The Impact of a Carry-Around Soft Commitment Device on Savings and Temptation Spending: Evidence from a Field Experiment in urban India

Janina Isabel Steinert[*,1,2] and Sebastian Vollmer[1]

[1]Department of Development Economics & Centre for Modern Indian Studies, University of Goettingen, Germany
[2]Department of Social Policy and Intervention, University of Oxford, United Kingdom

4 September 2019

Abstract

We study the impact of a portable "soft" commitment device on the financial behavior of low-income slum dwellers in urban Pune, India. The portable device may add value to existing commitment designs by activating its binding appeal precisely at the point in time when spending decisions are made. 1650 individuals will be randomly allocated to receive either a zip purse and a lockbox (treatment arm) or a lockbox only (control arm). Both groups are asked to formulate a savings goal and commit to a step-by-step individualized savings plan. We will estimate the causal impact of receiving the portable device on total savings amounts and temptation spending. Further, we will compare study arms with regards to their borrowing activity, expenditure levels, resilience to economic shocks, financial self-efficacy, and female empowerment. Findings from this study can enhance the current understanding of how commitment devices may help alleviate behavioral biases and temptations.

---

[*]corresponding author, janina.steinert@uni-goettingen.de

1   Introduction

Saving has important welfare-enhancing functions for the poor. First, saving plays an important role for the poor's cash-flow management as it can help to smooth consumption over their irregular incomes (Steinert et al., 2018; Karlan, Ratan & Zinman, 2014; Deaton, 1989). Second, savings can be used as a safety buffer to cope with unanticipated income shocks and may thus partly substitute formal insurances, which are often unavailable to the poor (Hulme et al., 2015). Third, accumulated savings can be used as investment capital for future-oriented purposes, such as furthering children's education or building up a family business (Rutherford & Arora, 2009; Collins et al., 2009).

Despite these inherent benefits of saving, empirical evidence suggests that the poor tend to under-save[1] (Karlan et al., 2014; Kast, Meier & Pomeranz, 2012). A first explanation for this lies in constrained access to formal bank accounts and restrictive and costly institutional regulations for low-income clients (Hulme et al., 2015; Brune et al., 2011). In addition, informal methods of saving such as holding money in a savings circle or at home may be considered as unattractive due to the increased risk of loss, theft or monetary depreciation due to high inflation (Avdeenko, Bohne, Frlich & Kemper, 2015; Wright & Mutesasira, 2001). Saving may also be disincentivized by social obligations to share disposable cash with family members or friends in need (Dizon, Gong & Jones, 2016; Dupas & Robinson, 2013; Ambec& Treich, 2007).

Another prominent saving constraint lies in psychological and behavioral biases. Such biases can be evidently expressed by a high prevalence of temptation expenditures. Accordingly, even among the very poor, alcohol and tobacco show up as prominent expense categories, although 44% of respondents indicate that they would like to reduce their expenditures on these items (Banerjee & Duflo, 2007). While behavioral biases can be exhibited by the rich and poor alike, they are more consequential for the poor who have fewer financial resources to absorb these (Banerjee & Mullainathan, 2010). More importantly, the poor live in contexts characterized by a high liquidity of cash holdings and poor institutional capacity. Therefore, they need to be *more* attentive than individuals in developed countries and exert a higher level of discipline, self-control, and patience (Lichand & Mani, 2016; Haushofer & Fehr, 2014; Banerjee & Mullainathan, 2010; Mullainathan & Shafir, 2009).

---

[1] Under-saving is defined as a lower level of savings than one would have in a world with perfect markets (perfect information, zero transaction costs, and perfect competition amongst financial institutions) and fully attentive, fully rational, fully consistent, etc., decision making (Karlan et al. 2014, p. 38).

To tackle the problem of temptation expenditures, several programs have introduced saving commitment tools. A commitment intervention is defined as an arrangement that fosters saving and financial self-discipline by making deviations from a savings goal or plan costly and unattractive (Bryan et al., 2010). It is thereby hypothesized that commitments increase a persons self-control, willpower, and intrinsic motivation, and, linked to this, help resist impulses and temptations. Literature makes a distinction between two types of commitment devices: a) hard commitments that are either associated with institutionalized flexibility constraints or economic penalties for deviations, and b) soft commitments that are primarily associated with psychological costs through instilling feelings of guilt or failure (Bnabou & Tirole, 2004). Examples for the former include bank accounts with specific withdrawal restrictions or lockboxes, in some cases distributed without keys and can thus only be opened by an external person (e.g., a program officer) (Aggarwal et al., 2019; Aker, 2018; Francis, 2018; Herskowitz, 2018; Dupas, Keats & Robinson, 2017; Karlan & Linden, 2014; Dupas & Robinson, 2013; Ashraf et al., 2010, Ashraf, Karlan & Yin, 2006b). Soft commitment interventions typically rely on self-imposed restrictions (promises to oneself), plans, and goals or can feature peer pressure elements in which saving success is closely monitored by a peer (Soman & Cheema, 2011; Benabou & Tirole, 2004).

Building on existing literature, our trial introduces and tests an innovation to current commitment product designs. The commitment intervention consists of a portable savings device, a zip purse, that is provided in addition to a stationary savings lockbox. We thereby extend on existing designs by activating the commitment function precisely at the point in time when spending decisions are made. We hypothesize that the portable device may activate an add-on effect to the lockbox via three mechanisms:

First, the portable device is carried during the day and physically present (and possibly visible) whenever spending decisions are made. It can thus serve as a **salient saving reminder** by bringing saving motivations to the top of mind (Karlan et al., 2016). In line with this, previous studies suggest that reminders can effectively increase savings rates. For instance, Karlan and colleagues (2016) demonstrate in a series of field experiments that study participants who received reminder messages were more likely to reach their individual savings goals and held significantly higher savings amounts in their bank accounts at post-test. Similarly, another randomized controlled trial illustrates how feedback text messages informing participants about their own and their peers saving performance almost tripled weekly deposit amounts in the treatment arm (Kast, Meier & Pomeranz, 2018). While this study also shows that treatment effects disappeared once text reminders were stopped, we contend that our device could work as a more cost-effective and sustainable reminder.

Second, the portable savings device allows for the **physical segregation** of liquid cash and can thus amplify mental accounting mechanisms in the day-to-day budgeting decisions. The concept of mental accounting has been popularized by Nobel prize winner Richard Thaler (1990) who argues, based on ethnographic evidence, that money is perceived as less fungible if it is mentally earmarked for a specific purpose (Benabou & Tirole, 2004). Corroborating this argument, a person may abstain from spending money on non-essential goods if the money is explicitly "reserved" for savings purposes (Karlan & Linden, 2014; Dupas & Robinson, 2013; Shafir & Thaler, 2006). By distributing portable savings purses, we will allow participants to keep money earmarked as savings physically separate from money that can be spend, thus reinforcing and materializing these mental rules.

Third, the portable savings device creates **temporal concurrency** between any psychological commitment effect and actual spending decisions. Since the portable device is carried during the day, it is physically present when most expenditures - including potential temptation expenditures - occur. Salience of its commitment function is thus more pronounced in comparison to a lockbox that is kept at home and is thus more distant and abstract. We hypothesize that for those carrying a portable commitment device, violations of saving intentions will *instantly* induce negative emotions and feelings of guilt (Shefrin & Thaler, 1988). Hence, any violations will likely impose higher perceived psychological costs. Conversely, in a scenario with a stationary device only, the potential future regrets associated with foregone deposits into a savings box are likely discounted at the point in time at which spending decisions are made.

Overall, this field experiment aims at generating new evidence on the viability of commitment savings devices. We introduce a low-cost addition to conventional commitment products that puts specific focus on behavioral biases and temptation spending. The effectiveness of this new commitment product is then tested in a low-income population where the prevalence of temptation spending is anecdotally high.

2 Research Design

The study is designed as an individually randomized controlled trial with 1650 slum residents in the city of Pune, located in India's western-central state Maharashtra.[2] Participants will be randomly allocated on a 1:1 ratio to either the treatment arm (825 participants) or control arm (825 participants). The goal of the trial is to evaluate the impact of a portable soft commitment device on savings rates and temptation spending.

2.1 Description of the Intervention

Treatment

The proposed intervention, named *Aaj bachat kara, udya khush raha* (Marathi for "Save today, be happy tomorrow"), consists of a portable savings device (a zip purse) that is provided in addition to a stationary savings box. It is thereby hypothesized that the carry-around savings device fulfils both a reminder and earmarking function that may instantly invoke feelings of guilt and failure if money is spent on temptation goods (Soman & Cheema, 2011). The device is thus an innovation to existing commitment interventions in that its binding and quasi-coercive power takes effect precisely at the point in time when spending decisions are made.

The stationary device is a metal box secured with a padlock. This is equivalent to the devices used in a savings trial in Kenya, which reported significant increases in participants health savings amounts (Dupas & Robinson, 2013). The key to the lock will be kept by the participants themselves. By allowing for this flexibility, money retains its liquidity and can be accessed at any time, most importantly, when a household is faced with an emergency. The intervention can therefore be conceived of as a soft nudge rather than a hard and fully coercive commitment (Burke et al., 2014). In support of this, the field experiment in Kenya exemplified that program effects were not increased if keys were held by an external program officer rather than the participants themselves (Dupas & Robinson, 2013).

The savings devices will be distributed to participants during home visits by local community workers trained as program facilitators (for the home visit script, see Appendix 2). With regards to the portable devices, participants will be encouraged to carry these with them whenever they leave their homes. They will further be instructed to empty money from their portable device into their stationary device on a

---

[2] Profiles of target slums are presented in more detail in Appendix 1.

regular basis for reasons of safety.

During the same visits, participants will be asked to formulate a savings goal and commit to an individualized savings plan that outlines specific steps on how to reach their savings goal. The goal and implementation plan will be visualized on a savings sheet (see Appendix 3). This intervention component is motivated by goal setting theory that postulates a direct link between conscious goals and action (Locke & Latham, 2002; Fiorill et al., 2014; Ryan, 1970). The link is explained by four mechanisms: Firstly, goal setting helps direct attention towards a specified goal (see Karlan et al., 2014); secondly, increases effort and enthusiasm and may thus help to overcome procrastination (see Rogers, Milman, John & Norton, 2016); thirdly, it motivates perseverance (Alan, Boneva & Ertac, 2016); and lastly, may lead to discovery of new skills, knowledge, and strategies that can help to reach the goal (Locke& Latham, 2002; 1990). Saving goals are combined with a concrete implementation strategy (i.e. how much money to deposit each week), following evidence pointing to higher goal attainment when implementation intentions were formed (Duckworth, Kirby, Gollwitzer & Oettingen, 2013; Soman & Zhao, 2011; Townsend & Liu, 2011).

(Active) Control

The control group will receive only the stationary and not the portable savings device. Similar to the treatment group, the device will be delivered during home visits and participants will be asked to formulate both a savings goal and a detailed savings plan.

2.2   Identification Strategy

To establish a causal relationship between the program and changes in defined outcomes, this study uses a randomized field experiment. We will randomly assign 1650 individuals to receiving either the lockbox and portable savings device (treatment group, 825 individuals) or the lockbox only (control group, 825 individuals). Randomization will be performed in Stata and stratified by participant sex, baseline savings,[3] and baseline levels of present bias.

---

[3]Here using a quartile split of total saving amounts.

The trial is set up with an active control group. Given that all participants will receive at least one savings device, we expect the risks of performance and expectancy bias to be reduced. Blinding of enumerators and program implementers will not be feasible.

2.3   Outcomes

This trial aims at understanding the impact of the carry-around commitment device on two primary outcomes: total savings and temptation spending. We will also collect data on secondary outcomes, including both psychological aspects and more distal welfare indicators. Firstly, we will measure participants' self-efficacy, which may help elucidate the potential psychological channels underlying changes in financial behavior. Secondly, we will examine wider aspects of household welfare, including the programs impact on levels of debt, coping with income shocks, household expenditures, and women empowerment. All measures were translated and back-translated from English into Marathi; and piloted with the target population. The set of outcomes and variable constructions are further detailed below, and individual items are presented in Appendix 4.

Primary Outcomes

Total Savings:
We will measure the total savings amounts (in Rupees) that are held in the stationary savings box. In addition to respondents' self-report, enumerators will manually count the amount of cash kept in the lockbox and potentially in the purse. To identify potential crowd-out effects, surveys will also collect self-reported information on savings held elsewhere, for instance in a formal bank account or savings group, documenting (a) total amounts, (b) past-month deposits, and (c) past-month withdrawals.

We hypothesize that, due to the additional savings device and savings reminder, total saving amounts in the treatment group will exceed those in the control group.

Total Temptation Spending:
We introduce a new measure of temptation spending by capturing past as well as desired future consumption of nine selected food and non-food items (e.g., alcohol, tobacco, gambling,...) (see Appendix 4). Temptation goods are said to differ from essential goods in that they provide utility when consumed, but not in anticipation of their consumption (Banerjee & Mullainathan, 2010; Benabou & Tirole, 2004). Based on this standard definition, we classify each of the selected items only as a tempta-

tion good if the reported amount for past expenses exceeds the desired future amount. For each respondent, these divergences will then be added up into a total amount of past-month temptation expenditures. This approach allows us to define a unique set of temptation goods for each individual respondent, without reliance on *a priori,* researcher-defined temptation categories.

We will further capture participants' self-rated propensity for temptation spending via three items specified in Appendix 4 (e.g., *I bought something and later regret that I did*). The three items will be aggregated into a temptation index based on principal-component analysis (PCA).[4]

We hypothesize that the portable savings device will decrease temptation spending more effectively than the lockbox only due to its physical presence during spending decisions and, linked to this, its earmarking and reminder function.

Secondary Outcomes

Self Efficacy:
Items for the psychological concept of self-efficacy were drawn from the *Internality, Powerful Others and Chance (IPC)* scale (Levenson, 1981) as well as from financial self-efficacy scales used previously in Steinert et al. (2018) and Lown (2011). Individual items will be aggregated into an index based on PCA.

Previous research suggests that the mere presence of self-defined goals and implementation plans can instill feelings of self-efficacy and control (Morisano et al., 2010). We therefore hypothesize that our intervention will have a direct impact on participants' perceived self-efficacy. Based on social cognitive theory, we further assume that self-efficacy ia crucial connectiong factor between saving and spending *intentions* and actual financial *behavior* (Fishbein & Yzer, 2003; Bandura, 1986, 1977). In line with this, recent empirical evidence has highlighted a direct link between a greater locus of control and improved savings outcomes (Avdeenko et al., 2015; Chatterjee et al., 2011). Building on this body of literature, we will therefore also test whether self-efficacy significantly mediates the program's impact on the primary outcomes of saving and temptation spending.

---

[4]More detail on the weighting procedure used is provided in Appendix 7.

Total Debts:

In addition to total savings amounts, we will also capture outstanding debts. The savings literature suggests that accumulated savings can be used for investment purposes and may thus be viewed as a possible substitute for loans. We hypothesize that - as an indirect consequence of increased savings - the intervention will also significantly reduce respondents levels of debt.

Resilience to Emergenies:

Further to this, our survey will measure respondents' capacity to cope with potential health or other emergencies. This will be determined based on whether participants have managed to cover the costs for potential medical treatment and medicine in the past month. We will further capture whether respondents have experienced some sort of income shock in the past six months and if so, how difficult it was for them to cope with this shock (see Appendix 4). Scholars have repeatedly argued that savings can be used as a quasi-insurance in settings where access to formal insurance markets is constrained. We hypothesize that - as an indirect consequence of increased savings - the intervention will also significantly improve participants resilience to income shocks.

Past-month Household Expenditures:

To assess the potential downstream impact of saving, we will also consider a more distal indicator of household welfare, namely past-month food and non-food expenditures. In line with previous studies (Dupas et al, 2017; Brune et al., 2015; Dupas & Robinson, 2013), we expect that higher savings rates may affect poverty-related outcomes through more effective protection from economic shocks or returns from business or human capital investments realized through accumulated savings.

Female Empowerment:

Previous financial inclusion literature has routinely assessed program impacts on female empowerment (e.g., Duvendack, Palmer-Jones & Vaessen, 2014). In line with this, we will measure female empowerment (for the female sample only) based on a PCA-weighted index composed of seven individual items (see Appendix 4). Items were drawn from Glennerster, Walsh and Diaz-Martin (2018) and further adapted to match the context of India.

Our intervention is targeted at couples (unless the housheold is single-headed), thus trying to both endorse collaboration and to actively promote female involvement in financial decision-making processes of a household. We hypothesize that the portable savings devices will provide participating women with more control over their own financial resources and thus also positively impact women's empowerment.

9

2.4 Qualitative Pilot Study

Prior to the study baseline, two focus group discussions (one consisting of twelve women and one consisting of eight men) were conducted in two slum locations in Pune. Focus group discussions served the two key purposes of 1) gaining a more nuanced understanding of the prevalence and nature of temptation spending and (under)saving in the targeted slum population, and 2) assessing the acceptability and appeal of the proposed commitment devices (lockbox and portable device).

Evidence from these discussions highlighted the following three key points:

1. *Target group*: Discussants in both groups voiced concerns about targeting individuals with very minimal or no incomes, arguing that these would likely lack sufficient financial slack to engage in both temptation spending and saving. There was unanimous agreement that the intervention should specifically be targeted at salaried workers such as rickshaw drivers, market vendors, cleaners or construction workers.

2. *Design of the portable devices:* Aternative designs of portable savings devices were presented to participants. The majority favored a zip purse made from traditional fabrics, arguing that these could be easily fixed or pinned to a belt or sari and thus protected from loss (see Appendix 5).

3. *Potentially exposing women to risk:* Participants in the female-only focus group pointed out that the proposed intervention could create possible conflict between female recipients of the savings devices and their male partner or spouse. This might particularly be the case if - as a consequence of the intervention - a female participant sought more involvement in household financial decision-making, tried to hide the lockbox and purse from her partner, or challenged her partner's engagement in temptation spending.

Based on the above findings as well as theoretical considerations, we defined the following criteria with regards to the targeting and delivery strategy of the intervention. Firstly, we added an eligibility threshold based on participants economic situation. That is, individuals were only considered eligible if they had a regular income, either from salaried work, remittances, or social security payments. Regular income was thereby defined as receiving either monthly payments or some payout at least four times per month. Secondly, to protect women from potential harm, the study team decided to deliver the intervention explicitly to couples rather than individuals.

While we will only conduct interviews with the main income earner of a household, the lockbox will be handed over to both adult household members jointly, both will receive their own key, and both will be encouraged to deposit their savings into the lockbox. In the treatment arm, both the male and female household head will receive a zip purse of their own.

While this measure may not fully eliminate the risk of conflict and disagreement over financial management decisions, it removes a certain imbalance that may have resulted from giving both information and device to only one of the partners. To rule out any potential iatrogenic effects, we will also add a survey item on potential conflict in participants' homes. This will allow us to assess whether the frequency of monetary conflict increased after delivery of the intervention.

2.5 Data

Eligibility

Our sample will be composed of male and female slum dwellers who are 18 years and older and indicate having some income - either through permanent employment, casual work or governmental cash transfers - at least once per week or on a monthly basis. While one of the household heads will be selected for the interview, both will receive the intervention and devices.

Sample Size

For the power calculations of this trial, the adequate minimum detectable effect size was determined based on related previous trials that were implemented in India and assessing the impact of some variant of a soft commitment intervention on savings. Accordingly, in one trial, Soman and Cheema (2011) tested an intervention with 146 day laborers in Indian slums in which participants were asked to label their savings for a specific purpose, such as children's education ("earmarking effect") and to physically segregate their money into differently labelled envelopes ("partitioning effect"). The intervention resulted in a standardized mean difference ranging from a Cohen's δ effect size of 1.39-2.52 for the "partitioning effect", and of $\delta = 0.61$-1.60 for the "earmarking effect". We conservatively stick to the lower-bound effect size of 0.61 for our sample size determination. A second study by Breza and Chandrasekhar (2018) implemented a social commitment intervention across 60 villages in rural In-

dia. The intervention was set up so that the savings progress of treatment group participants was closely monitored by a matched person from their village, thus likely to (i) increase the salience of saving through putting a person's social reputation at stake and (ii) create temporal proximity between a commitment effect and savings decision by communicating savings progress to monitors on a bi-weekly basis. The study documented significantly higher savings among monitored particiants, with a standardized effect size of δ=0.18. In a last related study, Somville and Vandewalle (2018) opened savings bank accounts for 442 villagers in east-central India. Participants were then randomly assigned to either receive their salary in cash or directly paid into their bank account (without any withdrawal restrictions), thus activating a physical separation effect. The authors reported significantly higher total savings in the bank payment group after five months, yielding an effect size of δ=0.33. In the following, we use the average of the three presented effect sizes (δ=0.37) to proceed with our power calculations.

Our power calculations account for the testing of seven hypotheses (i.e. seven primary and secondary outcomes) by using conservative Bonferroni corrections, reducing the alpha level from 0.05 to 0.007. Further, since power analyses are conventionally based on trial main effects, RCTs are often not sufficiently powered to detect heterogeneity in treatment effects (Karlan & Appel, 2016). Against this backdrop, we aim for a target sample size that will allow us to carry out a range of subgroup analyses (see Section 3.4 below). We utilized PowerUp![5] software to determine the required sample size for detecting treatment-by-moderator effects in this trial (Dong et al., 2018; Dong & Maynard, 2013; Spybrook et al., 2016). Accordingly, with a minimum detectable effect size of 0.37 and an alpha level of 0.007, we would still ensure 80 percent power with a sample size of 1500 participants. We anticipate an attrition rate of 10% and will therefore oversample by 150 participants, thus yielding a final target sample size of 1650.

Data Collection and Processing

Data will be collected at baseline (completed: November 2018 - January 2019) and endline (August - October 2019), six months after delivery of the intervention. Contingent on project funding and endline results, we will also collect follow-up data from twelve months after the intervention delivery (tentatively scheduled for February-

---

[5]See https://www.causalevaluation.org/power-analysis.html. Note that while this power calculation tool is set up for cluster RCTs, sample size for individual RCTs can be determined by setting ICC to 1, level-1 variance proportion to 0, n to 1 and J to the trial sample size.

April 2020). Data collection will be considered complete as soon as the target sample size is reached.

Once recruited and consented into the study, interviews with study participants will be conducted via standardized questionnaires administered on mobile computers (tablets). Questionnaires will be programmed using OpenDataKit, an open-source, freely available data collection software. Computer-assisted data collection was opted for in order to a) improve data quality by programming built-in skip-patterns, reminders, and consistency checks to prevent item non-response or selection errors, and b) reduce respondent fatigue through programming visually appealing questionnaires including vignettes and pictures. Surveys will be made available both in English and Marathi and each question was translated and back-translated and piloted for cultural adequacy. Individual interviews will last on average 45 to 60 minutes.

Collected data will be checked for interview length, plausibility, and consistency on a weekly basis. If reported answers appear implausible or odd, the acting project manager and field supervisor will follow up with the corresponding enumerator and see whether any corrections ought to be made. This procedure corresponds to evidence from an experimental trial finding that the most accurate data was produced by researchers with extensive field experience and when edits to the survey data were informed by fieldworkers (Sana & Weinreb, 2008). Survey data will be backed up on a daily basis and stored on a password-protected online cloud in addition to the ODK server. All data will be fully anonymized for analysis. No identifiable information related to individuals will be reported in publications and presentations.

3 Empirical Analysis

3.1 Randomization Verification

To establish experimental integrity, we will compare the treatment group to the control group on a range of socioeconomic characteristics as well as key outcome variables measured at baseline. The following variables will be assessed for balance:

Socioeconomic Characteristics:

1. Participant Age

2. Participant Sex

3. Educational Level

4. Employment

5. Income (including social security payments) [6]

6. Household Size

Financial Behavior:

1. Past-month Savings (see Appendix 4)

2. Temptation expenditures (see Appendix 4)

3. Loan activity (see Appendix 4)

4. Coping with health shocks (see Appendix 4)

5. Female empowerment (see Appendix 4)

6. Time preference (see Appendix 6)

7. Financial literacy (see Appendix 6)

3.2 Differential Attrition

We assess the potential threat from attrition using three approaches. First, we test whether the magnitude of attrition is different for treatment and control households:

$$attrit_i = \alpha + \beta T_i + \omega_{ij} \tag{1}$$

Second, we assess whether attrition households differ on a comprehensive set of baseline characteristics:

$$y_i = \alpha + \beta \times attrit_i + \omega_{ij} \tag{2}$$

Third, we examine whether the baseline characteristics of attrition households in the treatment group are significantly different from the control group, restricting the sample to attriting respondents only:

$$(y_i | attrit = 1) = \alpha + \beta T_i + \omega_{ij} \tag{3}$$

If there are any concerns with regards to differential attrition, we will employ Lee bounds (see Behaghel et al., 2009).

---

[6]Examples include: Old Age Pension, Disability Pension, Mahatma Gandhi National Rural Employment Guarantee Scheme.

3.3 Estimation of Treatment Effects

We will estimate the average effect of being assigned to the treatment group, the intent-to-treat effect (ITT), on each outcome variable Y by means of the following regression:

$$Y_i = \alpha + \beta T_i + \gamma Y_{i(t-1)} + \delta S_i' + \epsilon X_i' + \omega_i \qquad (4)$$

where T_i is an indicator variable for treatment assignment, equal to 1 if individual i has been assigned to receive the lockbox and the mobile savings device, $Y_{i(t-1)}$ the lagged outcome (at baseline), S_i' is a vector of stratification variables, X_i' is a vector of baseline covariates, and ω_{ij} is an error term. For each outcome, we will estimate three different specifications of the above regression: (1) a fist specification using only the treatment assignment and stratifying variable as predictors, (2) a second specification including the lagged outcome $Y_{i(t-1)}$ as additional control, and (3) a third specification including additional baseline controls X' (participant age, marital status, educational level, employment, household size, baseline poverty level). By conditioning on the baseline level of outcomes and additional controls in the AN-COVA specifications (2) and (3), statistical power and precision of estimates can be increased (Bruhn & McKenzie, 2009).

3.4 Heterogeneous Effects

We will test whether the impact of the program varies with some pre-determined individual characteristics. We will explore heterogeneity in treatment effects using the following specification:

$$Y_i = \alpha + \beta T_i + \theta TRAIT_i' \times T_i + \gamma Y_{i(t-1)} + \delta S_i' + \epsilon X_i' + \omega_{ij} \qquad (5)$$

where $TRAIT_i$ is a vector of baseline characteristics for which we assume heterogeneity in the effectiveness of the treatment (note that each individual trait is also included in the vector Xi). The average treatment effect for the subgroup of people holding a respective trait is given by the sum of the coefficients $\beta + \theta$ for that trait. All tests considered here will be two-sided.

Heterogeneous effects will be explored along the following four dimensions:

1. *Participant Sex*
   Based on findings from a recent meta-analysis of saving promotion interventions revealing significantly reduced program effects for female participants (Steinert et al., 2018b), we expect to observe smaller effect sizes for female respondents.

2. *Female Involvement in Household Financial Decision-Making*
   Following the above, reduced program impact for female beneficiaries may stem from their constrained involvement and limited bargaining power with regards to household financial management (Schaner, 2015; De Mel, McKenzie & Woodruff, 2009; Ashraf, 2009). We will therefore also analyze whether program effectiveness is lower for female participants who indicate being not or very little involved in financial decision making.[7] *(Further information on the measurement approach is provided in Appendix 6).*

3. *Income Levels*
   We will examine whether program effectiveness varies with different income levels. We hypothesize that participant with higher earnings might also be more likely to engage in temptation spending and could thus benefit more from a commitment intervention.

4. *Present Bias*
   Present bias refers to the the tendency to overweigh the present relative to the future and attribute more utility to immediate rewards. Present-biased time preferences are a key explanatory variable for temptation spending and under-saving (World Development Report, 2015; Bernheim, Ray & Yeltekin, 2015; Karlan et al., 2014). Our intervention intends to reduce the time lag between present utility gains that are associated with indulging in any temptations and future utility losses in the form of guilt and shame. We expect participants with high baseline levels of present bias to have higher demand for commitment products (Guerty, 2007; Ashraf et al., 2006a) and thus benefit most from the proposed intervention. *(See Appendix 6).*

3.5 Multiple Hypothesis Testing

We will account for multiple hypothesis testing by using False Discovery Rate (FDR) adjusted q-values (see Banerjee et al., 2015; Anderson, 2008; Benjamini et al., 2006; Benjamini & Hochberg, 1995). We utilize the Benjamini-Hochberg method as it is considered as less conservative than the Bonferroni adjustment, particularly when working with a range of outcomes that are likely correlated. For each tested outcome as well as tested effect size moderators (i.e. the *treatment x trait* interactions), we will report both unadjusted p-values as well as q-values corrected for multiple testing.

---

[7]Following Bernard et al. (2018), we will determine whether decision-making processes in a household follow a dictator model in which the male partner or spouse makes all (financial) decisions.

3.6 Treatment-on-the-Treated Effects

We will additionally estimate the average treatment-on-the-treated (TOT) program effect using an instrumental variable approach. Specifically, we instrument self-reported usage of the portable savings device with being assigned to the treatment. The TOT estimate is given by:

$$A_i = a + bT_i + cY_{i(t-1)} + dS'_i + eX'_i + w_i \qquad (6)$$

$$Y_i = \alpha + \beta A_i + \gamma Y_{i(t-1)} + \delta S'_i + \epsilon X'_i + \omega_i \qquad (7)$$

whereby A_i i is an indicator for whether individual i has made at least three deposits into the mobile savings device over the observation period.

3.7 Understanding (Barriers to) Usage of Savings Devices

We will also collect detailed usage data for both the savings box and the portable savings device (see Appendix 8). This will include self-reported usage of each device, number of household members who have access to each device, frequency of deposits and withdrawals from both devices, reported purposes of any withdrawals from each device, and lastly the self-reported willingness to pay for each device. This contextual information will help us gain a more nuanced understanding of the uptake and usage frequency of each device or to unpack potential null effects.

3.8 Dealing with Missing Data

In this study, data can be missing for the following three reasons:

- Missing data can occur if participants were interviewed at baseline but lost to follow-up surveying at post-test. This number is not expected to exceed 10% of all study participants.

- Item non-response can occur if enumerators mistakenly skip a question or if respondents refuse to provide an answer to a particular question. Based on initial baseline data checks, this number is not expected to exceed 0.5% of all study participants.

- Data is lost due to technical issues. This number is not expected to exceed 0.5% of all study participants (for collected baseline surveys 3/1677 interviews were not submitted to the online server and accidently deleted from the tablet).

In case missing data exceeds 5% of all study participants, we will include robustness checks based on multiple imputation (Rubin, 1987). More specifically, we will utilize multivariate imputation by chained equations (MICE). The procedure involves repeated estimation of missing values, thus creating multiple complete datasets. The process is iterative, and utilizes Markov Chain Monte Carlo (MCMC) techniques. MICE assumes that data is missing at random and that the likelihood of a missing value is only a function of observed characteristics. Therefore, it is crucial to establish which variables are likely predictive of the missing values and should therefore be included in the imputation model. Our imputation model is set up so as to include all variables that are used in the main outcome analysis as well as additional auxiliary variables (e.g. socioeconomic characteristics) that can yield improved imputation (Azur et al., 2011).

3.9   Dealing with Potential Outliers

Collected data will be checked for plausibility and consistency on a weekly basis in order to capture and rectify any potential data entry errors. This procedure will help reduce the number of implausible outliers.

For the outcomes of total savings, debts, and household expenditures we will include three types of robustness checks. First, variables will be windzorized at the top/ bottom 1 percent. Second, inverse hyperbolic sine transformations will be applied to the respective outcome variables (Burbridge et al., 1988). Third, quantile regressions will be estimated. If the pattern of results remains similar to the main outcome analyses, we can rule out that treatment effects are driven by only few influential observations.

4   Challenges and Limitations

A first possible limitation is highlighted by anecdotal evidence gathered during participant interviews. In several instances, female interviewees noted that their male spouse or partner frequently engages in temptation spending, with alcohol showing up as the most prominent expense categories. Linked to this, female participants voiced concerns about the possible effectiveness of the program if keys are also handed out to male household members and their partner thus receives access to the savings box. However, the study team deemed that withholding information and access rights from male household members could pose a threat to womens safety and physical integrity (in case male partners were to realize that information was deliberately concealed).

A second potential challenge for this trial is associated with the proximity and social connectedness between participating households. Therefore, it is possible that participants assigned to the control group might learn about the distributed purses from their neighbors or other community members. While we do not anticipate that purses will be passed on to another person and thus do not see any spillover threat, it is possible that control group participants may feel disadvantaged and "left out". This could potentially hamper their willingness to participate in the study's endline survey. If we were to encounter this issue during the endline phase, options of compensating control group participants in another way, for instance by giving out food vouchers or snacks, will be explored.

Another challenge is linked to aspects of implementation fidelity (Durlak & DuPre, 2008). That is, program effectiveness may partly depend on facilitator performance during the distribution of savings devices. For instance, we may expect that participants likely ascribe a higher value to the received devices if their use is enthusiastically and convincingly advocated by the facilitator. In response to this, our research team has made efforts to standardize the intervention delivery protocol, thus aiming to reduce heterogeneity in the program's implementation quality. To this end, a three-day training was carried out during which facilitators had to memorize and practice a standardized home visit script. In addition, throughout the entire implementation period, two field supervisors conducted unannounced sit-ins during home visits in order to assess whether facilitators adhered to the pre-defined protocol.

The study's data collection schedule has been carefully planned according to the Marathi/Hindi festive calendar. Our endline data collection will have to be terminated before mid-October 2019 in order to avoid any overlap with the region's most prominent festival, Diwali. Yet, in consequence, most of this data collection phase will thus coincide with the country's rainy season, which is characterized by heavy monsoons. It is possible that some of the study areas will therefore be hard to access. More importantly, some participants might be forced to relocate temporarily due to occasional flooding of slum areas.

References

Aggarwal, S., Brailovskaya, V., & Robinson, J. (2019). Saving for Multiple Financial Needs: Evidence from Malawi. Unpublished.

Aker, J.C., Sawyer, M., Goldstein, M., OSullivan, M. & McConnell, M. (2018). Just a Bit of Cushion: The Role of a Simple Savings Device in Meeting Planned and Unplanned Expenses in Rural Niger. Unpublished.

Alan, S., Boneva, T., & Ertac, S. (2016). Ever Failed, Try Again, Succeed Better: Results from a Randomised Educational Intervention on Grit. HCEO Working Paper.

Ashraf, N., Gons, N., Karlan, D. S., & Yin, W. (2003). A review of commitment savings products in developing countries. New Haven: Innovation for Poverty Action.

Ashraf, N., Karlan, D., & Yin, W. (2010). Female Empowerment: Impact of a Commitment Savings Product in the Philippines. World Development, 38(3), 333344.

Ashraf, N., Karlan, D., & Yin, W. (2006a). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. The Quarterly Journal of Economics, 121(2), 635672.

Ashraf, Nava, Dean Karlan, and Wesley Yin. (2006b). Household Decision Making and Savings Impacts: Further Evidence from a Commitment Savings Product in the Philippines. Economic Growth Center Discussion Paper No. 939.

Avdeenko, A., Bohne, A., Froelich, M., & Kemper, N. (2015). Overconfidence, Underconfidence, and the Use of Persuasive Messages in the Attainment of Savings Goals. Retrieved from https://ideas.repec.org/p/zbw/vfsc15/112982.html (last accessed 27/07/2018).

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple Imputation by Chained Equations: What is it and how does it work? International Journal of Methods in Psychiatric Research, 20(1), 4049.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. Psychological Review, 84(2), 191215.

Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. The Journal of Economic Perspectives: A Journal of the American Economic Association, 21(1), 141167.

Banerjee, A., & Mullainathan, S. (2010). The shape of temptation: Implications for the economic lives of the poor. Working Paper No.15973, National Bureau of Economic Research.

Benabou, R. & Tirole, J. (2004). Willpower and Personal Rule. Journal of Political Economy, 112(4), 84886.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society, 57(1), 289300.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive Linear Step-up Procedures That Control the False Discovery Rate. Biometrika, 93(3), 491507.

Bernard, T., Doss, C., Hidrobo, M., Hoel, J., & Kieran, C. (2018). Ask Me Why: Using Vignettes to Understand Patterns of Intrahousehold Decision Making in Rural Senegal. IFPRI Discussion Paper 01784.

Bernheim, B. D., Ray, D., & Yeltekin, S. (2015). Poverty and Self-Control. Econometrica, 83(5), 18771911.

Breza, E., & Chandrasekhar, A. G. (2019). Social Networks, Reputation, and Commitment: Evidence From a Savings Monitors Experiment. Econometrica, 87(1), 175216.

Brune, L., Gin, X., Goldberg, J., & Yang, D. (2015). Facilitating savings for agriculture: Field experimental evidence from Malawi. Economic Development and Cultural Change, 64(2), 187220.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. Annual Review Economics, 2(1), 671698.

Burke,J., Luoto, J., & Perez-Arce , F. (2014). Soft versus Hard Commitments. A Test on Savings Behaviors. WR-1055, RAND Labor & Population.

Bruhn, M., & McKenzie, D. (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. American Economic Journal: Applied Economics, 1(4), 200232.

Burbidge, J. B., Magee, L., & Robb, A. L. (1988). Alternative Transformations to Handle Extreme Values of the Dependent Variable. Journal of the American Statistical Association, 83(401), 123127.

Chatterjee, S., Finke, M., & Harness, N. (2011). The impact of self-efficacy on wealth accumulation and portfolio choice. Applied Economics Letters, 18 (7), 627631.

Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2009). Portfolios of the Poor: How the Worlds Poor Live on $2 a Day. Princeton, N.J.: Princeton University Press.

Deaton, A. (1989). Saving in Developing Countries: Theory and Review. The World Bank Economic Review, 3(1), 6196.

Dizon, F., Gong, E., & Jones, K. (2015). Mental Accounting and Mobile Banking: Can labeling an M-PESA account increase savings. Unpublished Manuscript.

Dong, N., & Maynard, R. A. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies. Journal of Research on Educational Effectiveness, 6(1), 24-67.

Dong, N., Spybrook, J., Kelcey, B., & Bulus, M. (2018). Power analyses for moderator effects with (non)random slopes in cluster randomized trials. Manuscript submitted for publication.

Duckworth, A. L., Kirby, T. A., Gollwitzer, A. & Oettingen, G. (2013). From Fantasy to Action Mental Contrasting With Implementation Intentions (MCII) Improves Academic Performance in Children. Social Psychological and Personality Science, 4(6), 745753.

Dupas, P., Karlan, D., Robinson, J., & Ubfal, D. (2017). Banking the Unbanked? Evidence from three countries. American Economic Journal: Applied Economics, 11.

Dupas, P., Keats, A., & Robinson, J. (2017). The Effect of Savings Accounts on Interpersonal Financial Relationships: Evidence from a Field Experiment in Rural Kenya. The Economic Journal.

Dupas, P., & Robinson, J. (2013). Why Dont the Poor Save More? Evidence from Health Savings Experiments. American Economic Review, 103(4), 11381171.

Durlak, J.A., & DuPre, E.P. (2008). Implementation matters: A review of research on the influence of implementation on programme outcomes and the factors affecting implementation. American Journal of Community Psychology, 41, 327350.

Duvendack, M., Palmer-Jones, R., & Vaessen, J. (2014) Meta-analysis of the impact of microcredit on womens control over household decisions: methodological issues and substantive findings, Journal of Development Effectiveness, 6(2), 7396.

Fiorill, A., Potok, L., & Wright, J. (2014). Applying Behavioural Economics to Improve Microsavings Outcomes. Ideas 24.

Fishbein, M., & Yzer, M. C. (2003). Using Theory to Design Effective Health Behaviour Interventions. Communication Theory, 13(2), 164183.

Francis, E. (2018). Paying to Repay? Experimental Evidence on Repayment Commitment. Unpublished.

Glennerster, R., & Takavarasha, K. (2013). Running Randomized Evaluations: A Practical Guide. Princeton University Press.

Glennerster, R. & Walsh, D. (2018). A Practical Guide to Measuring Womens and Girls Empowerment in Impact Evaluations. J-PAL Research Resources.

Guerty, M. K. (2007). You Cant Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya. Economic Development and Cultural Change, 55, 251282.

Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. Science, 344(6186), 862867.

Herskowitz, S. (2018). Gambling, Saving, and Lumpy Expenditures: Sports Betting in Uganda. Unpublished.

Hulme, D., Moore, K., & Barrientos, A. (2015). Assessing the insurance role of microsavings. In R. Vos, N. Islam, & M. Koparanova (Eds.), Financing for Overcoming Economic Insecurity. Bloomsbury Publishing.

John, A. (2018). When Commitment Fails Evidence from a Field Experiment. Mimeo.

Karlan, D., & Appel, J. (2016). Failing in the Field: What We Can Learn When Field Research Goes Wrong. Princeton University Press.

Karlan, D., & Linden, L. L. (2014). Loose Knots: Strong versus Weak Commitments to Save for Education in Uganda. Working Paper No. 19863, National Bureau of Economic Research.

Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2016). Getting to the Top of Mind: How Reminders Increase Saving. Management Science, 62(12), 33933411.

Karlan, D., Ratan, A. L., & Zinman, J. (2014). Savings by and for the Poor: A Research Review and agenda. Review of Income and Wealth, 60(1), 3678.

Karlan, D. & Zinman, J. (2014). Price and control elasticities of demand for savings. Journal of Development Economics, 130, 145-159.

Kast, F., Meier, S., & Pomeranz, D. (2018). Saving more in groups: Field experimental evidence from Chile. Journal of Development Economics, 133, 275294.

Levenson, H. (1981). Differentiating among Internality, Powerful Others, and Chance. In H. M. Lefcourt (Ed.), Research with the Locus of Control Construct, pp. 1563. New York: Academic Press.

Lichand, G., & Mani, A. (2016). Cognitive Droughts. Retrieved from https://econpapers. repec.org/paper/cgewacage/298.htm (last accessed 27/07/2018).

Locke, E. A.& Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. American Psychologist, 57, 705717.

Locke, E.A.,& Latham, G.P. (1990). A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice-Hall.

Lown, J.M. (2011). Development and Validation of a Financial Self-Efficacy Scale. Journal of Financial Counseling and Planning, 22(2), 54-63.

Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. Journal of Applied Psychology, 95(2), 255264.

Mullainathan, S., & Shafir, E. (2013). Scarcity: Why Having Too Little Means So Much. New York: Henry Holt & Company.

Reuben, E., Sapienza, P., & Zingales, L. (2010). Time discounting for primary and monetary rewards. Economics Letters, 106(2), 125127.

Rogers, T., Milman, K. L., John, L. K., & Norton, M. I. (2016). Making the Best Laid Plans Better: How Plan-Making Increases Follow-Through. Retrieved from $http://scholar.harvard.edu/files/todd_rogers/files.pdf$ (last accessed 27/07/2018).

Rutherford, S.,& Arora, S. S. (Eds.). (2009). The Poor and their Money: Microfinance from a Twenty-first Century Consumers Perspective. Bourton on Dunsmore, Rugby, Warwickshire England: Practical Action Publishing.

Ryan, T. A. (1970). Intentional Behavior. New York: Ronald Press.

Sana, M., & Weinreb, A. A. (2008). Insiders, Outsiders, and the Editing of Inconsistent Survey Data. Sociological Methods & Research, 36(4), 515541.

Shafir, E., & Thaler, R. H. (2006). Invest now, drink later, spend never: On the mental accounting of delayed consumption. Journal of Economic Psychology, 27(5), 694712.

Shefrin, H. M., & Thaler, R. H. (1988). The Behavioral Life-Cycle Hypothesis. Economic Inquiry, 26(4), 609643.

Soman, D., & Cheema, A. (2011). Earmarking and Partitioning: Increasing Saving by Low Income Households. Journal of Marketing Research, 48, 1422.

Soman, D., & Zhao, M. (2011). The Fewer, the Better: Number of Goals and Savings Behavior. Journal of Marketing Research, 48 (6), 944957.

Somville, V., & Vandewalle, L. (2018). Saving by default: Evidence from a field experiment in rural India. American Economic Journal: Applied Economics, 10(3), 39-66.

Spybrook, J., Kelcey, B., & Dong, N. (2016). Power for detecting treatment by moderator effects in two and three-level cluster randomized trials. Journal of Educational and Behavioral Statistics, 41(6), 605-627.

Steinert, J.I., Cluver, L. D., Meinck, F., Doubt, J., & Vollmer, S. (2018a). Household economic strengthening through financial and psychosocial programming: Evidence from a field experiment in South Africa. Journal of Development Economics, 134, 443466.

Steinert, J. I., Zenker, J., Filipiak, U., Movsisyan, A., Cluver, L. D., & Shenderovich, Y. (2018b). Do saving promotion interventions increase household savings, consumption, and investments in Sub-Saharan Africa? A systematic review and meta-analysis. World Development, 104, 238256.

Townsend, C., & Liu, W. (2011). Is Planning Good for You? The Differential Impact of Planning on Self Regulation. Rochester, NY: Social Science Research Network.

World Development Report (2015). Mind, Society, and Behaviour. Retrieved from http://www.worldbank.org/en/publication/wdr2015. (last accessed 27/07/2018).

Wright, G., & Mutesasira, L. (2001). The Relative Risks to Poor Peoples Savings. Journal of Small Enterprise Development, 12(3), 3345.

Appendix 1. Baseline Data Collection: Community Description

Data Collection Date	Community Name	Community Profile
17 December 2018	Dandekar Bridge	The community is geographically scattered, with a mixed caste population and a majority of Hindus and Buddhists.The residents belong to low-income groups, but employment is fairly regular. Both men and women work, mostly in occupations such as domestic help, peons, etc.
18 December 2018	Ambil Odha	This community is mostly comprised of scheduled caste/scheduled tribe (SC/ST) households. The majority of the population is Hindu, with a small Muslim population as well. Most houses are temporary. Women work as food vendors, while men work as laborers in coal and timber markets.
19- 20 December 2018	Panmala	Most community members are Buddhists. Women typically do household chores, while men work as drivers and helpers amongst other occupations. There is a lot of ongoing NGO work in this area, so the response was very good.
21 December 2018	Phule Wada	Most community members are Telugu and have migrated in search of work from Andhra Pradesh and Karnataka. They typically work in small industries; women often as bidi (cigarettes) makers at home due to restricted mobility and men as hospital workers or helpers in small stores. The response was low as most residents were busy with work.
22-26 December 2018	Minatai Thakare	Minati Thakare is an audyogik vasahat (industrial community). The community is quite large with around 10,000 households. Most of the residents belong to the SC/ST categories. Women typically work as house helpers with relatively regular incomes, and men as industrial labourers in small industries, with almost no fixed income.
27 December 2018	Laxmi Nagar	This community is characterized by widespread unemployment and most residents work as daily wage labourers. The composition of the community is dominated by Lamani people who typically do not speak and understand Marathi and Hindi. Most Lamani people are rural migrants in search of work. There is a history of opium use amongst this community, especially given to children during the day while the parents go in search of wage labour. Although there are 8000 homes, people were usually not at home and the response was low.
28 December 2018	Prem Nagar	Most residents belong to SC/ST categories and are Buddhists. People usually work as domestic help and hospital staff. The houses here are more permanent, and the response was enthusiastic. Alcohol is not sold anywhere in this community.
31 December 2018	Jaybhavani Nagar	Most residents belong to the Maratha community and are quite economically stable. The response was low as community members did not feel the need for savings assistance.
2-3 January 2019	Chaitraban	Community residents typically work as house helpers. The community is physically divided into two parts, one of which belongs to resource-poor households (mostly Dalits) and the other to economically better-off residents (mostly Marathas).

Data Collection Date	Community Name	Community Profile
4-7 January 2019	Wadarwadi	Most residents belong to the Wadar caste. There are two parts of the community: one consisting of Pune Municipal Corporation workers with permanent government jobs, and another consisting of daily wage laborers. Women here reported that cash savings were hard to maintain due to their partners alcoholism that drained household funds; another reported shock to cash savings was demonetisation. The percentage of unemployment is quite high and alcohol use is widespread.
8-9 January 2019	Minatai	Thakare Vasahat This community was specifically added to make the sample more religiously inclusive; most residents here are Muslims. There is physical segregation between Hindu and Muslim residential areas in this community. Most of the men work as industrial/mill workers, drivers etc. Women here reported that they are discriminated against when searching for jobs because of their religion.
10-12 January 2019	Sane Guruji Vasahat	The community is largely composed of Pune Municipal Corporation (PMC) workers with permanent government jobs, mostly Dalits. Many households are female-led, mostly by widows of deceased workers, with deaths linked to inadequate safety measures at work and highly prevalent alcoholism.
13-16 January 2019	SRA Building, Dandekar Bridge	The residents here mainly work in the Municipal Corporation. Many of these residents, however, are second generation occupants, which means those who work in PMC are allotted the houses, but their children work in other occupations as well. Thus, the economic condition is much better compared to other low-income communities in the sample. However, in households where the PMC workers are the primary breadwinners and their children do not work, the financial situation is worse.
17-21 January 2019	Rajendra Nagar	The community houses many PMC workers and their families. Residents in each building differ caste wise. In five buildings residents were exclusively Brahmin and excluded from the sample (due to high incomes). In remaining buildings, residents worked as cleaners and mostly belong to lower castes.
22-24 January 2019	Hanuman Tekdi	Most residents are daily wage laborers. There is also caste-wise segregation within the community. The Maratha residents have structurally sounder homes, while the lower castes are much worse off.
25-31 January 2019	Shastri Nagar	Most residents belong to the Maratha caste. The occupational cross-section is mixed: women work mostly as domestic helpers, while men work in general stores.
1-4 February 2019	Wadarwasti	Most residents work as daily wage laborers and migrated to Pune three decades ago in search of employment.

Data Collection Date	Community Name	Community Profile
5 February 2019	Ganesh Mala	Residents here have many different occupations: men work as helpers, and women work in domestic help. Most of the young men are uneducated and unemployed. The survey team was subjected to street sexual harassment by these young men.
6 February 2019	Dattawadi	Due to a death in the community on this particular field day, it was almost impossible to get any response from the community members. The team moved back to Wadarwasti after obtaining 3 interviews from this community.
7-8 February 2019	Gosavi Wasti	Most of the people in this community are unemployed, and the ones who are employed work in domestic help. The majority of residents belong to the ST category.

Appendix 2. Intervention Protocol

Introduction:
[Same for both arms]

Facilitator note: Make sure to involve all adult household members (male and female) (i.e. the couple heading the household). If not all of them are there, make sure to specify that this is important for all adults in the home and that information (and devices) should be passed on to those who are currently not present.

Facilitator: "Today we are here to talk to you about saving and help you make saving a bit easier for you and your family. Saving is important for many reasons. It can help you through times where income and money is very little, for instance if you suddenly lose your job, if someone in your family falls sick, or if your business is not going well. Saving is also important when it comes to big events, like Diwali celebrations or even a wedding. If you have money as savings, you can avoid borrowing lots of money and having all these extra costs for the interest you pay."

Distributing the devices:
[Treatment arm]

Facilitator note: Hand out one lockbox to the family and one key each to the male and the female household head. Also hand out a portable device to both the male and the female household head.

Facilitator: "In order to help you save money, we bring you two different savings devices today."

1. Lockbox: "This is a lockbox that you should keep somewhere at home. You should think of a place to hide this box."

2. Purse: "This is a savings device that you should always carry with you. You can fix it to your belt or pin it to your Sari or wear it around your neck. You should take this purse when you go to town. It will always remind you of your purpose to save money. It is especially important in moments where you might want to give in to your temptations. In these moments, for instance when you feel like spending money on tobacco or alcohol, the purse will be with you and you can instead put the money inside the purse, with a promise to yourself to not take it out again until you put into your saving box at home, at the end of your day."

[Control arm]

Facilitator note: Hand out one lockbox to the family and and one key each to the male and the female household head.

Facilitator: "In order to help you save money, we bring you this savings devices today. This is a lockbox that you should keep somewhere at home. You should think of a place to hide this box."

Budgeting:
[Same for both arms]

Facilitator note: Use a blank sheet of paper to draw down expense categories with the family and speak about needs and wants.

Facilitator: "Now we want to think about how you usually spend your money. Can you say what you spent money on in the past week (Facilitator: write down different categories). Now, looking at this would you say that all of these are needs? Would it be possible to cut spending on some of these things?" *(Facilitator: cross out these items)* "Can you think of times where you bought something that was a want rather than a need?"

Setting a Saving Goal:
[Same for both arms]

Facilitator note: Hand out the Savings Plan Printout

Facilitator: "This will be your personal savings plan. Let us first think of a saving goal you both have and would like to save for. Write it down or draw it."

Making a Savings Plan:
[Same for both arms]

Facilitator: "Today this goal might seem very far away. Some days you may even loose hope and think that you will never reach the goal. Therefore, it is important that you take small steps. You think from one week to the next. And you make little goals for each week. If you are successfully reaching this small goal in one week, you should always remember that this brings you closer to your big goal!"

Facilitator note: Discuss and decide on the following things and write this down in the savings plan:

- How many Rupees do you want to save each weak?
- How many Rupees do you want to save each month?
- Which are the things that you want to spend less money on?
- How will you support each other to reach your goal?

Facilitator: "This is already the first important step to reach your goal together. Put this sheet somewhere on your wall so that it reminds you every day of what you are saving for and why it is important and will help you live a happier and better life."
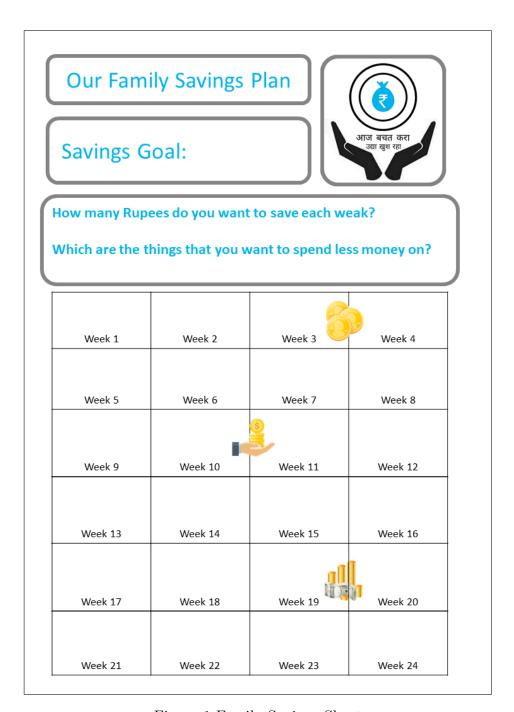
Appendix 3. Family Savings Plan

Figure 1 Family Savings Sheet

Appendix 4. Specification of Outcome Measures

Variable	Individual Items	Response Option	Aggregation
Total Savings	What is the total amount of money that you currently keep in: • your savings box? [based on self-report and hand count] • your savings account? • your accounts in post offices/national savings centres? • cash savings at home (other than those kept in the lockbox)? • cash savings with relatives or friends? • your mobile phone money account? • any informal savings association?	Total amount in rupees	Sum of total savings amounts for each saving method
Temptation Spending	In the past month, how much money did you spend on [sugar, meat, cola/lemonade, alcohol, fried snacks, cake, gambling, tobacco, toys] ? In next month, how much money would you like to spend on this item?	Total amount in rupees	For all items where: past amount > desired future amount, the difference (in rupees) is calculated and added up into a total amount of past-month temptation expenditures
	Self-rated temptation: • In the past month, I spent money on things that I didnt really need. • In the past month, I bought something and later regret that I did. • In the past month, I found it difficult to really control on how I spend my money	5-point Likert scale from never-very often	PCA-weighted index aggregating three individual items

Outcome	Individual Items	Response Option	Aggregation
Self-efficacy/LoC	- When I make plans, I am almost certain to make them work - When I get what I want, its usually because I worked hard for it - My life is controlled by other powerful people - I am confident that I will not run out of money before the next payday - I am confident that I can plan carefully in advance how to use my money during each week	Rated on a 1-10-point Likert scale, ranging from very much disagree to very much agree	PCA-weighted index aggregating five individual items
Female empowerment	- Boys should not be allowed to get more opportunities and resources for education than girls. - Boys should be fed first and given more food compared to girls. - A husband should be more educated than his wife. - Daughters should have a similar right to inherited property as sons. - It would be a good idea to elect a woman as the President of India again. - Do you get in trouble for leaving the house without informing your husband or another household member? - Do you get in trouble for making unescorted outings such as visiting your parents, friends, going to the market?	Rated on a 1-10-point Likert scale, ranging from very much disagree to very much agree Last two items are binary and answered by female respondents only	PCA-weighted index aggregating seven individual items

Outcome	Individual Items	Response Option	Aggregation
Resilience to Emergencies	- Did you experience any kind of emergency in the past six month? - If yes: How difficult was it for you and your family to find enought money to cope with that emergency? - Imagine an emergency would happen tomorrow. How difficult would it be for you and your family to find ten thousand INR to cope with this emergency? - Was there a time in the last 4 weeks when you needed to be admitted at the hospital but didnt because you didnt have enough money? - Was there a time in the last 4 weeks when you needed to buy medicine from a chemist but didnt because you didnt have enough money?	Rated as very difficult, somewhat difficult, not difficult at all	Total sum of counts for both items
Household Expenditures	How many Rupees did you spend in the last month for [rice, dal, cleaning utensils, insurances, transport....]	Amount in rupees	PCA-weighted index aggregating individual items
Total debt	Are there any outstanding loans that you have to pay back? How much money do you expect to pay for any loan in the next month?	Amount in rupees	Total sum of money owed

Appendix 5. Product Designs

Figure 2 Portable Commmitment Devices

Figure 3 Stationary Commitment Device

Appendix 6. Specification of Effect Moderators

Moderator	Individual Items	Response Option	Aggregation
Present Bias	- Would you prefer 4000 rupees now or 5000 rupees tomorrow? - Today is more important than tomorrow. - I am impatient. - I easily give in to my temptations. - It is difficult for me to avoid eating a snack food I enjoy if it is easily available, even if I am not hungry.	Today or tomorrow Rated on a 1-10-point Likert scale, ranging from very much disagree to very much agree	PCA-weighted index aggregating five individual items
Female Decision making	- Amol and Devika are married. Amol decides how to spend the money because he makes all decisions for the family. *(Dictator Vignette)* - Aarav and Rucha are married. Aarav decides how to spend the money because he is the person in the household who earns the money *(Contribution Vignette)* - Rohan and Sunita are married. Rohan decides how to spend the money because he makes these decisions while Sunita makes other decisions for the family *(Separate Spheres Vignette)*	For each vignette: Do you resemble this couple? - Completely different - Somewhat different - Somewhat similar - Completely similar	Categorical variable: dictator, contribution, separate spheres, norms, most informed

Moderator	Individual Items	Response Option	Aggregation
Female Decision making	• Navin and Shilpa are married. Navin decides how to spend the money because most men in the community make these decisions *(Norms Vignette)* • Sai and Saanvi are married. Sai decides on how to spend the money because he has the most information on different prices and sellers and knows where to get the best bargain *(Most informed Vignette)* • Are you involved in decisions about money in your home?	For each vignette: Do you resemble this couple? • Completely different • Somewhat different • Somewhat similar • Completely similar	Categorical variable: dictator, contribution, separate spheres, norms, most informed

Appendix 7. Index Weighting

We will use principal component analysis to determine item weights.

In principal component analysis, variables are expressed as the linear combination of a set of underlying components for each respondent j:

$$a_{1j} = v_{11} \times A_{1j} + v_{12} \times A_{2j} + ... + v_{1N} \times A_{Nj}$$

$$a_{Nj} = v_{N1} \times A_{1j} + v_{N2} \times A_{2j} + ... + v_{NN} \times A_{Nj} \tag{8}$$

where A_N denotes the components and v_N the coefficients on each component for each variable.

Principal component analysis is used to find the linear combination of the individual variables with maximum varianceyielding the first principal component A_{1j} and then finding a second linear combination with the maximum of the remaining variance, and so forth. The scoring factors are then retrieved by inverting the structure of Equation (8), which then produces estimates for the N principal components:

$$A_{1j} = f_{11} \times a_{1j} + f_{12} \times a_{2j} + ... + f_{1N} \times A_{Nj}$$

$$A_{Nj} = f_{N1} \times a_{1j} + f_{N2} \times a_{2j} + ... + f_{NN} \times a_{Nj} \tag{9}$$

Ultimately, the index for each respondent is given by the expression:

$$A_{aj} = f_{11} \times (a*_{1j} - a*_1)/(s*_1) + f_{1N} \times (a*_{Nj} - a*_N)/(s_N) \tag{10}$$

whereby $a*_{1j}$ to $a*_{Nj}$ represent N items for individual j, $a*_1$ the mean of $a*_{1j}$ across respondents and $s*_1$ the standard deviation.

Appendix 8. Usage Data

Outcome	Individual Items	Response Option
Device Usage	- Do you use this box/purse? - Do you currently have any savings in the box/purse? If so, how much ? *[count with respondend]* - The last time you went out of the house, for example to the market or to town, did you take your purse with you? - How often did you put money in your box/purse in the past month? - How often did you take money out of your box/purse in the past month? - How often did you put money from the purse in your savings box in the past month? - For what purpse did you take money out of your box/purse?	never in the past month, once, every second week, weekly more often than weekly, daily, more than daily
Benefits of the purse	- When I see the purse, it reminds me of the importance to save money. - When I am tempted to buy something I do not really need, the purse helps me to resist my temptations. - When I go out of the house, I usually have my purse with me. - This purse helps me to keep money for myself and not to give it to other people (my partner, children, friends) - The purse helps me to better control how I spend my money. - If this purse had not been given to you for free, how much Rupees would you have been willing to pay to have it?	Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly disagree
Explanation of Null	- When did you stop using the box/purse? - Did you give or sell the box/purse to someone else? - Why was this purse not useful to you?	I never used it from the beginning, I used it for about a month and then stopped, I used it for several months and then stopped; last item open question