**Mobile Financial Services for Female Business Owners**

**Pre-Analysis Plan**

**March 16, 2019**

# Introduction

The available evidence indicates that Indonesian women business owners (WBOs) are substantially under-banked relative to male business owners (MBOs), and possibly as a consequence, have lower savings, fewer business assets and lower business incomes. Evidence from several studies suggests that improved access to savings accounts, through lowering transactions costs, leads to increased savings and, ultimately, to increased business investment and higher incomes among WBOs. However, the literature also indicates that there are several important constraints to the use of financial services and savings, including both pecuniary and non-pecuniary transactions costs, lack of trust in financial institutions, regulatory barriers, information and knowledge gaps, social constraints and behavioral biases. Several constraints affect women in particular, who tend to have less free time and reduced mobility, less education, and whose financial resources are more vulnerable to competing demands from spouses. Although many women save, they are more likely to use informal channels, such as ROSCAs, or to save in the form of assets (e.g., household durables, jewelry, livestock) that provide lower returns and greater risks.

To promote financial inclusion, the Government of Indonesia has encouraged several large banks to develop branchless banking services, including basic savings accounts that are supported by village-based agents and that can be accessed using mobile phones. However, little is known about the most effective ways to promote branchless banking services. The Mobile Financial Services for Female Business Owners trial helps fill that gap by supporting both supply-side and demand-side interventions designed, initially, to increase access to, take up and use of formal savings accounts and, ultimately, to increase WBOs’ savings and incomes. The interventions were randomized within an experimental design that makes it possible to assess the effectiveness and cost effectiveness of both the supply and demand-side interventions, both separately and in combination. Because the trial is limited to villages in which branchless banking services are available, the impact evaluation does not address the impact of establishing branchless banking services in villages in which they do not already exist.

This pre-analysis plan (PAP) defines the methodology and data that will be used to assess the endline impact of the trial. The first section of the PAP provides an overview of the impact evaluation, including a description of the interventions, the sampling process, take-up and compliance rates of the interventions, data and sources. The second section of the PAP describes the hypotheses that will be tested and the variables that will be used for this purpose. The final section of the PAP discusses the estimation methods that will be used in the analysis.

# Overview of the study

## Background

The trial was conducted in 400 predominantly rural villages of five districts of East Java province, Indonesia in which branchless banking services were available.[[1]](#footnote-1) The Indonesian model of branchless banking (*Laku Pandai* or LP)[[2]](#footnote-2) uses village-based agents and mobile telephones to provide basic banking services, including basic savings accounts, credit and financing for micro businesses, micro insurance and other government-approved financial products. The duties of branchless banking agents include: (1) promoting the take up and use of branchless banking products in their villages, (2) identifying and enrolling new clients, and (3) supporting existing clients. The agents participating in the trial (47% of whom are women) are mostly existing shop-owners who are authorized to accept cash deposits and make cash disbursements from customers’ savings accounts, functions that are normally performed at branch offices or by ATMs. The LP basic savings accounts are free of account opening fees, have a maximum balance of Rupiah (Rp.) 20 million (equivalent to roughly US$1,400), a monthly maximum cash withdrawal or transfer of Rp. 5 million and are interest bearing. They are intended to supplement digital wallet (LKD)[[3]](#footnote-3) products already offered by several banks and retail firms. Credit and financing for LP customers have a maximum loan period of one year and a maximum loan value of Rp. 20 million.

Indonesia’s branchless banking system has the potential to reduce substantially both the pecuniary and non-pecuniary transactions costs of opening and using savings accounts. Users are able to check balances and to make payments and transfers using their mobile phones, and they are able to use the expanding network of village agents to open accounts and to make cash deposits and withdrawals without having to travel to bank branches. Agents also assist users to prepare the paperwork necessary to open an account and are responsible for transmitting the completed paperwork to branch offices for processing. Only the account holders and their banks have access to account balances, which makes them less vulnerable to the demands of spouses and other social claimants.

Five large Indonesian banks (one of which was a partner in the trial) have already been approved to provide branchless banking, and more are expected to be approved in the future. Although branchless banking services have been established in some areas, take up has been slow. Constraints to take up include: (1) poor and/or limited internet services in many rural areas, (2) the limited internet and financial skills of potential customers, and (3) insufficient training and lack of motivation of branchless banking agents. The trial was designed to increase the take up and utilization of basic savings accounts, particularly among currently under-banked WBOs in predominantly rural areas through mutually supportive supply- and demand-side interventions. On the supply side, the trial provided training and mentoring to all branchless banking agents while varying randomly the level of their financial incentives for enrolling new clients in a basic savings account. On the demand side, the trial provided training, mentoring and follow-up contacts to randomly-selected female business owners in the 400 sample villages.

## Description of the Interventions

All agents included in the 400 trial villages were trained and mentored by Mercy Corps Indonesia (MCI). The agent training was provided in one individual session that averaged about 3 hours in length (but that varied from 2.5 to 4 hours) in which the agent learned how to use the on-line branchless banking software and the features and relative advantages of both the basic savings (LP) account and the digital wallet (LKD) account.[[4]](#footnote-4) The agent training also included a module on marketing that emphasized the potential value of marketing to under-banked groups, particularly women. Following the agents’ initial training, MCI and partner bank staff provided one-on-one mentoring in three subsequent visits during which monitoring data were also collected.

Agents earn a fee for each new LP savings account client identified as well as a fee for each client’s deposits and withdrawals. The standard agent fee for identifying a new LP client that deposits at least Rp. 20,000 (approximately US$1.54) and who maintains an average savings balance of Rp. 20,000 over two weeks is Rp. 2,000 (about US$0.15). This is the “low-incentive treatment” in the trial. The supply-side “high-incentive treatment” is Rp. 10,000 (about US$0.77) for each new client identified.[[5]](#footnote-5) The randomization is at the village level.[[6]](#footnote-6)

The demand-side interventions were targeted exclusively to randomly selected WBOs in each trial village. Unlike the supply-side incentives, the randomization is within village at the WBO level. The training focused on personal financial management (tracking income and expenses, setting priorities, the importance of saving, financial planning), business financial management (basic bookkeeping, cash flow planning, record-keeping), and description of the branchless banking products (LP and LKD). The initial training was conducted in village groups and lasted for about 3 hours (but ranging from 1 to 4 hours) and was followed by three group mentoring sessions that focused on addressing any questions from the trainees and on actual practices using their own individual businesses as cases.

We define an individual as having “signed up” for an account if she opens one and keeps a minimum balance of Rp. 20,000 (1.42 USD) for at least two weeks. The latter condition has been added to avoid collusion between the customer and the agent.[[7]](#footnote-7) Assuming that the median agent signs up 10 clients per month — irrespective of the treatment they are assigned to — the difference in the monthly pay of a high- vs a low-incentives agent is 80,000 IDR, roughly 9.1% of the average monthly food consumption in East Java (Central Bureau of Statistics).

## Description of the sampling process

In each sample village, the bank recruited a branchless banking agent using its standard selection criteria (the selected agents were in many cases clients with a good credit history).[[8]](#footnote-8) The trial villages are rural or semi-urban villages in which the trial’s partner bank planned to establish branchless banking services. The bank’s branchless banking savings accounts earn an annual interest rate of 0.15%. In order to increase take-up, an additional criterion for village selection was that, to the extent possible, another bank was not providing branchless banking products in the same village at the time.

Agents and female business owners were randomly assigned to their respective treatments prior to the corresponding baseline survey interviews as the agents were recruited. In the case of agents, the first step was to list the sample agents in strata defined on the basis of three criteria: (1) number of households in the village (based on 2011 PODES data), (2) distance of the village from the nearest bank branch (based on 2011 PODES data), (3) and the number of agent competitors in the village.[[9]](#footnote-9) The agents were ranked within each stratum on the basis of the value of a random number generated for each agent from a uniform distribution (using Stata software). 205 agents were assigned to the low-incentive treatment, while 195 agents were assigned to the high-incentive treatment. Similarly, the sample WBOs were listed by village and ranked within each village on the basis of a similarly generated random number. 1,599 WBOs were assigned to receive the training and mentoring (819 in low-incentive villages and 780 in high-incentive villages), while 1,234 WBOs were assigned not to receive any training or mentoring (628 in low-incentive village and 606 in high-incentive villages). No adjustments were made to the original random assignments.

## Take-up rate and respect of treatment assignment

All 400 agents received their initial training and 98% participated in all three mentoring sessions, including those who replaced agents who resigned or were terminated. Among the 1,599 WBOs who were randomly assigned to receive the training, 91% participated in the initial training while 80% participated in all three mentoring sessions. None of the WBO controls received any training or mentoring. The original plan was to provide the mentoring at one-week intervals immediately following the initial training. However, the fact that the WBO training and mentoring was provided in group sessions made it difficult to schedule the follow-up mentoring. Among the 98% of all agents who completed the training and three mentoring visits, only 19% completed it within the planned 21 days of completing their initial training while 30% completed it in 43 or more days after completing their initial training. Among the 80% of WBOs who completed the training and three mentoring visits, only 15% completed it within the planned 21 days while 29% completed it in 43 or more days after completing their initial training.

## Description of the data sources

Baseline and two follow-up surveys collected data on agents and both female and male business owners. Baseline data were collected for 4,828 business owners and 475 agents in a Baseline Survey that was conducted in two phases (November 2016-February 2017 in 107 villages and July-November 2017 in 293 villages).[[10]](#footnote-10) The Baseline Survey collected extensive data on both business owners and agents. The business owner data were collected in a single household questionnaire that required about 1.5 hours to administer and that included modules on (1) basic background characteristics (e.g., age, sex, education, marital status, number of children), (2) mobile phone usage, (3) relationships (if any) with the village bank agents, (4) knowledge and use of mobile and other financial services, (5) sources of income, (6) business assets, (7) household assets and housing characteristics, (8) savings and credit, (9) intra-household decision-making, and (10) feelings and perceptions about the agents and financial institutions.

Similarly extensive data on agents were collected in the Baseline Survey. The agent questionnaire also required about 1.5 hours to administer and included modules on (1) basic background characteristics, (3) mobile phone usage, (4) agent job expectations and perceptions, (5) agent work activities, (6) trust in financial institutions, (7) knowledge and use of mobile and other financial services, (8) personal liquidity, (9) sources of income, (10) voluntary activities, (11) business assets, and (12) cognitive skills.

There are several features of the survey data that make it a unique resource. First, it includes data on both female and male business owners (although the male business owners received no training or mentoring). The availability of similar data on female and male business owners will make it possible to measure differences in the effects of higher agent incentives on their outcomes. Second, the availability of detailed data on agents will make it possible to assess the effects of higher agent incentives on their behavior as an important link in the trial’s causal chain.

# Hypotheses and variable definitions

## Groups of Hypotheses:

A: Impact of the treatments on primary outcomes

B: Impact of the treatments on secondary outcomes

C: Process and Mechanisms

D: Heterogeneity of impacts

## Group A: Impact of the treatments on primary outcomes

### Hypothesis A1: Both treatments will have a positive effect on the profits of treated WBOs’ primary and secondary businesses

This hypothesis will be tested by estimating the treatment effects on the following indicators. However, the winsorized average monthly profit of the primary business (indicator #7) is the primary profits indicator. The others will be used as robustness checks.

1. Average monthly profit of primary business during the past year (H22, column H1A)
2. Average monthly profit of secondary business during the past year (H22, column H1B)
3. Total average monthly profits of primary and secondary businesses combined during the past year (H22, sum of columns H1A and H1B)
4. Average monthly profit of primary business during the past year, transformed to inverse hyperbolic sine values (H22, column H1A)
5. Average monthly profit of secondary business during the past year, transformed to inverse hyperbolic sine values (H22, column H1B)
6. Total average monthly profits of primary and secondary businesses combined during the past year, transformed to inverse hyperbolic sine values (H22, sum of columns H1A and H1B)
7. Average monthly profit of primary business during the past year, winsorized at the 99th percentile (H22, column H1A)
8. Average monthly profit of secondary business during the past year, winsorized at the 99th percentile (H22, column H1B)
9. Total average monthly profits of primary and secondary businesses combined during the past year, winsorized at the 99th percentile (H22, sum of columns H1A and H1B)

Note: All reported profits of the primary business will be recoded to zero if the WBO no longer has a business. When calculating total average monthly profits in primary and secondary businesses (indicators 3, 6 and 9), reported profits of the secondary business will be recoded to zero if the WBO does not have a secondary business.

### Hypothesis A2: Both treatments will have a positive effect on the savings of treated WBOs

This hypothesis will be tested by estimating the treatment effects on the following indicators and a standard z-score index (described in the methodology section):

1. WBO personally saved during the past 12 months (G01)
2. Total amount saved in the last 12 months (sum of G03A, G03I, G03J, G03B, G03C, G03D, G03E, G03F, G03G, G03H, G03V)
3. Amount saved in the last 12 months in a formal bank account or electronic savings account (sum of G03A, G03I, G03J, G03B)
4. Amount saved in the last 12 months in an electronic savings account (sum of G03I, G03J, G03B)
5. Total amount saved in the last 12 months, transformed to inverse hyperbolic sine values (sum of G03A, G03I, G03J, G03B, G03C, G03D, G03E, G03F, G03G, G03H, G03V)
6. Amount saved in the last 12 months in a formal bank account or electronic savings account, transformed to inverse hyperbolic sine values (sum of G03A, G03I, G03J, G03B)
7. Amount saved in the last 12 months in an electronic savings account, transformed to inverse hyperbolic sine values (sum of G03I, G03J, G03B)
8. Total amount saved in the last 12 months, winsorized at the 99th percentile (sum of G03A, G03I, G03J, G03B, G03C, G03D, G03E, G03F, G03G, G03H, G03V)
9. Amount saved in the last 12 months in a formal bank account or electronic savings account, winsorized at the 99th percentile (sum of G03A, G03I, G03J, G03B)
10. Amount saved in the last 12 months in an electronic savings account, winsorized at the 99th percentile (sum of G03I, G03J, G03B)
11. Proportion of total savings in last 12 months in a formal bank account or electronic savings account (Proportion = [G03A+G03I+G03J+G03B] / [G03A+G03I+G03J+G03B+G03C+G03D+G03E+G03F+G03G+G03H+G03V] )
12. Proportion of total savings in last 12 months in an electronic savings account (Proportion = [G03I+G03J+G03B] / [G03A+G03I+G03J+G03B+G03C+G03D+G03E+G03F+G03G+G03H+G03V] )
13. Total current savings balance (sum of G03aA, G03aI, G03aJ, G03aB, G03aC, G03aD, G03aE, G03aF, G03aG, G03aH, G03aV)\*
14. Total current savings balance in a formal bank account or electronic savings account (sum of G03aA, G03aI, G03aJ, G03aB)\*
15. Total current savings balance in an electronic savings account (sum of G0aI, G03aJ, G03aB)\*
16. Total current savings balance, transformed to inverse hyperbolic sine values (sum of G03aA, G03aI, G03aJ, G03aB, G03aC, G03aD, G03aE, G03aF, G03aG, G03aH, G03aV)\*
17. Total current savings balance in a formal bank account or electronic savings account, transformed to inverse hyperbolic sine values (sum of G03aA, G03aI, G03aJ, G03aB)\*
18. Total current savings balance in an electronic savings account, transformed to inverse hyperbolic sine values (sum of G03aI, G03aJ, G03aB)\*
19. Total current savings balance, winsorized at the 99th percentile (sum of G03aA, G03aI, G03aJ, G03aB, G03aC, G03aD, G03aE, G03aF, G03aG, G03aH, G03aV)\*
20. Total current savings balance in a formal bank account or electronic savings account, winsorized at the 99th percentile (sum of G03aA, G03aI, G03aJ, G03aB)\*
21. Total current savings balance in an electronic savings account, winsorized at the 99th percentile (sum of G03aI, G03aJ, G03aB)\*
22. Proportion of Total current savings balance in a formal bank account or electronic savings account (Proportion=[G03aA+G03aI+G03aJ+G03aB] / [G03aA+G03aI+G03aJ+G03aB+G03aC+G03aD+G03aE+G03aF+G03aG+G03aH+G03aV] )\*
23. Proportion of Total current savings balance in an electronic savings account (Proportion=[ G03aI+G03aJ+G03aB] / [G03aA+G03aI+G03aJ+G03aB+G03aC+G03aD+G03aE+G03aF+G03aG+G03aH+G03aV] )\*
24. Formal bank account or electronic savings account cited as a savings instrument primarily used when saving for emergencies (G06A, G06B, G06K, G06L: any indicator=1)
25. Electronic savings account cited as a savings instrument primarily used when saving for emergencies (G06B, G06K, G06L: any indicator=1)

Note: Indicators marked with an asterisk (\*) have no baseline values. Indicators 2-10 will be coded zero if indicator 1=3 (“No”). Indicators 11, 12 (23 and 24) will be coded to zero if indicator 2 (13) is equal to zero. A z-score index reflecting “saving” will be calculated and analyzed for all of the above indicators, with the exception of indicators 11, 12 and 22-25. A separate z-score index will be calculated for indicators 11, 12 and 22-25 representing a possible shift to formal saving from other forms of saving.

## Group B: Impact of the treatments on secondary outcomes

A simple form of a firm’s production function can be defined as:

Y = *f* (K, L, A, E),

where K=Capital, L=Labor, A=Productivity and E=Business owner’s skills

A WBO is assumed to choose the appropriate level of K and L, depending on A and E such that the marginal productivity is equal to the market interest rate and the market wage rate respectively.

### Hypothesis B1: Both treatments may increase the levels of K and L that the entrepreneur chooses

This hypothesis will be tested by estimating the treatment effects on the following indicators of K and L and on standardized z-score indices (described in the methodology section):

A z-score index will be calculated for the following capital (K) indicators:

1. Value of the capital stock in all businesses owned, winsorized at the 99th percentile (N02, sum of rows 1-6)
2. Capital investments during the past 12 months in primary and secondary businesses combined, winsorized at the 99th percentile (EH02, sum of columns H1A and H1B, with EH02 recoded to zero if response to EH01 is “No”)\*
3. Increase in the value of stocks in primary and secondary businesses combined, winsorized at the 99th percentile (EH04, sum of columns H1A and H1B, with EH04 recoded to zero if response to EH03 is “No”)\*
4. Value of investments in the past 12 months in relation to the value of total business assets (sum of recoded EH02 and EH04 over columns H1A and H1B / divided by the value of the capital stock in all businesses, as defined in #1 above)\*
5. Value of investments in the past 12 months in relation to total business profits in the past 12 months (sum of recoded EH02 and EH04 over columns H1A and H1B / divided by the sum of H22 over columns H1A and H1B multiplied by 12)\*
6. Invested in physical capital of primary or secondary business in past 12 months (EH01=1 for either column H1A or H1B)\*
7. Increased value of stocks in primary or secondary business in past 12 months (EH03=1 for either column H1A or H1B)\*

The primary labor (L) indicator will be a z-score based on the following indicators:

1. Number of days worked by WBO in a typical month in her primary business (H11, column H1A)
2. Number of days worked by WBO in a typical month in her secondary business (H11, column H1B)
3. Number of hours worked by WBO in a typical day in her primary business winsorized at the 99th percentile (H12, column H1A)
4. Number of hours worked by WBO in a typical day in her secondary business winsorized at the 99th percentile (H12, column H1B)
5. Number of hours worked by WBO in a typical month in her primary business winsorized at the 99th percentile (H11\*H12, column H1A)
6. Number of hours worked by WBO in a typical month in her secondary business winsorized at the 99th percentile (H11\*H12, column H1B)
7. Number of unpaid workers in the primary business winsorized at the 99th percentile (H13, column H1A)
8. Number of unpaid workers in the secondary business winsorized at the 99th percentile (H13, column H1B)
9. Number of paid workers in the primary business (H14, column H1A)
10. Number of paid workers in the secondary business (H14, column H1B)
11. Total number of paid and unpaid workers in the primary business winsorized at the 99th percentile (sum of indicators #s 7 and 9)
12. Total number of paid and unpaid workers in the secondary business winsorized at the 99th percentile (sum of indicators #s 8 and 10)

Note: Indicators marked with an asterisk (\*) have no baseline values. The z-scores will exclude indicators referring to secondary businesses if no secondary business is reported.

### Hypothesis B2: The training and mentoring of WBOs may increase their use of recommended business practices

This hypothesis will be tested by estimating the treatment effects on a simple average of the following indicators of recommended business practices:

1. Asked a supplier about which products are selling well in WBO’s industry (P10)
2. Used a special offer to attract customers in the last three months (P11)
3. Done any form of advertising in the last six months (P12)
4. Done anything to measure the effectiveness of the advertising (P13, equal to zero if P12 is “No”)
5. Attempted to negotiate with a supplier for lower prices on raw materials in the last three months (P14)
6. Has a record-keeping system which allows WBOs to know their stock of goods to sell or raw materials on hand (P16)
7. Keeps written business records (P17)
8. Records every purchase sale made by the business (P18, equal to zero if P17 is “No”)
9. Able to use records to see easily how much cash is on hand at any point in time (P19, equal to zero if P17 is “No”)
10. Regularly uses records to know whether sales of a particular product are increasing or decreasing from month to month (P20, equal to zero if P17 is “No”)
11. Has worked out the cost of each main product sold (P21)
12. Has a written budget for business expenses (P23)
13. Has records needed to apply for a bank loan (P26)
14. Keeps business money separate from household money (EP27)\*
15. Uses bank account or branchless banking account for the business (EP28)\*
16. WBO’s primary business is registered with the Government (H08, column H1A)

Note: Indicators marked with an asterisk (\*) have no baseline values.

### Hypothesis B3: Both treatments may have a positive impact on women’s agency

This hypothesis will be tested by estimating the treatment effects on two alternative standardized z-score indices (described in the methodology section) based on subsets of the following indicators:

1. WBO alone decides or participates in decision whether or not to buy an appliance for the home (O01a: responses=1, 2 or 4)
2. WBO alone decides or participates in decision in what way household members may work outside the home (O01b: responses=1, 2 or 4)
3. WBO alone decides or participates in decision whether to support other family members (O01c: responses=1, 2 or 4)
4. WBO alone decides or participates in decision whether to save for the future (O01d: responses=1, 2 or 4)
5. WBO alone decides or participates in decision whether to sign up for a new banking product (O01e: responses=1, 2 or 4)
6. Proportion of WBO’s business earnings NOT known to spouse (Proportion=1-[O05/100000], Not Applicable if no spouse)
7. WBO alone has access to the money made in her business (O09=”A”)
8. WBO alone decides how money from her business will be spent (O010=”A”)
9. WBO has some money that she has sole control over (O12=1)
10. WBO alone decides or participates in decision about how money spouse earns will be spent (O013=2-5, 6=Not Applicable due to no spouse or no spouse income)

Note: Two z-score indices will be calculated with the above indicators. The first, representing empowerment in decision-making, will be based on indicators 1-5 and 9-10. The second z-score index, representing non-cooperative households and control over business earnings, will be based on indicators 6-8. If WBO respondent does not have a business or her spouse is not present (O04x) or if WBO’s spouse does not have income, so that some indicators are Not Applicable, the z-score indices (with values ranging from zero to one) will be based only on the applicable indicators.

### Hypothesis B4: Both treatments may have a positive effect on access to credit

This hypothesis will be tested by estimating the treatment effects on a standardized z-score index (described in the methodology section) based on the following indicators:

1. WBO has a registered bank account under her name (F18A-F18V: any indicator=1)
2. WBO uses bank credit (F17=C, F, G, H, or J)
3. WBO has a loan with a bank (G12)
4. WBO has borrowed money from a money lender during the past 12 months (EG01A=1 and EG02A is not equal to 3)\*
5. WBO has borrowed money from a non-bank financial institution during the past 12 months (EG01B=1 and EG02B is not equal to 3)\*
6. WBO currently has a loan with the partner bank agent (EG03)\*
7. Last amount borrowed from the partner bank agent (EG04, coded zero if EG03=”No”)\*

Note: Indicators marked with an asterisk (\*) have no baseline values.

### Hypothesis B5: Both treatments may have a positive impact on household welfare

This hypothesis will be tested by estimating the treatment effects on the following indicator:

1. First principal component of the ownership of 20 household durables (Ma-Mt)

## Hypothesis group C: Process and mechanisms of change

This section focuses on how the treatments could generate changes in the primary outcomes of interest (Hypotheses: Group A). The causal chain describing the process by which the trial is expected to affect business owners’ savings and business outcomes is depicted in Figure 1 and involves the following steps:

* Step 1: Agents are recruited and participate in the agent training and mentoring
* Step 2: The performance-based agent incentives improve treated agent performance
* Step 3: Treated WBOs are better informed about the branchless banking services available in their villages
* Step 4: Treated WBOs have more favorable perceptions of agents and banks and are more open to taking up and using branchless banking services. This step will be tested by co-investigators in a separate paper.
* Step 5: Treated WBOs have increased knowledge and agency and improved business practices
* Step 6: Treated WBOs increase their saving and borrowing, leading in turn to increased business inputs
* Step 7: Increased business inputs, utilization of branchless banking services, increased WBO knowledge and improved WBO business practices increase WBO business profits, leading ultimately to higher household welfare

Figure 1. Causal chain



There are two external conditions that must be fulfilled in order for the causal chain to hold:

### Condition 1: Branchless banking services remain available to WBOs in their villages

This condition will be reviewed using the MCI monitoring data on agent deployment, which indicate periods in which an agent was not available in a village due to resignation or discharge (with replacements in some cases) and similar data from Section A (questions EA01-EA05) of the Endline Survey Agent Book.

### Condition 2: Agents and WBOs have access to the internet

This condition will be reviewed with respect to the following indicators:

1. Agents have mobile signal coverage every day (Agent B10a=3)
2. Agents are able to send/receive text/SMS every day (Agent B10b=3)
3. Agents never receive SMS messages late (Agent B10c=3)
4. Agents never lose service while on call (Agent B10d=3)
5. Agents can access the internet every day (Agent B10e=3)
6. Agent has internet access at work (Agent B17)
7. Agent usually has no internet network problems (Agent B19=0)
8. WBOs have mobile network signal every day (B10a=3)
9. WBOs can access the internet every day (B10e=3)
10. Indicators 1-8 are all equal to one (indicator 9 is not included in this summary indicator because it was not reported for 64% of sample WBOs)

### Hypothesis C-S2 (testing step 2): Performance-based agent incentives improve agent performance

This hypothesis will be tested by estimating the treatment effects on the following *agent* indicators and a standardized z-score index (variable references refer to the Agent Questionnaire):

1. Agent promotes LP to women (Agent ED13, responses include C)
2. Agent promotes LP to business owners (Agent ED13, responses include G)
3. Agent promotes LKD to women (Agent ED14, responses include C)
4. Agent promotes LKD to business owners (Agent ED14, responses include G)
5. Number of hours per week worked by agent in agent job (Agent ED15d)
6. Percentage of agent’s working time in agent job spent promoting LP products in agent’s shop (Agent ED16a)
7. Percentage of agent’s working time in agent job spent promoting LP products outside agent’s shop (Agent ED16b)
8. Percentage of agent’s working time in agent job spent educating client about financial product (Agent ED16f)
9. Agent has made investments to improve her/his performance (Agent ED18)
10. Amount of commission earned by agent from partner bank job in typical month (Agent ED19)
11. Amount of commission earned by agent from partner bank job in October 2018 (Agent ED23)
12. Agent’s satisfaction with partner bank job (Agent ED36)
13. Agent’s assessment of competence in partner bank job (Agent ED39)
14. Agent has personally adopted LKD product (Agent AK09)
15. Agent has personally adopted LP product (Agent AK10)
16. Hours spent by agent doing voluntary activities in a typical month (Agent P01, P03)
17. Amount of money contributed by agent to any social institution/organization in the past year (Agent P04)
18. Agent sees self as doing a thorough job (Agent V01.1, response=5)
19. Agent sees self as considerate and kind to almost everyone (Agent V01.11, response=5)
20. Agent sees self as doing things efficiently (Agent V01.12, response=5)
21. Agent sees self as outgoing, sociable (Agent V01.13, response=5)

### Hypothesis C-S3a (testing step 3): The constraints faced by WBOs are associated with their participation in the training and mentoring

This hypothesis will be explored using regression analysis to estimate the partial relationships between the following left-side indicators of participation on the following possible right-side constraints to participation:

Left-side indicators of participation (from the MCI monitoring data):

1. WBO participated in initial training
2. WBO completed all training and mentoring
3. Number of days WBO needed to complete the training and mentoring
4. WBO participated in the financial literacy training supported by the trial (EF01a=1 & EF01c=”A” or “B”)\*

Note: Indicators marked with an asterisk have no baseline values. WBOs are considered to have participated in the initial training if the MCI WBO training database reports a date for her initial training session as well as a date for her first mentoring session. WBOs are considered to have completed all training and mentoring if the MCI WBO training data base reports a date for all of her training and mentoring sessions. The number of days WBOs needed to complete the training is defined, for WBOs who have completed all training and mentoring sessions, as the number of days elapsed between the date of her third mentoring session and the date of her initial training session.

Right-side indicators of constraints (from Baseline Survey):

1. WBO’s age at baseline (Baseline AR03)
2. WBO’s is currently married at baseline (Baseline A04=2)
3. Baseline value of the WBO’s agency z-score defined as in Hypothesis B3 (Baseline O01a-O01e, O05, O09, O10, O12 and O13)
4. Baseline number of children in the WBO’s household (Baseline AR10, coded zero if Baseline AR09 is “No”)
5. Household size (Baseline A08)
6. Highest level of schooling completed by the WBO (Baseline A06, recoded per Note below)
7. Baseline cognitive score (Baseline X01, X02, X03, X04: sum of correct responses)
8. WBO uses smart phone at baseline (Baseline B06)
9. First principal component of the baseline ownership of 20 household durables (Baseline Ma-Mt: “Yes”=1) and housing characteristics (L01-L10, coded as explained in the Note below)
10. Total average baseline profits of primary and secondary businesses during the past year (Baseline H22, sum of columns H1A and H1B)
11. WBO knows the agent’s name (unprompted) at baseline (Baseline C01)
12. Proportion of other sample WBOs in same village who are known to the WBO at baseline (Baseline K02)
13. WBO belongs to a business-related association or group at baseline (Baseline K01)

Note: the estimated relationships are only associations due to endogeneity. Variable A06 (schooling) will be recoded as follows: 1, 14, 90, 95=None/kindergarten; 2, 11, 72=Primary; 3, 4, 12, 73=Lower secondary; 5, 6, 15, 74=Upper-secondary; 60-63=Tertiary; 98=Non-response. Variables L01-L10 are recoded as follows: L01 (all responses converted to dummy variables), L02 (total number of rooms), L03 (number of rooms for sleeping), L04 (all responses are converted to dummy variables), L05 (all responses are converted to dummy variables), L06 (“Yes”=1), L07 (all responses are converted to dummy variables), L08 (“Yes”=1), L09 (all responses are converted to dummy variables), and L10 (“Yes”=1).

### Hypothesis C-S3b (testing step 3): Treated WBOs know more about the branchless banking services in their villages

This hypothesis will be tested by estimating the treatment effects on the following indicators and a standardized z-score index (described in the methodology section):

1. WBO knows (unprompted) the name of the MM agent (C01a)
2. WBO knows (prompted with name) the MM agent (C01)\*
3. WBO knows about branchless banking services (F01)
4. WBO thinks that the training provided enough information about LP and LKD (F01a)\*
5. WBO knows that there is an agent in her village promoting/selling LKD and/or LP (EF03)\*
6. WBO knows about LKD (F02A)\*
7. WBO knows about LP (F02D)\*
8. WBO knows about other MM products ([sum of F02B, F02C, F02E, F02F, F02G, F02H, F02I, F02J, F02K, F02V]/10)\*
9. WBO knows the interest rate offered with LKD (F03A)\*
10. WBO knows the interest rate offered with LP (F03D)\*
11. WBO knows the fee for a deposit with LKD (F04aA)\*
12. WBO knows the fee for a deposit with LP (F04aD)\*
13. WBO knows the fee for a withdrawal with LKD (F04aA)\*
14. WBO knows the fee for a withdrawal with LP (F04bD)\*

Note: Indicators marked with an asterisk have no baseline values.

### Hypothesis C-S4b (testing step 4): Both treatments motivate WBOs to enroll in and use branchless banking services

This hypothesis will be tested by estimating the treatment effects on the following indicators and a standardized z-score index (described in the methodology section):

1. WBO has enrolled in LKD product of partner bank (EF16)\*
2. WBO has enrolled in LP product of partner bank(EF16)\*
3. WBO has enrolled in the MM product of any non-partner bank (“Yes” response for any non-partner bank in EF16)\*
4. WBO has registered LKD product of partner bank (EF19)\*
5. WBO is currently using LKD product of partner bank (F13)
6. WBO has registered LP product of partner bank (EF19)\*
7. WBO is currently using LP product of partner bank (F13)
8. WBO is currently using the MM product of any non-partner bank ( “Yes” response for any non-partner bank in F13)
9. WBO uses bank savings account (F17, responses include A)
10. WBO uses bank to obtain business loans (F17, responses include F)
11. WBO uses LKD product of partner bank to save money (F14, response=D)
12. WBO uses LP product of partner bank to save money (F14, response=D)\*
13. WBO uses LP product of partner bank to take a loan (F14, response=I)\*
14. WBO uses the MM product of any non-partner bank to save money (response of “D” for any non-partner bank in F14)
15. WBO uses the MM product of any non-partner bank to take a loan (response of “I” for any non-partner bank in F14)\*

Note: Indicators marked with an asterisk have no baseline values.

## Hypothesis group D: The treatments will affect WBOs differently depending on their baseline characteristics

For each of the hypotheses D1-D10 listed below, the estimated treatment effects on the primary outcomes described in hypothesis Group A above will be tested for their possible heterogeneity. For the characteristics that have statistically significant heterogeneous effects on primary outcomes, the heterogeneity of the estimated treatment effects on the secondary outcomes described in hypothesis group B and the mechanisms described in hypothesis group C will also be tested.

### Hypothesis D1: WBOs and male business owners are affected differently by performance-based agent incentives

Tests of this hypothesis will be tested for all outcomes under groups A-C.

Note: The sub-sample of male business owners will be pooled with the WBO sub-sample for this test.

### Hypothesis D2: WBOs are affected differently by both treatments depending on their stated interest in enrolling in branchless banking at baseline (R01)

The WBO sample will be divided according to whether WBO’s indicate that they are (1) sure to enroll (R01=1), (2) not sure about enrolling (R01=2), or (3) sure not to enroll (R02=3).

### Hypothesis D3: WBOs are affected differently by both treatments depending on their baseline work experience

The WBO sample will be divided into quintiles according to their baseline years of experience working in their primary business (Baseline H10, column H1A).

### Hypothesis D4: WBOs are affected differently by both treatments depending on their baseline adherence to recommended business practices

The WBO sample will be divided into quintiles according to their baseline business practices score (defined as under hypothesis B2).

### Hypothesis D5: WBOs are affected differently by both treatments depending on their baseline profitability

The WBO sample will be divided into quintiles according to their baseline primary business profitability (Baseline H22, column H1A).

### Hypothesis D6: WBOs are affected differently by both treatments depending on their completed levels of schooling at baseline

The WBO sample will be divided according to levels of schooling defined as in the Note to C-S3a, but with None/kindergarten (N=33) and Other (N=3) combined with Primary (N=688) and with Tertiary (N=149) combined with Upper secondary (N=957).

### Hypothesis D7: WBOs are affected differently by both treatments depending on their baseline household assets

The WBO sample will be divided into quintiles according to the baseline values of the household asset index (defined as in the Note to C-S3a).

### Hypothesis D8: WBOs are affected differently by both treatments depending on their district of residence

The WBO sample will be divided according to district of residence. The percentages of the sample WBOs by district are as follows: Tuban (17.7%), Bojonegoro (17.9%), Ngawi (25.0%), Lamongan (35.4%) and Gresik (4.0%). Because of its small sample size, Gresik will be combined with neighboring Lamongan.

### Hypothesis D9: WBOs are affected differently by both treatments depending on their baseline agency

The WBO sample will be divided into quintiles according to the baseline values of the z-score calculated as in Hypothesis B3.

### Hypothesis D10: WBOs are affected differently by both treatments depending on the number of children in their household.

The baseline number of children in the WBO’s household (Baseline AR10, coded zero of Baseline Ar09 is “No”) will be interacted with the treatment dummies in order to test this hypothesis (the baseline number of children range from 0 to 7, but 90% of the reported numbers are 2 or below).

# Methodology

## Identifying the treatment effect

For outcomes for which baseline values are available, intention-to-treat effects will be estimated using the following regression model:

Yij = β0 + β1Highj + β2Trainij + β3Highj\*Trainij + γY(t=0)ij + **Z**ij**δ** + εij (1)

where

Yi is the post-treatment outcome of individual i in village j

Highj, Trainij and Highj\*Trainij are dummy variables referring respectively to high agent incentives only, WBO training and mentoring only, and the combination of high agent incentives and WBO training/mentoring.

β1, β2 and β3 are fixed parameters measuring intention-to-treat effects

Y(t=0)ij is the baseline value of Yij

γ is the fixed coefficient of the baseline value of Yij

**Z**ij is a row vector of covariates, including dummy variables referring to the strata used in the random assignment and other baseline covariates[[11]](#footnote-11)

**δ** is a column vector of the fixed coefficients of the covariates

εij is the random error term = cj + uij, where cj is an unobserved village-level effect and uij is the idiosyncratic error

In the case of outcomes for which no baseline value is available, the following regression model will be used:

Yij = β0 + β1Highj + β2Trainij + β3Highj\*Trainij + **Z**ij**δ** + εij (2)

where the definitions of the variables are the same as in equation (1).

For both equations, the estimation method will be pooled OLS with the estimated standard errors adjusted for heteroscedasticity and for clustering at the village level (reflecting the village-level treatment of high agent incentives).

In addition to calculating the intention-to-treat effects, estimates of the treatment effects on the treated will also be obtained and reported by instrumenting WBO participation in the training and mentoring on the random assignment to the treatment group. This estimate will make it possible to control for WBO non-compliance with the treatment assignment.

## Hypothesis tests

1. There is no treatment effect (β1=β2=β3=0)
2. High agent incentives have no effect (β1=0) and (β3=0)
3. The training and mentoring of WBOs has no effect (β2=0) and (β3=0)
4. There is no interaction effect of high agent incentives with WBO training and mentoring (β3=0)

## Dealing with multiple outcomes

There are two challenges when working with a large number of outcome variables. First, the volume of outcome variables can make it more difficult to interpret the results. Second, analysis of multiple outcomes can increase the risk of Type I errors (rejecting the null hypothesis of no effect when it is true) if the significance tests are not adjusted appropriately.

Two alternative strategies are used to address these risks:

a. The PAP clearly identifies a set of primary outcomes in hypothesis group A

b. The methodology of Kling, Katz and Liebman (2007) will be used to test the significance of related groups of outcomes using a single aggregate (a standardized z-score index of the related outcomes).[[12]](#footnote-12)

## Estimating heterogeneous outcomes

To test for the heterogeneity of effects, the variable(s) of interest will be interacted with the treatment variables in equation (2) and with the treatment variables and lagged dependent variable in equation (1).

## Addressing missing data and questions with limited variation

It is not yet clear to what extent missing data will be a problem. However, it is noted that response rates were high for most variables in the Baseline Survey.

## Survey attrition

If Aij represents whether individual i in village j attrits from the study because the individual cannot be found or refuses to participate, we will estimate the following regression model to test whether survey attrition is related to treatment status:

Aij = β0 + β1Highj + β2Trainij + β3Highj\*Trainij + **Z**ij**δ** + εij (3)

where the definitions of the variables are the same as in equation (1).

If treatment status does not affect survey attrition at the 5% significance level, then we will not adjust the estimates for attrition.

If treatment status does have a statistically significant effect on survey attrition, we will test the robustness of our results using Lee bounds (Lee 2008).

## Outliers

As mentioned in the variable specifications, we will test for the sensitivity of the results on savings and profits to outliers by estimating winsorized versions of these variables at the 99% percentile.

1. The districts are: Tuban, Bojonegoro, Ngawi, Lamongan and Gresik, with a combined 2010 population of 5.56 million. [↑](#footnote-ref-1)
2. “Laku Pandai” is an Indonesian acronym for the provision of banking and other financial services without branch offices. [↑](#footnote-ref-2)
3. LKD is an acronym for *layanan keungan digital*, Indonesian for “digital financial services,” which is an electronic payment and transfer system already in wide use. [↑](#footnote-ref-3)
4. Unregistered LKD accounts have a maximum balance of Rp. 1 million and can only be used for payments. However, when registered by an agent, the maximum balance is Rp. 5 million, and the account can also be used for transfers and withdrawals. [↑](#footnote-ref-4)
5. In addition, all agents earn Rp. 1,000 for each cash deposit of minimum Rp. 10,000 and Rp. 2,500 for cash withdrawals under Rp. 200,000 and Rp. 4,000 for cash withdrawals of Rp. 200,000 and above. [↑](#footnote-ref-5)
6. In addition, a public-private information treatment was cross-randomized to equal shares of the high and low-incentive agents. In the “private information” treatment, the sample business owners are given information about the new product but are not informed about the compensation of the agent. In the “public information” treatment, the business owners are informed not only about the product but also about the compensation of the agent (i.e., the payment received for each new customer). This public-private information treatment is orthogonal to the other treatments and is part of a separate evaluation (https://www.socialscienceregistry.org/trials/3167/history/32248). [↑](#footnote-ref-6)
7. This was done to minimize the likelihood of a client signing up for the account with the only intention to raise the agent’s pay, and then closing it immediately after. As explained below, our data will provide us with detailed information on each client's account and will allow us to assess whether any client “games” the system by withdrawing their money and/or closing their account exactly two weeks after she opens it. [↑](#footnote-ref-7)
8. The standard criteria are: (1) the owner is a previous borrower from the bank, (2) the business is in a central location in the village, (3) the owner is mostly present at the business premises, (4) the owner has a good reputation among villagers (as confirmed by the village chief), (5) the owner is able to demonstrate sufficient financial liquidity, (6) the owner is not an agent for another bank, and (6) the owner is willing to participate as an agent. [↑](#footnote-ref-8)
9. The actual strata were defined on the basis of the following three binary variables: (1) one if the distance of the agent’s village to the nearest bank branch exceeded the median value in our 400 sample villages, and zero otherwise, (2) one if the number of households in the village exceeded the median value for nearby villages, and zero otherwise, and (3) one if there were one or more agent competitors in the same village, and zero otherwise. [↑](#footnote-ref-9)
10. The 475 surveyed agents included both “primary” agents (N=400) and “secondary” agents (N=75). The primary agents are those that the baseline survey team determined were actually doing the agent job and were the only agents subsequently trained and mentored. [↑](#footnote-ref-10)
11. These covariates will include the following variables: the WBO’s age (and age-squared) and categorical variables referring to her highest completed level of schooling, the baseline household asset index (as described in the note to C-S3a), the baseline value of business capital, and the baseline household size. In addition, other baseline covariates may be included when the baseline value of the dependent variable is not available (e.g., baseline savings during the last year as an additional covariate in a model determining the value of current savings balances). [↑](#footnote-ref-11)
12. This methodology involves: (1) converting all outcomes so that the sign of the effect of all variables in a family goes in the same direction, (2) calculating the z-score of each variable by subtracting the control group mean and dividing by the control standard deviation, and (3) taking an average of the z-scores in each family. [↑](#footnote-ref-12)