneurs in Indonesia might hence additionally be inhibited by a lack of managerial capital.

2.2 The Intervention

To test this hypothesis, the ILO partnered with 13 FSPs in East Java and West Java. More than 170 loan officers of the 13 participating FSPs were formally trained to become trainers and/ or counselors. They then gave classroom training and/ or individual business counseling to their clients. The intended increase in profitability of the clients' enterprises is expected to benefit (1) the clients directly through greater profits and increased household expenditures, (2) the FSPs through improved loan repayment and larger loan sizes as well as (3) the broader economy through job creation.

⁴ Most famously, Bank Rakyat Indonesia is one of the largest microfinance service providers in the world with more than 4,400 units serving the rural population



In the ToT approach, the ILO relied on its existing methodology, the SIYB. Now one of the largest business training programs worldwide, it has its roots in the 1970s and has experienced constant revision and refinements for its implementation around the world (International Labour Organization 2013). Its ToT was an intensive 10-day course which focuses on concepts such as generating a business idea, determination of costs and prices, bookkeeping, financial planning, stock control, and purchasing/ buying. The training also taught adult learning methodologies, including the delivery of practice sessions by the participants, various simulation games, and exercises for presentation skills, concluding with a certification test. The ToT schedule followed in the intervention is attached in Appendix A.1.1.

The newly trained trainers then conducted classroom training courses for their clients who are individually invited by phone or in person. Up to 20 clients participated in a training session held in a location close to the clients' residences. Given the priorities set by the FSPs trainers focused on financial management and marketing in their courses for clients.⁵ An ILO master trainer visited a random selection of the training courses to check whether the loan officers were closely adhering to the guidelines and content of the SIYB. The loan officers trained the clients only once in a two-day training session which ends by asking clients to prepare their business plan. In the "training only" treatment arm no further advice to carry out their plans was provided.

In contrast, the ToC ran for five days as an adaptation of the SIYB specifically undertaken for this intervention. In addition to the technical skills of marketing and financial planning, the participants learned the fundamentals of adult learning and counseling, but were not trained to become classroom teachers. In Appendix A.1.2 we include the schedule of the five-day ToC training. Counseling of clients happened semi-structured, starting with a review of the MSE's past activities compared to the ones in the initially agreed-upon business plan. Subsequent counseling included advice and encouragement as needed.

2.3 Intended Effects

The intervention builds on the following rationale: Participation in training and/or counseling increases knowledge in the focus topics: marketing and financial planning. The increased knowledge leads to behavioral changes in business practices among participants. We refer to increased knowledge and changed practices as intermediate outcomes. These outcomes are then expected to affect business outcomes, which in turn lead to better lives for the entrepreneurs and their households, as well as to improved loan behavior, which also benefits the FSPs. We refer to business and households outcomes as well as loan behavior as final outcomes.

Regarding average treatment effects, to our knowledge, De Mel, McKenzie, and Woodruff (2014) and Fiala (2013) provide the only two rigorous evaluations of applications of SIYB, in Sri Lanka and Uganda respectively. De Mel et al. (2014) find changes in business practices, but changes in profitability only for new start-ups. Fiala (2013) finds significant increases in profits six months after the intervention ended, but only for male participants and only when combined with loans and not grants. While their evaluation gives us reason to believe that the intervention will positively impact our participants' lives, there are fundamental differences in the respective intervention set-ups. In Sri Lanka, the treatment groups consisted of women

5 The ILO supported the FSPs in conducting (not necessarily representative) client surveys in 2016 in order to understand and prioritize the needs of their clients. A total of 2,405 clients were interviewed to this purpose. A staggering 82% believed business support services to be important, yet only a small part had received training and mostly only related to repayment of loans; a topic they themselves deemed of little importance. In contrast, clients prioritized business support services regarding marketing, product quality improvement and financial management. The self-reported willingness to pay for these services is, however, low, with more than half of the respondents stating that they would not pay any fee.

running subsistence enterprises and women out of the workforce. They were assigned randomly to a classroom training treatment or a classroom training plus cash grant treatment. In Uganda, the treatment group also received either loans or grants, and some of the participants had no prior experience with loans. In contrast, our evaluation focuses on existing enterprises, both female- and male-owned, which have access to microfinancial services and are hence not severely cash-constrained. In fact, we are considering clients who are already borrowing at commercial rates, which may imply that they are already established businesses.

Our treatment also differs from De Mel et al. (2014) and Fiala (2013) by including individual counseling. We investigate whether this additional service can strengthen the positive effects found in the previous study. This might be a promising approach to more development as various authors point out it is impossible to separate the question whether business skills impact business outcomes from the question whether such skills can be transmitted in a specific format (Drexler et al. 2010; Bruhn et al. 2013). While treatment training treatment arm (T) is likely to improve knowledge, treatment consulting treatment arm (C) might have a more direct effect on business practices. The most intensive treatment arm, training and consulting treatment arm (TC), is expected to bring the highest benefit to the entrepreneurs.

As shown in previous studies, personal and business characteristics are linked to the direction and magnitude of effects. The intervention provides rather basic training which is aimed at entrepreneurs with little previous knowledge or experience. It is intended to help those entrepreneurs who face less favorable starting conditions, i.e., those entrepreneurs who are younger, less educated, less experienced and a lower business perception index, and whose business exhibit lower baseline profit.

Prior to the follow-up survey, we formulated 16 hypotheses which we summarized in a PAP, registered on the American Economic Association (AEA) website. The hypotheses are listed in Appendix B.1 for the reader's reference.



3.

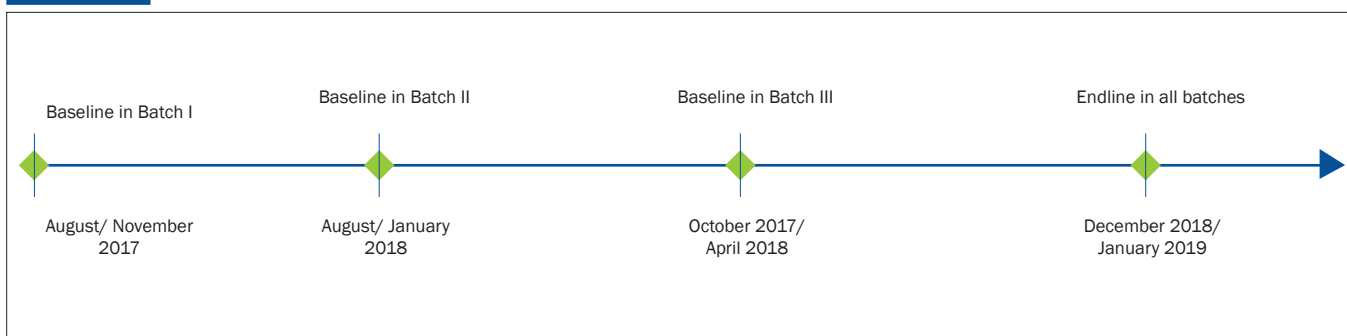
EXPERIMENTAL DESIGN AND TREATMENT IMPLEMENTATION



WE EVALUATE the impact of the intervention using an RCT implemented in partnership with 13 different FSPs. Following a baseline survey, we randomized clients into different treatment or a control group. Subsequently, after the intervention period which varied from six to eight months, we conducted an endline survey interviewing 4,292 clients.

As the partnerships between the ILO and the various FSPs were formalized at different times, the FSPs were divided into three batches.⁶ We completed the randomization for Batch I in January, for Batch II in February and for Batch III in April 2018. The implementation of the intervention started shortly after. Figure 1 illustrates the timing of the baseline survey, the randomization with subsequent start of implementation, and the endline survey.

Figure 1: Timeline of Baselines and Randomization



3.1 Participating FSPs and clients' baseline characteristics

The ILO selected partner FSPs based on an assessment of (1) the geographical and sectoral coverage (i.e., manufacturing MSEs in East and West Java), (2) their mission and governance structure (including commitment to the double bottom line approach, i.e., maximizing profits and clients' welfare), (3) operational management

⁶ The six FSPs of Batch I are the savings and loan cooperatives, the four FSPs of Batch II and the three FSPs of Batch III are the banks.



(including social performance management), and (4) product and services offered (including the willingness and capacity to experiment with new non-financial products). FSPs are savings and loan cooperatives, rural banks, and regional development banks. Whereas some institutions were just over 10 years old at the time of the intervention, one FSP was established more than 60 years ago. Their client base varies substantially from 1,000 to 433,000 active borrowers with 30% to 100% female clients. Several FSPs offer both group and individual loans, which are repaid either under conventional interest-paying or sharia-compliant profit-sharing agreements. Group loans do not necessarily follow the Grameen Methodology with mutual guarantees for individual loans, but are used for channeling funds and as means for social intermediation. In most institutions, the number of active depositors exceeds the number of borrowers. As part of their core business, some FSPs also offer some form of non-financial services, albeit on a limited scale. In conclusion, participating FSPs exhibit similarities in their mission and scope of operations, but are considerably diverse regarding their client outreach and professional experience. Tables A.1 and A.2 in the Appendix contain some of the key characteristics of the FSPs.

In our study, the number of participating clients varies between FSPs, from 178 to 582 clients in treatment and control group. Where possible FSPs targeted all the clients from the manufacturing sector. In cases where the FSPs did not have sufficient number of clients in this sector clients from trading or services were added in the sampling frame.

Table 1: Baseline Characteristics of Estimation Sample (All FSPs)

Panel A: Client Characteristics	Mean	Stand Dev.	Min. Value	Max. Value
Share of female clients	0.651	0.477	0	1
Average age of client	44.137	9.096	19	65
Total nr. of people in the household (HH) btw 18-65yrs. (incl. the client) ^{99p*}	2.801	1.052	1	6
Share of clients having no education	0.031	0.172	0	1
Share of clients having primary school education	0.289	0.453	0	1
Share of clients having secondary education	0.197	0.398	0	1
Share of clients having vocational education	0.359	0.480	0	1
Share of clients having university education	0.125	0.330	0	1
Total nr. of people in the HH that additionally earn income	1.265	0.987	0	8
Share of clients having no add. income	0.587	0.492	0	1
Share of clients having add. income from another business	0.250	0.433	0	1
Share of clients having add. income from full time job	0.057	0.232	0	1
Share of clients having add. income from oth. sources	0.039	0.193	0	1
Share of clients covering the HH against unforeseen exp.	0.881	0.323	0	1
Panel B: MSE Characteristics				
Nr. of yrs. the business exists ^{99p*}	11.647	9.000	0	39
Nr. of yrs. with FSP ^{99p*}	4.569	4.751	0	24
Share of MSEs not registered	0.789	0.408	0	1

Panel A: Client Characteristics	Mean	Stand Dev.	Min. Value	Max. Value
Share of MSEs registered as single merchant	0.146	0.353	0	1
Share of MSEs selling directly at the market	0.528	0.499	0	1
Share of MSEs selling through agents	0.172	0.377	0	1
Share of MSEs that are part of the HH	0.676	0.468	0	1
Share of MSEs active throughout the year	0.908	0.289	0	1
Nr. of family worker/s at the start of the business ^{99.5p*}	1.167	1.165	0	7
Share of MSEs having a written contract with workers	0.035	0.185	0	1
Share of MSEs paying workers regularly	0.548	0.498	0	1
Panel C: Business Practices (M -related)				
Share of clients understanding the 4Ps	0.352	0.478	0	1
Share of clients stating that marketing is important	0.861	0.346	0	1
Share of clients having a customer identification strategy	0.728	0.445	0	1
Share of clients having a strategy to make customers like their product	0.799	0.401	0	1
Share of clients having a competition strategy	0.753	0.431	0	1
Share of clients preparing a business/ financial plan	0.397	0.489	0	1
Share of clients keeping record of all transactions	0.370	0.483	0	1
Share of clients keeping business and HH finances separately	0.529	0.499	0	1
Share of clients investing profit into the business	0.651	0.477	0	1
Panel D: Business Outcomes (Y -related)				
Profit generated in the last 30 days in mln IDR ^{99p**}	5.397	7.720	0	50
Revenue in the last 30 days in mln IDR ^{99p**}	19.280	36.774	0	270
Cost/day for raw material in mln IDR ^{90p**}	0.950	1.412	0	4
Cost/day for workers salaries in mln IDR ^{90p**}	0.197	0.342	0	1
Cost/day for transport in mln IDR ^{95p**}	0.036	0.068	0	0
Cost/day for equipment or leasing in ths. IDR ^{90p}	0.081	0.273	0	1
Total nr. of non-family, permanent workers ^{99p*}	1.185	2.735	0	17
HH exp. on non-durables in the last 30 days in mln IDR ^{99p**}	2.923	2.643	0	16
HH exp. on educ. for kids in the last 30 days in mln IDR ^{99p*}	0.498	0.730	0	5
HH exp. on food in the last 30 days in mln IDR ^{99p*}	1.350	1.001	0	6
Share of clients having positive spending on durables	0.402	0.490	0	1
Panel E: Loan Behavior				
Average number of loans per year	25.341	55.584	0	400
Last loan amount in mln IDR ^{99p**}	0.692	0.462	0	1
Share of clients having their last loan as business/ individual loan				



Panel A: Client Characteristics	Mean	Stand Dev.	Min. Value	Max. Value
Share of clients reporting being late with the loan payment at the FSP	0.178	0.382	0	1
Share of clients want to borrow more for business	0.716	0.451	0	1
Panel F: Attitudes and Perceptions				
Share of clients stating finding market as the main barrier	0.243	0.429	0	1
Share of clients stating expensive raw materials as the main barrier	0.334	0.472	0	1
Share of clients stating little experience as the main barrier	0.045	0.208	0	1
Share of clients stating tough competition as the main barrier ^{NR}	0.270	0.444	0	1
Share of clients not stating any business barrier ^{NR}	0.189	0.392	0	1
Share of clients stating that business definitely brings high income	0.299	0.458	0	1
Share of clients stating that business gives security	0.517	0.500	0	1
Share of clients stating that business is definitely complicated	0.221	0.415	0	1
Share of clients stating that business is definitely risky	0.266	0.442	0	1
Share of clients stating that business definitely brings respect	0.446	0.497	0	1
Share of clients stating that business definitely brings satisfaction	0.554	0.497	0	1

▲ Note: This table contains summary statistics from the baseline survey. The table shows all variables which were used for randomization on the left, the different descriptive statistics on top.

▲ Sample: Estimation sample, i.e., clients from all FSPs who we were able to re-interview in the endline survey (N = 4,292). ▲ Source: Baseline survey (2017 - 2018).

▲ Columns (1) - (2) display the mean and standard deviation. Columns (3) - (4) display the minimum and maximum values.

▲ The superscript np. indicates the winsorizing level. We winsorized variables per batch prior to randomization and following an auto-mated rule to define percentiles. The stars indicate deviation from this rule: * indicates that we defined the winsorizing level manually, * indicates that we re-ran the winsorizing routine for all batches combined after the randomization.

▲ Table A.5 in the Appendix provides the definitions and further cleaning steps for the individual variables. Table A.6 presents the summary statistics for the final estimation sample, along with a number of tests for imbalances between treatment [arms] and control group.

Table 1 contains summary statistics from the baseline survey, averaged over all the 4,292 clients who we were able to re-interview in the endline survey. That is, the table describes the baseline characteristics of those clients for whom we estimate effects, which we refer to as the estimation sample. Columns (1) and (2) present the mean and standard deviation of each variable, followed by the minimum (3) and the maximum value (4). For example, the second line says that the average age in our estimation sample is 44 years with a standard deviation of 9, a minimum of 19 and a maximum of 65 years. We summarize here all the variables which were used for randomization; Table B.2 in the Appendix provides the definitions and cleaning steps for the data related to individual variables.

Approximately two thirds of our interviewed clients are female and the average household comprises of three adults among whom two earn an income. In Indonesia, participation rates in the education system are high and this fact is reflected in our survey: Only 3.1% did not attain any official education degree, whereas nearly 12.5% have a university degree. Clients who completed a vocational training form the largest group in our sample with 35.9%.

Notably, almost 60% of the interviewed clients earn no additional income themselves, besides their business. These businesses can be described as micro or small enterprises. Most of them have no written contracts for their workers, operating from the same location as the household of the owner, and mainly selling their products directly at the market. Whereas the average business has been operational for 11.6 years, there is a lot of variation. Our sample covers businesses between 0 and 39 years of age with a median of 9 years. Regarding business practices, more than 70% of clients reported having some marketing-related strategies in place, but less than 40% prepare a business plan or keep records of all the transactions. The most common loan type is a business/ individual loan. On an average, they borrow once in a year from the respective FSP. 17.8% report that they were late in making a loan payment to the FSP and slightly more than 70% plan to borrow more for their businesses.

On average, a client employs 1.2 non-family, permanent workers in their business, and costs for raw material outweigh workers' salaries by a factor of five. A significant proportion of clients (33.4%) report expensive raw materials as the main barrier to business. The average client claims revenues of 19.3 million IDR from their businesses and the final profit lies at around 25% of the reported revenues. However, all quantitative measures here show a high standard deviation. This might suggest that businesses vary in size or MSEs have difficulty reporting financial data. Given that we observe inconsistencies between reported profits and difference between reported revenues and costs, the latter explanation seems more likely. The data presented was cleaned including removal of outliers. For the results of the impact evaluation, we further discuss this in the robustness section (Section 7).

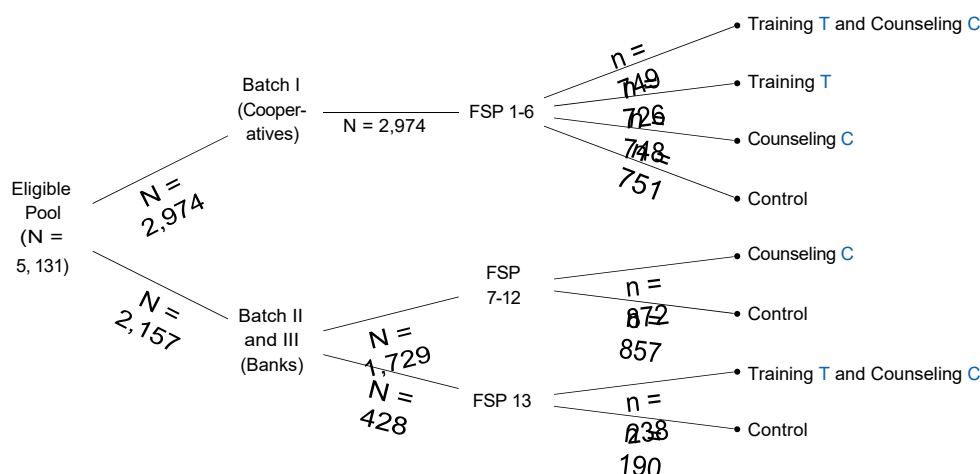
3.2 Randomization

We offered the participating FSPs a choice of different treatment arms: (1) T, (2) C and (3) TC. Whereas the six FSPs of Batch I chose all the three treatment, the FSPs of Batch II and Batch III were offered only the C treatment, except for FSP 13, which was also offered all the three treatments. Given that baseline data arrived in successive batches, each batch was successively randomized. We randomly allocated individuals to one of the treatment arms or to the control group by respective FSP.⁷

Figure 2 illustrates the assignment of observations to the different treatment arms. In particular, note that within Batch I there is approximately an equal number of clients assigned to the three different treatment arms and to the control group. In Batch II and III without FSP 13, the number of clients in the C treatment is approximately equal to the number of control clients. Overall, we have 726 clients in the T treatment, 1,620 in the C treatment, 987 clients in the TC treatment and 1,798 clients in the control group.

In allocating treatments, we followed a re-randomization approach whereby we retained the first randomization vector that passed a balancing threshold. As explained above, we did this subsequently for each batch, but our results are estimated for all the FSPs together. We also looked at specific FSPs separately. In Appendix Table A.6, we, therefore, show summary statistics, testing the balancing of the randomization variables between treatment and control groups for all the batches jointly. We show t-test and F-test statistics for the full baseline sample and the sample net of attrition, and find no systematic difference in observable baseline characteristics between treatment and control group.

⁷ While randomization is at the individual level, we applied a two-stage procedure: First, to vary geographical treatment intensity, we allocated the treatment probability at the branch/ cashpoint level. (These are operational offices of the FSPs. Usually, a cashpoint is one unit below the branch and consists of five loan officers and a manager. However, the size of a branch or cashpoint depends on the size of the respective FSP, e.g., a relatively large FSP might have a cashpoint that has the size of a branch of a smaller FSP.) Second, we allocated individuals within a branch/ cashpoint to treatment. As a consequence of the random assignment of treatment probabilities, we allocated a larger number of clients to treatment in some branch/ cashpoints than in others. For more details on the randomization procedure, see Appendix B.3.

**Figure 2: Assignment of Clients to Different Treatment Arms per Batch**

3.3 Implementation fidelity

We monitored the implementation process via an online platform starting in April 2018. Loan officers were requested to complete a questionnaire directly after each training/ counseling session for each participating client separately. Through the questionnaire, we gathered information on the session in general (such as when and where the training and/ or counseling took place, how many clients participated, if the officer faced difficulties, etc.) and on the specific participating client (such as how attentive the client was, if he/she asked any questions, etc.). It is important to note that the loan officers filled the monitoring questionnaire on a voluntary basis and, hence, our data does not adequately reflect every session. As such, we have monitoring data on half of our initial treatment group, specifically on 430 clients who received at least one training session, 1,071 clients who received at least one counseling session, and 313 clients who received at least one training and one counseling session. In addition to the online monitoring data, we gathered information on treatment offer in the endline survey. Specifically, we asked whether clients were aware of the intervention and whether they had been invited to participate. In the following, we first present evidence using both sources of information (Section 3.3.1) and then look in more depth into the self-reported data in the endline survey (Section 3.3.2). We present additional descriptive statistics from the monitoring data in Appendix A.2.

3.3.1 Participation: Evidence from the monitoring data and the endline survey

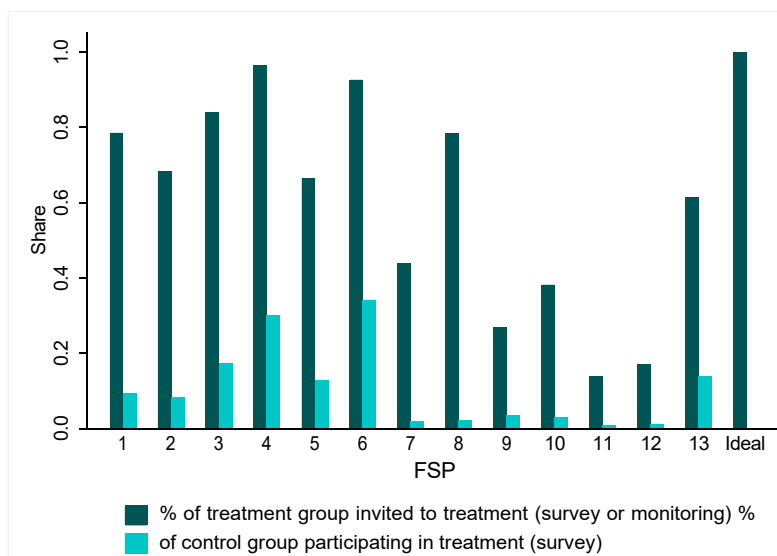
In the following paragraphs, we combine information from the monitoring data and survey data into one indicator, capturing ‘any’ reported participation in the intervention. The degree to which the treatment was reportedly implemented (i.e., the intervention offered to clients) and the assignment to treatment groups adhered to varies greatly across FSPs. Figure 3 illustrates this fact by showing the compliance with the treatment protocol for each FSP (x-axis). The first bar indicates the degree to which the FSP offered treatment to the treatment group.

The second bar indicates the degree of contamination within the control group.⁸ For example in the first FSP, 78.60% of clients assigned to the treatment group either report in the endline survey that they were offered treatment or our monitoring data confirms that they took part in at least one

⁸ Some FSPs offer their clients some basic financial literacy training when they take out their first loan. To make sure that respondents did not confuse this training with our treatment, we also asked for the date of training. We count as self-reported participation in treatment only instances where the training was received within the intervention’s timeframe.

treatment (first bar). For the same FSP, 9.50% of clients assigned to the control group report in the endline survey that they participated in the treatment (second bar). Perfect compliance with the treatment protocol would have led to the bars displayed in the far right labeled “Ideal”. These numbers reflect the participation as reported. Given the voluntary basis of monitoring data entry and limitations in the survey (such as recall error, non-response, etc.), we believe that actual participation was higher, as there are certainly unrecorded cases.⁹ The latter in particular holds for participation in the C treatment. Here, we find that loan officers rarely reported treatment in the monitoring system. An exception is FSP 8, which shows relatively high rates of participation according to the monitoring data, but very low rates according to the survey data. Upon our request, the ILO followed up on this divergence with the respective FSP, which again confirmed the monitoring data. This gives reason to believe that our endline survey questions were not able to fully measure participation in the counseling treatment, hence large measurement errors in this aspect are likely.¹⁰

Figure 3: Reported Implementation Fidelity by FSP: Monitoring and Survey Data Evidence



Note: The figure displays the compliance with the treatment protocol by the 13 FSPs. The first bar indicates the degree to which the FSP offered treatment to the treatment group. The second bar indicates the degree of contamination within the control group.

▲ Source: Monitoring data and endline survey (2018 - 2019).

From Figure 3 it is evident that some FSPs reportedly were more engaged in the intervention than others. On the one hand, FSPs 7 and 9 to 12 offered the treatment to less than 50% of their assigned treatment clients. On the other hand, FSPs 4 and 6, which invited over 90% of their treatment clients, also extended treatment to about a third of their control clients. This finding is in line with evidence presented in the work by Karlan and Valdivia (2011). Here, the authors implement an entrepreneurship training intervention at different village banks of FINCA Peru. They report heterogeneity of treatment intensity across the village banks, with only half of them reaching the 17 out of 22 envisioned sessions within two years. The authors also report that these delays are typical for similar interventions and conclude that the analysis should not further account for these differences to avoid selection bias. While we agree with this conclusion, we would like to emphasize the importance of selecting the right partner institutions in multi-stakeholder interventions

⁹ For the final analysis, we will mainly rely on the assignment indicator, i.e., if a person was assigned to the treatment we consider him as treated, independent of whether this is confirmed in our monitoring or survey data. We will also make use of the participation indicator based on monitoring or endline data in a separate set of analyses.

¹⁰ In the endline, after introducing the intervention, we first ask whether the respondents were aware of the intervention, then whether they were invited, and only conditional on invitation we asked about participation. However, it is likely that clients were not formally invited to the counseling and hence were not asked about participation.



such as this. Therefore, in addition to analyzing the whole sample, we will also show results for FSP 8 separately. This FSP shows the highest implementation fidelity, measured by the difference in height of the two bars in Figure 3. That is, FSP 8 shows the highest jump in treatment probability between treatment and control group.

Qualitative interviews led by the ILO shed light on the reasons for the variation in implementation fidelity. During a roundtable workshop in March 2019, the ILO elicited feedback from the FSPs on the challenges in intervention implementation. Six of the 13 FSPs cited the large geographic spread of the clients as a key challenge, mostly in combination with the limited time available to loan officers. Four FSPs reported that clients were little motivated and too passive during counseling, which should be improved by better targeting. Three FSPs explained low treatment compliance with high drop out rates of clients who finished their last loan cycle. Finally, some FSPs found the number of trained loan officers insufficient, in particular as loan officers changed their duty station or resigned after the ToT/ ToC. Still, the feedback received does not explain the different level of compliance illustrated in Figure 3. That is, FSP 8 named more challenges than any other FSP, whereas FSP 10 named the most observed benefits (such as improved reputation and increase loan sizes). Instead, it is rather noteworthy that FSPs 1 to 6 are cooperatives, whereas 7 to 13 are the rural and development banks. The latter have, on average, more employees and pay loan officers about 1.5 to twice the amount paid in the cooperatives. They also have larger loans outstanding and more active borrower and depositors. In short, on average, the larger FSPs were less likely to comply with the treatment protocol than the smaller FSPs, with FSP 8 being the exception.¹¹ Overall, we conclude that participation rate among invited clients was relatively high, yet varied by FSPs. Declining participation was related mostly to time constraints and almost never to a limited interest in the service. Evidence from the endline survey allows us a deeper analysis of the clients' self-reported experience and perceptions about the intervention for those individuals who participated in the endline. These are presented in Section 3.3.2.

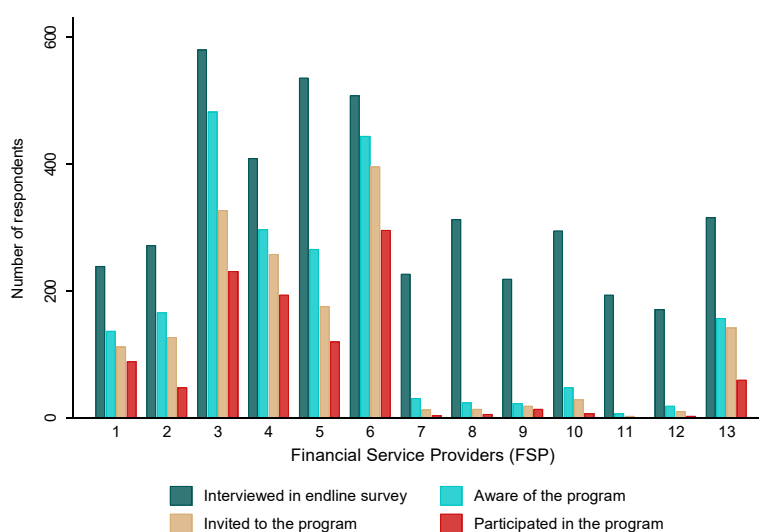
Lastly, wherever monitoring data are available, we can analyze the nature of the services delivered. The intervention design with respect to its content and structure was mostly adhered to, except for the number of counseling sessions, which did not average the intended five. Two thirds of the participants were actively engaged and questions were answered satisfactorily by the trainer. The endline survey shows that most of the respondents find the intervention helpful, and that training sessions were regarded more favorably than counseling services. Survey evidence also confirms that the topics covered in classroom training and counseling sessions were largely similar.

11 Table A.1 and A.2 in Appendix A show that cooperatives and banks hardly differ regarding the share of rural borrowers. Whereas the sample of cooperatives clients contains more women than the sample of bank clients (82.78% versus 41.10%), the cooperatives FSP 2 and FSP 5 also only have a low share of female clients and this does not seem to explain implementation fidelity.

3.3.2 Self-reported program participation: Evidence from the endline survey

This section presents self-reported evidence on the experience and perceptions of the program as reported by the clients in the endline survey. The endline survey contained a module on experiences with the program, which we posed to treatment and control group of clients at the end of the survey. In this module we asked all clients: “Are you aware of the existence of services such as classroom training and individual counseling that [FSP name] is providing?”¹² Only if the respondent said “Yes”, we asked the next question, which was: “Were you offered such training or counseling on business development as aforementioned?” And only if the answer was affirmative, we asked: “Did you participate in classroom training or individual counseling sessions or both?”

Figure 4: Distribution of Endline Survey Respondents over FSPs



Note: Figure 4 describes how the respondents are distributed over FSPs. We further indicate whether they were aware of the program, invited to the program and participated in the program.

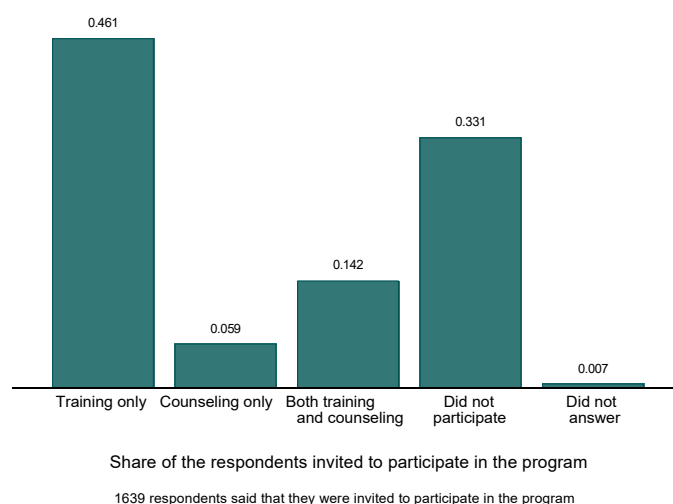
▲ Source: Endline survey (2018 - 2019).

In the endline survey, we were able to re-interview 4,292 entrepreneurs who had taken part in the baseline survey. Figure 4 shows respondents distributed over FSPs and the responses to the above-mentioned filtering questions. Conditional on reporting of having been invited to treatment, the overall participation rate as reported in the survey lies at 66.2%. Notably, the degree to which the FSP offer was accepted varies between 40% and 82% across FSPs.

Figure 5 illustrates the self-reported participation of respondents in the different components of the program (treatment arms) as share of all respondents who received an invitation.¹³ Whereas 46% of the invited clients report participation in only classroom training, 5% in only individual counseling sessions and 14% in both, one third of invited clients report that they did not participate in the program.

¹² Note that the FSP name was inserted from the baseline survey. If the respondent changed the FSP he/ she indeed might not be aware of training/ counseling offered at his/ her old FSP.

¹³ Note, not every client who was initially assigned to the program received an invitation, and some control group clients also received an invitation.

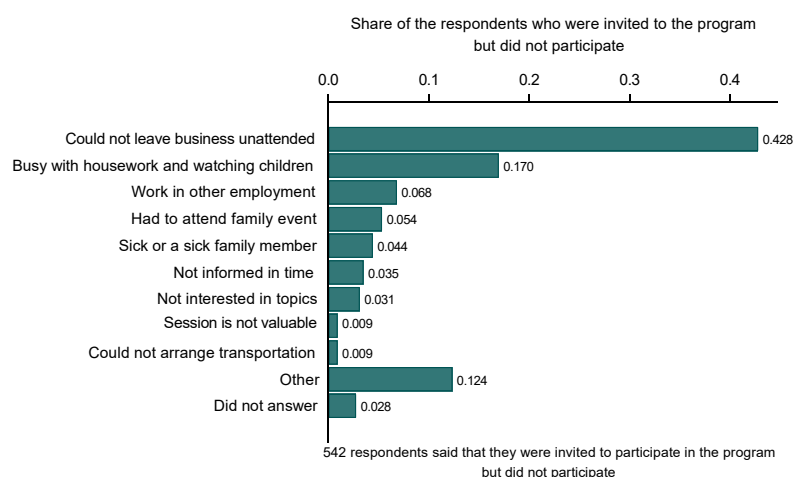
**Figure 5: Program Participation by Treatment Arms: Evidence from the Endline Survey**

Note: Figure 5 illustrates the participation of respondents in the different treatment arms as share of all respondents who received an invitation.

▲ Original survey question (# E9_3): “Did you participate in classroom training or individual counseling sessions or both?”

▲ Source: Endline survey (2018-2019).

Figure 6 shows the stated reasons for declining the invitation. The most common self-reported reason for not participating in the program was that the clients could not leave their businesses unattended (see Figure 6). The second most frequent answer was that they were busy with household chores and childcare. These two reasons alone account for 60% of clients who declined to participate. Importantly for the future rollout of the program, only 4% of the respondents said they were not informed about advisory services on time. Only 3% of the respondents reported that they were not interested in the topics while a small proportion of 1% who were offered advisory services stated that it would not bring any value for them.

Figure 6: Reported Reasons for Non-Participation in the Program

Note: Figure 6 displays the most common reasons for not participating in the program.

▲ Original survey question (# E9_5): “What was the main reason for not participating?”

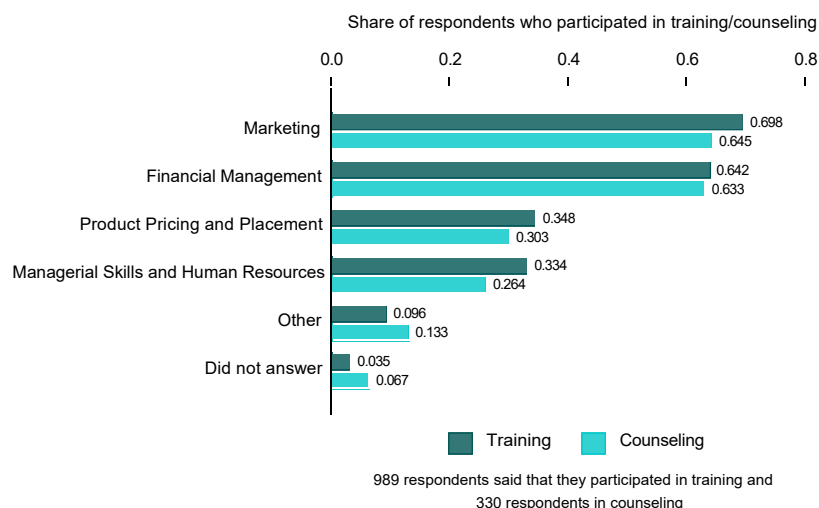
▲ Source: Endline survey (2018 - 2019).

If the respondent answered positively to the question on participation (“yes, classroom training only”, “yes, individual counseling only” or “yes, classroom training and individual counseling”) we continued the survey with further questions. Among the 1,639 respondents who reported having been invited, 1,086 respondents (66.26%) reported participation in training, counseling or both. Accordingly, the rest of the questions in this module were addressed to this group. This is a selective group, given that many more individuals were invited to participate and also invited to respond to the endline survey. Dissatisfied clients might have rejected both invitations and would thus be underrepresented in the following statistics. Moreover, in what follows, we consider all the individuals who responded, irrespective of whether they were originally assigned to the treatment or the control group. Thus, the descriptive statistics in this section need to be treated with caution and do not reflect the impact of the program.

For the sample of respondents whose experiences will be described in more detail below, we report key background characteristics of the respondents. Nearly 86% of the clients who report having participated in the program were women. Conversely, female clients were much more likely to participate: 75% of female clients invited, accepted the offer, as opposed to 52% among male clients.¹⁴ When considering the respondents’ educational background, the clients with no education comprised 2.7% of the participants, while nearly half of the clients who report having participated in the program had at least vocational education. Approximately 65% of the participants reported having no additional income besides their businesses (for more information see Table A.7 in the Appendix).

As Figure 7 shows, the topics covered were largely similar in classroom training and counseling sessions. Respondents could provide multiple responses to this question. The most common topics reportedly discussed during training or counseling were marketing and financial management, mentioned by almost two thirds of the respondents who participated in training or counseling. This is in line with the program objectives and shows that the program design was largely adhered to. On average, participants report that they attended two days of training and/ or had three counseling sessions. In most of the cases, the class size during training exceeded 25 people, while approximately in half of the counseling sessions the respondents were

Figure 7: Topics Discussed during Training or Counseling Sessions



Note: Figure 7 displays the topics reportedly covered in the training or counseling sessions.

▲ Original survey question (# E9_1_8/E9_2_7): “What were the topics that you have covered in the classroom training/counseling sessions?” Multiple responses were permitted. Respondents answered either regarding training or counseling or both.

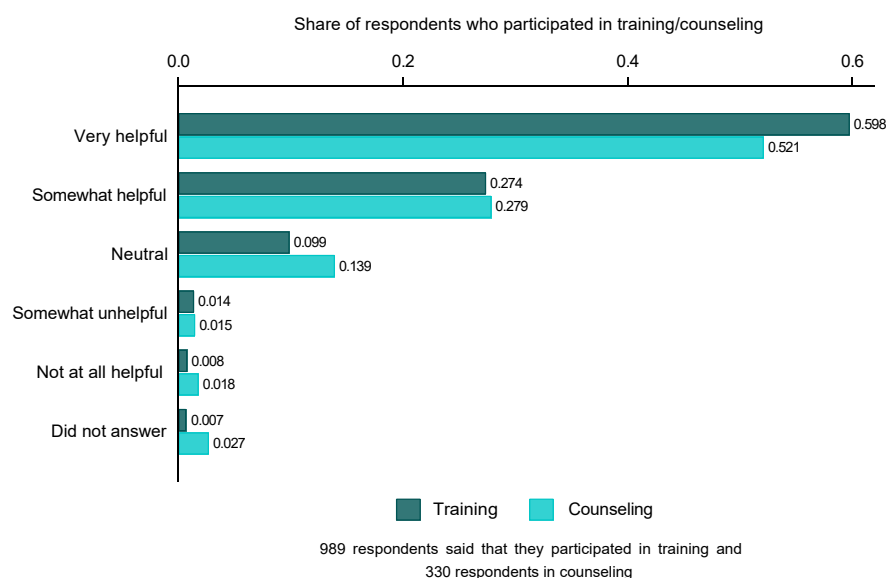
▲ Source: Endline survey (2018 - 2019).

¹⁴ Notably, we cannot confirm the finding by Valdivia (2015), where women with young children are less likely to participate than women without young children.



not accompanied by anyone. The average duration of training was six hours, whereas counseling sessions lasted for approximately two hours. Around 62% of the respondents said they asked questions during the training and 65% during the counseling sessions. In most cases, all questions during training and counseling were answered satisfactorily. This was reported by 90% of the respondents who participated in the program.

Figure 8: Respondent Opinion on the Helpfulness of Classroom Training and Counseling Sessions



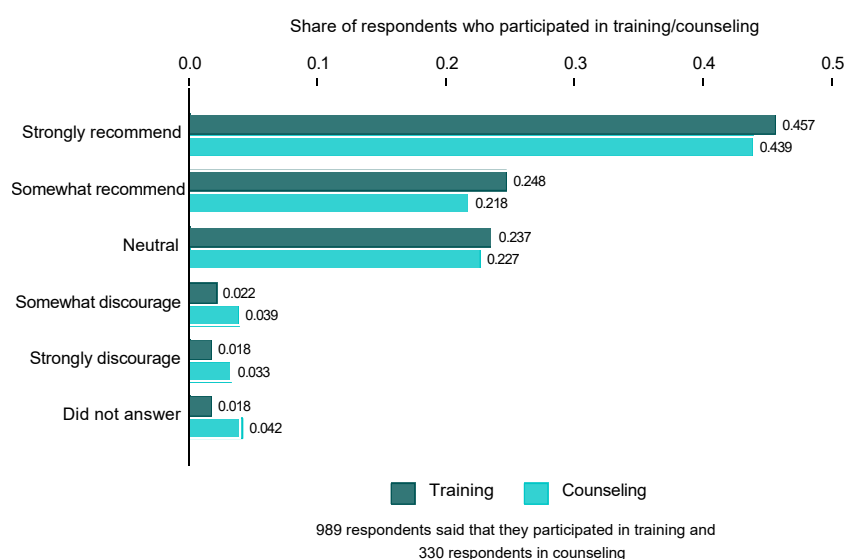
Note: Figure 8 describes the helpfulness of classroom training and counseling sessions as reported by the respondents.

▲ Original survey question (# E9_1_9/E9_2_8): "To what extent was the training program/ counseling helpful to you?" Respondents answered either regarding training or counseling or both.

▲ Source: Endline survey (2018 - 2019).

Turning to the perceptions of the respondents towards the program, Figure 8 describes the helpfulness of classroom training and counseling sessions as reported by the respondents. Overall, the results show that approximately 87% of the respondents considered training as helpful, while 80% reported the same for counseling sessions. In particular, a little more than half of the respondents said training or counseling sessions were very helpful. However, a higher share of the respondents considered training as very helpful, compared to counseling sessions. Accordingly, the share of the respondents who reported counseling as unhelpful was slightly larger for training.

Figure 9: Respondent Recommendation of Classroom Training and Counseling Sessions



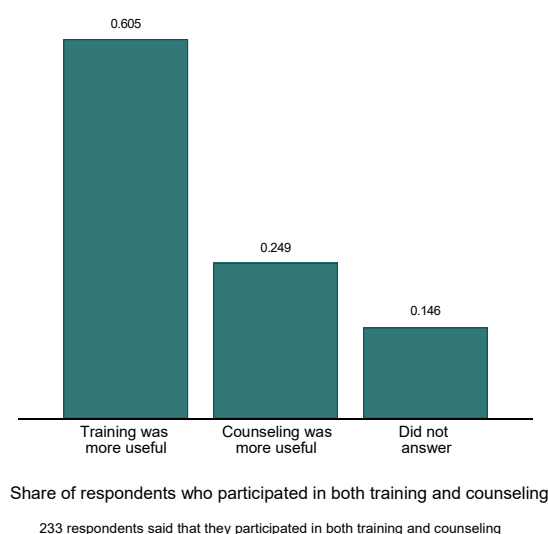
September 26,
2019

Note: Figure 8 describes the helpfulness of classroom training and counseling sessions as reported by the respondents.

- ▲ Original survey question (# E9_1_9/E9_2_8): "To what extent was the training program/ counseling helpful to you?" Respondents answered either regarding training or counseling or both.
- ▲ Source: Endline survey (2018 - 2019).

Figure 9 shows to what extent the respondents who participated in the program recommend such advisory services to other individuals. Approximately two thirds of the respondents said that they would strongly or somewhat recommend training and counseling. Similar to the helpfulness of the program, the share of those who chose to discourage others from training was smaller than those who would discourage receiving counseling services.

Figure 10: Usefulness of Classroom Training and Counseling Sessions



Note: Figure 10 reports the self-perceived comparison of classroom training vs. counseling sessions.

- ▲ Original survey question (# E9_3_1): "What did you find more useful, training or counseling sessions?" Respondents answered either regarding training or counseling or both.
- ▲ Source: Endline survey (2018 - 2019).



Lastly, among the respondents who participated in both training and counseling, nearly 61% of the respondents said that classroom training was more useful than individual counseling sessions. The opposite was reported by 25% of the respondents.

4.

DATA AND ESTIMATION METHOD



4.1 Survey data

WE USE data from a baseline survey before the randomization and an endline survey up to 1.5 years after the baseline, in addition to the monitoring data. The baseline and endline surveys were conducted in computer assisted personal interviews (CAPI) mode by two local universities, the University of Padjajaran (UNPAD) and Institut Teknologi Sepuluh Nopember (ITS).¹⁵

The baseline survey captures client characteristics including socio-demographic background and household finances, and business-related information. Regarding the latter, we collected information on the use of financial services, business performance, business planning, and marketing practices. The endline survey was more extensive, additionally covering knowledge questions and questions about the treatment participation.

Cleaning of baseline data involved the removal of duplicates, observations with missing contact information or non-response in more than six items or in age or gender. If less than seven items were missing, we imputed the answer using regression imputation based on the set of complete observations. We accounted for outliers and measurement errors by winsorizing at the 90th, 95th or 99th percentile, depending on the initial variation. In the endline survey we winsorized at the same threshold as in the baseline.¹⁶ The full list of outcome variables and covariates, including their definition, cleaning and processing steps is in Appendix B.2. We provide further details on the survey methodology and data cleaning in Appendix B.4.

After the baseline we randomized over 5,131 MSEs. In the endline, we were able to re-interview 4,292 clients, as the rest (839) were not found or refused the interview.¹⁷ Although an attrition rate of 16.35% seems quite high at first glance, it is common (McKenzie and Woodruff, 2013 for similar interventions.

¹⁵ Before the endline survey, supervisors and enumerators from the universities received extensive training from our research staff and we also conducted multiple pre-tests. During data collection, we monitored the quality daily via an automated.

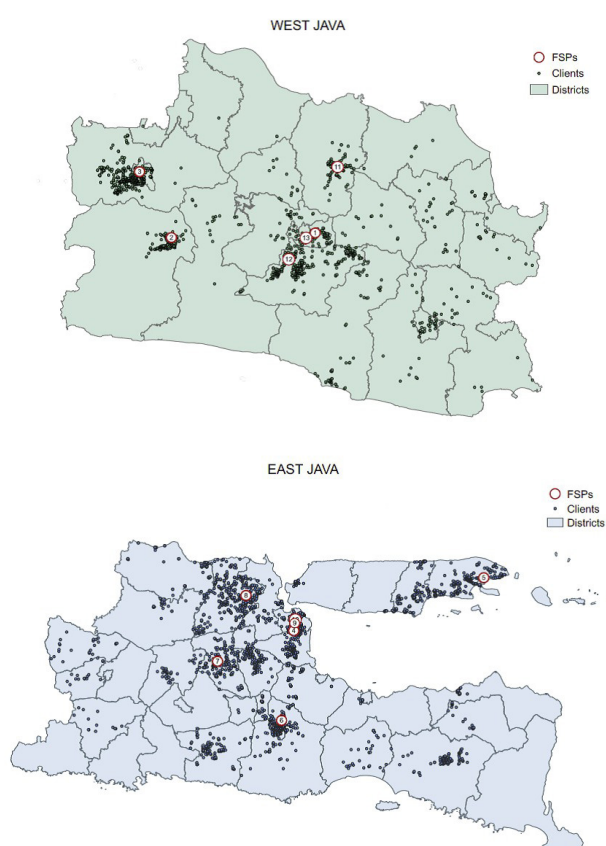
¹⁶ We winsorized variables per batch prior to randomization and following an automated rule to define winsorizing levels. Where outliers remained, we re-ran the winsorizing routine for all batches combined again after the randomization, or adjusted the winsorizing level manually to account for individual outliers.

¹⁷ After the randomization, we realized that there were 37 duplicates and 28 clients without sufficient information for the follow-up survey. Eventually, we used the baseline sample 5,066 clients used for the endline survey.

Studies show that attrition rate range from as low as 5.3% in Field et al. (2010) and 8% in de Mel et al. (2008) to 24% in Karlan and Valdivia (2011) and 26% in Calderon, Cunha, and De Giorgi (2013). We find a higher attrition in East Java, of around 18.22%, while in West Java we have an attrition rate of 13.56%. As most FSPs in Batch II are located in East Java, while only half of the FSPs in Batch I are in East Java, this regional difference in attrition also translates into a differential attrition across batches.

In Table A.6 in the Appendix, we analyze differences between the treatment group(s) and control group as well as the attrition rate. We find that clients from the control group were 3.6 percentage points less likely to participate in the endline survey than clients from the treatment group. Still, when we look at the difference in baseline characteristics between treatment and control group, we find no systematic imbalances after controlling for FSP fixed effects, age, and gender.¹⁸ The only difference occurs in the average number of loans per year, but this difference was already prevalent at baseline. Yet, other imbalances disappear in the estimation sample, i.e., in the sample net of attrition. Overall, the F-test of joint significance yields an F-statistic of 0.865 in the baseline sample [N = 5,131] and 0.912 in the estimation/ endline sample [N = 4,292], respectively. This gives us confidence that there was no systematic pre-treatment or selection-out-of-the-sample differences. Table A.6 also displays the p-value of an F-test for joint significance for the three treatment arms, i.e., the treatment arms regressed on the balance variables finds only four cases where the p-value is < 0.1.

Figure 11: Location of Clients



Note: The maps in Figure 11 illustrate the geographic coverage of our endline survey within East and West Java.

▲ Source: Endline survey data (2018 - 2019).

¹⁸ In all our subsequent analysis we will control for FSP fixed effects, age, and gender.

4.2 Econometric approach

We calculate average treatment effects to test the above specified intended effects using primarily the survey data. We focus on calculating intention-to-treat (ITT) effects. Here, our treatment groups are comprised of those FSP clients assigned to the respective treatment arms and control group, where selection is determined by the randomization procedure.

We regress the relevant outcome variable on a dummy variable indicating treatment assignment, a set of control variables, and two types of fixed effects, leading to the following main model specification:

$$Y_{i,t=1} = \beta_{ITT} D_i + X_i \beta_5 + X_i \beta_4 + \text{Miss} X_i \beta_3 + \mu_{FSP} \quad \beta_2 + \mu_{Enu} \beta_1 + \beta_0 + \epsilon_i. \quad (1)$$

Here, $Y_{i,t=1}$ is the outcome variable for entrepreneur i at endline ($t = 1$), D_i is a vector indicating treatment status (assigned to any of the three treatment groups or control group), and the parameter of interest is the vector β_{ITT} , the intention to treat effect. It gives the average difference in means between the combined treatment group and the control group and is interpreted as the impact of being officially eligible for the training and/ or counseling sessions.

We include four types of covariates: First, the set X_i comprises the covariates age and gender, which we control for in all regressions. Second, collecting additional individual entrepreneur's pre-randomization characteristics, we denote by X_i the vector of randomization strata. While we randomized a total of 65 baseline variables, we include in the regression twenty baseline variables that show the highest imbalance within the respective sample considered.¹⁹ This implies that different covariates are used for different subsamples.²⁰ The randomization variables are listed in the Appendix in Table A.5.²¹ To test robustness, we also calculate a number of different model specification, in particular one specification with ten control variables and one with thirty control variables (see Section 7).

19 We have used 65 baseline variables for the randomization, i.e. we ensured that treatment and control group are balanced on these 64 variables. However, this was done by batch, so when throwing the whole sample together and taking out attrited households, there may be some imbalances on baseline characteristic. We cannot control for all of them as these are too many covariates. Instead, we automatically select those twenty covariates which have the greatest normalized mean difference between treatment and control group.

20 However, age and gender are always included. For our full sample which pools all FSPs we additionally control for: indicator for female clients, indicator for clients having add. income from full time job, last loan amount in mln IDR^{99p**}, number of loans per year, indicator for clients having their last loan as business/ individual loan, indicator for clients stating that business definitely brings high income, indicator for MSEs having a written contract with workers, indicator for clients having positive spending on durables, indicator for clients having no add. income, indicator for clients want to borrow more for business, indicator for MSEs not registered, indicator for MSEs active throughout the year, indicator for MSEs that are part of the HH, cost/day for workers salaries in mln IDR^{90p*}, indicator for clients stating that business definitely brings respect, nr. of family worker/s at the start of the business^{99.5p*}, indicator for clients having university education, indicator for clients keeping record of all transactions, indicator for MSEs selling directly at the market, indicator for clients stating expensive raw materials as the main barrier. In the sample of FSP 8 clients, we control for: indicator for female clients, indicator for clients having add. income from full time job, last loan amount in mln IDR^{99p**}, number of loans per year, indicator for clients having their last loan as business/ individual loan, indicator for clients stating that business definitely brings high income, indicator for MSEs having a written contract with workers, indicator for clients having positive spending on durables, indicator for clients having no add. income, indicator for clients want to borrow more for business, indicator for MSEs not registered, indicator for MSEs active throughout the year, indicator for MSEs that are part of the HH, cost/day for workers salaries in mln IDR^{90p*}, indicator for clients stating that business definitely brings respect, nr. of family worker/s at the start of the business^{99.5p*}, indicator for clients having university education, indicator for clients keeping record of all transactions, indicator for MSEs selling directly at the market, and indicator for clients stating expensive raw materials as the main barrier.

21 In the Appendix in Table A.5 all variables named with "R_" indicate the full set of balancing variables that were used during randomization. To decide upon the twenty most imbalanced variables, we compare the treatment and control group, calculating the T-statistic and sorting all variables by the absolute value of T-statistic in the respective sample. The first twenty variables are chosen.



Third, there are a few cases of non-responses in some baseline randomization variables in the raw data, i.e., there were missing values for some of the variables. In order not to lose observations with less than seven missing items in baseline, we impute the missing values by estimating a regression model and include the imputed values in the vector X_i . Thus, to account for this imputation, we include the vector $MissX_i$ which is a dummy equal to 1 if the baseline value for a covariate is missing for individual i due to non-response and has been imputed. To avoid collinearity, we group the baseline variables and set $MissX_i$ equal to 1 if any of the variables in the respective group has been imputed. We define seven different groups as listed in Table A.5. As a robustness check, we also repeat the analysis on the subsample of observations with only non-missing values (see Section 7), in which case $MissX_i$ is dropped.

Fourth, we include fixed effects for FSPs and enumerators. The former accounts for variation within the FSP, which stem from differential commitment to the intervention, loan officer qualifications, client retention rates, or other unobservable factors. The latter accounts for reporting differences across enumerators. In our main specification with FSP fixed effects, we present robust standard errors, ϵ_i . In an alternative specification, we omit the fixed effects and cluster the standard errors on the FSP level.²² In addition to comparing the combined treatment group to the control group, we also estimate average treatment effects for each treatment arm separately and test whether there are significant differences in these effects. We also examine whether treatment affected different population subgroups differently. Denoting by R_i an interaction covariate, i.e., an indicator of the individual, business or FSP characteristics, we estimate the following general regression model for each of our outcomes:

$$Y_{i,t=1} = \beta_{ITT} D_i + \gamma_1 R_i + \gamma_{ia} D_i \times R_i + X_i \beta_5 + X_i \beta_4 + MissX_i \beta_3 + \mu_{FSP} \beta_2 + \mu_{Enui} \beta_1 + \beta_0 + \epsilon_i. \quad (2)$$

In this model, the parameter of interest is γ_{ia} , which provides the differential impact of treatment for different values of the interaction covariate. For example, if R_i is a dummy variable for gender, equaling one for female clients, γ_{ia} gives the additional treatment effect for women as compared to men. A variable denoted as R_i is a credibly exogenous pre-treatment variable.

²² The above model specification slightly deviates from the one specified in the PAP. Here, we had included batch fixed effects and also intended to use the full set of randomization variables. We present all results in the robustness section. The reason we had to deviate from the PAP model is the stark, not anticipated variation in the implementation of the treatment between the FSPs. By controlling for 13-1 FSP fixed effects instead of only 3-1 batch fixed effects, we believe that we can capture more variation within each FSP, reducing a potential omitted variable bias, and thereby estimating a more restrictive model. The enumerator fixed effect also make a difference in results and should hence be included in all estimations.

5.

MAIN RESULT



5.1 Intermediate and final outcomes

OUR MAIN analysis tests whether assignment to any of the three treatment arms led to different outcomes compared to the control group. As previously described, we distinguish between the intermediate outcomes (knowledge, practices) and the final outcomes (business and household outcomes, loan behavior).

Overview. For both groups of outcomes, we estimate our main results for three different model specification, which are contained in Table 2 for intermediate and in Table 3 for final outcomes. In both tables, columns (3) and (4) contain the ITT point estimates and standard errors from a model which includes age, gender and the twenty most imbalanced baseline variables as covariates, and cluster standard errors at FSP level, as laid out in Section 3. Columns (5) and (6) contain the ITT point estimates from our main specification with FSP fixed effects and robust standard errors. For the estimates in columns (7) and (8) we restrict our analysis to the sample of FSP 8 clients, which showed the highest implementation fidelity (see Section 3.3).

We run the same specification as in columns (3) and (4), yet using the most imbalanced baseline variables for this sample. Additionally, columns (1) and (2) show the mean and standard deviation of the control group. For example, the first line of Table 2 tells us that the average client in the control group correctly answered 0.383 (s.d. 0.311) questions on marketing knowledge (columns (1) and (2)). The outcome variables are displayed on the left and grouped by the hypotheses groups. We display the original indicators (such as in row of hypothesis 1a), but also the individual indicators used to calculate the final index (rows of hypothesis 1b). For example, for the indicator “Share of marketing knowledge questions answered correctly”, the regression model estimates an insignificant coefficient of 0.000 (s.e. 0.012) (columns (3) and (4)), which remains largely unchanged if we include FSP fixed effects (columns (5) and (6)). Considering only FSP 8, we obtain a larger and now highly significant coefficient of 0.08 (s.e. 0.028) (columns (7) and (8)). Finally, the very first column tells us that this outcome tests hypothesis 1a from the PAP (see Appendix B.1).

Overall, the tables illustrate that in the short-run we find no robust evidence for the hypotheses that the intervention achieved significant effects on intermediate outcomes averaged across all FSP, and this is independent of the model specification. The notable exception is an increase in clients who include cash flow statements in their business plan by 3.1 percentage points, which is a robust result. Consistently, we also find no significant effects on final outcomes. Meanwhile, FSP 8 achieved some positive and significant results on



intermediate outcomes. Most notably, the share of marketing questions answered correctly increased by 7.9 percentage points and the share of clients preparing a business/ financial plan by a remarkable 18 percentage points in this FSP. Additionally, we find reduced costs, but also reduced profits for their clients: the total cost of business in the last 30 days decreased by 10.28 million IDR and the profit generated decreased in the last 30 days by 2.70 million IDR in the treatment group compared to the control group, all else equal.

In what follows, we display the results graphically and provide a more in-depth discussion for selected indicators and results.

Table 2: Intermediate Outcomes

H	Outcome	N ^{al}	Control		All		All + FSP FE		FSP 8	
			Mean (1)	SD (2)	β^{all} ITT (3)	SE (4)	$\beta^{all, FEIT}$ (5)	SE (6)	$\beta^{FSP8, ITT}$ (7)	SE (8)
Panel A: Marketing Knowledge										
1a	Share of marketing knowledge questions answered correctly	4277	0.383	0.311	-0.000	0.012	0.000	0.010	0.079	0.028 ***
1b	Ans. cor.: First thing to do in marketing	4233	0.467	0.499	-0.005	0.018	-0.004	0.016	0.072	0.052
1b	Ans. cor.: Total profit depends on	4248	0.348	0.477	0.003	0.020	0.006	0.016	0.084	0.051 *
1b	Ans. cor.: A good salesman is	4256	0.334	0.472	0.003	0.019	0.001	0.015	0.082	0.041 **
Panel B: Financial Management Knowledge										
2a	Share of financial management knowledge questions answered correctly	4283	0.652	0.315	-0.001	0.007	-0.001	0.009	0.011	0.028
2b	Ans. cor.: Purpose of business plan	4252	0.520	0.500	-0.006	0.011	-0.007	0.015	0.012	0.034
2b	Ans. cor.: Accurate business records require	4222	0.586	0.493	0.017	0.012	0.019	0.015	0.028	0.051
2b	Ans. cor.: More liabilities than assets means	4248	0.852	0.355	-0.015	0.010	-0.015	0.011	-0.009	0.043
2c	Share of clients not knowing the payment type	3610	0.111	0.315	-0.014	0.006 **	-0.015	0.010	0.003	0.031
Panel C: Marketing Practices										
3a	Marketing practices index n.m.	4292	1.515	1.594	0.020	0.052	0.009	0.047	0.274	0.132 **
3b	Nr. of different marketing forms used n.m.	4292	0.250	0.518	0.025	0.021	0.021	0.016	0.083	0.043 *
3b	Share of MSEs exporting outside Indonesia	3936	0.173	0.378	-0.011	0.013	-0.015	0.012	0.028	0.029
Panel D: Financial Management Practices										
4a	Share of clients keeping business and HH finances separately	3922	0.438	0.496	-0.014	0.017	-0.013	0.016	-0.071	0.054
4b	Share of clients investing profit into the business	3909	0.664	0.472	0.016	0.018	0.008	0.015	0.051	0.056
4c	Share of clients preparing a business/ financial plan	3917	0.446	0.497	0.019	0.019	0.020	0.016	0.180	0.045 ****

H	Outcome	N ^{al}	Control		All		All + FSP FE		FSP 8	
			Mean (1)	SD (2)	β^{all} ITT (3)	SE (4)	$\beta^{all,FEIT}$ (5)	SE (6)	$\beta^{FSP8ITT}$ (7)	SE (8)
4c	Business includes plan how to improve a product	1720	0.628	0.484	0.038	0.031	0.036	0.025	-0.205	0.104 *
4c	Business plan includes plan how to reach a new market	1720	0.485	0.500	-0.046	0.031	-0.043	0.026 *	0.004	0.112
4c	Business plan includes expected sales and costs	1720	0.291	0.455	-0.020	0.017	-0.022	0.022	-0.015	0.095
4c	Business plan includes cash flow	1720	0.066	0.249	0.031	0.008 ***	0.033	0.013 ***	0.002	0.040
4d	Financial management practices indexn.m.									
FSP FE			No		Yes		No			

Note: This table shows intermediate outcome variables on the left, the different models and statistics on top. The first column refers to the hypothesis as numbered in PAP. The same clients are followed over time. Δ Sample: The estimation sample is represented in columns (3)-(6) (N = 4,292), FSP 8 in columns (7)-(8) (N = 314).

Δ Source: Endline survey (2018 - 2019).

Δ Columns (1)-(2) display the control group mean and standard deviation. In columns (3)-(8) we present regression results. In all regressions we control for age, gender, the twenty most imbalanced baseline covariates, dummies indicating imputation in baseline variables, and enumerator fixed effects. In columns (3)-(4) we cluster the standard errors at the FSP-level. In columns (5)-(6) we add FSP fixed effects and present robust standard errors. In columns (7)-(8) we restrict our regressions to clients of FSP 8, which showed the highest compliance with the randomization protocol. Additionally, we display the sample size for each of the outcome variables for the full estimation sample (N^{al}), and for the FSP 8 (N^{FSP8}).

Δ The superscript n.m. indicates that missing values were interpreted as zero to generate index variables.

Δ The statistical significance is given as follows: * indicates $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Δ See Table A.14 in the Appendix for results for each FSP individually. Δ We also tested for changes in taxation behaviour as per PAP but are not reporting the results here.

Table 3: Final Outcomes

H	Outcome	N ^{al}	Control		All		All + FSP FE		FSP 8	
			Mean (1)	SD (2)	β^{all} ITT (3)	SE (4)	$\beta^{all, FEIT}$ (5)	SE (6)	β^{FSP8IT} (7)	SE (8)
Panel A: Business Financials										
5a	Revenue in the last 30 days in mln IDR ^{99p}	2971	21.214	45.177	1.192	2.058	1.117	1.532	-5.964	5.165
5b	Nr. of permanent workers	3920	2.188	5.295	-0.145	0.136	-0.151	0.144	-1.048	0.650
5c	Nr. of permanent and casual workers (sum) ^{99p}	3942	4.323	5.950	0.135	0.224	0.138	0.183	-1.114	0.896
6a	Total cost of business in the last 30 days in mln IDR ^{99p}	4292	15.780	39.772	-0.361	1.467	-0.349	1.143	-10.281	5.179 **
6b	Profit generated in the last 30 days in mln IDR ^{99p}	3233	6.012	9.728	0.216	0.355	0.254	0.345	-2.704	1.282 **
6b	Profit in the worst month mln IDR ^{99p}	3303	3.955	6.851	0.181	0.191	0.207	0.243	-1.143	0.916
6b	Profit in the best month in mln IDR ^{99p}	3309	9.859	15.953	-0.046	0.518	-0.001	0.524	-3.857	2.158 *
6c	Share of business earnings covers the exp. of this business	3900	0.911	0.285	-0.011	0.006 *	-0.012	0.010	-0.015	0.037
6d	Profit increased during the last 6 months	3887	0.325	0.469	-0.000	0.019	0.000	0.016	0.066	0.053
7a	New business started in 2018 (in addition to current business)	3941	0.084	0.278	-0.002	0.010	-0.005	0.009	0.003	0.022
Panel B: Business Attitudes										
8a	Business perception index (standardized score)	4248	0.006	0.987	-0.003	0.035	-0.000	0.031	-0.061	0.095
8b	Nr. of barriers for business	4292	1.148	0.874	-0.003	0.029	-0.001	0.026	-0.056	0.098
Panel C: Household Financials										
9a	Total savings from all sources in mln IDR ^{90p}	2151	11.531	20.133	-1.022	0.891	-1.002	0.726	0.717	2.918
9b	Share of clients reporting increase in HH exp. in the last 6 months	4239	0.457	0.498	0.013	0.015	0.012	0.016	0.050	0.055
9b	HH exp. on non-durables in the last 30 days in mln IDR ^{99p}	4275	3.038	3.431	-0.147	0.133	-0.133	0.104	-0.174	0.380
9b	HH exp. on durables in the last 6 months in mln IDR ^{99p}	4288	9.454	31.861	-1.595	1.025	-1.640	0.950 *	-0.347	1.682

H	Outcome	N ^{al}	Control		All		All + FSP FE		FSP 8	
			Mean (1)	SD (2)	β^{all} ITT (3)	SE (4)	$\beta^{all, FEITT}$ (5)	SE (6)	$\beta^{FSP8ITT}$ (7)	SE (8)
Panel D: Life Satisfaction										
10a	Life satisfaction (1=worst, 10=best)	4191	7.342	1.781	0.017	0.075	0.008	0.055	0.152	0.161
10b	Difference between life satisfaction now and two yrs. ago	4165	0.354	1.740	-0.035	0.088	-0.050	0.057	-0.015	0.146
Panel E: Loan Behaviour										
11b	Current/last loan used for productive purposes	3610	0.729	0.445	-0.012	0.012	-0.013	0.015	-0.025	0.053
11c	Share of clients want to borrow more for business	1418	0.790	0.408	-0.038	0.021 *	-0.037	0.024	0.069	0.107
12a	Currently in loan default or behind with repayments for any loan	2937	0.189	0.391	-0.021	0.016	-0.019	0.015	-0.097	0.062
12b	Share of clients reporting being late with the loan payment at the FSP	3958	0.181	0.385	-0.017	0.009 *	-0.011	0.012	-0.004	0.048
FSP FE			No		Yes		No		No	

Note: This table shows final outcome variables on the left, the different model and statistics on top. The first column refers to the hypothesis as listed in PAP. The same clients are followed over time. Standard errors are clustered at the FSP level.

▲ Columns (1)-(2) display the control group mean and standard deviation. In columns (3)-(8) we present regression results. In all regressions we control for age, gender, twenty most imbalanced baseline covariates, dummies indicating imputation in baseline variables, and enumerator fixed effects. In columns (3)-(4) we cluster the standard errors at the FSP-level. In columns (5)-(8) we add FSP fixed effects and present robust standard errors. In columns (7)-(8) we restrict our regressions to clients of FSP 8, which showed the highest compliance with the randomization protocol. Additionally, we display the sample size for each of the outcome variables for the full estimation sample (Nall), and for the FSP 8 (NFSF8).

▲ The superscript np. indicates the winsorizing level.

▲ The statistical significance is given as follows: * indicates $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

▲ See Table A.15 in the Appendix for results for each FSP individually. ▲ We also tested for changes in taxation behaviour as per PAP but are not reporting the results here.

Intermediate Outcomes. The following graphs illustrate selected results from Table 2, columns (3) and (4). To test whether the treatment improved knowledge in the training topics, we asked a set of three questions on marketing and financial planning each. The questions were taken directly from the training material used by the loan officers and masters trainers who trained them.

For each of the six knowledge questions, Figure 12 shows the share of clients in treatment and control group in all FSPs, who answered the respective question correctly. As the graph shows, this share increased among the treatment group for three out of six questions, but these increases are insignificant. Given that the shares also decreased for three questions, these differences are likely due to chance only.

Figure 12: Answers to Knowledge Questions in Treatment and Control Group



Note: Figure 12 depicts the share of clients who correctly answered the knowledge questions, the average share of questions correctly answered, and the share of clients knowing their payment type, in treatment and control group. The red bar indicates the ITT effect according to the specification in Table ??, columns (3) and (4).

▲ Source: Endline survey (2018 - 2019).

The main outcome variables of interest are the questions answered correctly in the two training topics. We find no significant effects on these outcomes, with the point estimates being very close to zero. It is important to note here that prior knowledge (before the intervention) of financial management was higher than marketing. The average control group client answered two questions on financial management correctly, but only one on marketing.

Interestingly, FSP 8 was able to significantly increase knowledge of marketing issues among treatment clients. The share of marketing questions answered correctly increased by 7.9 percentage points and this increase is significant at the 1% level. Knowledge of financial management remained unchanged even for this FSP.

We further tested whether treated clients are more likely to know their loan conditions, in particular their payment modalities (interest rate, profit sharing agreement). In the control group, 11.1% of clients were not able to answer this question, and in the treatment group this fraction is 1.4 percentage points lower. While the effect is significant in the selected specification, it is insignificant when accounting for FSP fixed effects.

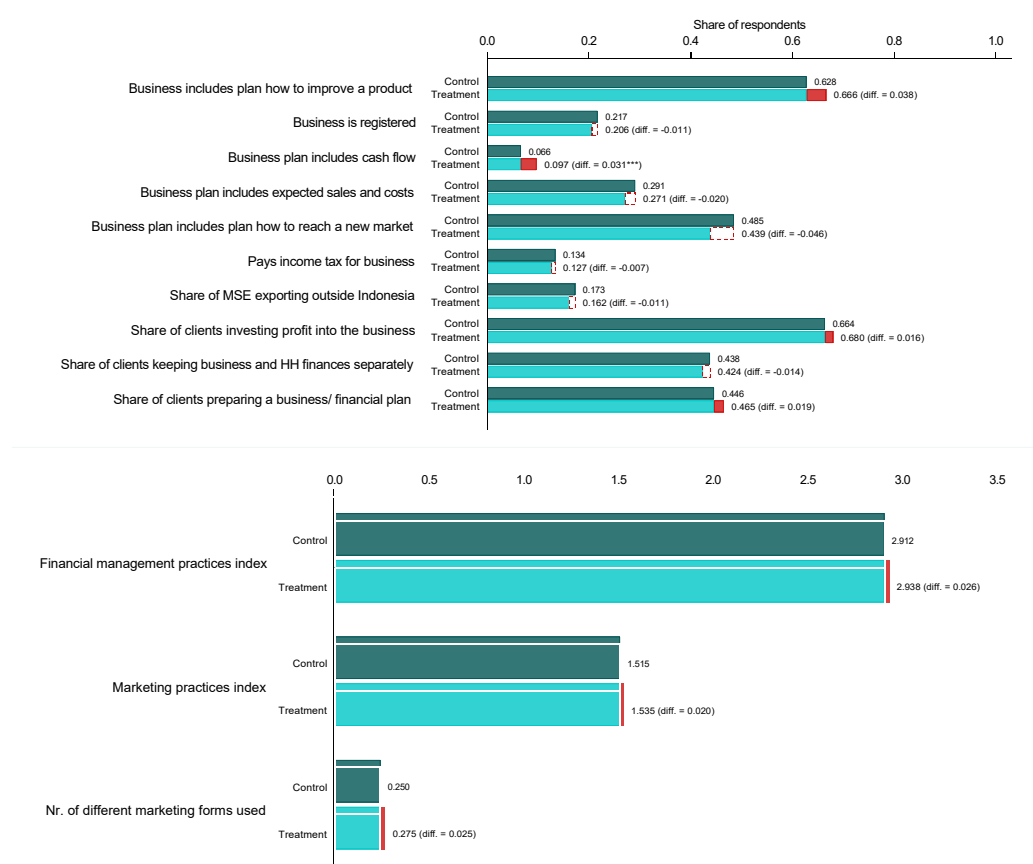
To test whether treatment led to behavioral changes in marketing and financial management practices, we asked two set of five questions each, regarding the two training topics. These questions were taken from De Mel et al. (2014), who evaluated the SIYB training intervention in Sri Lanka, albeit with a different target group. The main outcome variables were related to best practices followed. Marketing behavior is summarized in the marketing practice index that ranges from 0 to 5 with higher values representing the use of more desirable practices (e.g., visiting competitors, asking customers, using any kind of advertisement).



Questions related to financial management (e.g., frequency of reviewing the financial performance of the business) are summarized in the financial management index that ranges from 0 to 7, with higher values being more desirable. Note that the number of best practices followed in financial management exceeds that in marketing in the control group (2.912 versus 1.515). Contrary to De Mel et al. (2014), we find no significant effect of the treatment on these practices across FSPs. Given that our trainers might have focused on different topics than in De Mel et al. (2014), we ask additional questions on marketing and financial management practices, such as the number of marketing channels used, whether the business is registered or paying taxes, but also fail to detect changes on most of these outcomes.

For each of the practice questions, Figure 13 shows the share of clients in treatment and control group in all FSPs, who follow the practice. Again, we can see some increases and some decreases, but almost all are insignificant and hence likely due to chance only.

Figure 13: Business Practices in Treatment and Control Group



Note: Figure 13 depicts the share of clients who answered that they use a particular business practice, the scores on the practice indexes, and the number of marketing channels used, in treatment and control group. The red bar indicates the ITT effect according to the specification in Table ??, columns (3) and (4).

▲ Source: Endline survey (2018 - 2019).

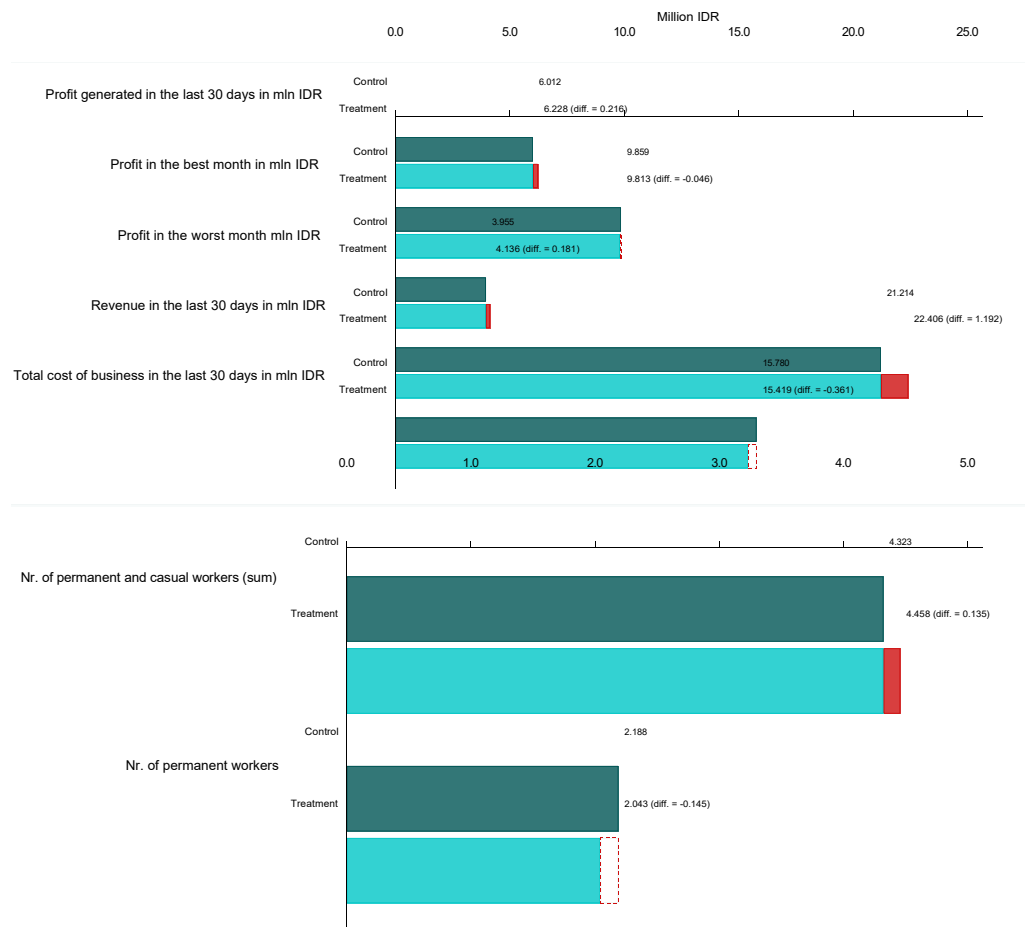
We find a significant and positive increase in the clients whose business plan includes cash flow analysis. This effect is robust across model specification when considering the full estimation sample, but this is not the case for FSP 8. In this FSP, however, we note significant and positive increase in the marketing practice index and the share of clients preparing a business/ financial plan. This is consistent with the increase in knowledge found in the treatment group of this FSP.

In summary, the intervention effect on intermediate outcomes is limited. We find no increase in marketing and financial management knowledge, except for FSP 8. We find improvement in business practices for the inclusion of cash flow in the business plan. The higher knowledge among treatment clients

in FSP 8, however, is consistent with changes in better marketing practices and business plan preparation, suggesting that the intervention has indeed the potential to affect client behavior.

Final Outcomes. Specifically, across model specification we find no effect on revenues during the last 30 days or on the number of permanent or total workers. The following graphs illustrate selected results from Table 3, columns (3) and (4). Figure 14 displays the averages for the quantitative financial measures for control and treatment group over all FSPs. Again, there is no consistent pattern which would explain the increases and decreases observed in different variables, and none of these are significant.

Figure 14: Profits, Revenues, Costs and Number of Workers in Treatment and Control Group

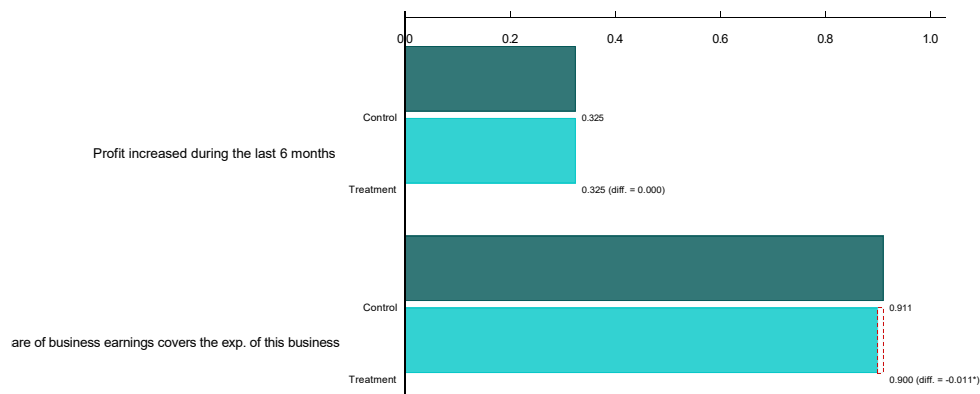


Note: Figure 14 depicts the average costs, revenues, profits and number of workers in treatment and control group. The red bar indicates the ITT effect according to the specification in Table ??, columns (3) and (4).

▲ Source: Endline survey (2018 - 2019).

Considering only FSP 8, we find a significant decrease in the cost of business in the last 30 days, but also a significant decrease in the profit generated in the last 30 days and in the profit in the best month when clients made the most profit. The latter is just weakly significant. The number of workers and revenues are also decreasing, but we lack significance here.

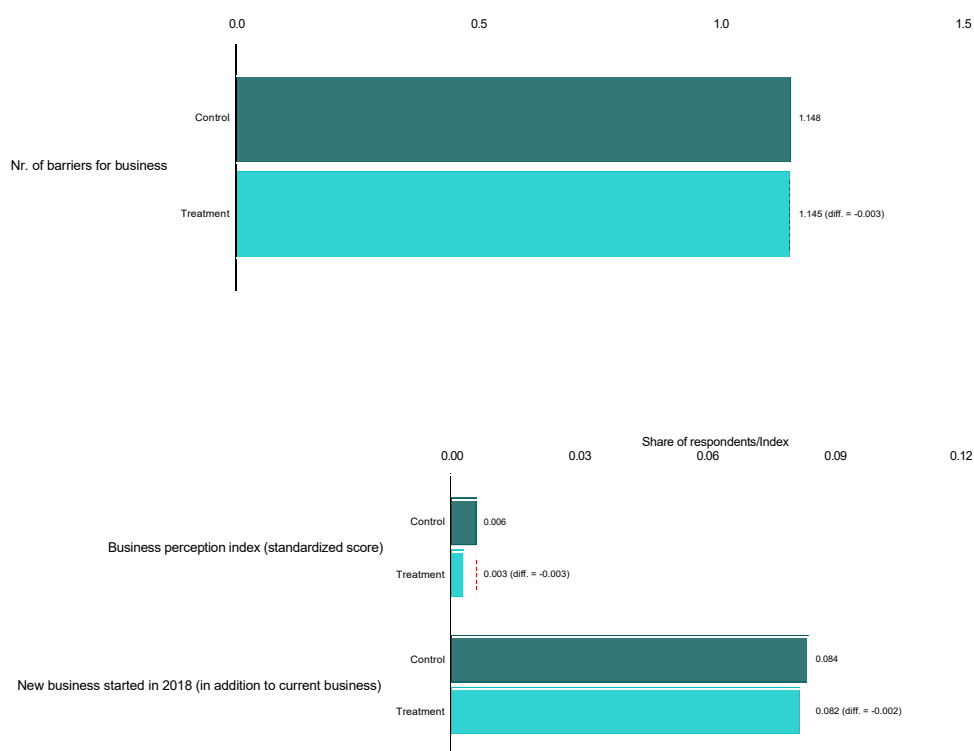
Our quantitative measurement of revenues, costs and profits are heavily affected by noise, which we already expected based on similar studies (De Mel et al. (2014)). To account for this, we also measured two binary outcomes: Whether earnings cover business expenses and whether profits increased in the last six months. Here as well we cannot display any intervention effect (see Figure 15). There is also no effect in these outcomes when considering only FSP 8.

**Figure 15: Increased Profits and Coverage of Business Expenses in Treatment and Control Group**

Note: Figure 15 depicts the share of clients who state that profits increased in the last six months and that earnings cover business expenses in treatment and control group respectively. The red bar indicates the ITT effect according to the specification in Table ??, columns (3) and (4).

▲ Source: Endline survey (2018 - 2019).

We further include a number of outcomes which proxy the optimism and entrepreneurial attitudes in our sample. The first item is whether clients started new businesses. We also asked about current attitudes towards running a business in Indonesia, i.e., whether it is perceived as complicated, risky, financially rewarding, earns respect, satisfaction or gives the feeling of security. We aggregated the answers into a standardized score, which we call the business perception index. Additionally, we asked about the perceived challenges and barriers to business development and counted the number of mentioned barriers. Figure 16 illustrates the result. Again, we observe no significant changes.

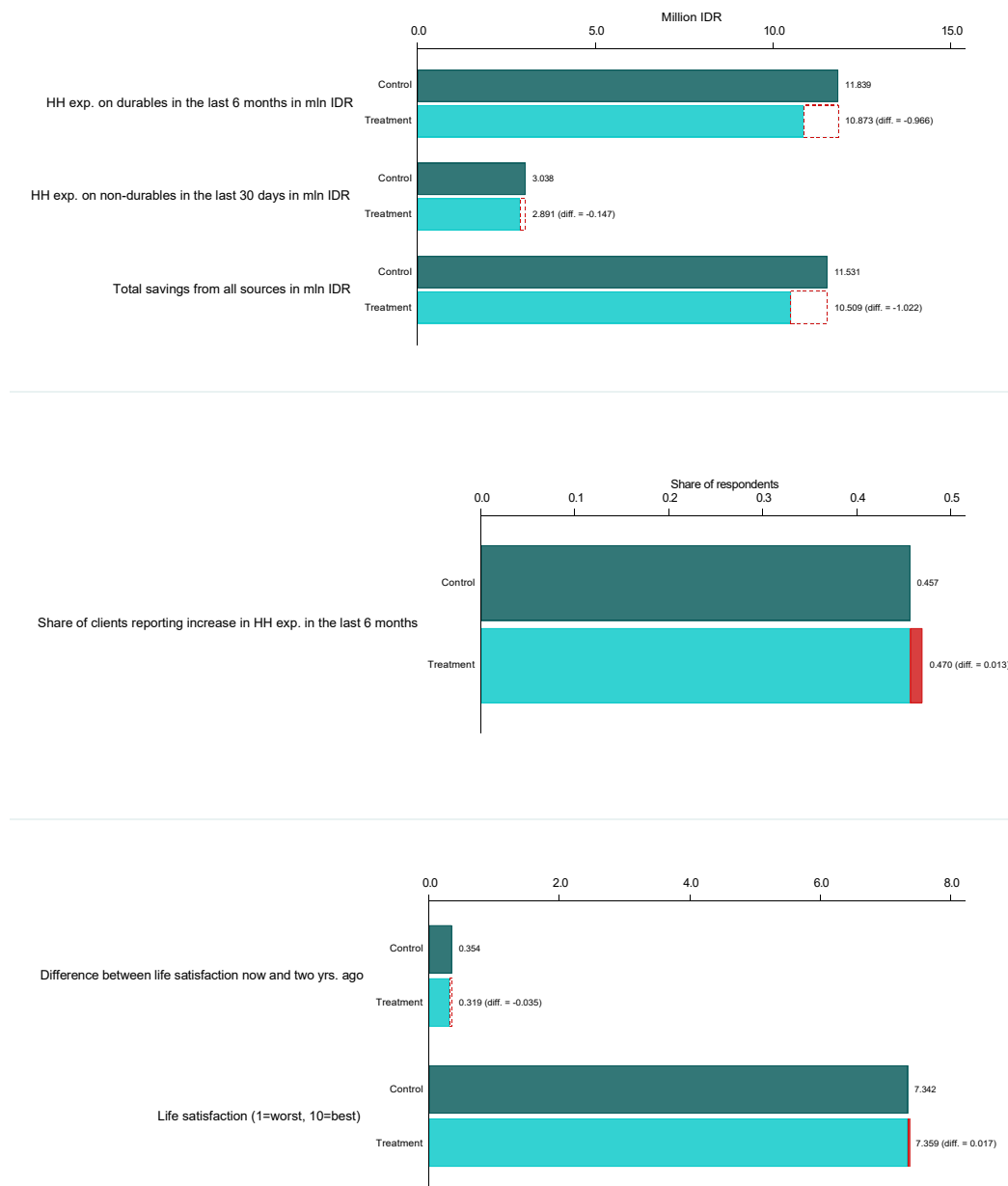
Figure 16: Entrepreneurial Attitudes and Business Start-ups in Treatment and Control Group

Note: Figure 16 depicts the share of clients having started a new business, the average business perception index and the number of barriers to business developments mentioned in treatment and control group. The red bar indicates the ITT effect according to the specification in Table ??, columns (3) and (4).

▲ Source: Endline survey (2018 - 2019).

Furthermore, we tested a number of outcomes, which relate to the clients' personal lives, i.e., total savings, household expenditures, life satisfaction and the change thereof during the past two years. Given the previous results, it is not surprising that we find no effects in these variables (see Figure 17).

Figure 17: Household Outcomes in Treatment and Control Group



Note: Figure 17 depicts the average total savings, household expenditures, life satisfaction and change thereof during the past two years, in treatment and control group. The red bar indicates the ITT effect according to the specification in Table ??, columns (3) and (4).

▲ Source: Endline survey (2018 - 2019).

Finally, we test whether treatment affected loan behavior. Majority of clients stated that they used their last loan for productive (business) purposes. Also, 72.9% plan to borrow more, but even a slightly larger proportion (79.0%) of clients in the control group want to borrow more. As illustrated in Figure 18, there is no statistically significant difference though.

**Figure 18: Loan Behavior Treatment and Control Group**

Note: Figure 16 depicts the share of clients having started a new business, the average business perception index and the number of barriers to business developments mentioned in treatment and control group. The red bar indicates the ITT effect according to the specification in Table ??, columns (3) and (4).

▲ Source: Endline survey (2018 - 2019).

In summary, we find no statistical evidence that the intervention affected the financial situation of the MSEs or the household, at least not in the short run. This result shows when averaging over all the FSPs. FSP 8 achieved significant decreases in cost and profit. The fact that profit also decreased might be explained by a more precise reporting due to the treatment. Other measurements such as optimism and entrepreneurial attitude, and loan behavior, show no change in any sample.

Local average treatment effects. As discussed before, implementation of the intervention as captured in the information from the online monitoring and from the endline survey varied. Some FSPs closely followed the original design protocol, while others appeared to have struggled. In additional analyses we adjust our estimates considering non-compliance. The effects are then to be interpreted as local average treatment effect (LATE), i.e., the average effects on the outcomes of interest for those clients who reportedly participated in the intervention. Thus, the effect is for the specific case of individuals who complied with the intervention and were also originally assigned to this treatment. In essence, the ITT results presented above are adjusted by the likelihood to actually participate in the intervention. The results are presented in the Appendix, Section C.3. As expected, we observe that the impacts are higher once we adjust the coefficients for the - certainly selective - decision to actually participate in the intervention. See Table A.8 for the results.²³

²³ Note: we were running linear models which is not always appropriate for binary outcome variables, but acceptable in a randomized setting

5.2 Comparison across treatment arms: Training vs. counseling

We estimate the effect for each of the three treatment arms separately and test for significant differences between them. We run the regression of our main specification on the three treatment arms separately. Since only the FSPs of Batch I opted to offer all three treatment arms, this part of our analysis is restricted to the subsample of Batch I clients, i.e., clients of saving and credit cooperatives. Results for other FSPs are reported in separate individual reports.

We regress the outcomes of interest simultaneously on being assigned to one of the three treatment arms and compare the results to a control group, which was assigned not to receive any treatment. Table 4 and 5 contain the results for intermediate and final outcomes respectively, estimating our main specification (with FSP fixed effects). In more detail, the columns describe the sample size, control mean and standard deviation in Batch I as well as the three ITT point estimates and standard errors of the T, C and TC treatments. For example, the first line in Table 4 tells us that the average control group client in Batch I answered 0.382 (s.d. 0.307) marketing knowledge questions correctly (columns (1)-(2)). Regressing the outcome “share of marketing knowledge questions answered correctly” on a treatment dummy indicating assignment to the training only arm yields a coefficient of -0.035 (s.e. 0.017) and this estimate is significantly different from zero with a p-value less than 0.05 (columns (3)-(4)). At the same time, the treatment dummy indicating assignment to the counseling only arm yields a coefficient of -0.028 (s.e. 0.017) (columns (4)-(5)), and on a treatment dummy indicating assignment to the training plus counseling arm a coefficient of -0.019 (s.e. 0.016) (columns (7)-(8)). For the last two coefficients we cannot confidently reject the hypothesis of equality to zero.

The stars in the table illustrate the p-values of testing the null hypothesis of no effect. There are a number of outcomes for which we estimate significant changes in one treatment arm but not in other arms. For example, the decrease in the share of marketing questions answered correctly in the T arm is weakly significant, but not in the other two treatment arms. However, the null hypothesis that any two treatment coefficients are equal cannot be rejected for any combination for this outcome. This is so owing to the three treatment coefficients being similar in magnitude, but there is least variance in the T arm.

Among intermediate outcomes, there are no effects that differ significantly across the treatment arms. Among the final outcomes, we find a significant decrease in the T treatment arm regarding the share of clients who want to borrow more for business and in the share of clients reporting being late in loan repayment. Both coefficients are, however, not significantly different from those estimated for the C or TC treatment arms respectively.

In summary, our hypothesis that we would observe the largest effects in the TC treatment is not supported by the survey data. However, it should be noted that it is likely that not all clients in the TC arm received both training and counseling. Instead, whereas some report participation only in the training course, others reported participating in the counseling session.

Table 4: Intermediate Outcomes by Treatment Arm (Batch I)

H	Outcome	N ^{al}	Control		Training		Counseling		Training & Counseling	
			Mean (1)	SD (2)	β^{all} ITT (3)	SE (4)	$\beta^{all,FEIT}$ (5)	SE (6)	β^{FSPITT} (7)	SE (8)
Panel A: Marketing Knowledge										
1a	Share of marketing knowledge questions answered correctly	2541	0.389	0.309	-0.035	0.017 **	-0.028	0.017 *	-0.019	0.016
1b	Answ. cor.: First thing to do in marketing	2513	0.457	0.499	-0.036	0.027	-0.040	0.027	-0.018	0.027
1b	Answ. cor.: Total profit depends on	2519	0.343	0.475	-0.023	0.027	-0.020	0.027	-0.024	0.027
1b	Answ. cor.: A good salesman is	2521	0.368	0.483	-0.043	0.027	-0.032	0.027	-0.015	0.027
Panel B: Financial Management Knowledge										
2a	Share of financial management knowledge questions answered correctly	2546	0.659	0.317	0.004	0.016	0.010	0.015	-0.006	0.016
2b	Answ. cor.: Purpose of business plan	2522	0.549	0.498	0.019	0.025	0.036	0.025	-0.000	0.026
2b	Answ. cor.: Accurate business records require	2508	0.577	0.494	-0.005	0.027	0.002	0.026	0.027	0.026
2b	Answ. cor.: More liabilities than assets means	2523	0.851	0.356	-0.001	0.020	-0.008	0.019	-0.043	0.021 **
2c	Share of clients not knowing the payment type	2000	0.147	0.354	-0.029	0.020	-0.004	0.021	-0.015	0.020
Panel C: Marketing Practices										
3a	Marketing practices index n.m.	2550	1.508	1.584	-0.015	0.079	0.022	0.080	-0.080	0.077
3b	Nr. of different marketing forms used n.m.	2550	0.284	0.527	-0.024	0.027	0.017	0.029	-0.001	0.027
3b	Share of MSEs exporting outside Indonesia	2376	0.175	0.380	-0.038	0.021 *	-0.007	0.021	0.005	0.021
Panel D: Financial Management Practices										
4a	Share of clients keeping business and HH finances separately	2374	0.468	0.499	-0.044	0.028	-0.022	0.027	-0.009	0.028
4b	Share of clients investing profit into the business	2360	0.646	0.479	-0.030	0.025	-0.002	0.025	0.012	0.025

H	Outcome	N ^{al}	Control		Training		Counseling		Training & Counseling	
			Mean (1)	SD (2)	β^{all} ITT (3)	SE (4)	$\beta^{all, FEIT}$ (5)	SE (6)	$\beta^{FSP, ITT}$ (7)	SE (8)
4c	Share of clients preparing a business/ financial plan	2364	0.420	0.494	-0.010	0.026	-0.021	0.026	0.001	0.026
4c	Business includes plan how to improve a product	967	0.646	0.479	0.044	0.040	0.051	0.041	0.061	0.041
4c	Business plan includes plan how to reach a new market	967	0.481	0.501	-0.070	0.045	-0.098	0.045 **	-0.017	0.045
4c	Business plan includes expected sales and costs	967	0.300	0.459	-0.054	0.036	-0.052	0.038	-0.029	0.038
4c	Business plan includes cash flow	967	0.093	0.291	0.022	0.022	0.043	0.024 *	0.020	0.023
4d	Financial management practices index n.m	2550	2.903	2.375	-0.054	0.120	-0.096	0.117	0.085	0.118

Note: This table shows intermediate outcome variables on the left, treatment arms on top. First column refers to the hypothesis as numbered in PAP. Same clients are followed over time. ▲ Sample: Batch I clients (N = 2,974)

▲ Source: Endline survey (2018 - 2019)

▲ Columns: (1)-(2) control group (N = 751). (3)-(4) training only arm (N = 748). (5)-(6) counseling only arm (N = 749). (7)-(8) training and counseling arm (N = 749).

▲ In all regressions we control for age, gender, twenty most imbalanced baseline covariates, dummies indicating imputation in baseline variables, and enumerator- and FSP fixed effects and present robust standard errors.

▲ The superscript n.m. indicates that missing values were interpreted as zero to generate index variables.

▲ The statistical significance is given as follows: * indicates $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

▲ Table ?? in the Appendix contains results without FSP fixed effects. ▲ We also tested for changes in taxation behaviour as per PAP but are not reporting the results here.

Table 5: Final Outcomes by Treatment Arm (Batch I)

H	Outcome	N ^{al}	Control		Training		Counseling		Training & Counseling	
			Mean (1)	SD (2)	$\beta_{all,ITT}$ (3)	SE (4)	$\beta_{all,FEIT}$ (5)	SE (6)	$\beta_{FSP,ITT}$ (7)	SE (8)
Panel A: Business Financials										
5a	Revenue in the last 30 days in mln IDR ^{99p}	1755	11.144	25.408	0.011	1.544	-1.507	1.503	-0.722	1.385
5b	Nr. of permanent workers	2377	1.504	3.383	0.144	0.265	-0.073	0.197	-0.219	0.166
5c	Nr. of permanent and casual workers (sum) ^{99p}	2381	3.234	4.210	0.008	0.226	-0.009	0.228	-0.287	0.204
6a	Total cost of business in the last 30 days in mln IDR ^{99p}	2550	8.779	27.135	-1.309	1.190	-1.107	1.254	-0.751	1.197
6b	Profit generated in the last 30 days in mln IDR ^{99p}	1957	4.106	7.150	-0.160	0.428	-0.308	0.412	0.264	0.441
6b	Profit in the worst month mln IDR ^{99p}	2008	2.579	4.794	0.059	0.295	-0.112	0.275	-0.076	0.301
6b	Profit in the best month in mln IDR ^{99p}	2012	6.226	11.304	-0.165	0.631	-0.629	0.636	0.205	0.673
6c	Share of business earnings covers the exp. of this business	2358	0.917	0.276	-0.019	0.017	-0.013	0.016	-0.029	0.017 *
6d	Profit increased during the last 6 months	2345	0.349	0.477	-0.022	0.027	0.022	0.027	-0.005	0.027
7a	New business started in 2018 (in addition to current business)	2380	0.073	0.261	0.019	0.016	-0.005	0.015	0.033	0.016 **
Panel B: Business Attitudes										
8a	Business perception index (standardized score)	2526	-0.096	1.026	0.046	0.053	0.044	0.053	0.024	0.055
8b	Nr. of barriers for business	2550	1.112	0.800	-0.028	0.043	-0.028	0.040	-0.014	0.041
Panel C: Household Financials										
9a	Total savings from all sources in mln IDR ^{99p}	1225	6.990	14.810	-0.190	1.075	-0.921	1.012	-0.772	1.004
9b	Share of clients reporting increase in HH exp. in the last 6 months	2520	0.454	0.498	0.026	0.028	-0.011	0.027	0.011	0.027
9b	HH exp. on non-durables in the last 30 days in mln IDR ^{99p}	2538	2.500	3.030	0.010	0.162	-0.173	0.154	-0.103	0.153

H	Outcome	N ^{al}	Control		Training		Counseling		Training & Counseling	
			Mean (1)	SD (2)	β^{all} ITT (3)	SE (4)	$\beta^{all;FEITT}$ (5)	SE (6)	$\beta^{FSP;ITT}$ (7)	SE (8)
9b	HH exp. on durables in the last 6 months in mIn IDR ^{99p}	2549	5.683	23.616	-0.872	1.263	-0.287	1.296	0.314	1.299
Panel D: Life Satisfaction										
10a	Life satisfaction (1=worst, 10=best)	2480	7.550	1.762	-0.024	0.095	-0.010	0.094	0.032	0.093
10b	Difference between life satisfaction now and two yrs. ago	2462	0.640	1.669	0.043	0.099	-0.136	0.101	-0.096	0.100
Panel E: Loan Behaviour										
11b	Current/last loan used for productive purposes	2000	0.725	0.447	-0.007	0.027	-0.011	0.027	0.008	0.027
11c	Share of clients want to borrow more for business	922	0.771	0.421	-0.083	0.042 **	-0.003	0.040	-0.044	0.040
12a	Currently in loan default or behind with repayments for any loan	1632	0.116	0.321	-0.006	0.020	0.006	0.020	0.025	0.021
12b	Share of clients reporting being late with the loan payment at the FSP	2280	0.117	0.322	-0.031	0.017 *	-0.006	0.018	0.005	0.018

Note: This table shows final outcome variables on the left, treatment arms on top. First column refers to the hypothesis as numbered in PAP. Same clients are followed over time. ▲ Sample: Batch 1 clients (N = 2,974)

▲ Source: Endline survey (2018 - 2019)

▲ Columns: (1)-(2) control group (N = 751), (3)-(4) training only arm (N = 748), (5)-(6) counseling only arm (N = 748), (7)-(8) training and counseling arm (N = 749).

▲ In all regressions we control for age, gender, twenty most imbalanced baseline covariates, dummies indicating imputation in baseline variables, and enumerator- and FSP fixed effects and present robust standard errors.

▲ The superscript np. indicates the winsorizing level.

▲ The statistical significance is given as follows: * indicates p < 0.1, ** p < 0.05, *** p < 0.01, **** p < 0.001.

▲ Table ?? in the Appendix contains results without FSP fixed effects. ▲ We also tested for changes in taxation behaviour as per PAP but are not reporting the results here.



6.

HETEROGENEOUS EFFECTS



THE RESULTS presented above might mask the fact that certain client subgroups benefited from the intervention while others did not. Therefore, we test whether the treatment effect differs across subgroups by adding an interaction term (γ_{ia} in Equation 2) to display whether the treatment effect is stronger or weaker for certain subgroups. Given the results of the previous section on different treatment arms, we will refer to the overall treatment of the intervention, not distinguishing between the groups.

In what follows, we give the subgroup example of being female. The main question here is: Did females benefit more or less from the intervention than men? One possibility is that female clients benefited more than male clients due to better targeting. The opposite may be true owing to less time available to participate in the training offered or counseling session. The presence of a significant interaction of treatment $\times$ being female indicates that the effect of the intervention on the response variable is different for males and females. The effect of being female is $\gamma_{ia} + \gamma_1$, while the effect of the intervention is $\beta_{ITT} + \gamma_{ia}$. Thus, γ_{ia} measures the difference in effects between being female and being male. Accordingly, its standard error or p-value indicates whether the effects are significantly different for the two types of groups.

In Table 6 we present the results for subgroups defined by characteristics of the entrepreneur or business, whereas in Table 7 we summarize results for subgroups defined by characteristics of the respective FSP. In more detail: Table 6 Part (1) analyses an indicator for female clients, Part (2) an indicator for age above 45, Part (3) an indicator for above secondary school education, Part (4) an indicator for entrepreneurs in the trading sector, and Part (5) an indicator for non-registered MSEs. Part (1) in Table 7 contains results for Batch I, Part (2) results for Batch III/24, Part (3) distinguishes FSPs by amount of loans disbursed in the last 12 months, Part (4) by the number of years FSP has been operating, and Part (5) by the share of female active borrowers. We present p-values for the null hypothesis that the effect is equal across the two respective groups (dark grey area). Additionally, we present the p-value, testing the null hypothesis that the intervention had no effect on the particular subgroup, i.e., of $\beta_{ITT} + \gamma_{ia} = 0$ (light gray areas).²⁵ Testing a large number of outcomes using our model specification independently increases the probability of rejecting a true null hypothesis for at least one outcome above the significance level used for each test. We reduce the number of tested hypotheses by restricting attention here to final outcomes only, with the intermediate results presented in the Appendix, Tables A.9 and A.10.

For example, the first line in Table 6 looks at the change in reported revenues in the last 30 days in million IDR. There is a positive increase for female clients by 1.325 million IDR (Part (1), “YES” column) and also a positive increase for male clients by 0.805 million IDR (Part (1), “No” column). Neither the difference between these two groups (Part (1), dark grey column) nor the total effect on female clients (Part (1), light



grey column) is significant. Thus, we cannot make the claim that the effects differ between men and women for this outcome variable. Part (2) of the table then tell us that there was a positive increase by 3.996 million IDR for clients older than 45 years (“YES” column), and a decrease by 1.630 million IDR for younger clients (“NO” column). This difference is significant with a p-value below 0.05, and also the total effect on older clients is significantly different from zero. The remaining columns can be interpreted analogously.

24 Batch II as reference category is not included in the table.

25 This sum tells us the size of the effect for the group of interest.

Table 6: Effects by Entrepreneur and Business Characteristics (All FSPs)

Group	(Part 1) Female				(Part 2) Age>45				(Part 3) More than secondary school education				(Part 4) MSEs in trading sector				(Part 5) MSEs not registered			
	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total
Panel A: Business Financials																				
Revenue in the last 30 days in mln IDR ^{29p}	1.325	0.805			3.996	-1.630	*	*	-1.393	3.326			0.963	1.454			0.124	4.463		
Nr. of permanent workers	0.021	-0.435			-0.204	-0.087			-0.508	0.144	**	*	0.055	-0.369			-0.041	-0.506		
Nr. of permanent and casual workers (sum) ^{30p}	-0.035	0.430			0.318	-0.048			-0.056	0.269			0.052	0.303			0.190	-0.029		
Total cost of business in the last 30 days in mln IDR ^{29p}	0.133	-1.152			1.458	-2.211	*		-1.362	0.620			0.062	0.220			-0.622	0.582		
Profit generated in the last 30 days in mln IDR ^{29p}	0.163	0.398			0.665	-0.159			-0.290	0.676			0.496	-0.160			-0.024	1.197		
Profit in the worst month mln IDR ^{29p}	0.094	0.419			0.280	0.131			0.137	0.275			0.233	-0.086			-0.143	1.409	**	
Profit in the best month in mln IDR ^{29p}	0.302	-0.501			0.132	-0.127			-1.298	1.001	**		-0.642	0.959			-0.041	0.139		
Share of business earnings covers the exp. of this business	-0.020	0.002			-0.002	-0.021			0.000	-0.021			-0.009	-0.007			-0.012	-0.011		
Profit increased during the last 6 months	-0.002	0.004			0.006	-0.003			-0.015	0.015			-0.009	0.011			0.008	-0.025		
New business started in 2018 (in addition to current business)	0.007	-0.024			-0.013	0.007			-0.006	-0.004			-0.002	-0.022			-0.001	-0.017		
Panel B: Business Attitudes																				
Business perception index (standardized score)	0.031	-0.053			-0.006	0.001			0.055	-0.046	*		0.007	-0.016			0.001	-0.007		
Nr. of barriers for business	0.013	-0.022			-0.004	0.006			-0.008	0.008			-0.001	-0.041			-0.004	0.011		
Panel C: Household Financials																				
Total savings from all sources in mln IDR ^{30p}	-1.364	-0.444			-1.107	-0.924			-2.571	0.301	**	**	-2.819	0.463	**	**	-1.280	0.109	*	
Share of clients reporting increase in HH exp. in the last 6 months	0.002	0.028			0.024	0.003			0.001	0.023			0.003	0.009			0.017	-0.006		
HH exp. on non-durables in the last 30 days in mln IDR ^{29p}	-0.045	-0.273			-0.081	-0.193			-0.133	-0.127			-0.293	-0.080	*		-0.149	-0.077		
HH exp. on durables in the last 6 months in mln IDR ^{29p}	-0.123	-4.189	*		-0.953	-2.296			-3.344	-0.228	**		-2.159	-2.762			-1.622	-1.702	*	
Panel D: Life Satisfaction																				
Life satisfaction (1=worst, 10=best)	0.022	-0.010			0.013	0.001			-0.196	0.176	****		-0.041	0.037			-0.003	0.043		
Difference between life satisfaction now and two yrs. ago	-0.102	0.037			0.001	-0.092			-0.186	0.066	**	**	-0.101	0.075			-0.034	-0.106		
Panel E: Loan Behaviour																				
Current/last loan used for productive purposes	-0.014	-0.011			-0.000	-0.026			-0.002	-0.024			-0.017	-0.020			-0.032	0.052	**	*
Share of clients want to borrow more for business	-0.032	-0.049			-0.029	-0.039			0.029	-0.075	**		-0.060	-0.030	*		-0.025	-0.093		

Table 6: Continued from previous page

Group	(Part 1) Female			(Part 2) Age>45			(Part 3) More than secondary school education			(Part 4) MSEs in trading sector			(Part 5) MSEs not registered						
	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total			
Currently in loan default or behind with repayments for any loan	-0.023	-0.013			-0.028	-0.011			-0.030	-0.010			0.019	-0.082	***			-0.015	-0.033
Share of clients reporting being late with the loan payment at the FSP	-0.012	-0.010			-0.009	-0.013			-0.014	-0.008			0.008	-0.057	**			-0.018	0.016
N	2827.000	1465.000			2126.000	2157.000			1931.000	2355.000			2026.000	1411.000				3385.000	907.000

▲ Note: Table 6 documents the outcome variables on the left. Parts (1)-(5) on the top indicate heterogeneous effects for different groups. ▲ Sample: Estimation sample (all FSPs) (N ≤ 4,292).

▲ Source: Endline survey (2018 - 2019).

▲ Part (1): Indicator for female clients; Part (2): Indicator for age above 45; Part (3): Indicator for above secondary school education; Part (4): Indicator for MSEs in trading sector as in baseline; Part (5): Indicator for MSEs not registered. Dark grey columns show significance level of the null hypothesis which assumes the effect to be equal across the two respective groups. Light grey columns show significance level of the null hypothesis of zero effect in the particular group.

▲ In all regressions we control for the twenty most imbalanced baseline covariates and dummies indicating imputation in baseline variables, enumerator- and FSP fixed effects, and, if applicable, for age and gender. The superscript np. indicates the winsorizing level.

▲ The statistical significance is given as follows: * p < 0.1, ** p < 0.05, *** p < 0.01, **** indicates p < 0.001.

Table A.9 in the Appendix reports the results for intermediate outcomes. ▲ We also tested for changes in taxation behaviour as per PAP but are not reporting the results here.

Table 7: Effects by FSP Characteristics (All FSPs)

Group	(Part 1) Batch I (Cooperatives)				(Part 2) Batch II (Banks)				(Part 3) Large FSP				(Part 4) Old FSP				(Part 5) Many female borrowers			
	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total
Panel A: Business Financials																				
Revenue in the last 30 days in mln IDR ^{9p}	-0.659	2.994			0.849	1.173			3.246	-1.496			4.345	-1.446	*		0.011	-2.043		
Nr. of permanent workers	0.018	-0.341			-0.180	-0.145			-0.094	-0.215			-0.406	0.042			-0.125	-0.216		
Nr. of permanent and casual workers (sum/99p)	-0.043	0.342			0.111	0.144			0.430	-0.196	*		0.228	0.070			-0.061	-0.217		
Total cost of business in the last 30 days in mln IDR ^{9p}	-1.355	0.730			2.288	-0.953			1.293	-2.294			0.551	-0.982			-1.078	-2.013		
Profit generated in the last 30 days in mln IDR ^{9p}	-0.172	0.740			1.972	-0.098	*		1.104	-0.735	***		0.621	-0.022			0.312	-0.058		
Profit in the worst month mln IDR ^{9p}	-0.122	0.585			1.243	-0.003			0.627	-0.271	*		0.296	0.140			-0.113	0.353		
Profit in the best month in mln IDR ^{9p}	-0.523	0.605			1.616	-0.326			0.988	-1.099	**		0.114	-0.085			0.262	-0.858		
Share of business earnings covers the exp. of this business	-0.019	-0.003			-0.010	-0.012			-0.004	-0.020			-0.007	-0.016			-0.026	-0.006		
Profit increased during the last 6 months	-0.001	0.002			-0.055	0.011			-0.005	0.007			-0.021	0.017			0.005	-0.009		
New business started in 2018 (in addition to current business)	0.014	-0.026	**		-0.057	0.005	**	***	-0.020	0.013	*		-0.025	0.011	*		0.009	-0.008		
Panel B: Business Attitudes																				
Business perception index (standardized score)	0.051	-0.056	*		0.014	-0.004			-0.025	0.028			-0.024	0.016			0.091	-0.037	*	*
Nr. of barriers for business	-0.020	0.020			-0.002	-0.000			-0.034	0.039			-0.003	0.001			-0.014	-0.032		
Panel C: Household Financials																				
Total savings from all sources in mln IDR ^{9p}	-1.134	-0.874			-1.175	-0.965			-1.157	-0.801			-2.029	-0.271			-0.516	-0.586		
Share of clients reporting increase in HH exp. in the last 6 months	0.011	0.013			-0.007	0.016			-0.003	0.030			0.017	0.009			0.005	0.016		
HH exp. on non-durables in the last 30 days in mln IDR ^{9p}	-0.118	-0.149			-0.270	-0.101			-0.306	0.071	*		-0.205	-0.082			0.088	-0.375	**	
HH exp. on durables in the last 6 months in mln IDR ^{9p}	-0.113	-3.283	*		-5.726	-0.703	*	**	-1.651	-1.628			-3.821	-0.107	*	**	-0.154	-2.359		
Panel D: Life Satisfaction																				
Life satisfaction (1=worst, 10=best)	-0.021	0.038			-0.024	0.015			0.088	-0.089			0.053	-0.025			0.028	0.009		
Difference between life satisfaction now and two yrs. ago	-0.086	-0.012			-0.119	-0.034			-0.032	-0.071			0.100	-0.157	**		0.033	-0.148		
Panel E: Loan Behaviour																				
Current/last loan used for productive purposes	-0.008	-0.018			-0.034	-0.008			-0.010	-0.018			0.008	-0.031			0.004	-0.036		
Share of clients want to borrow more for business	-0.037	-0.036			-0.053	-0.034			-0.032	-0.042			-0.003	-0.060			-0.029	-0.035		
Currently in loan default or behind with repayments for any loan	0.004	-0.040			-0.050	-0.012			-0.018	-0.020			-0.022	-0.016			0.033	-0.071	****	**

Table 7: Continued from previous page

Group	(Part 1) Batch I (Cooperatives)				(Part 2) Batch II (Banks)				(Part 3) Large FSP				(Part 4) Old FSP				(Part 5) Many female borrowers			
	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total	Yes	No.	Diff.	Total
Share of clients reporting being late with the loan payment at the FSP	-0.008	-0.013			-0.030	-0.006			-0.019	-0.000			0.007	-0.024			0.017	-0.037	**	
N	2550.000	1742.000			684.000	3608.000			2246.000	2046.000			1664.000	2628.000			1740.000	1808.000		

▲ Note: Table 7 documents the outcome variables on the left. Parts (1)-(5) on the top indicate heterogeneous effects for different groups. ▲ Sample: Estimation sample (all FSPs) (N ≤ 4,292).

▲ Source: Endline survey (2018 - 2019).

▲ Part (1): Indicator for Batch I; Part (2): Indicator for Batch II; Part (3): Indicator for loans disbursed in 12 months IDR is above median; Part (4): Indicator for numbers of years FSP is operating is above median; Part (5): Indicator for share of female active borrowers is above median. Dark grey columns show significance level of the null hypothesis that the effect is equal across the two respective groups. Light grey columns show significance level of the null hypothesis of zero effect in the particular group.

▲ In all regressions we control for the twenty most imbalanced baseline covariates, dummies indicating imputation in baseline variables, enumerator- and FSP fixed effects, and, if applicable, for age and gender. The superscript np. indicates the winsorizing level.

▲ The statistical significance is given as follows: * p < 0.1, ** p < 0.05, *** p < 0.01, **** indicates p < 0.001.

▲ Table A.10 in the Appendix reports the results for intermediate outcomes. ▲ We also tested for changes in taxation behaviour as per PAP but are not reporting the results here.

Effects by Entrepreneur and Business Characteristics. As has been addressed in previous studies, we test whether treatment effects differ across gender, age, and educational background. Regarding age and gender, we find no difference in effects, but there are substantial differences between clients with an education above or below secondary school education (the median): for more educated clients, the number of permanent workers decreased by 0.508 due to the treatment. The profit in the best month (when the profit is highest in the period) seems to have decreased as well; however, whereas the difference to the less educated clients is significant, the overall effect is not. Total savings decreased by 2.571 million IDR among more educated clients, and this is significantly different from zero and from the effect on less educated clients. Interestingly, more educated clients seem to experience a slight decrease in life satisfaction and the difference to less educated clients is significant with a p-value below 0.01 (a decrease of 0.2 on a 1-10 item scale). Among less educated clients, the intervention also had a negative impact on the willingness to borrow more for business: here, the share of clients who want to borrow more for business decreases by 7.5 percentage points among the treated clients, whereas the coefficient is positive, but insignificant, for more educated clients. Overall, there are only few differences in the impact of the intervention across entrepreneurs, but more educated clients seem to experience some adverse effects due to the intervention.

In addition to the client characteristics, we also analyze whether differences in the MSE characteristics lead to different effects of the intervention. Here, too, we find an adverse effect on total savings among MSEs in the trading sector compared to other sectors. There is a decrease in the share of clients who are in loan default or behind with repayments for any loan of 8.2 percentage points for businesses outside the trading sector (Column “NO”), but the effect is insignificant for the trading sector.²⁶ Among non-registered businesses, there are only few and only weakly significant results. Registered businesses seem to profit from the intervention through an increase in the profit in the worst month by 1.409 million IDR and an increase in the share of clients who used their current/ last loan for productive purposes by 5.2 percentage points. From the point of view of the FSPs, therefore, it may be more promising to offer the intervention to registered businesses outside the trading sector.

Effects by FSP Characteristics. FSPs that participated in the intervention was quite diverse including rural banks, development banks, and saving and credit cooperatives. FSPs also differ in their size in terms of outstanding loan volume and in their geographic focus. On the one hand, it is not unreasonable to assume that they also differ in their degree of operational soundness, i.e., larger and longer established rural and development banks, which also operate in urban areas are likely to have more formalized procedures, better risk management practices and better trained human resources than the cooperatives. Whereas Section 3.3 suggests that they were less likely to follow the randomization protocol or implement the treatment to a sufficient degree, this deviance is, at least to a certain degree, likely due to the fact that counseling was not identified as treatment by the clients in the endline survey. The notable exception is FSP 8 (rural bank), where loan officers diligently filled out the monitoring tool, thus treatment implementation is confirmed through this additional data, and for which we displayed separate results in the main tables above. On the other hand, it is common that higher tier financial institutions focus on breadth of outreach rather than depth, i.e., banks commonly focus more on the number of clients, whereas cooperatives would focus on reaching particular target groups which would otherwise be financially excluded.

Hence, the cooperatives focusing on rural areas might serve the more marginalized clients and be better positioned to address their needs than the rural and development banks.

²⁶ We asked for the sector of the MSE retrospectively in the endline survey. Not all clients answered this question, hence the smaller sample size. Of the estimation sample, 50.05% are in the trading sector, 17.24% in the manufacturing sector, 9.81% in services, and 5.06% in agriculture.



Indeed, we find some differences across the types of institutions. The following results will need to be considered not only with the differences in the type of institutions, but also the differences in the experimental design: Batches II and III commenced later and had fewer treatment arms. The cooperatives in Batch I achieved positive, but insignificant impact on the share of new businesses started in 2018, the business perception index, and the household expenditures on durables in the last six months. The rural and development banks of Batch II and III achieved significant but adverse effects on these variables: The share of new businesses decreased by 2.6 percentage points and the household expenditures on durables decreased by 5.910 million IDR in the last six months. The next column reveals that these negative effects are largely driven by the banks of Batch III: the coefficients in column “NO” Part (1) are smaller in magnitude than the coefficients of “YES” Part (2), i.e., Batch III alone generated larger decreases than Batch II and III jointly. The negative effects in Batch III are larger in magnitude and highly significant. This result is surprising given that the banks in Batch II and III are rather similar and we cannot explain the deviance in terms of FSP characteristics. This suggests that the success of the intervention hinges on FSP characteristics which are unobservable to us.

However, we find some difference in effects between small and large FSPs where we define as large any FSP, which disbursed more than 100 billion IDR in loans in the last 12 months. These large FSPs are FSPs 5 and 6 from Batch I, FSPs 9 and 10 from Batch II, and all three FSPs from Batch III. For these FSPs, we estimate larger and significantly positive effects on the three profit variables, with the profit generated in the last 30 days increasing by 1.104 million IDR among treated clients in large FSPs. FSPs that have been running for longer (>19 years) are FSP 6 from Batch I, FSPs 7, 8 and 10 from Batch II, and FSP 13 from Batch III. Clients in these FSPs show significant negative effects in household expenditure on durables (-5.916 million IDR in treatment compared to control group). In younger FSPs, the intervention led to a significant decrease in life satisfaction compared to two years ago. Finally, FSPs with more than 50% female borrowers are FSPs 1, 3, 4 and 6 from Batch I. These FSPs have more than 90% female borrowers, whereas the remaining FSPs have less than 45% female borrowers. However, we lack information on this variable for three FSPs, hence the lower total sample size here. In FSPs with largely female borrowers, the intervention led to an increase in the share of clients currently in loan default by 3.3 percentage points, which is sizable and significantly different from zero. In contrast, the share decreased by 7.1 percentage points in FSPs with largely male borrowers.

Overall, there is little conclusive evidence which would differentiate FSPs. In FSPs from Batch III we estimate some negative effects, but clients from larger FSPs (which include Batch III) experience positive effects on profits. FSPs with mostly male borrowers experience a decrease in loan default due to the treatment, whereas it increases for FSPs with mostly female borrowers. This was seen only at the aggregated FSP level, whereas we find no difference in effects on female or male clients individually.

Keeping in mind the differences in the experimental design, the wide range of results suggests that care must be taken when judging the external validity of such studies, not only across countries or population groups, but also across partner institutions serving similar target groups.

Additional Results: Selected Heterogeneous Effects. Further, we identify the subgroups of clients that benefited from the program the most by estimating the treatment effect for more than 100 different combinations of subgroups. Table 8 and Table 9 list the results for the intermediate and final outcomes respectively. For each outcome variable in both tables, we list the ITT point estimates for the full estimation sample and the subgroups of clients for which the estimated treatment effect is statistically significant. For example, in Table 8, the first listed outcome - the share of marketing questions answered correctly - has an insignificant coefficient of 0.000 for the full estimation sample. But for the clients of FSPs operating for above average number of years, the owners of MSEs with above average number of non-family, permanent workers, the clients from Batch 2, and the clients of FSP 8, the coefficients are positive and statistically significant (at least at the 5% significance level). The largest coefficient corresponds to the sub-sample of

clients from FSP 8. Please note that we report results only for those outcomes that have a statistically significant coefficient for at least one subgroup of clients (analyzing the full sample of clients in our study) and that show meaningful coefficients. All results are available upon request. Overall, out of the results that are highly significant (at the 0.1% significance level), we observe that clients from the trading sector had a higher likelihood to know their payment type (difference by 5.9 percentage points to others), that the individuals from Batch II and especially individuals from FSP 8 had a higher likelihood to prepare a business or financial plan (higher by 9.6 and 18.6 percentage points). A larger number of clients now include cash flows, including FSPs with above average amount of loans disbursed in the last 12 months (by 0.047 percentage points), FSPs with above average number of active borrowers (by 0.054 percentage points), and clients aged between 35 and 50 (by 0.065 percentage points). Results for the final outcomes are significant at the 1% significance level, which implies that some results are likely to be captured by chance given the high number of regressions we were running. At the 1% significance level we observe that the revenue of clients from FSP 10 increased by 22.331 million IDR, that the profit generated in the last 30 days was higher by 1.667 million IDR for treated clients reporting tough competition as the main business barrier compared to others. The same group also experienced higher profits in the worst month (by 1.157 million IDR) and higher profits in the best months (by 2.540 million IDR). While overall we find zero effects on the likelihood that profits increased during the last six months prior to the endline survey, we find that clients from FSP 9 were more likely to benefit (by 0.165 percentage points) and clients reporting “struggle with management” as the main business barrier (by 0.166 percentage points). On a scale from 1 to 10, life satisfaction improved by up to 0.3 percentage points for clients with primary education due to the program. Also, life satisfaction as compared to two years ago improved by 0.775 percentage points for treated clients having no education and by 0.34 for clients from development banks. A number of client groups seems to have benefitted in a way that they now show a smaller likelihood to be in loan default or being behind with repayments for any loans. In particular, these are clients who reported that business brings high income (a reduction by 7.2 percentage points), clients from the manufacturing sector (a reduction by 9.0 percentage points), and clients with more personnel support (a reduction by 11.4 percentage points). Finally, beneficiaries who already prepared business plans at baseline were 5.3 percentage points less likely to report late repayment at FSP.²⁷

26 Please note, further additional results for a range of subgroups are presented in the Appendix in Section C.4. Here, we split the sample by the following additional indicators: Indicator for the number of daily hours devoted to business being above median; Indicator for days per week devoted to business being above median; Indicator for the amount of last loan being above median. Indicator for entrepreneurs who selected >5 for willingness to take financial risk in the 1-10 scale (1-not willing, 10-very willing); Indicator for MSEs which exist for more than 2 years; Indicator for opportunity entrepreneurs; Indicator for reporting that marketing is important; Indicator for applying marketing practices; Indicator for preparing a business/ financial plan; Indicator for clients reporting no business barrier; Indicator for clients reporting finding market as the main barrier; Indicator for clients reporting low prices as the main barrier; Indicator for clients reporting expensive raw material as the main barrier; Indicator for clients reporting little experience as the main barrier; Indicator for clients reporting struggle with management as the main barrier; Indicator for clients reporting low competence of workers as the main barrier; Indicator for clients reporting access to finance as the main barrier; Indicator for clients reporting tough competition as the main barrier.

**Table 8: Selected Heterogenous Effects - Intermediate Outcomes**

Variable / Group	βITT	
Share of marketing knowledge questions answered correctly:		
All clients	0.000	
FSPs operating for above average number of years	0.033	**
MSEs with above average number of non-family, permanent workers	0.035	**
Clients from Batch 2	0.043	**
Clients of FSP 8	0.088	***
Answ. cor.: First thing to do in marketing:		
All clients	-0.004	
FSPs operating for above average number of years	0.049	**
Clients from Batch 2	0.062	**
Clients reporting no business barrier	0.077	**
Answ. cor.: Total profit depends on:		
All clients	0.006	
Clients with business in service sector	0.150	***
Clients of FSP 11	0.167	**
Answ. cor.: A good salesman is:		
All clients	0.001	
Clients having secondary education (baseline)	0.069	**
Clients with at least one person who can replace in business	0.084	**
Clients of FSP 8	0.097	**
Clients reporting above average nr. of participants in last training session	0.132	**
Clients reporting above average nr. of months since last training session	0.143	**
Clients of FSP 12	0.156	**
Answ. cor.: Accurate business records require:		
All clients	0.019	
Clients reporting above average amount of last loan	0.043	**
MSEs having generated above average profit in the last 30 days	0.054	**
Clients from Batch 2	0.067	**
Clients having secondary education (baseline)	0.069	**
Share of clients not knowing the payment type:		
All clients	-0.015	
Clients who started the business to make more money	-0.028	**
MSEs existing for above average nr. of yrs.	-0.030	**
Clients with above average business perception index	-0.043	***
Clients having secondary education (baseline)	-0.053	**
Clients not knowing the payment type	-0.058	**
Clients with business in trading sector (as in baseline)	-0.059	****
Clients having secondary education (endline)	-0.062	***
Clients reporting above average nr. of participants in last counseling session	-0.152	**
Marketing practices indexn.m.:		
All clients	0.009	
Clients older than 50	0.281	**
Clients of FSP 8	0.284	**

Table 8 - Continued from previous page

Variable / Group	βITT	
Nr. of different marketing forms used ^{a,m} :		
All clients	0.021	
Clients investing profit into the business	0.041	**
Clients spending above average time on looking after dependents/ household chores	0.052	**
Clients reporting above average amount of last loan	0.056	**
Clients reporting tough competition as the main business barrier	0.064	**
Clients reporting that business brings high income	0.077	**
Clients of rural banks	0.079	***
Clients of FSP 9	0.149	**
Clients reporting above average nr. of participants in last training session	0.169	**
Clients reporting above average nr. of participants in last counseling session	0.318	**
Clients reporting above average nr. of counseling sessions participated	0.462	***
Share of clients investing profit into the business:		
All clients	0.008	
FSPs with above average nr. of employees	0.042	**
Clients with above average business perception index	0.044	**
Clients reporting above average nr. of yrs. with FSP	0.046	**
Clients whose family members work	0.058	**
Share of clients preparing a business/ financial plan:		
All clients	0.020	
FSPs with above average nr. of employees	0.045	**
Clients of rural banks	0.073	**
Clients having secondary education (endline)	0.081	**
Clients reporting finding market as the main business barrier	0.085	***
Clients having university education (baseline)	0.090	**
Clients from Batch 2	0.096	****
Clients of FSP 8	0.186	****
Business includes plan how to improve a product:		
All clients	0.036	
Clients using marketing practices	0.052	**
Clients reporting that marketing is important	0.053	**
Clients having a competition strategy	0.057	**
Clients who want to borrow more for business	0.061	**
Clients investing profit into the business	0.062	**
Clients having more than one loan	0.065	**
Clients reporting that business brings satisfaction	0.066	**
MSEs having generated above average profit in the last 30 days	0.073	**
Clients with above average age	0.075	**
Clients reporting that business brings respect	0.076	**
FSPs with above average amount of loans disbursed in the last 12 months	0.078	**
Clients preparing a business/ financial plan	0.086	**
Clients reporting that business brings high income	0.094	**
Clients who started the business to make more money	0.095	***
Clients reporting struggle with management as the main business barrier	0.188	**



Table 8 - Continued from previous page

Variable / Group	β_{ITT}	
Clients of FSP 12	0.208	**
Business plan includes cash flow:		
All clients	0.033	***
Clients of FSP 1-12	0.027	**
Clients having customer identification strategy	0.029	**
Clients recorded in monitoring data	0.029	**
Clients having no additional income	0.031	**
Clients reporting that marketing is important	0.032	**
Clients reporting above average business/marketing practice index	0.033	***
Not registered MSEs	0.033	**
Clients covering the HH against unforeseen exp.	0.033	**
Clients having more than one loan	0.033	**
Clients having a strategy to make customers like their product	0.034	**
Clients reporting that business brings satisfaction	0.034	**
Clients reporting above average nr. of days per week devoted to business	0.036	**
Clients reporting above average nr. of yrs. their businesses exist	0.036	***
Clients with business in trading sector (as in baseline)	0.036	**
Clients whose spouse has a business	0.036	**
Clients reporting above average amount of last loan	0.036	**
Clients keeping record of all transactions	0.037	**
Clients having a competition strategy	0.037	***
Clients reporting cost of raw materials higher than cost of workers	0.038	***
Clients with above average education level	0.038	**
Clients reporting risk-taking behaviour	0.038	**
MSEs existing for above average nr. of yrs.	0.039	**
Clients mainly responsible for looking after children	0.039	***
Clients reporting that business gives security	0.039	**
Clients using marketing practices	0.039	***
FSPs with above average attrition rate	0.040	**
Clients who want to borrow more for business	0.040	***
Clients reporting profits with above average bias	0.044	**
FSPs operating for above average number of years	0.044	***
FSPs with above average amount of loans disbursed in the last 12 months	0.047	****
FSPs with above average nr. of employees	0.049	***
FSPs with above average nr. of active borrowers	0.054	****
Clients reporting tough competition as the main business barrier	0.057	**
Clients from Batch 3	0.061	**
Clients having primary education (endline)	0.064	**
Clients aged between 35 and 50	0.065	****
Clients reporting above average nr. of hours per day devoted to business	0.066	***
Clients of development banks	0.071	***
Clients with business in manufacturing sector	0.075	**
Clients of FSP 13	0.080	**
Clients with at least one person who can continue running business	0.083	***

Table 8 - Continued from previous page

Variable / Group	β_{ITT}	
Clients having university education (baseline)	0.093	***

Note: This table documents the estimates of treatment effects for all clients in estimation sample and the groups for which the estimated effect is statistically significant. The table shows the outcome variables (in bold) and the sample groups on left.

▲ Source: Baseline survey (2017 - 2018) and endline survey (2018-2019).

▲ Columns (1)-(2) shows the regression results for the selected sample groups and significance of the estimated effect. ▲ In all regressions we control for the twenty most imbalanced baseline covariates and dummies indicating imputation in baseline variables, enumerator- and, if applicable, FSP fixed effects, age and gender.

▲ The superscript np. indicates the winsorizing level.

▲ * indicates $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

**Table 9: Selected Heterogenous Effects - Final Outcomes**

Variable / Group	βITT	
Revenue in the last 30 days in mln IDR ^{99p} :		
All clients	1.117	
Clients having no education (endline)	7.123	**
Clients of development banks	12.785	**
Clients of FSP 10	22.331	***
Nr. of permanent and casual workers (sum) ^{99p} :		
All clients	0.138	
FSPs with above average nr. of active borrowers	0.481	**
Clients with above average business perception index	0.521	**
Clients reporting that business is risky	0.731	**
Clients having no education (baseline)	1.463	**
Total cost of business in the last 30 days in mln IDR ^{99p} :		
All clients	-0.349	
Clients of FSP 8	-10.451	**
Profit generated in the last 30 days in mln IDR ^{99p} :		
All clients	0.254	
Opportunity entrepreneurs	0.856	**
FSPs with above average attrition rate	1.072	**
FSPs with above average amount of loans disbursed in the last 12 months	1.104	**
Clients reporting tough competition as the main business barrier	1.667	***
Profit in the worst month mln IDR ^{99p} :		
All clients	0.207	
Clients having primary education (baseline)	0.730	**
Clients reporting tough competition as the main business barrier	1.157	***
Profit in the best month in mln IDR ^{99p} :		
All clients	-0.001	
Clients reporting above average nr. of participants in last training session	2.261	**
Clients reporting tough competition as the main business barrier	2.540	***
Clients having no education (endline)	3.046	**
Profit increased during the last 6 months:		
All clients	0.000	
Clients of FSP 9	0.165	***
Clients reporting struggle with management as the main business barrier	0.166	***
New business started in 2018 (in addition to current business):		
All clients	-0.005	
Clients of FSP 1, 2 ,3 or 5	0.029	**
Clients reporting finding market as the main business barrier	0.037	**
Business perception index (standardized score):		
All clients	-0.000	
Clients having vocational education (endline)	0.111	**
Nr. of barriers for business:		
All clients	-0.001	
Clients reporting above averge nr. of participants in last counseling session	-0.472	***

Table 9 - Continued from previous page

Variable / Group	βITT	
Total savings from all sources in mln IDR90p:		
All clients	-1.002	
Clients of FSP 9	5.799	**
Share of clients reporting increase in HH exp. in the last 6 months:		
All clients	0.012	
Clients reporting above average nr. of hours per day devoted to business	0.063	**
Clients of FSP 1	0.176	**
Life satisfaction (1=worst, 10=best):		
All clients	0.008	
Clients having primary education (baseline)	0.222	**
Clients having primary education (endline)	0.300	***
Clients of FSP 9	0.444	**
Difference between life satisfaction now and two yrs. ago:		
All clients	-0.050	
Clients of development banks	0.344	***
Clients of FSP 13	0.394	**
Clients having no education (baseline)	0.775	***
Currently in loan default or behind with repayments for any loan:		
All clients	-0.019	
Clients spending above average time on looking after dependents/ household chores	-0.050	**
Clients preparing a business/ financial plan	-0.051	**
Clients reporting tough competition as the main business barrier	-0.069	**
Clients reporting that business brings high income	-0.072	***
Clients with business in manufacturing sector	-0.090	***
Clients with at least one person who can continue running business	-0.114	***
Clients of FSP 8	-0.116	**
Clients having no education (baseline)	-0.206	**
Share of clients reporting being late with the loan payment at the FSP:		
All clients	-0.011	
Clients whose family members work	-0.044	**
Clients preparing a business/ financial plan	-0.053	***

▲ Note: This table documents the estimates of treatment effects for all clients in estimation sample and the groups for which the estimated effect is statistically significant. The table shows the outcome variables (in bold) and the sample groups on left.

▲ Source: Baseline survey (2017 - 2018) and endline survey (2018-2019).

▲ Columns (1)-(2) shows the regression results for the selected sample groups and significance of the estimated effect. ▲ In all regressions we control for the twenty most imbalanced baseline covariates and dummies indicating imputation in baseline variables, enumerator- and, if applicable, FSP fixed effects, age and gender.

▲ The superscript np. indicates the winsorizing level.

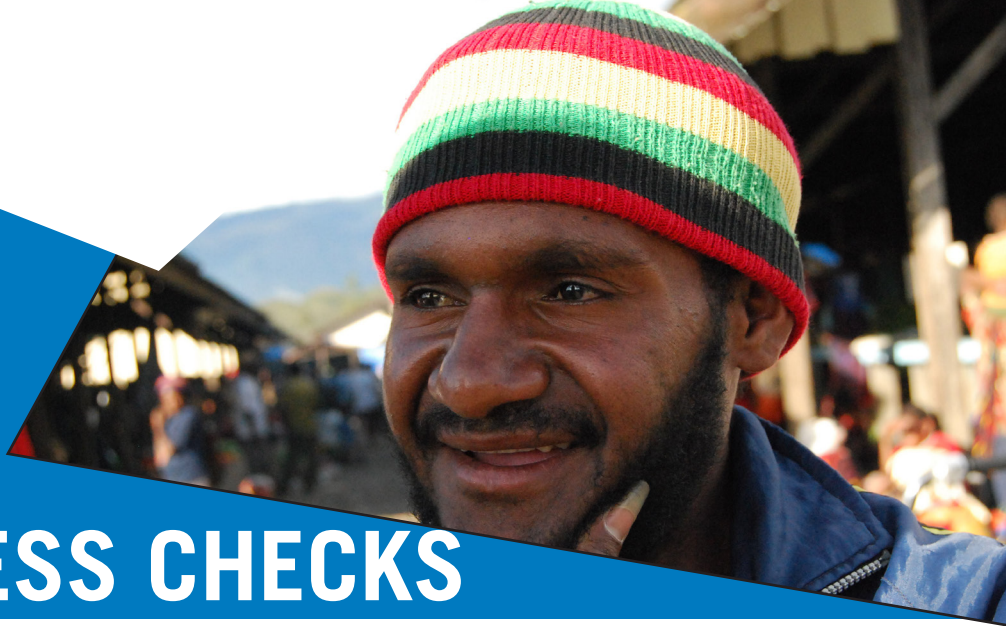
▲ * indicates $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.



Additional Results: Regressing Actual Participation on Outcomes. Recognizing the limitation that participants self-select and that this process, which can be driven by unobservable client characteristics is hard to predict, we present additional results whereby showing correlations between actual participation and the outcomes of interest. Here the treatment is endogenous as other factors (confounders) that are unobserved may influence both the participation in question as well as the outcomes. The explanatory variable is not whether someone was randomly assigned to the program, but rather information from survey or monitoring on whether a client participated in the program within the timeframe of the intervention. This naive estimation produces biased estimates for the effect of the endogenous variable on outcomes. Still, the results might be interesting and are thus presented in Section C.7 of the Appendix.

7.

ROBUSTNESS CHECKS



IN OUR main model specifications, we attempt to account for the main factors, which affect the results, namely, differences in baseline characteristics and differences across FSPs. Nevertheless, there are additional factors that can influence the estimation results. In this section, we analyze how robust our results are against alternative model specifications. For this purpose, we estimate 13 different models. These deviate from the main model either in the covariates or fixed effects included in the regression, in the unit of clustering for standard errors, or in the observations included in the sample. We present these results in Table 10.

In this table, the first column contains the results from the main specification as above, i.e., controlling for the 20 most imbalanced covariates, including FSP fixed effects and robust, unclustered standard errors (column (1)). In the second model, we remove enumerator level fixed effects (column (2)). In the third model, we run the main specification (column (1)) and add a dummy control variable which equals one if we have no reported participation in treatment in the branch/ cashpoint cluster (column (3)). In the fourth model, we add - to the main specification - a control variable indicating whether the clients changed the FSP between baseline and endline survey (column (4)). In the fifth model, we additionally control for the distance between client and FSP using GPS data gathered during endline (column (5)). In the sixth model, we remove the FSP fixed effects, but cluster standard errors at the FSP level (column (6)). In the seventh model, we exclude observations with missing values in baseline variables, which we had imputed for the main model (column (7)). In the eighth model, we run the main specification and now add an indicator whether the respondent lives in East Java or West Java. In the ninth model, we run the same model as in column (8), but this time cluster the standard errors at the FSP level (column (9)). In the tenth model, we cluster standard errors at the branch/ cashpoint level (column (10)). It should be noted that the models lead to the same coefficient estimates, but the significance level may change due to different standard errors. In the eleventh model, we add fixed effect for batch instead of FSP (column (11)). In the twelfth model, we include the 30 most imbalanced control variables instead of 20 (column (12)). In the thirteenth model, we include only the ten most imbalanced control variables instead of twenty (column (13)).



Due to the number of columns and size of the table, we abbreviate the outcome variable on the left with the number of the hypothesis as listed in the PAP and also in the main results tables above.²⁸

As would be expected, the individual estimates differ slightly across the different specifications. The important thing to note is that there is little change in the significance of the estimates. Whereas some models yield additional weak significance for some outcomes, none of these are robust against changes in the models. Notably, the increase in the share of clients whose business plans include a cash flow analysis is highly significant in all but one model specification and similar in magnitude (between 3.1 and 3.3 percentage points). Overall, Table 10 confirms the results from the main specification presented above.

28 1-1a-Share of marketing knowledge questions answered correctly; 2-1b-Answ. cor.: A good salesman is; 3-1b-Answ. cor.: Total profit depends on; 4-1b-Answ. cor.: First thing to do in marketing; 5-2a-Share of financial management knowledge questions answered correctly; 6-2b-Answ. cor.: More liabilities than assets means; 7-2b-Answ. cor.: Accurate business records require; 8-2b-Answ. cor.: Purpose of business plan; 9-2c-Share of clients not knowing the payment type; 10-3a-Marketing practices index, non-missing; 11-3b-Share of MSEs exporting outside Indonesia; 12-3b-Nr. of different marketing forms used, non-missing; 13-4a-Share of clients keeping business and HH finances separately; 14-4b-Share of clients investing profit into the business; 15-4c-Business plan includes Cash Flow; 16-4c-Business plan includes expected sales and costs; 17-4c-Business plan includes plan how to reach a new market; 18-4c-Business includes plan how to improve a product; 19-4c-Share of clients preparing a business/ financial plan; 20-4d-Financial management practices index, non-missing; 21-4e-Pays income tax for business; 22-4e-Business is registered; 23-5a-Revenue in the last 30 days in mill. IDR^{99p}; 24-5b-Nr. of permanent workers; 25-5c-Nr. of permanent and casual workers (sum), win^{99p}; 26-6a-Total cost of business in the last month (30 days) in the IDR, win^{99p}; 27-6b-Profit in the best month in the IDR, win^{99p}; 28-6b-Profit in the worst month in the IDR, win^{99p}; 29-6b-Profit generated in the last 30 days in mill. IDR^{99p}; 30-6c-Share of business earnings covers the expenses of this business; 31-6d-Profit increased during the last 6 months; 32-7a-New business started in 2018 (in addition to current business); 33-8a-Business perception index (standardized score); 34-8b-Nr. of barriers for business; 35-9a-Total savings from all sources in the IDR, win^{99p}; 36-9b-Aggregate HH spending in the IDR, win^{99p}; 37-9b-HH expenditure in the last 30 days in mill. IDR^{99p}; 38-9b-Change in HH expenditure in the last 6 months; 39-10a-Satisfaction with current life: 1=worst, 10=best; 40-10b-Change in life satisfaction (now - two yrs. ago); 41-11b-Current/ last loan used for productive purposes; 42-11c-Share of clients want to borrow more for business; 43-12a-Currently in loan default or behind with repayments for any loan; 44-12b-Share of clients reporting being late with the loan payment at the FSP.

Table 10: Robustness Checks: All FSPs

#	H	Main	Alternative Specifications											
			(1)	(2) -En.FE	(3) +No Impl. CI	(4) Change FSP	(5) +Distance	(6) CI(FSP)	(7) No miss.	(8) No incomplete	(9) +East/West	(10) CI(CashP)	(11) B.FE	(12) 3OC
1	1a	0.000	0.006	0.000	-0.005	0.000	-0.000	-0.005	-0.006	-0.000	0.000	-0.000	-0.000	0.001
2	1b	-0.004	0.008	-0.004	-0.016	-0.004	-0.005	-0.007	-0.010	-0.005	-0.004	-0.004	-0.005	-0.005
3	1b	0.006	0.008	0.006	-0.001	0.006	0.003	0.000	-0.001	0.003	0.006	0.003	0.005	0.006
4	1b	0.001	0.005	0.001	0.005	0.001	0.003	-0.008	-0.008	0.003	0.001	0.001	0.001	0.002
5	2a	-0.001	0.007	-0.001	-0.001	-0.001	-0.001	-0.002	-0.004	-0.001	-0.001	-0.003	-0.001	-0.000
6	2b	-0.007	0.003	-0.007	-0.008	-0.006	-0.006	-0.014	-0.014	-0.006	-0.007	-0.009	-0.007	-0.006
7	2b	0.019	0.026	0.019	0.018	0.020	0.017	0.023	0.019	0.017	0.019	0.017	0.020	0.020
8	2b	-0.015	-0.009	-0.015	-0.011	-0.015	-0.015	-0.014	-0.015	-0.015	-0.015	-0.016	-0.015	-0.013
9	2c	-0.015	-0.017	-0.016	-0.013	-0.016	-0.014 **	-0.017	-0.018 *	-0.014 **	-0.015 *	-0.016	-0.015	-0.015
10	3a	0.009	0.024	0.009	0.017	0.009	0.020	-0.018	-0.020	0.020	0.009	0.007	0.010	0.014
11	3b	0.021	0.020	0.021	0.022	0.021	0.025	0.025	0.025	0.025	0.021	0.021	0.021	0.025
12	3b	-0.015	-0.012	-0.016	-0.021	-0.015	-0.011	-0.015	-0.013	-0.011	-0.015	-0.010	-0.015	-0.015
13	4a	-0.013	-0.013	-0.013	-0.013	-0.011	-0.014	-0.011	-0.015	-0.014	-0.013	-0.014	-0.013	-0.010
14	4b	0.008	0.011	0.008	0.004	0.009	0.016	0.010	0.007	0.016	0.008	0.012	0.007	0.008
15	4c	0.020	0.022	0.020	0.016	0.020	0.019	0.023	0.022	0.019	0.020	0.018	0.021	0.022
16	4c	0.036	0.043 *	0.036	0.034	0.034	0.038	0.047 *	0.048 *	0.038	0.036	0.036	0.035	0.038
17	4c	-0.043 *	-0.044 *	-0.043 *	-0.051 *	-0.044 *	-0.046	-0.043	-0.043	-0.046	-0.043 *	-0.044 *	-0.041	-0.042
18	4c	-0.022	-0.052 **	-0.022	-0.024	-0.022	-0.020	-0.015	-0.016	-0.020	-0.022	-0.024	-0.023	-0.020
19	4c	0.033 ***	0.011	0.033 ***	0.033 **	0.033 ***	0.031 ***	0.031 **	0.032 **	0.031 ***	0.033 **	0.032 **	0.032 **	0.033 ***
20	4d	0.018	0.061	0.018	0.017	0.020	0.026	0.008	-0.003	0.026	0.018	0.013	0.012	0.026
21	5a	1.117	1.078	1.172	0.933	1.252	1.192	1.511	1.259	1.192	1.117	1.270	1.265	1.064
22	5b	-0.151	-0.228	-0.153	-0.118	-0.142	-0.145	-0.222	-0.247 *	-0.145	-0.151	-0.118	-0.141	-0.161
23	5c	0.138	0.123	0.144	0.024	0.147	0.135	0.070	0.054	0.135	0.138	0.208	0.148	0.096
24	6a	-0.349	-0.272	-0.337	-0.398	-0.255	-0.361	-0.125	-0.321	-0.361	-0.349	-0.198	-0.370	-0.501

Table 10: Continued from previous page

#	H	Main	Alternative Specifications											
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
			-En.FE	+No Impl. CI	Change FSP	+Distance	CI(FSP)	No miss.	No incomplete	+East/West	C(CashP)	B.FE	30C	10C
25	6b	0.254	0.374	0.247	0.118	0.273	0.216	0.532	0.533	0.216	0.254	0.268	0.228	0.233
26	6b	0.207	0.330	0.199	0.180	0.229	0.181	0.342	0.344	0.181	0.207	0.214	0.201	0.195
27	6b	-0.001	0.334	-0.010	-0.035	0.037	-0.046	0.368	0.375	-0.046	-0.001	0.007	0.007	-0.012
28	6c	-0.012	-0.011	-0.012	-0.012	-0.012	-0.011 *	-0.017	-0.017	-0.011 *	-0.012	-0.011	-0.012	-0.011
29	6d	0.000	-0.001	0.001	-0.004	0.001	-0.000	-0.005	-0.004	-0.000	0.000	0.000	0.001	0.001
30	7a	-0.005	-0.003	-0.005	-0.004	-0.005	-0.002	-0.002	-0.002	-0.002	-0.005	-0.003	-0.005	-0.005
31	8a	-0.000	0.011	0.001	-0.007	0.001	-0.003	0.009	0.006	-0.003	-0.000	-0.001	0.000	-0.000
32	8b	-0.001	-0.016	-0.000	0.003	-0.002	-0.003	0.007	0.005	-0.003	-0.001	-0.005	-0.000	0.000
33	9a	-1.002	0.008	-1.010	-1.003	-0.952	-1.022	-0.451	-0.485	-1.022	-1.002	-0.937	-0.957	-0.875
34	9b	0.012	0.012	0.012	0.011	0.011	0.013	0.006	0.008	0.013	0.012	0.012	0.014	0.013
35	9b	-0.133	-0.098	-0.133	-0.082	-0.126	-0.147	-0.051	-0.052	-0.147	-0.133	-0.127	-0.130	-0.124
36	9b	-1.640 *	-1.296	-1.614 *	-1.354	-1.606 *	-1.595	-1.450	-1.527	-1.595	-1.640 *	-1.531	-1.614 *	-1.716 *
37	10a	0.008	0.018	0.006	0.008	0.006	0.017	-0.004	-0.001	0.017	0.008	0.013	0.007	0.016
38	10b	-0.050	-0.029	-0.052	-0.072	-0.049	-0.035	-0.045	-0.048	-0.035	-0.050	-0.045	-0.052	-0.045
39	11b	-0.013	-0.007	-0.013	-0.020	-0.015	-0.012	-0.011	-0.012	-0.012	-0.013	-0.011	-0.015	-0.015
40	11c	-0.037	-0.015	-0.036	-0.037	-0.037	-0.038 *	-0.029	-0.027	-0.038 *	-0.037	-0.032	-0.036	-0.041 *
41	12a	-0.019	-0.026 *	-0.019	-0.019	-0.019	-0.021	-0.012	-0.010	-0.021	-0.019	-0.019	-0.020	-0.019
42	12b	-0.011	-0.018	-0.010	-0.011	-0.010	-0.017 *	-0.005	-0.003	-0.017 *	-0.011	-0.013	-0.011	-0.011
43	20a	0.547	0.761	0.559	0.719	0.553	0.727	0.598	0.631	0.727	0.547	0.561	0.578	0.544
44	20a	-0.019	-0.001	-0.020	0.031	-0.017	-0.019	0.045	0.047	-0.019	-0.019	-0.008	-0.013	-0.011
45	20a	-0.006	0.003	-0.006	0.003	-0.005	-0.007	-0.004	-0.005	-0.007	-0.006	-0.005	-0.005	-0.005
46	20a	-0.014	-0.011	-0.014	-0.015	-0.014	-0.014	-0.009	-0.009	-0.014	-0.014	-0.013	-0.013	-0.014
47	20a	-0.065	-0.062	-0.065	-0.084	-0.064	-0.066	-0.076	-0.076	-0.066	-0.065 *	-0.063	-0.066	-0.062
48	20b	0.018	0.025 *	0.018	0.019	0.018	0.017	0.019	0.017	0.017	0.018	0.018	0.018	0.019

Table 10: Continued from previous page

#	H	Main (1)	Alternative Specifications											
			(2) -En.FE	(3) +No Impl. CI	(4) Change FSP	(5) +Distance	(6) C/(FSP)	(7) No miss.	(8) No incomplete	(9) +East/West	(10) C/(CashP)	(11) B.FE	(12) 30C	(1) 10C
49	20b	0.003	0.004	0.003	0.002	0.006	0.005	0.002	0.003	0.003	0.003	0.003	0.003	
50	20b	-0.008	-0.005	-0.008	-0.007	-0.006	-0.006	-0.008	-0.008	-0.007	-0.008	-0.008	-0.008	
51	20b	-0.001	0.000	0.000	-0.001	0.000	-0.000	-0.001	-0.001	-0.002	-0.001	-0.001	-0.001	
52	20b	-0.005	0.002	-0.002	-0.005	-0.005	-0.005	-0.006	-0.005	-0.005	-0.006	-0.006	-0.006	
53	20b	-0.002	0.001	0.001	-0.002	-0.002	-0.002	-0.001	-0.002	-0.002	-0.002	-0.002	-0.003	
54	20b	-0.015	-0.010	-0.023 *	-0.015	-0.014	-0.019	-0.014	-0.015	-0.015	-0.016	-0.015	-0.015	
55	20b	-0.009	-0.004	0.003	-0.008	-0.003	-0.004	-0.008	-0.009	-0.009	-0.008	-0.008	-0.008	
56	20b	-0.005	-0.004	-0.003	-0.005	-0.006	-0.006	-0.005	-0.005	-0.004	-0.005	-0.004	-0.004	
57	20b	0.003	0.007	-0.002	0.003	0.001	0.001	0.004	0.003	0.004	0.002	0.002	0.002	
58	20c	0.020	0.019	0.021	0.021	0.024	0.023	0.021	0.020	0.021	0.020	0.018	0.018	
59	20c	0.015	0.003	0.015	0.017	-0.006	-0.007	0.024 *	0.015	0.022	0.018	0.011	0.011	
60	20c	-0.011	-0.011	-0.008	-0.013	-0.006	-0.008	-0.023	-0.011	-0.017	-0.015	-0.002	-0.002	
61	20c	-0.002	0.001	0.000	-0.001	0.005	0.004	0.000	-0.002	-0.004	0.000	-0.004	-0.004	

▲ Note: Table ?? documents the hypotheses on the left▲ Sample: All FSPs. Sample size varying.

Column (1) includes the main specification.

Column (2) relative to main specification in column (1); Removes enumerator fixed effects.

Column (3) relative to main specification in column (1); Additionally adds a dummy for zero implementation in a cluster.

Column (4) relative to main specification in column (1); Additionally controls for whether the FSP was not changed between baseline and endline. Column (5) relative to main specification in column (1); Additionally controls for the distance to the FSP.

Column (6) relative to main specification in column (1); Removes the fixed effects for the FSP and instead clusters the standard errors at the FSP-level. Column (7) relative to main specification in column (1); Excludes observations with missing values at baseline.

Column (8) relative to main specification in column (1); Additionally includes an East/ West Java dummy.

Column (9) relative to main specification in column (1); Adds an East/West indicator and clusters the standard errors at the FSP-level.

Column (10) relative to main specification in column (1); Clusters the standard errors at the cash point.

Column (11) relative to main specification in column (1); Adds batch fixed effects instead of FSP fixed effects.

Column (12) relative to main specification in column (1); Adds thirty most imbalanced control variables instead of twenty.

Column (13) relative to main specification in column (1); Adds ten most imbalanced control variables instead of twenty.



8.

COST BENEFIT ANALYSIS



THE BUNDLING services (combination of loans and business development) provided in this intervention are part of the double bottom line approach of participating FSPs. The services are not only supposed to directly benefit the clients, but can also lead to better financial performance of the FSP by improving repayment behavior and client retention. This cost-benefit analysis aims to validate or refute this argument.

The cost-benefit analysis is conducted from the standpoint of the FSPs (as opposed to the standpoint of an external stakeholder). We distinguish between two different types of cost-benefit analyses: financial and social, which are presented in the following subsections. The financial cost-benefit analysis compares the costs and benefits incurred at FSP level, i.e., benefits such as larger loan sizes, client retention, or less default, versus costs incurred for providing the training or counseling.

The social cost-benefit analysis compares the costs and benefits incurred at MSE level, i.e., benefits such as increased profits or household spending, versus costs incurred for participating in the training or counseling. Throughout this section, we neglect the costs incurred by the ILO for training of trainers and consultants. In addition, whereas the direct costs and opportunity costs for participation in the ILO program (such as for feasibility assessment, client survey, participation at workshops, etc.) are non-negligible, they are in effect sunk costs which do not inform the FSP decision to continue the business support services.

Furthermore, we assume that clients are not paying any additional costs and that the fee or interest rate for loans is the same for clients who get bundled services and those who only borrow. That is, we assume that the FSPs have not passed on the costs of the treatment to their clients. We also build our calculations on the assumption that each training course took two full days, and counseling a total of 10 hours (i.e., five sessions at two hours each). To our knowledge, this corresponds to the design of the intervention and is largely consistent with the participants' reports in the endline survey.²⁹ Our estimated costs are hence calculated per client per complete training and/ or counseling treatment. In what follows, we describe the underlying assumptions, sources of information, and calculations. We display results aggregated over all FSPs in this section and for each FSP individually in the tables in the respective appendix. These tables can be interpreted analogously to the aggregate results explained below.

²⁹ In the endline data, the reported duration of a counseling sessions averages 2.5 hours but varies between 30 minutes and 12 hours, with the latter likely a measurement error. As explained before, the endline survey did not identify all the counseling participants. We hence deem it appropriate to base our calculations on the original design, as per assignment into the treatment vs. control.



A main restriction in our analysis is the limited timeframe which we can cover in this evaluation. It is possible that more benefits develop over time. Also, we do not consider spill-over effects, i.e., benefits occurring within the control group. Put differently, we cannot include the long-term and broader benefits that arise from higher trained loan officers at FSP or client level. If further effects along the lines were to be expected, the estimates presented would therefore reflect lower bounds.

8.1 Financial cost benefit analysis

Assumptions. For the financial cost benefit analysis, we furthermore rely on the following assumptions:

- FSPs do not incur any opportunity costs beyond what can be proxied through the loan officers' salaries. In practice, loan officers who provide training and/ or counseling spend less time on identifying new clients or following up on delinquent borrowers. We assume that the lost income to the FSPs is largely captured in the loan officers' salaries.
- FSPs did not incur any costs for material development. That is, we assume that loan officers are equipped with standardized training material.

Data sources. We calculate costs from data provided to us by the FSPs. In our understanding, FSPs were partly or fully reimbursed for these costs. Nevertheless, we consider them in this financial cost benefit analysis as FSPs will need to cover them in future if they decide to continue with the intervention. We estimate benefits from the endline survey, but we also asked FSPs to provide additional data from their Management Information System (MIS). Only one FSP provided such data.

Costs to FSP. At FSP level, we see four main cost factors for the training course: (1) transportation for trainers (including a trainer's fee if applicable), (2) accommodation and catering during training, (3) material printing and dissemination, and (4) loan officer salaries. For the counseling sessions, the main cost factors are as follows: (1) transportation, meals and communication costs for loan officers, and (2) loan officer salaries. This information was provided to us by two FSPs.³⁰ Salaries were provided as monthly salaries for cooperatives, rural banks and development banks separately. Material costs were provided per client. The unit used for the remaining training cost positions is per client per day. To calculate client cost, the actual number of training participants was used. To add loan officers' salaries, we divide the daily salary by 20 to get a cost per client per day estimate. That is, we assume that 20 clients participated in one training course at the same time. Actual average costs will be higher than our estimates if less than 20 clients participated in the training.

Another limitation of our estimation is that the costs of training are only incurred upon participation. However, we calculate benefits as ITT effects, i.e., averaged over all clients who were offered to participate. At the same time, our estimates for participation rates are crude for the counseling sessions. Therefore, we provide two different cost estimates: first, assuming a participation rate of 100%, i.e., we use the costs incurred per client as explained above; second, assuming the participation rates described in Section 3.3, we calculate the expected costs incurred per client offered to participate.

Benefits to FSP. In principle, benefits to the FSPs are the estimated income the FSP receives from a client offered to participate in the intervention minus the estimated income the FSP would have received without the offer of the intervention. To estimate this difference, we use the main regression specification.

³⁰ For FSPs 9 and 10, we will hence not be able to present cost estimates in the individual tables in the appendix. We derived the estimates in Table 11 by averaging over the eleven FSPs that reported costs, weighted by the share of treated clients in the FSP relative to the total number of treated clients.

In particular, we calculate again ITT effects, i.e., the effect of being offered to participate rather than the effect of having participated. We do this because, also in future, participation rates will be less than 100%, and because our data does not allow us to clearly identify the counseling participants.

In the absence of MIS data, we can only analyze those benefits for which we collected data in the endline survey. These are changes in (1) the loan volume, (2) the interest payments, (3) the amount of savings held at the FSP, (4) the client retention rate, and (5) the default rate.³¹ The first three benefits are monetary in nature, but the monetary values we present are not net-benefits as we lack information on refinancing rates, interests paid on savings and other operating costs of FSPs. This is even more clear when considering the last two variables, which are shares, not monetary values. That is, from the tables in the appendix, each FSP can independently determine monetary profits associated with the changes in any of these variables.

In the impact evaluation above we considered similar outcomes variables as we do now, the former relate to the current or last loan held at any financial institution. We now focus on the loans held with the respective FSP. That is, we define the loan size as zero if the client changed the FSP before the endline survey. For this reason, the figures below deviate from the results in the main text, in Table 3. Also, note that some of these loans might have been taken before the intervention started, in particular for those FSPs which started last with the treatment implementation. If this applies to a substantial part of the loans we consider here, our figures present only rough estimates. In addition, as mentioned before, we estimate only one-time changes, whereas the real benefits might be accumulated over repeated loan cycles.

Results. Table 11 shows the financial costs and benefits as explained above, averaged over all FSPs which offer the respective treatment arms. In Appendix D, we present result tables for each FSP individually. These should be read in the same way as the table below, which is as follows:

The first line of the table displays the costs of treatment per client for the training treatment (Column (1)), the counseling treatment (Column (2)) and the TC treatment (Column (3)). Note that the latter naturally is the sum of the first two columns. Comparing training and counseling, we find that training is by far the more expensive treatment. This is because the cost for room rent and catering outweighs the time needed for counseling sessions. Adjusting the costs for actual participation, which for these FSPs averages 63.75%, the expected costs of offering treatment to a client is displayed in the second line.

The financial benefits of the intervention are displayed in the following five lines. For example, the training only treatment decreased the loan size by 1.666 million IDR among treated clients compared to the control group (Column (1)). The counseling only treatment decreased the loan size by 1.942 million IDR, and the training and counseling treatment by 1.096 million IDR. The depicted stars indicate whether the effects are statistically significant. For example, the decreases in loan size are significantly different from zero for the training only and the counseling only treatment, but might be due to randomness in measurement in the combined treatment arm. Similarly, a decrease in the interest payment by 0.935 million IDR in the training only group is weakly significant and consistent with the decrease in loan size. Whereas these results are rather adverse from the point-of-view of the FSPs, a positive effect is that the client retention rate increased significantly by 4.4 percentage points among clients who received training compared to the control group.

³¹ The choice of variables was inspired by Karlan and Valdivia 2011, who focus on client retention, loan size, accumulated savings, perfect repayment and drop-out rates.



8.2 Social cost benefit analysis

Assumptions. For the social cost benefit analysis, we rely on the following assumptions:

Table 11: Financial Costs and Benefits

	Treatment arm		
	T (1)	C (2)	TC (3)
Costs of treatment per client			
assuming full participation (mln IDR)	0.519	0.113	0.608
assuming actual participation (mln IDR)	0.331	0.072	0.388
Benefits			
Change in loan size (mln IDR)95p	-1.666 **	-1.942 **	-1.096
Change in savings held at FSP (mln IDR)95p	0.164	0.285	-0.048
Change in interest payments (mln IDR)95p	-0.935 *	-0.394	-0.543
Change in client retention rate (percentage points)	0.044 **	-0.018	0.005
Change in repayment rate (percentage points)	-0.029	-0.017	0.005

Note: This table shows the financial costs and benefits per treatment arm, averaged over all FSPs that offer the respective treatment arm. Cost and benefit positions on the left, treatment arms on top.

▲ Sample: Costs: FSPs without FSPs 9 and 10 (N=11). Benefits: Estimation sample for benefits (N=4,292)▲ Source: Costs: ILO. Benefits: Endline survey (2018 - 2019).

▲ Columns (1) presents estimate for the T treatment (training only), column (2) for the C treatment (counseling only), column (3) for the TC treatment (training and counseling). Costs are per client, either assuming perfect participation or actual participation rates, and comprise costs for trainer's transportation, room rent and catering, material, and loan officers' salary. Costs cover two full days of training and/ or five counseling sessions à two hours each. Costs are calculating averaging over the eleven FSPs that provided information and weighted by the share of treated clients. Actual participation rate in this sample is 63.75%. Benefits are estimated as ITT effects via linear regressions. In all regressions we control for age, gender, the twenty most imbalanced baseline covariates, FSP fixed effects, and dummies indicating imputation in baseline variables. All benefit outcomes pertain only to loans associated with the respective FSP and to one loan cycle only.

▲ The superscript np. indicates the winsorizing level.

▲ The statistical significance is given as follows: * indicates $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$. ▲ Result tables for each FSP individually are presented in the respective FSP appendix.

- We assume that counseling sessions are conducted at the client's premise during routine loan collection visits. Consequently, we neglect any direct costs for counseling sessions for clients.
- We assume that monetary outcomes at client level (e.g. higher profits) accumulate over a time frame of one year and stay constant over this time. It is possible that the benefits last longer, hence we consider our estimates lower bounds. Given that loan cycles typically also last one year, this is in line with our above assumption that effects on loans are considered one-time benefits for the FSP.
- We discount these future benefits by using the treasury bond rates of Indonesia as proxy for discount rates. The annual interest rate at the time of the endline survey was 7.847%
- Even though endline results were collected immediately after the implementation, since our assumption of benefits being constant we treat costs and benefits as if they occurred simultaneously.

Costs at client level. We consider direct and indirect costs, which occur at client level for the participation in the intervention. Direct costs are monetary costs, which include cost for transportation, and possibly for material and accommodation. Since these costs were covered by the FSPs or reimbursed through the program, we can use the per client cost information provided to us. In the endline survey, we also asked about any costs incurred for treatment participation, but find that the reported costs are negligible, averaging 0.20 USD per client in the treatment group. We therefore base our calculations on the data provided and the amount deemed to be sufficient to cover the participation costs of the clients.

For indirect monetary costs we calculate the beneficiaries' opportunity costs in terms of forgone earnings during the time of treatment. As proxy measures, we use estimated means of monthly profits from the control group of the considered FSPs for opportunity cost calculation and assume the above mentioned duration of two days for training and/ or 10 hours for counseling.

Benefits to clients. The benefits to the clients are all the variables we have discussed in Section 5. These include monetary as well as non-pecuniary outcomes for the beneficiaries of the intervention. Monetary outcomes include higher profits, increases in household expenditures as well as increased savings. The non-pecuniary benefits include the start-up of new businesses, increases in business perception or current life satisfaction.

For monetary outcomes which we measured for a period of 30 days (revenues, costs, profit, expenditures on non-durables), we now also provide the net present value for a period of one year, discounted, as explained above. Note that the difference of revenues and costs does not equal the profit in most cases. This is due to measurement errors, which is common for quantitative measurements. We believe that direct profit measurements are more reliable than the difference between revenues and costs.

For non-monetary benefits, we do not extrapolate or discount these as it is likely that benefits such as an increase in life satisfaction are one-time events rather than monthly repeating benefits. These benefits can hence be directly taken from the main results table. To obtain some indication for their monetary valuation, we instead provide the willingness to pay clients for the training and counseling services. We obtained these in the endline survey by directly asking participants and non-participants about the maximum amount of money they would be willing to pay for training or counseling services provided by their FSP. Below, we present results for participants and non-participants separately.

Results. Table 12 contains the costs, monetary discounted benefits and willingness to pay estimates aggregated over all FSPs. The first panel displays the costs of participation for clients in million IDR. The first line contains the direct costs, the second line the opportunity costs (average profit in control group), and the last line the sum of the two. The second panel contains the net present value of monetary benefits to the client, discounted over a period of one year. These are the changes in revenue, costs, profits and expenditures on non-durables.³² The third panel displays the reported willingness to pay, disaggregated by participants and non-participants. For all outcomes, we distinguish between the three treatment arms, training (Column (1)), counseling (Column (2)), and training with follow-up counseling (Column (3)). We display the asterisks indicating significance after the point estimates. For willingness to pay, we instead display the standard errors. We do this because this is a descriptive statistic only, and we do not need to test for differences between treatment and control group.

Averaging for all clients, we find that the opportunity costs of training surpass direct costs by a factor of four. Given that training took longer than counseling sessions, it is clear that training is the more costly treatment alternative from the client's point-of-view. The net present values of monetary benefits are all insignificant (no stars), although the increase in revenues is sizable in magnitude. The significance corresponds to that from the main results, in Table 3. The willingness to pay for training and counseling services is surprisingly similar in this sample, suggesting that clients are indifferent between the two treatment alternatives. Their magnitude corresponds roughly to the direct costs covered. Participants are willing to pay slightly less than non-participants, but this difference is not significant and might be due to entitlement effects, which imply that people are less willing to pay for services that they have previously obtained for free. The overall willingness to pay is low, which is consistent with the limited evidence on

³² Note that there are other monetary benefits, such as changes in total savings. These, however, we did not measure for a 30 day period, hence, we do not provide net present values here. For further benefits, including non-monetary ones, see table 3.



benefits and the existence of opportunity costs. Yet, it is noteworthy that the willingness to pay is significantly different from zero, suggesting that clients do perceive some value of the services even in the presence of opportunity costs.

With these cost and benefit estimates, one can calculate differences (e.g., subtracting total costs from changes in profits) to obtain net benefits, or ratios (e.g., changes in profits over total costs) to obtain the benefit per dollar invested. However, given that our benefit estimates are insignificant and mostly negative, our numbers suggest that participation in the treatment is not profitable from the point-of-view of the clients, at least if considering only the short-term benefits measured in the endline survey.

9.

CONCLUSION



WITH THIS work we contribute to a body of evidence aiming to answer the question: Can business training and counseling have a transformative impact on the lives of microentrepreneurs? The appeal of this approach is apparent: Classroom training is cost-effective and easily replicated, and even individual counseling could be cost-effective if delivered through loan officers of FSPs who themselves benefit from improved loan behavior of their clients.

Using an RCT, we evaluate a pilot training and counseling intervention as part of a pilot in Indonesia. In this pilot 13 FSPs from East Java and West Java took part.

Table 12: Social Costs and Benefits

	T (1)	Treatment arm	
		C (2)	TC (3)
Costs of participation for clients			
Direct cost per client (mln IDR)	0.103	0.000	0.103
Opportunity cost per client (mln IDR)	0.401	0.250	0.651
Total cost per client (mln IDR)	0.504	0.250	0.754
Monetary benefits per year			
Change in yearly revenues (NPV, mln IDR) ^{99p}	13.484	3.866	16.648
Change in yearly cost of business (NPV, mln IDR) ^{99p}	-14.278	-10.125	0.449
Change in yearly profit (NPV, mln IDR) ^{99p}	-2.428	-2.520	5.350
Change in yearly HH exp. on non-durables (NPV, mln IDR) ^{99p}	-1.070	-2.876	-1.519
Willingness to pay	s.e.	s.e.	
Non-participants (mln IDR)	0.101	0.008	0.101
Participants (mln IDR)	0.088	0.007	0.081



- ▲ Note: This table shows the costs, monetary benefits and willingness to pay per treatment arm, averaged over all FSPs that offer the respective treatment arm. Cost and benefit positions on the left, treatment arms on top. All estimates are presented in million IDR.
- ▲ Sample: Costs: FSPs without FSPs 9 and 10 (N=11). Benefits and willingness to pay: Estimation sample (N=4,292). ▲ Source: Costs: ILO. Benefits and willingness to pay: Endline survey (2018 - 2019).
- Columns (1) presents estimate for the T treatment (training only), column (2) for the C treatment (counseling only), column (3) for the TC treatment (training and counseling). Costs are per client, either as reported by the ILO or in terms of estimated opportunity costs (foregone profits) or the sum of both. Costs cover two full days of training and/ or five counseling sessions à two hours each. Costs are calculating averaging over the eleven FSPs that provided information and weighted by the share of treated clients. Benefits are estimated as ITT effects via linear regressions. In all regressions we control for age, gender, the twenty most imbalanced baseline covariates, FSP fixed effects, and dummies indicating imputation in baseline variables. We show net present values (NPV) of benefits for one year, assuming constant monthly benefits over this timeframe, and discount these monthly with 5-year treasury bond of Indonesia, at currently 7.847%. Willingness to pay is winsorized at 99%, i.e., at 1 million IDR. We did not measure the willingness to pay for combined training and counseling services. For willingness to pay, we display standard errors of the mean instead of stars for significance.
- ▲ The superscript np. indicates the winsorizing level.
- ▲ The statistical significance is given as follows: * indicates $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$. ▲ Result tables for each FSP individually are presented in the respective FSP appendix.

The intervention built the capacities of loan officers of the FSPs to provide classroom training and individual counseling to their clients. We randomly assigned 5,131 clients to the control group or to one of the three treatment arms: (1) only classroom training, (2) only individual counseling, or (3) classroom training combined with individual counseling in follow-up visits. In our endline survey, we managed to re-interview 4,292 clients after the intervention, which lasted between six and eight months.

We find no significant effects of the training and/ or counseling intervention at the impact level when averaged over all FSPs within the first year. There are some weakly significant effects on intermediate outcomes. A robust and highly significant outcome is a 3.3 percentage point increase in the share of clients whose business plans include a cashflow analysis. The control group mean for this variable is 6.6%, implying that the effect is sizable in magnitude.

Given the large number of participating institutions, we take a closer look at the FSP with the highest reported implementation fidelity separately and find more significant results in this sample subset. In particular, clients assigned increased their marketing knowledge, improved marketing practices, and are more likely to prepare a business plan. For the same FSP, we then also find a reduction in costs but also in profits generated in the last 30 days. We present a number of robustness checks for our results, varying the main specification.

A comparison of the treatment effects across the three different treatment arms does not reveal any difference. In fact, some estimates on intermediate outcomes are negative. The notable exception is a decrease in the share of clients not knowing their payment types, which is highly significant in the T only treatment arm. In general, we cannot confirm the hypothesis that the largest treatment intensity (TC) achieved greatest results.

We take a closer look at the characteristics of different clients assigned to the intervention. There are only few differences in the impact of the intervention across different entrepreneur types with more educated clients seemingly experiencing some adverse effects due to the intervention, such as a decrease in life satisfaction, permanent workers, and total savings. Notably, we cannot replicate the effect that men benefit more than women, which similar studies have found. Also, business characteristics have little influence on the impact. However, the intervention might lead to lower loan default rates among businesses outside the trading sector and might increase the share of clients who used their loans for productive purposes among registered businesses. Given the large number of hypothesis tested, the few significant results might also be due to chance, and we cannot decipher a client group who would benefit most from the intervention in the data.

An important finding of our study is, however, that effects vary across the partner institutions. Yet its importance has been little discussed in the literature. Whereas previous studies predominantly cooperate with only one institution, our intervention targeted 13 FSPs that are quite different from each other. The degree to which these FSPs implemented the intervention and adhered to the randomized assignment varies greatly. Thus, organizations supporting similar intervention should take adequate care while selecting partners. Whereas a random selection of partner institutions will be unfeasible in most of the cases, further research is needed to understand the institutional characteristics of successful intervention partners. In this study, saving and credit cooperatives achieved positive but insignificant results, whereas some of the banks showed adverse effects in a number of final outcomes. In contrast, clients from larger banks, with a larger volume of disbursed loans, experienced increased profits. Clients from FSPs which predominantly serve male borrowers are less likely to default on their loans due to the intervention.

The latter results reflect only on the aggregated FSP level, whereas we find no difference in effects on female or male clients individually. This suggests that unobservable factors, which are only proxied by the share of female clients per FSP, play a role in explaining the heterogeneity across the FSPs.



References

- Aldianto, L., Rudito, B., Mirzanti, I., Situmorang, B., & Larso, D. (2010). The development of center of entrepreneurship and business incubator in Pangalengan, west Java - Indonesia. *Technology Management for Global Economic Growth (PICMET)*, 2010 Proceedings of PICMET.
- Atmadja, A. S., Su, J.-J., & Sharma, P. (2016). Examining the impact of microfinance on microenterprise performance (implications for women-owned microenterprises in Indonesia). *International Journal of Social Economics*, 43 (10), 962-981.
- Ayyagari, M., Beck, T., & Demirguc-Kunt, A. (2007, Dec). Small and medium enterprises across the globe. *Small Business Economics*, 29 (4), 415-434.
- Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The miracle of microfinance? evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 7 (1), 22-53.
- Berge, L., Bjorvatn, K., & Tungodden, B. (2015). Human and financial capital from microenterprise development: Evidence from a field and lab experiment. *Management Science*, 61 (4), 707-722.
- Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Robert, J. (2013, February). Does management matter? Evidence from India. *Quarterly Journal of Economics*, 128 (1).
- Bruhn, M., Karlan, D. S., & Schoar, A. (2013). The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico. *World Bank Policy Research Working Paper*.
- Calderon, G., Cunha, J., & De Giorgi, G. (2013). Business literacy and development: Evidence from a randomized controlled trial in rural Mexico. *Working Papers, Banco de Mexico* (2013-24).
- de Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment*. *The Quarterly Journal of Economics*, 123 (4), 1329-1372.
- De Mel, S., McKenzie, D., & Woodruff, C. (2014). Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka. *Journal of Development Economics*, 106, 199-210.
- De Mel, S., McKenzie, D. J., & Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made? *Journal of Development Economics*, 88 (1), 19-31.
- Drexler, A., Fischer, G., & Schoar, A. (2010, July). Keeping it simple: Financial literacy and rules of thumb. Fiala, N. (2013, December). Stimulating microenterprise growth: Results from a loans, grants and training experiment in Uganda. doi: <http://dx.doi.org/10.2139/ssrn.2358086>
- Field, E., Jayachandran, S., & Pande, R. (2010). Do traditional institutions constrain female entrepreneurship? a field experiment on business training in India. *The American Economic Review*, 100 (2), 125-129.
- Gine, X., & Mansuri, G. (2014, June). Money or ideas? a field experiment on constraints to entrepreneurship in rural Pakistan. *The World Bank Policy Research Working Paper* (6959).

- International Labour Organization. (2013). Start and improve your business programme. (<http://www.ilo.org/empent/areas/start-and-improve-your-business/lang-en/index.htm>; accessed on 2018.02.04)
- International Labour Organization. (2018). Start and improve your business programme. (http://www.ilo.org/jakarta/whatwedo/projects/WCMS_444105/lang-en/index.htm); accessed on 2018.02.04)
- Karlan, D., Knight, R., & Udry, C. (2012). Hoping to win, expected to lose: theory and lessons on micro enterprise development. NBER Working Paper Series. (Working Paper 18325)
- Karlan, D., Knight, R., & Udry, C. (2015). Consulting and capital experiments with microenterprise tailors in Ghana. *Journal of Economic Behavior and Organization*, 118, 281-302.
- Karlan, D., & Valdivia, M. (2011, May). Teaching entrepreneurship: impact of business training on micro-finance clients and institutions. *The Review of Economics and Statistics*, 93 (2), 510-527.
- Lafortune, J., Riutort, J., & Tessada, J. (2017, February). Are micro-entrepreneurs constrained by their lack of knowledge or motivation? lessons from a randomized experiment in Chile. (working paper).
- McKenzie, D., & Woodruff, C. (2013). What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, 29 (1), 48-82.
- Rosengard, J., & Prasetyantoko, A. (2011, 12). If the banks are doing so well, why can't I get a loan? regulatory constraints to financial inclusion in Indonesia. *Asian Economic Policy Review*, 6 (2), 273-296.
- Statistics Indonesia. (2018). Percentage of poor people 2007 to 2018. (accessed on 2018.07.28)
- Valdivia, M. (2015). Business training plus for female entrepreneurship? short and medium-term experimental evidence from Peru. *Journal of Development Economics*, 113, 33-51.



Appendices

