Pre-analysis plan: Increasing Deposits with Banking Agents at Scale - Experimental Evidence from India*

Luke Butcher† Karan Nagpal‡ Andres Parrado§

April 01, 2020

Contents

*__Contact__: andres.parrado@idinsight.org. **Current version**: April 01, 2020. **Preliminary draft. Please do no cite or circulate without permission.** We thank Pawan Bakshi, Jeremy Shapiro, Divya Nair, Sasidhar Thumuluri, Dinabandhu Bharti, Syed Maqbool, and all the staff from BASIX Sub-K for their invaluable support and feedback and to the Bill & Melinda Gates Foundation for their support. All errors are our own.

†Associate, IDinsight

‡Economist, IDinsight

§Manager, IDinsight

1 Project Overview

1.1 Introduction

This pre-analysis plan (PAP) outlines the context, research questions, and empirical strategy to measure the impact of a composite intervention, consisting primarily of providing savings calendars to customers, on banking agents in southern India. Our PAP is being submitted to the AEA RCT registry after the implementation of our intervention is complete, but before the first dataset with results is available and before any data analysis takes place.

1.2 Context

While India has made substantial progress towards financial inclusion,[1] much remains to be done to ensure everyone utilizes, and benefits from, the increased access to the formal financial system. Take, for instance, India's effort to extend bank accounts to all of its population. The Pradhan Mantri Jhan Dan Yojana (PMJDY), a flagship financial inclusion program whose main feature is zero-balance accounts, has opened more 220 million bank accounts as of February 2020.[2] More bank account ownership, however, does not equal more usage. Many accounts lie dormant, with only 39% of Indians reporting that they have made any deposits or withdrawals in the last 12 months (Demirguc-Kunt et al. 2018).

These problems are particularly pronounced in rural and semi-urban areas, where there are fewer bank branches. Banks might not have incentives to set up the necessary infrastructure because of the high fixed costs and potentially low marginal returns to serving the poor (Burgess and Pande 2005). To address this problem, various countries like India have experimented with "Cash-in, Cash-out" banking agents. These agents are generally local microentrepreneurs such as shopkeepers, who provide bank account openings, deposits, withdrawals, and other core banking services (Mehrotra et al. 2018). In many developing countries, they are the frontline workers of financial inclusion.

However, several factors may limit the extent to which the presence of these agents promotes financial inclusion in target areas, as highlighted in figure 1 at the end of this section. First, behavioral barriers may prevent people from saving as much as they could and saving at the agent point. Our reading of the behavioral sciences literature, as well as our human-centered design sessions and qualitative work, inform this hypothesis. People often do not make concrete savings goals or plan for how to meet such goals, as they lack the salience of present concerns (Fiorillo et al. 2014). Moreover, even if customers do save, they may not make deposits in their accounts. They may believe that only large sums of money need to be deposited in a formal bank account. Plus, some customers open accounts only to receive government benefits or loans, rather than save, and lack sufficient awareness about the benefis of banking services (Johnson and Meka 2010). For example, they may not know the cost and security benefits of formal saving, instead relying on informal options such as keeping cash at home, or with informal saving groups. If these behavioral barriers can be mitigated, customers may start saving more, and they may also start depositing their savings in their bank accounts.

[1] See for instance https://www.cgap.org/blog/india-moves-toward-universal-financial-inclusion

[2] The details of the PMJDY program along with the latest numbers can be found here: https://www.pmjdy.gov.in/account

Second, even if customers want to save money in bank accounts, they may not have sufficient trust and confidence in the banking agents, so the increased desire may not translate into deposits. This might happen because agents often come from different socioeconomic groups as the low-income customers; in the social context of rural India, lack of trust between socioeconomic groups can often undermine the customer-agent relationship (Anderson, Francois, and Kotwal 2015). Promoting a positive relationship between the customers and the agents may mitigate the effects of these factors.

Third, like frontline workers in low-income countries, banking agents may not have sufficient incentives to provide high-quality services. After all, banking agents are not bank employees: even though they are always associated with a specific bank, they are hired, trained, and managed by banking agent network companies. Given the scale of these companies' operations, principal-agent problems may emerge – both among the agents and their managers. A well-established body of work suggests that employees show more commitment to an employer that shows commitment to them (for example Eisenberger et al. 1986, 2001; Rhoades and Eisenberger 2002; Cropanzano and Mitchell 2005). Without such commitment from an employer, agents might lack intrinsic motivation and could even engage in fraud (Mudiri 2013). These facts are also in line with more recent work in development economics suggesting that frontline workers care about non-monetary incentives, such as social recognition (Ashraf, Field, and Lee 2014; Rasul and Rogger 2018). Thus, if network companies provide agents more resources and recognition, agents may perform better.

Our project tests a composite intervention to address these three barriers, with the aim of increasing formal savings. We partner with BASIX Sub-K, a large banking agent network company managing over 8,000 banking agents in India. We implement a randomized control trial (RCT) with 794 banking agents across 36 districts in the Indian states of Telangana and Andhra Pradesh, with 402 assigned to treatment and 392 to control. The agents who comprise the treatment group receive calendars that encourage regular savings and distribute them to their customers. These agents are also trained to explain the calendar to their customers, to encourage customer savings, and to track the distribution of the calendars. Finally, they receive SMS messages that remind them about the content of the training. Managers who supervise these banking agents also receive this training and similar SMS reminders. After 7 weeks (when the national lockdown in India began in response to COVID-19) and after three months, we will evaluate the impact of this intervention on monthly deposits, net monthly deposits (deposits minus withdrawals), and transactions at the level of the banking agents. In the calendar, we incorporate design elements similar to the ones described in (Fiorillo et al. 2014). Treatment agents with as much or more customers than the median number of customers per agent received 240 calendars, while treatment agents with fewer customers than the median received 80 calendars. In the aggregate, we distributed a total of 64,000 calendars to the banking agents to further distribute to their customers. Agents were asked to distribute at least half the calendars to women customers. Below we describe each component of the intervention in more detail.

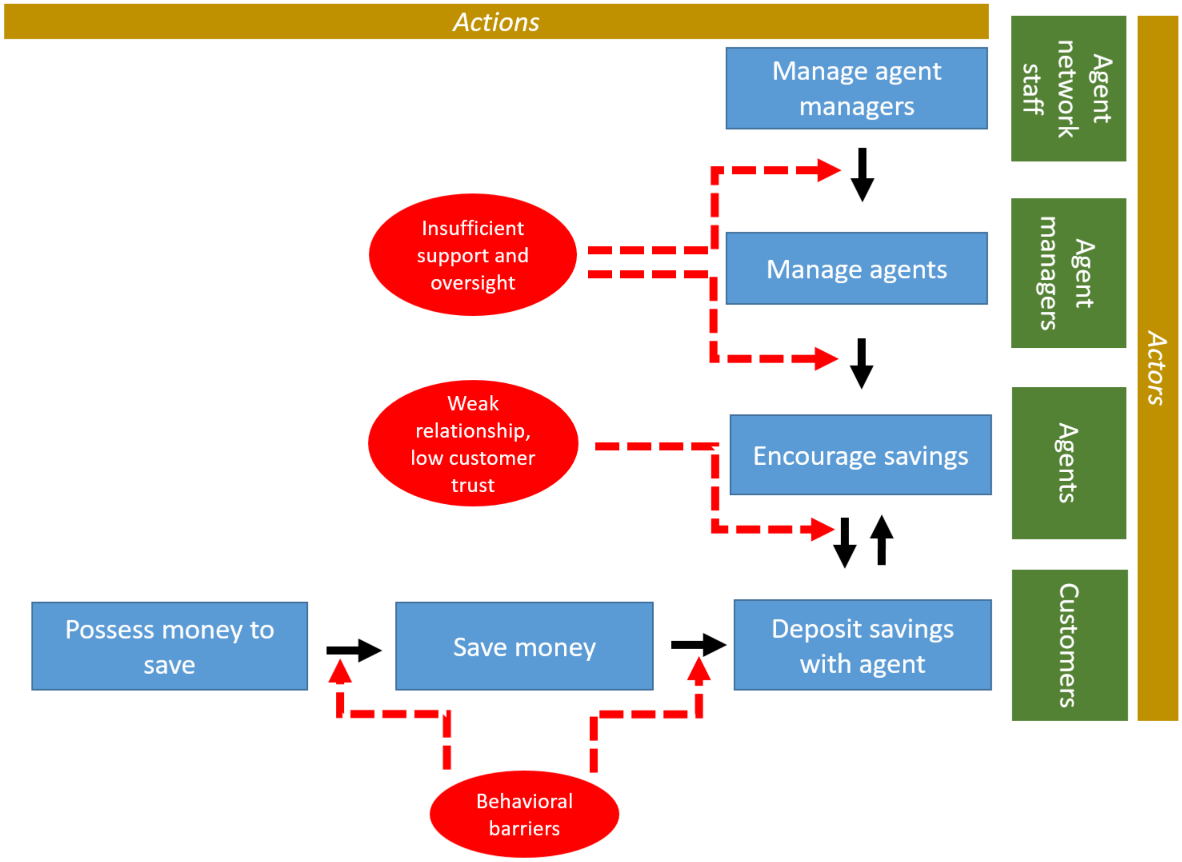

Figure 1: Three causes of low formal saving with agents

1.3 Intervention Description

Savings Calendars

Savings calendars are a potential avenue to address behavioral barriers to savings (Fiorillo et al. 2014) while also increasing the volume and value of transactions that the customers do with the banking agent. This tool can create a sense of ownership by specifying a personal savings goal (Haynes et al. 2012); allow customers to create a concrete plan and keep track of their daily, weekly, and monthly savings amounts ((Rogers and Bazerman 2008); and improve the relationship with the banking agent who provided the calendar as a 'gift' (Haisley and Loewenstein 2011).

Against this background, distributing free calendars to customers could be understood as business-to-customer gift giving. Recent literature suggests that it is precisely in this type of scenario where business gift giving generates an implicit request for reciprocity that is well-received by customers (Bodur and Grohmann 2005; Marchand et al. 2017; Lee and Yi 2019). In other words, the very act of providing a gift might strengthen the relationship between the agent and the customers, thus addressing another barrier to formal saving. Customers may be more likely to increase their transactions with the banking agent who has given them this gift. Moreover, agents may feel more valued by the agent network and therefore exert more effort.

We designed the calendars in three stages:

- First, we conducted human-centered design (HCD) workshops in Uttar Pradesh, India with more than 100 respondents in November 2018 based on the HCD kit created by IDEO (2019). These workshops presented images and design features from which respondents could choose to make their own calendars. We then asked them structured questions about their choices.
- Next, we created a mock calendar based on the most popular pictures and themes in the first stage. We piloted these calendars in Andhra Pradesh in November 2019 and collected qualitative data from a variety of respondents using semi-structured in-depth interviews.
- Based on the feedback from the qualitative work, we created a final version of the calendar in December 2019.

The final version of the calendar contains a placeholder for customers to write their name and savings goal; a space for customers to write down their daily, weekly, and monthly savings; and imagery related to the customers' most common savings goals, such as medical emergencies and saving for their children's education.

Training for agents and agent managers

In order to distribute the calendars, we conducted a total of 10 training sessions for the banking agents. Each session was attended by anywhere between 14 and 60 agents. These sessions were accompanied by Sub-K sessions on operating banking software (The control agents did not receive this additional Sub-K training). The savings calendar session focused on three components:

1. The team explained every component of the savings calendars to agents and their managers and highlighted the importance of the calendars as an avenue to build better customer relationships.
2. The team provided advice to agents and their managers on addressing the aforementioned behavioral barriers to saving that customers often face. Agents were asked to communicate to customers the importance of saving for achieving their personal goals, the security benefits of saving with the agent, and the fact that customers should deposit even small amounts into their accounts. These three recommendations were also printed on a savings encouragement sheet that was distributed to the agents.
3. The team also explained the process for monitoring the distribution of the calendars. We provided a distribution sheet to the banking agents to document which customers received calendars. Agents had to send a copy of this sheet twice a week to their managers, who would collate them for the research team.

Following the training, each agent received their set of calendars, the savings encouragement sheet, and the calendar distribution sheet. 77% of agents completed the in-person training session. Sub-K separately trained and delivered materials to agents who did not complete the training session. Moreover, in the time between randomization and training, 6 agents assigned to treatment stopped working as agents, lowering the number to be trained from 402 to 396. As of 1 April, monitoring data from 82% of these 396 agents reached the research team. Agents reported distributing 39,180 of 64,000 calendars (61%).

SMS reminders for agents and agent managers

Lastly, we are sending SMS reminders to the treatment agents and their managers. These messages remind the agents to distribute the calendars, in particular to distribute at least half their stock of calendars to women; the three recommendations for encouraging savings; and to carefully track the distribution.

During the first two months, we sent SMS reminders three times a week. During the last month, we will send these reminders twice a week. All the SMS reminders are in Telugu, the local language in both Andhra Pradesh and Telangana.

1.4 Contributions to the Literature

Our project contributes to a number of distinct literatures. First, our results will speak to the growing body of work that studies the use of nudges and other low-touch behavioral interventions to further public policy goals. Many of these focus on addressing behavioral barriers in the context of financial inclusion (For instance Karlan, Morten, and Zinman 2012; Choi et al. 2017; Aggarwal, Brailovskaya, and Robinson 2019), including ones that prevent people from increasing formal savings (such as Karlan et al. 2016; Aker et al. 2020). A key difference in our study is that we provide these calendars to banking agents for them to distribute to their own customers. Therefore, we focus our attention on the aggregate effects of a behavioral intervention that has successfully increased savings at the individual level elsewhere, and on whether this intervention translates into more deposits at the agent level.

Second, we will generate evidence on how to motivate frontline workers in developing countries to provide better services. Recent work has studied what types of interventions are effective in motivating workers in sectors such as nutrition and education (a thorough review is presented in Finan, Olken, and Pande 2017). However, evidence on what might work for the frontline workers of financial inclusion is far more limited and has focused on how to improve demand and inventory (Balasubramanian and Drake 2015; Acimovic et al. 2019). We provide this evidence, especially on plausible mechanisms for non-monetary incentives. These aforementioned mechanisms include addressing customers' behavioral barriers, improving agent-customers relations, and making agents feel more supported.

Finally, this project contributes to the nascent work around experimentation at scale. The literature highlights that smaller policy "pilots" might not capture general equilibrium effects or accurately reflect the logistical challenges that come with larger implementations (Muralidharan and Niehaus 2017). In fact, a recent scale up of a famous policy intervention underscores this point. While the initial results in Bryan, Chowdhury, and Mobarak (2014) showed large gains for encouraging seasonal migration in Bangladesh, a subsequent scale up exercise found less promising results.[3] Additionally, capturing aggregate effects might be of particular importance in the context of financial inclusion. Ogden (2019) points out that while there is a link between financial sector development and economic

[3] The following link contains a more detailed explanation of why "No Lean Season," the NGO that tried to scale up the first pilot, did not work as intended: https://www.evidenceaction.org/why-test-at-scale-no-lean-season/

growth, the evidence on the effects of formal financial products on individuals and households is mixed, perhaps because household- and individual-level evaluations are often not set up to capture spillovers or aggregate effects, especially not when the treatment effect might be very small. Our intervention will provide evidence on the aggregate effects of a low-touch, behaviorally-informed savings tool. While the sample size for the RCT is limited to 794 banking agents, this could serve as an imperfect proxy for the aggregate effects of distributing more than 60,000 calendars across two Indian states. A less ideal proxy would take into consideration all formal savings options (e.g., bank branches or other network companies' agent points).

2 Impact Evaluation Design

Below we outline the core research questions that we will answer with the RCT. We also describe exploratory research questions for which we will provide suggestive evidence through a qualitative phone survey.

2.1 Research Questions

2.1.1 Primary Research Question

Q1.1: Does a composite intervention with banking agents that consists of savings calendars for their customers, training them on the best way to encourage their customers to save, and SMS reminders, increase the average amount deposited in the agents' accounts, compared to the control group?

This question only concerns the impact of the intervention on the average amount (in INR) deposited with the banking agent. While customers may deposit money in other formal sources or withdraw money from their accounts within the same time period, this outcome variable is the closest proxy for a change in formal savings.

2.1.2 Secondary Research Questions

Q1.2: Does a composite intervention with banking agents that consists of savings calendars for their customers, training them on the best way to encourage their customers to save, and SMS reminders, increase net average amount deposited in the agents' accounts, compared to the control group?

For this question, the outcome variable is defined as deposits less withdrawals, both in INR, at the banking agent level. This question is relevant from a business standpoint. An intervention that increases deposits but not net deposits may benefit customers, but may not be cost-effective for the banking agent network company to implement at scale.

Q1.3: Does a composite intervention with banking agents that consists of savings calendars for their customers, training them on the best way to encourage their customers to save, and SMS reminders, increase the average number of transactions at the banking agent level, compared to the control group?

This question concerns general activity at the agent point. Unlike the other two indicators, this one includes fund transfers.

2.1.3 Exploratory Research Questions - Savings

Q2.1: Do customers who receive savings calendars regularly use them?

Q2.2: Do customers who receive savings calendars increase their amount of formal savings?

Q2.3: What customer characteristics are correlated with receiving a savings calendars?

Q2.4: What agent characteristics are correlated with distributing more savings calendars?

Q2.5: What agent characteristics are correlated with distributing more savings calendars to women?

2.1.4 Exploratory Research Questions - Mechanisms

Q3.1: Do savings calendars distributed through banking agents for free to customers improve agent-customer relations?

Q3.2: Do agents who receive savings calendars for customers, a company-sponsored training, and SMS reminders feel more supported by their employer?

2.2 Evaluation Methodology

To answer the primary research question, we conduct a randomized control trial (RCT) at the banking agent level.

2.2.1 Unit of Analysis

For our primary and secondary research questions, both the unit of analysis and the unit of randomization are the agents. This allows us to estimate the causal impact of the intervention on agent-level transactions. In contrast, for the exploratory research questions, we do not aim to make any causal claims or claim to be representative of the customer population. These additional research questions will supplement the primary and secondary research questions and allow us to hypothesize the mechanisms through which the intervention will or will not have worked.

2.2.2 Treatment Arms

Our study has one treatment arm receiving the composite intervention and one control group.

2.2.3 Outcomes of Interest

The main outcomes of interest for this study are **deposits**, **net deposits**, and **number of transactions**. All variables are measured at the level of the agent. In our main regression specifications, we will use the logarithmic transformation of these variables to account for outliers, which we observed in the baseline data. We will also estimate regression specifications without the logarithmic transformation as a robustness check.

Given the spread of COVID-19 in India approximately 1.5 months after the training, the planned time horizon for our outcome variables has changed. Instead of focusing primarily on the 3-month and 6-month marks, we will measure outcomes 7 weeks after training (up until the national lockdown) and 3 months after the training.

The variable names and their respective definitions are as follows:

- **deposits_cum_t:** Log of the total value of cumulative deposits in INR t weeks after January 2020. After 7 weeks, t=7 and the outcome variable is **deposits_cum_7**. After 13 weeks (3 months), t=13 and the outcome variable is **deposits_cum_13**.
- **netdeposits_cum_t:** Log of the total value of net cumulative deposits (i.e., deposits less withdrawals) t weeks after January 2020. After 7 weeks, t=7 and the outcome variable is **netdeposits_cum_7**. After 13 weeks (3 months), t=13 and the outcome variable is **netdeposits_cum_13**.
- **trxn_cum_t:** Log of the total number of transactions t weeks after January 2020. After 7 weeks, t=7 and the outcome variable is **trxn_cum_7**. After 13 weeks (3 months), t=13 and the outcome variable is **trxn_cum_13**.
- **deposits_1 through deposits_13:** Log of the total value of deposits at the agent point in INR in a given week after the intervention.
- **netdeposits_1 through netdeposits_13:** Log of the total value of net deposits at the agent point in INR in a given week after the intervention.
- **trxn_1 through trxn_13:** Log of the total number of transactions in a given week after the intervention.

2.2.4 Multiple Hypothesis Testing

To account for the fact that we are testing multiple hypotheses, we will control for the false detection rate (FDR) by using the procedure described in Benjamini, Krieger, and Yekutieli (2006) and operationalized in Anderson (2008). Where relevant, we will present the FDR-corrected q-values. These corrections will be based on testing the following hypothesis:

- What is the impact of the intervention up until the national lockdown (7 weeks after training) and three months (13 weeks) after training? We will make the correction using the two p-values for these two hypotheses.

Since we are pre-specifying the outcomes of interest before the analysis, we will not make any corrections for the number of outcome variables tested.

2.2.5 Stratification

We stratified the sample on the district in which the agent is located, the bank with which the agent is affiliated, and whether the number of customers who bank with the agent is higher or lower than the median. Agents in the study were affiliated with one of two banks. For the high or low customer indicator, for each agent we calculated the average number of unique customers per month from June to October 2019; agents with an average below the median were categorized as having a low number of customers. Agents at and above the median were categorized as having a high number of customers.

2.2.6 Agent-level Controls

We will use the following controls at the agent level:

- **Pre-intervention outcomes:** These controls are the pre-intervention values for the given outcome variable 1, 2, and 3 months before the intervention (i.e., December, November, and October). Data from January is not included as the training sessions were conducted over the last two weeks of this month.
- **Received calendars:** This control accounts for whether the agent received the calendars.
- **Demographics:** These controls include gender, age, and time spent as a Sub-K agent.

2.2.7 Fixed Effects

In addition, we include the following fixed effects:

- **Training:** These fixed effects account for the ten centralized training sessions that agents attended and the alternative of individual training.
- **Managers:** These fixed effects account for the agent's manager. Approximately 23 agents are managed by one CRO.

2.2.8 Treatment Effects

Intent to Treat

We will estimate the effect of the composite intervention using the following ANCOVA specification:

$$Y_{ist} = \beta_1 T_{is} + \alpha_1 Y^B_{is-1} + \alpha_2 Y^B_{is-2} + \alpha_3 Y^B_{is-3} + X'_{is}\delta + \lambda_s + \lambda_r + \lambda_m + \epsilon_{ist} \quad (1)$$

for agent i in stratum s t weeks after the intervention.

The variables are defined as follows:

- Y_{ist} is the main outcome variable for agent i in strata s t weeks after the intervention.
- T_{is} is a dummy variable that indicates whether agent i in stratum s received the treatment
- Y^B_{is-1}, Y^B_{is-2}, and Y^B_{is-3} indicate the baseline value of the outcome variable for agent i in stratum s 1, 2, and 3 months before the intervention, respectively.
- X_{is} is a vector of agent-level controls (demographics and receipt of calendars) for agent i in stratum s
- λ_s denotes fixed-effects for strata
- λ_r denotes fixed-effects for training
- λ_m denotes fixed-effects for manager
- ϵ_{ist} represents the error term for agent i in stratum s t weeks after the intervention.

β_1 is our main parameter of interest. It measures the percentage change in the outcome variable for treatment agents compared to control agents.

Treatment on the Treated

We will estimate the complier causal effect of the intervention for the main outcomes using 2SLS. We will use the treatment dummy as an instrument for program take-up.

2.2.9 Heterogenous Treatment Effects

In order to estimate heterogenous treatment effects, we will rely on machine learning methods as described in Athey and Imbens (2016) and Athey et al. (2019). We will use random forests to estimate treatment effect heterogeneity, to test hypotheses about the differences between the effects on the subpopulations identified by the machine learning algorithm, and to identify characteristics associated with greater impacts. Using this approach, we will analyse heterogeneity of impacts across the all of our baseline characteristics.

We will carry out the general following procedure:

- We will estimate hetereogenous treatment effects with the `causal_forest` command from the r package `grf`.
- We will use most of the default tuning paramters, except `num.trees`. We will use 2,500 trees for a reasonably low variance in our estimates.[4]

[4] We base our choice in the reference materials developed by Athey and co-authors: https://grf-labs.github.io/grf/

- Once we have a distribution of heterogenous treatment effects, we will assess heterogeneity by dividing the distribution in four quartiles and comparing the conditional average treatment effect in each quartile.
- We will report the relative variable importance as identified by the algorithm.

2.3 Sample Size

Based on logistical constraints from our implementation partner, we could not choose a sample size that exceeded 800 agents across Telangana and Andhra Pradesh. Our final sample size is 794 agents. Therefore, we conducted power calculations in order to find out the minimum detectable effect size (MDES) for this intervention. We estimated the MDES based on the number of pre- and post-intervention rounds of baseline/endline data. We are powered to detect a minimum change of 6% in average monthly deposits at the agent point.

The following table presents the parameters for our calculations:

Parameter	Description
alpha	0.05. This is the statistical convention
Power (1- beta)	0.8. This is the statistical concention
Pre-intervention rounds	3 - Based on available data
Post-intervention rounds	3 - Based on available data
Correlation between rounds	0.92. Calculated based on available data
Control mean - deposits	INR 280,000. Value based on previous transactions
Standard Deviation	INR 390,000. Value based on previous transactions
Treatment/control Ratio	One - Optimal sample size ratio
Total sample size	794 No of agents in AP and Telangana
Minimum detectable effect size	6%. MDES that emerges from a sample size of 794 agents

2.4 Sampling and Randomization

As we mentioned in previous sections, we will estimate the intent-to-treat (ITT) and the treatment-on-the-treated(ToT) effects of our intervention on deposits, net deposits, and number of transactions at the agent point. The ITT will be our primary specification.

2.4.1 Population

The population of interest for this study is the banking agents who work for BASIX Sub-K in Telangana and Andhra Pradesh, India. All of these agents are associated with one of two public sector banks.

articles/ci_and_num.trees.html

2.4.2 Sampling Criteria

We only consider agents whose monthly average deposits are below the 99th percentile or above the 1st percentile of the agent distribution. We also do not include agents whose number of deposits were zero at baseline. Finally, we excluded agents with whom we had conducted qualitative research. This leaves us with a total sample of 794 agents that were randomized into control and treatment groups.

2.4.3 Randomization

We used the statistical software package Stata to conduct our randomization using the following procedure:

1. **Use baseline administrative data provided by Sub-K:** Prior to the randomization, we acquired administrative data capturing transactions and basic agent demographics.
2. **Stratified random assignment:** To increase statistical power, we stratify the randomization on bank, the district in which the agent operates, and whether the agent has a high or low number of customers.

We conducted the randomization after all the administrative data for the baseline had been properly cleaned.

2.4.4 Balance Test

To ensure that our randomization procedure generated groups that are balanced on observable characteristics, we conduct t-tests for the differences of means on relevant baseline values. For the state indicator, 0=Telangana and 1=Andhra Pradesh.

As the balance table shows, there are no statistically significant differences in the values of observable charactersistics between the treatment and the control groups.

Variable	Control.Mean	Control.SD	Treatment.Mean	Treatment.SD	T.test.Diff
Customers	403.19	22.55	456.73	24.12	-53.54
Deposits	357,000	23,221.05	400,000	24,412.63	-43,100
Transactions	3,080,000	212,000	3,520,000	230,000	-440,000
Gender	0.2	0.02	0.18	0.02	0.01
State	0.67	0.024	0.7	0.02	-0.03
N	392		402		

Note:
F-stat: 0.908, F-test no Obs: 794; * sig. at $p<0.1$, ** sig. at $p<0.05$, *** sig. at $p<0.01$

2.5 Data Collection and Validation

This section describes all the types and sources of data for the research project.

2.5.1 Types of Data

Administrative Data

We will rely on administrative data provided by Sub-K to estimate equation (1) and answer our primary and secondary research questions. Administrative data has a few advantages. The data is relatively high quality because it represents real transactions that occur at the agent point along with their amounts. They also provide reliable information about the geographical location of the agent, simple demographic characteristics such as age and gender, and how long the agent has been working with Sub-K. Furthermore, administrative data allows us to use up to six rounds of monthly pre-intervention outcomes to construct our baseline and helps us keep track of post-intervention agent attrition.

All the transaction data is collected as part of Sub-K's routine business operations while administrative data on the implementation of the RCT was collected jointly with the research team at IDinsight.

Monitoring Data

We will collect calendar distribution data sent from agents to agent managers to the research team. This respondent-level data includes name, gender, method of receiving calendar (agent service point or door-to-door), date of calendar receipt, phone numbers, and consent to be contacted.

Qualitative Phone Survey Data

Initially, we had planned to conduct in-person semi-structured interviews to better understand the mechanisms of the intervention. However, due to safety considerations arising from the outbreak of COVID-19, we have cancelled in-person qualitative research. We will instead conduct qualitative work through phone surveys in April 2020, two months into the intervention.

Our sample will include approximately 40 customers and 20 banking agents. As the sample of customers will be drawn from data reported by the banking agents, we do not expect the sample to be representative. However, we will select customer respondents purposefully to ensure variation in terms of gender. Similarly, we will ensure variation among agent respondents in terms of the number of unique customers (high/low) and compliance with the intervention. All respondents will vary across type of geographical location (rural/semi-urban).

We will build the questionnaire based on previous qualitative work we conducted in November 2019. The data will provide some indication of the demographic characteristics of the customers who received calendars and suggestive evidence about the intervention's impact on calendar usage, financial behavior, and agent-customer interactions.

2.5.2 Data Quality

Given that the data for our main analysis, the estimation of equation (1), will rely solely on administrative data, we do not foresee substantial problems with data quality. After all, this data is routinely used for business decisions and captures real transactions that are validated through Sub-K's infrastructure.

Our phone surveys will be conducted through our field-management staff, who have experience conducting these types of surveys.

2.6 Technical Risks

We believe that our study design will allow us to produce internally valid estimates of the effects of the intervention. Nevertheless, we have identified a few potential challenges that might bias our estimates. The following sections highlight these challenges as well as the steps we will take to overcome them.

2.6.1 Attrition

Some banking agents will drop out from the study between baseline and the time we measure the outcomes. We consider three possible cases:

- Attrition could be random, i.e. there are no systematic differences between agents who remain and agents who are dropped. This does not affect internal validity, but reduces the statistical power of the study. As of April 01, 2020, Sub-K has terminated close to 4% of the sample across both treatment and control.
- Attrition may be non-random, i.e. the attriters may differ significantly from the non-attriters, changing the composition of our sample. If this attrition is balanced across treatment and control groups, internal validity is preserved, though the generalizability of our results diminishes. According to Sub-K, the termination of agents did not have anything to do with their participation (or lack thereof) in this study. However, agents are terminated when they perform few monthly transactions. It is thus likely that attrition is non-random.
- Finally, most damaging for internal validity, we may have differential attrition between the treatment and control groups.

To test whether attriters are statistically different from non-attriters, we will compare baseline characteristics between the two groups. These characteristics include the outcome variables and all control variables used in our analysis. In case attrition is non-random, we will employ inverse probability weighting to address this, using all available information on the attriters from baseline.

We will also test whether attrition is differential between the treatment and control groups, a major risk to the internal validity of our estimates. We will conduct this analysis by regressing a binary outcome variable equal to 1 if the agent is present at baseline and endline, and 0 otherwise, on treatment status, and baseline values of our covariates. We will additionally test for differential attrition by regressing attrition on a dummy variable for treatment, all baseline covariates, and the treatment dummy interacted with baseline covariates.

In case attrition is significantly different across treatment and control groups, we will follow the approach suggested in Wooldridge (2010), which comprises the following:

- Estimate a probit specification for the probability of being present, which is a dummy variable that takes the value 1 if the agent is employed by Sub-K at both baseline and endline, and 0 otherwise. From this specification, we will use the inverse of the predicted probability as the probability weights or sample weights in all the regressions.
- We then re-estimate the equation using the other outcome variables on the left-hand side and these estimated weights.

2.6.2 Spillovers

It is highly unlikely that any of the agents in the control group received the intervention by mistake. The distribution of calendars, attendance to the training sessions, and sending the SMS reminders have all been carefully tracked by the research team. We know of only one agent in the control group who participated in half a training, but was not provided with calendars and is not receiving SMS reminders. However, it is still possible for spillovers to occur. For instance, control agents may hear about the training from other agents in the district, or they may come across a customers who received calendars. Finally, agent managers may encourage savings among control agents.

However, we believe that spillovers are not likely to pose a major threat to the study for the following reasons:

1. **Control agents are unlikely to find out about the treatment in the first place.** Agents are geographically dispersed across rural and semi-urban areas, which makes substantial inter-agent communication unlikely.
2. **Sub-K routinely conducts trainings with only a subset of agents.** This means that agents are on average used to being part of some trainings and not in others. Therefore, it is unlikely that agents would change their behavior if they found out that they weren't part of a training.
3. **The treatment does not represent a substantial monetary gain in and of itself.** Relatively recent work suggests that RCT attrition rates are differential when the treatment group receive a high incentive, like a large grant.[5] The composite intervention does not entail

[5] See for instance: https://blogs.worldbank.org/impactevaluations/attrition-rates-typically-aren-t-different-control-group-treatment-group-really-and-why

a big monetary gain for the agents, as the calendars are supposed to be distributed free of cost to their customers. Furthermore, there were some transit costs associated with agents getting to the training locations. This means the value of the treatment is not a priori "large and clear," implying that it is unlikely that agents would perform differently conditional on learning about others participating in the treatment.

4. **Training of agent managers was framed around monitoring calendar distribution and savings encouragement among treatment agents only.** Agent managers primarily played a monitoring role in the intervention, being required to follow-up with treatment agents and collect distribution data. For this reason, it is unlikely that managers changed how they engaged with control agents.

2.6.3 Evaluation-Driven Effects

Evaluation-driven effects refer to changes in the outcome variable that are related to evaluation activities independent of the actual evaluation. For example, simply asking agents questions about their relationship with their customers might have an impact on the effort they put into generating transactions. These types of effects are more likely to occur when conducting large-scale survey activities.

In our case, we think that it is unlikely for evaluation-driven effects to bias our estimates. For our main research questions, we will rely only on administrative data, which means there won't be effects stemming from survey activities. Furthermore, ours is a composite intervention, so all activities including communicating with agents about the training and following up on the calendar distribution are part of the intervention.

3 References

Acimovic, Jason, Chris Parker, David Drake, and Karthik Balasubramanian. 2019. "Show or Tell? Improving Agent Decision Making in a Tanzanian Mobile Money Field Experiment." *Working Paper*.

Aggarwal, Shilpa, Valentina Brailovskaya, and Jonathan Robinson. 2019. "Cashing in (and Out): Experimental Evidence on the Effects of Mobile Money in Malawi."

Aker, Jenny C, Melita Sawyer, Markus Goldstein, Michael O'Sullivan, and Margaret McConnell. 2020. "Just a Bit of Cushion: The Role of a Simple Savings Device in Meeting Planned and Unplanned Expenses in Rural Niger." *World Development* 128: 104772.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

Anderson, Siwan, Patrick Francois, and Ashok Kotwal. 2015. "Clientelism in Indian Villages." *American Economic Review* 105 (6): 1780–1816.

Ashraf, Nava, Erica Field, and Jean Lee. 2014. "Household Bargaining and Excess Fertility: An Experimental Study in Zambia." *American Economic Review* 104 (7): 2210–37.

Athey, Susan, and Guido Imbens. 2016. "Recursive Partitioning for Heterogeneous Causal Effects." *Proceedings of the National Academy of Sciences* 113 (27): 7353–60.

Athey, Susan, Julie Tibshirani, Stefan Wager, and others. 2019. "Generalized Random Forests." *The Annals of Statistics* 47 (2): 1148–78.

Balasubramanian, Karthik, and David Drake. 2015. "Service Quality, Inventory and Competition: An Empirical Analysis of Mobile Money Agents in Africa." *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, nos. 15-059.

Benjamini, Yoav, Abba M Krieger, and Daniel Yekutieli. 2006. "Adaptive Linear Step-up Procedures That Control the False Discovery Rate." *Biometrika* 93 (3): 491–507.

Bodur, H Onur, and Bianca Grohmann. 2005. "Consumer Responses to Gift Receipt in Business-to-Consumer Contexts." *Psychology & Marketing* 22 (5): 441–56.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica* 82 (5): 1671–1748.

Burgess, Robin, and Rohini Pande. 2005. "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment." *American Economic Review* 95 (3): 780–95.

Choi, James J, Emily Haisley, Jennifer Kurkoski, and Cade Massey. 2017. "Small Cues Change Savings Choices." *Journal of Economic Behavior & Organization* 142: 378–95.

Cropanzano, Russell, and Marie S Mitchell. 2005. "Social Exchange Theory: An Interdisciplinary

Review." *Journal of Management* 31 (6): 874–900.

Demirguc-Kunt, Asli, Leora Klapper, Dorothe Singer, Saniya Ansar, and Jake Hess. 2018. *The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution*. The World Bank.

Eisenberger, Robert, Stephen Armeli, Barbara Rexwinkel, Patrick D Lynch, and Linda Rhoades. 2001. "Reciprocation of Perceived Organizational Support." *Journal of Applied Psychology* 86 (1): 42.

Eisenberger, Robert, Robin Huntington, Steven Hutchison, and Debora Sowa. 1986. "Perceived Organizational Support." *Journal of Applied Psychology* 71 (3): 500.

Finan, Frederico, Benjamin A Olken, and Rohini Pande. 2017. "The Personnel Economics of the Developing State." In *Handbook of Economic Field Experiments*, 2:467–514. Elsevier.

Fiorillo, Alexandra, Louis Potok, Josh Wright, Julie Peachey, and Kimberly Davies. 2014. "Applying Behavioral Economics to Improve Microsavings Outcomes." In *Ideas42*.

Haisley, Emily, and George Loewenstein. 2011. "It's Not What You Get but When You Get It: The Effect of Gift Sequence on Deposit Balances and Customer Sentiment in a Commercial Bank." *Journal of Marketing Research* 48 (1): 103–15.

Haynes, Laura, Ben Goldacre, David Torgerson, and others. 2012. "Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials| Cabinet Office."

IDEO. 2019. "The Human-Centered Design Toolkit."

Johnson, Doug, and Sushmita Meka. 2010. "Access to Finance in Andhra Pradesh." *Available at SSRN 1874597*.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science* 62 (12): 3393–3411.

Karlan, Dean, Melanie Morten, and Jonathan Zinman. 2012. "A Personal Touch: Text Messaging for Loan Repayment." National Bureau of Economic Research.

Lee, Shinhyoung, and Youjae Yi. 2019. "'Retail Is Detail! Give Consumers a Gift Rather Than a Bundle': Promotion Framing and Consumer Product Returns." *Psychology & Marketing* 36 (1): 15–27.

Marchand, André, Michael Paul, Thorsten Hennig-Thurau, and Georg Puchner. 2017. "How Gifts Influence Relationships with Service Customers and Financial Outcomes for Firms." *Journal of Service Research* 20 (2): 105–19.

Mehrotra, Aakash, Akhand Tiwari, M. P. Krthick, Minansa Khanna, and Vivek Khanna. 2018. "State of the Agent Network, India 2017."

Mudiri, Joseck Luminzu. 2013. "Fraud in Mobile Financial Services." *Rapport Technique, MicroSave*, 30.

Muralidharan, Karthik, and Paul Niehaus. 2017. "Experimentation at Scale." *Journal of Economic Perspectives* 31 (4): 103–24.

Ogden, Timothy. 2019. "What Is the Impact of Investing in Financial Systems?" *CDC Insight*, no. 003.

Rasul, Imran, and Daniel Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *The Economic Journal* 128 (608): 413–46.

Rhoades, Linda, and Robert Eisenberger. 2002. "Perceived Organizational Support: A Review of the Literature." *Journal of Applied Psychology* 87 (4): 698.

Rogers, Todd, and Max H Bazerman. 2008. "Future Lock-in: Future Implementation Increases Selection of 'Should'choices." *Organizational Behavior and Human Decision Processes* 106 (1): 1–20.

Wooldridge, Jeffrey. 2010. "Censored Data, Sample Selection, and Attrition." *Econometric Analysis of Cross Section and Panel Data*, 777–852.

4 Appendix

4.1 Code Examples

The following code provides an example of how the data will be analyzed using the statistical software R

```
# List of controls
controls_list <- c(
  "age",
  "time_subk",
  ...
)
controls                  <- paste(controls_list, collapse = "+")

# Treatment
treatment_list            <- "in_treatment"

# Outcome
outcome_deposits          <- "deposits_t"

# Equations
eq_deposits               <- as.formula(paste(outcome_deposits, "~",
                                            treatment_list, "+",
                                            controls))
# Main model
reg_deposits              <- lm(eq_deposits, data = main_data)
```