TUM School of Governance TUM School of Management

Munich Papers in Political Economy
Working Paper No. 4/2020

Commitment or Concealment? Impacts and Use of a Portable Saving Device: Evidence from a Field Experiment in Urban India

Janina Isabel Steinert
Rucha Vasumati Satish
Felix Stips
Sebastian Vollmer

September 2020

Commitment or Concealment? Impacts and Use of a Portable Saving Device: Evidence from a Field Experiment in Urban India

Janina Isabel Steinert*[1,2], Rucha Vasumati Satish[3], Felix Stips[4], and Sebastian Vollmer[3]

[1]TUM School of Governance, Technical University of Munich, Germany
[2]Department of Social Policy and Intervention, University of Oxford, UK
[3]Chair of Development Economics, University of Goettingen, Germany
[4]Centre for Evaluation (CEval), Germany

Abstract

We study the impact of a portable "soft" commitment device on the financial behavior of low-income slum dwellers in Maharashtra, India. 1525 individuals were randomly allocated to receiving either a zip purse and a lockbox (treatment arm) or a lockbox only (control arm). Based on self-reported measures and hand counts of money held in the distributed saving devices, we document an 81% increase in total savings in the treatment group. We do not find significant reductions in temptation spending, thus suggesting that increases in savings were not primarily realized through improvements in self-control. Instead, we suggest that reduced sharing obligations are driving the effect. In additional analyses, we document a 35% decrease in past-month transfers of cash to other household members. Hence, our findings suggest that saving can be more effectively promoted by alleviating access-related rather than behavior-related constraints, and particularly by giving women access to a saving device of their own.

Keywords: Saving, Temptation Spending, Commitment Device, RCT
JEL codes: D14, D15, D91, I31, O12, O16
Study pre-registration: Social Science Registry, ID: AEARCTR-0003682

*first and corresponding author

1 Introduction

Saving has crucial welfare-enhancing functions for the poor. First, it plays an important role for the poor's cash-flow management and can help to smooth consumption over irregular incomes (Steinert et al., 2018; Karlan, Ratan & Zinman, 2014; Deaton, 1989). Second, savings can partly substitute formal insurances and protect the poor against unanticipated income shocks (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, 2018). 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, Frölich & 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, evidently expressed by a high prevalence of temptation spending among the poor (Banerjee & Duflo, 2007). While behavioral biases affect 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 exert a higher level of self-control and patience than individuals in developed countries (Haushofer & Fehr, 2014; Banerjee & Mullainathan, 2010; Mullainathan & Shafir, 2009).

To tackle the problem of temptation spending and thus promote saving behavior, several programs have introduced commitment tools. A commitment intervention is defined as an arrangement that fosters saving and financial self-discipline by making deviations from a savings goal costly and unattractive (Bryan et al., 2010). Literature

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

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 (Benabou & Tirole, 2004). Examples for the former include bank accounts with specific withdrawal restrictions or lockboxes (sometimes with the key held by an external agent) (Aggarwal et al., 2020, 2018; Aker et al., 2020; Herskowitz, 2020; Karlan & Linden, 2014; Dupas & Robinson, 2013; Ashraf et al., 2010, Ashraf, Karlan & Yin, 2006). 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 this literature, we introduce an innovation to existing soft commitment products. Precisely, our intervention consists of a portable saving device, a zip purse, that is provided in addition to a stationary lockbox. We argue that our portable device may add value to existing designs by serving as a reminder and activating the binding appeal precisely at the point in time at which spending decisions are made (Karlan et al., 2017; Shafir & Thaler, 2006). To test our hypothesis, we conduct a field experiment with 1525 low-income slum dwellers in India's Maharashtra province. Half of the participants were randomly assigned to obtain a lockbox and the portable savings device and thus compose the treatment group. The remaining half was randomly assigned to obtain a lockbox only and serve as the (active) control group. Our analyses are based on data from endline surveys with 1421 participants, conducted six months after delivery of the savings devices.

There are three main findings. Firstly, we observe significantly higher total savings in the treatment group relative to the control group, corresponding to an 81% increase in average savings amounts. The treatment effect on savings was primarily realized through significantly lower withdrawal rates – rather than higher deposit rates. That is, treatment group participants were more likely to keep the money – once deposited – in their savings device.

Secondly, we do not find any evidence of reductions in temptation spending among those who received the portable saving device. The observed higher savings amounts are therefore not a direct consequence of diminished temptation expenditures. This finding contradicts our initial hypothesis suggesting that the zip purse would primarily function as a soft commitment device and help reduce participants' impulsive spending. Instead, our quantitative and qualitative data point to an alternative chan-

nel: participants appear to use the portable device with the intention to hide private savings from others, most likely from their spouse or other family members. This motivates a reinterpretation of the purpose of the portable saving device as a *hiding* rather than *commitment* tool.

Thirdly, we reveal treatment effects on secondary outcomes, namely on female empowerment and levels of debt. The former effect could reflect increases in women's financial autonomy, linked to their access to a saving tool of their own. Against this backdrop, the intervention may have enabled female participants to make financial decisions without seeking their spouses' approval and to keep money as savings that would otherwise have been consumed by their spouse (Fiala, 2017). Apart from this, the significant decreases in participants' levels of debt may imply that accumulated savings can (partly) substitute loans as investment capital. This could have important long-term effects by counteracting possible debt spirals and utility losses linked to high interest rates for borrowing(see Steinert et al., 2018).

Findings from this trial feed into a growing body of scholarship that discards a unitary household model and its underlying assumption that preferences of household members, and specifically the preferences of husband and wife, can be regarded as homogeneous (Anderson et al., 2017). Instead, scholars have argued that prevailing differences in spouses' relative decision-making power are highly relevant for the allocation of household resources. Hence, they can determine how much money is invested into children's future, business and entrepreneurship activities, and how much money is saved (Jayashandran, 2019; Anderson et al., 2017; Iversen et al., 2011; Bobonis, 2009; Duflo & Udry, 2004; Duflo, 2003). Women with low bargaining power are particularly likely to "lose" potential profits and savings to their husband or partner (Bernhardt et al., 2019; Fiala, 2017; De Mel, McKenzie & Woodruff, 2009). In consequence, there is a high demand among women for privately held financial tools (Castilla, 2019; Schaner, 2015, Anderson & Balland, 2002). In a similar vein, women in our sample seem to benefit from the portable savings device in that it enables them to hide money from others and increase their financial autonomy. This is most evidently expressed in the following quote by one of our participants: *"I keep some money with me in the purse. If he [husband] wants money he takes it from the box or asks me. I give him the money from the box but he does not know that I have more money with me in my purse"*.

The remainder of this paper proceeds as follows: the next section describes the trial's setting, sample, data, and experimental approach. Section 3 presents the main results and complier as well as heterogeneity analyses. Section 4 seeks to identify possible channels underlying the treatment effect on savings, before the concluding section discusses caveats and policy implications.

2 Experimental design and data collection

2.1 Study setting

The study took place in slum communities surrounding the cities of Pune and Pimpri-Chinchwad, which are both located in India's western-central state Maharashtra. Maharashtra is the second most populous state and one third of its population live below the poverty line (World Bank, 2017; Desai et al., 2010). India is one of the lowest-ranking countries on the Gender Inequality Index (World Development Report, 2018) and gender discrimination is thus also prevalent in Maharashtra. For instance, literacy rates in the state are ten percentage points lower for women relative to men, health outcomes for women are substantially worse (e.g., 48% of women are anemic vs. 17% of men), and one third of ever married women have experienced physical or sexual violence by a partner (International Institute for Population Sciences, 2018). Potential implications of prevailing gender inequalities were carefully considered for the intervention design, selection of outcome measures, and interpretation of findings. Financial inclusion in India is currently at 80% of the adult population (Demirgüç-Kunt et al., 2018). While this rate is relatively high compared to other low- and middle-countries, 48% of account holders have not made any deposits or withdrawals in the previous year and account inactivity is thus more prevalent than in other countries (ibid). In addition, rates of financial inclusion are lower among women (only 45% have a bank account of their own) and members of backward casts. Overall, India is still home to 190 million unbanked individuals (International Institute for Population Sciences, 2018; Demirgüç-Kunt et al., 2018). We therefore expected demand for alternative saving devices to be high among low-income slum dwellers.

2.2 Sampling

The sample of this study consisted of 1525 female (82%) and male (18%) slum dwellers who were 18 years and older. Eligibility criteria were defined to ensure that participants' financial resources exceeded their subsistence needs and thus gave them the

potential to save money. Accordingly, a person was considered eligible for this study if he/she indicated having some income at least once per week or on a monthly basis – either through permanent employment, casual work, remittances or governmental cash transfers. Enumerators conducted door-to-door community visits and were instructed to recruit the main income earner of each household. While interviews were conducted with only one spouse, intervention and saving devices were delivered to both spouses. We opted for this joint delivery strategy to alleviate the risk of potential conflicts and power imbalances that may arise from withholding putative intervention benefits from one partner. Informed consent was obtained from the main respondent during the first home visit, prior to the baseline interview. We additionally sought informed consent from the participants' spouse prior to delivery of the intervention. We did not provide any monetary incentives for participation in this study.

2.3 The intervention

The intervention, named *"Aaj bachat kara, udya khush raha"* (Marathi for *"Save today, be happy tomorrow"*), was developed with the intention to reduce temptation spending and promote saving. The core feature of the intervention was a portable commitment device – a zip purse – that participants received in addition to a stationary savings box. Our intervention was thereby built on the guiding hypothesis that the zip purse would serve as a reminder and exert a psychological function by penalizing temptation spending with instant feelings of guilt and failure (Soman & Cheema, 2011). The presumed innovation vis-à-vis existing commitment interventions was that the binding force of the device would take effect precisely at the point in time when spending decisions were made.

The stationary device was a metal box secured with a padlock. We thereby built on several field experiments that have endorsed lockboxes as effective saving instruments, particularly in settings where more formal saving infrastructure is not accessible (e.g., Dupas & Robinson, 2013). In this study, each spouse was provided with a key to the padlock to ensure that money retained its liquidity and could be accessed in case of emergencies.

Both the portable and stationary saving device were distributed to participants by local community workers who were trained as program facilitators. Facilitators visited participants at home and encouraged them explicitly to carry the portable device with them whenever they left their homes. They further advised participants to move money from their portable into their stationary device on a regular basis for reasons

of safety. The full home visit script is provided in the pre-analysis plan (see AEA trial registry).

After explaining the purpose and use of both saving devices, facilitators asked participants to formulate a savings goal. Together with participants, they then outlined an individualized savings plan, which specified daily or weekly targets and a timeline to reach the savings goal. The savings goal and implementation plan were visualized on a savings sheet that participants could put on their walls as an additional reminder (see pre-analysis plan). This intervention component was motivated by goal setting theory that postulates a direct link between conscious goals and action (Locke & Latham, 2002; Fiorill et al., 2014; Ryan, 1970). 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 et al., 2016); and thirdly, it motivates perseverance (Alan, Boneva & Ertac, 2016). Building on evidence pointing to higher goal attainment when implementation intentions were formed, our intervention combined the savings goals with a concrete implementation strategy (i.e. how much money to deposit each week) (Duckworth et al., 2013; Soman & Zhao, 2011; Townsend & Liu, 2011).

(Active) Control

Control group participants received only the stationary and not the portable saving device. All other intervention components were kept identical. That is, similar to the treatment group, the lockbox was delivered during home visits, keys were distributed to the participant and her/his spouse, and participants were asked to formulate a savings goal and a detailed savings plan.

2.4 Experimental design and timeline

The study randomly assigned 1525 individuals to receiving either the lockbox and portable saving device (treatment group, n= 771) or the lockbox only (control group, n=754). Randomization was performed in Stata and stratified by participant sex, baseline savings, and baseline levels of present bias. The trial and a pre-analysis plan were registered in the American Economic Association's registry for randomized controlled trials (ID AEARCTR-0003682). Blinding of participants and program implementers was not feasible. However, considering that the trial was set up with an active control group and that all participants received at least one saving device, we expect the risks of performance and expectancy bias to be reduced.

Recruitment and baseline surveys were carried out from November 2018 to January 2019. The saving devices were delivered to treatment and control group participants between February and April 2019. Endline surveys were administered between August and October 2019, with some unanticipated delays due to monsoon-related flooding of several slum locations in Pune.[2] In addition, we held four focus group discussions with program participants in October 2019.

2.5 Data

Baseline and endline data were collected via standardized questionnaires that were administered on tablets. Computer-assisted data collection was opted for in order to (i) improve data quality by programming built-in skip-patterns, reminders, and consistency checks to prevent item non-response or selection errors, and (ii) reduce respondent fatigue through programming visually appealing questionnaires including vignettes and pictures. Questionnaires were available in both English and Marathi, and each item was piloted with the local research team for cultural adequacy and accurate understanding. Enumerators were recruited from local communities and had to be fluent in Marathi. The local research team received a five-day training focused on interview techniques, research ethics, and familiarization with the survey. Individual interviews lasted between 30 and 45 minutes and were typically conducted in participants' homes. The research team made substantial efforts to guarantee a private and comfortable interview atmosphere and avoid the presence of other household members (especially of the spouse or parents-in-law) during the interview. Ethical approval for this study was obtained through the University of Goettingen.

The survey captured essential sociodemographic information, including household composition, caste, religion, and a number of wealth indicators, namely education, employment, and income as well as asset ownership. The study's primary outcomes were total savings and total temptation expenditures. First, total savings were measured by hand-counting money held in the stationary savings box, and, if applicable, money kept in the portable saving device. Given that participants were not informed about the date and time of their endline interview prior to the visit, we expect this measure to be relatively immune to potential reporting and social desirability biases. To rule out any potential crowd-out effects, we also collected detailed self-reported information on savings held elsewhere, for instance in a formal bank account or savings group, and documented (a) total amounts, (b) past-month deposits, and (c) past-

[2]From June to September 2019, India recorded the most extreme level of monsoonal rainfall in the past 25 years. 28.000 people living in Pune had to be evacuated in consequence of flooding.

month withdrawals. Second, we designed a new measure of temptation spending, which defined a unique set of temptation goods for each individual without reliance on a priori, researcher-defined temptation categories. Theoretical literature characterizes temptation goods as goods that provide utility when consumed, but not in anticipation of their consumption (Banerjee & Mullainathan, 2010; Benabou & Tirole, 2004). Capitalizing on this standard definition, we captured the past as well as desired future consumption for nine selected food and non-food items (e.g., alcohol, tobacco, gambling). Items were only classified as temptation goods if the reported amount for past expenses exceeded the desired future amount. For each respondent, divergences were then added up into a total amount of past-month temptation expenditures.

In addition, six secondary outcomes were included in the survey. First, we measured respondents' self-efficacy by drawing on selected items from the *Internality, Powerful Others and Chance (IPC)* scale (Levenson, 1981) as well as from financial self-efficacy scales used previously by Steinert and colleagues (2018) and Lown (2011). We expected increases in self-efficacy based on previous research suggesting that self-defined goals and implementation plans can instil feelings of self-efficacy and control (Morisano et al., 2010). Second, acknowledging gender biases as a crucial contextual factor in our study population, we assumed that distribution of saving devices (i.e., the zip purses) for individual usage may have the potential to affect participants' gender attitudes. In line with existing financial inclusion literature (e.g., Duvendack, Palmer-Jones & Vaessen, 2014), we therefore included five items on the roles and rights of men and women (boys and girls) in society, thus aiming to elicit participants' gender attitudes. Third, we also included a more direct measure of female empowerment. For this, we focused exclusively on the sub-sample of women and collected data from our female respondents only. Items were drawn from Glennerster, Walsh and Diaz-Martin (2018) and adapted to match the context of India.[3]

Fourth, survey items were included to assess respondents' financial resilience to potential health or other emergencies. This was determined based on participants' self-reported capacity to cover the costs for potential medical treatment and medicine if needed, and – contingent on whether they had experienced some sort of income shock in the previous six months – their ability to financially cope with an unfore-

[3]Note that gender attitudes and female empowerment were listed as one overarching measure (female empowerment index) in the pre-analysis plan. However, factor analysis of individual items pointed to a two-factor solution and thus suggested the importance of two distinct underlying concepts, namely gender attitudes (reported by men and women) and female empowerment (reported by women only). We therefore decided to report these as separate outcomes in the subsequent analyses.

seen emergency. Fifth, assuming that accumulated savings could substitute loans for investment purposes, we captured respondents' total outstanding debts. Lastly, we collected data on past-month expenditures for six selected food and seven selected non-food items (excluding temptation goods). This more distal welfare indicator enabled us to examine whether higher savings rates may have beneficial downstream impacts and help alleviate poverty, for instance through more effective protection from economic shocks or returns from business or human capital investments that were realized through accumulated savings (Dupas et al, 2018; Brune et al., 2015; Dupas & Robinson, 2013).

Lastly, we collected qualitative data in order to complement our quantitative findings and elucidate possible mechanisms of change. For this purpose, we conducted four focus group discussions with an average of ten participants per focus group. Three focus groups were conducted with treatment group participants (two gender-segregated and one mixed) and one with control group participants (mixed). Discussion guides included open-ended questions probing participants to reflect upon any changes (positive or negative) that they and their families had experienced as a result of receiving the lockbox and, for treatment group participants, the zip purse. Discussion guides also covered several questions on intra-household decision-making and sought to shed light on the relational dynamics between spouses that shape household financial management. In addition, our enumerators kept written records ("field journals") of their experiences and observations during home visits throughout the implementation and endine phase. Recordings from focus group discussions and enumerator observations were transcribed and translated to English and then coded using thematic analysis (Braun and Clarke, 2006). Qualitative statements that were conceptually similar were categorized into overall themes and then discussed and validated with a second coder (RVS).

2.6 Attrition

Attrition in this study was at 6.8%. In order to test whether attrition was differential, we first regressed the attrition dummy on the treatment dummy, then added additional controls, and lastly interaction terms between the treatment and control variables. We show that attrition was not significantly associated with treatment status ($p = 0.62$) (see Table A1). However, attritors were on average more likely to be female, employed, and have higher incomes.

2.7 Estimation Strategy

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

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

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' a vector of stratification variables, X_i' a vector of baseline covariates, and ω_i 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 $Y_{i(t-1)}$ to the previous specification, and (3) a third specification including additional baseline controls X_i' , namely participant age, marital status, educational level, employment, household size, and baseline poverty level. We condition on the baseline level of outcomes and additional controls in the ANCOVA specifications (2) and (3) to increase statistical power and precision of estimates. Our coefficient of interest was β, which indicates the intent to treat (ITT) effect.

We also estimated the average treatment-on-the-treated (TOT) program effect by using an instrumental variable approach. The TOT estimate was given by:

$$A_i = a + bT_i + cY_{i(t-1)} + dS_i' + eX_i' + w_i \tag{2}$$

$$Y_i = \alpha + \beta A_i + \gamma Y_{i(t-1)} + \delta S_i' + \epsilon X_i' + \omega_i \tag{3}$$

whereby we instrumented A_i, namely self-reported usage of the portable saving device in the study observation period, with being assigned to the treatment.

We computed False Discovery Rate (FDR) adjusted q-values to account for multiple hypothesis testing (see Benjamini et al., 2006; Benjamini & Hochberg, 1995). We utilized the Benjamini-Hochberg method, which is considered as less conservative than simple Bonferroni adjustments and particularly suitable when working with a range of outcomes that are likely correlated. Adjustments were made across primary and secondary outcomes separately. In our results section, we report both unadjusted

p-values as well as q-values corrected for multiple testing for each outcome.

Lastly, we tested for heterogeneous effects based on the following specification:

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

where $TRAIT_i$ was a vector of baseline characteristics for which we assumed heterogeneity in the effectiveness of the treatment. These baseline variables were prespecified in the analysis plan and included (i) participant sex, (ii) female involvement in household financial decision-making, (iii) present bias, and (iv) income levels. The average treatment effect for a subgroup of people with a respective trait (i.e. those below median income) is then given by the sum of the coefficients $\beta + \theta$ for that trait.

3 Results

3.1 Summary statistics and orthogonality verification of randomization

Baseline characteristics of study participants are summarized in Table 1. More than 80% of the sample were female, which likely reflects the higher willingness of women to participate in our study. Most participants reported being Hindus while only 9.7% were Buddhists and 5.4% were Muslims and thus belonged to the country's religious minority groups. More than 80% of our respondents were married and around 40% had not completed any form of primary or secondary schooling.

Our study sample was characterized by relatively high levels of poverty. Almost half of our participants reported belonging to a backward caste or scheduled tribe and thus to a historically marginalized social group. Although discrimination of the so-called "untouchables" is prohibited by the Indian constitution, a person's caste is still a strong predictor of poverty levels today and continues to determine, for example, land ownership, access to public goods, and social capital (Lastrapes & Rajaram, 2016; Borooah, 2008; Gang et al., 2008; Mehta & Shah, 2003). In addition, one third of our respondents were unemployed and the average monthly income (across both study arms) was at 15,589.00 INR (equivalent to approx. 218.00 USD). However, baseline savings rates were already quite high. Almost 80% of our participants indicated holding some form of savings and the average total savings amount was at 7,660.00 INR (equivalent to approx. 96.00 USD).

At the same time, participants reported strikingly low baseline levels of temptation spending. Only 40% of participants reported having purchased some sort of temptation good in the previous month and the average monthly amount spent on these was very low at 127.00 INR (equivalent to approx. 1.60 USD), less than 1% of respondents' monthly incomes. Given that we had to rely on participants' self-reports for our measure of temptation spending, it is possible that the small amount was partly a function of social desirability, and thus under-reporting of expenditures on stigmatized behaviors such as gambling and alcoholism. Another possible explanation is prompted by our enumerators' field notes, documenting that many female participants indicated that their husbands were mainly engaging in temptation spending and that they had little control over this. Possible gender disparities with regards to temptation spending were also illustrated by our quantitative data: total baseline temptation expenditures were significantly higher among male participants (average: 169.35 INR) relative to female participants (average: 117.50 INR), however still represent only a small share of other consumption.

We used a joint orthogonality F-test to assess baseline balance across both study arms and confirmed randomization as effective (F=0.765). The treatment and control group were balanced along all but one characteristics (see Table 1, Column (3)). The only significant difference was a slightly higher rate of married women in the treatment group (p=0.02). We control for marital status in our regression specification III and further add robustness by controlling for the baseline value of each respective outcome variable in specifications II & III.

Table 1. Baseline Balance

	Control (N=754)	**Treatment (N=771)**	**t-test Difference (1)-(2)**
Female	0.82 (0.01)	0.81 (0.01)	0.00
Age	35.59 (0.46)	36.37 (0.54)	-0.78
Belongs to scheduled/backward caste or tribe	0.47 (0.02)	0.45 (0.02)	0.02
Married	0.81 (0.01)	0.85 (0.01)	-0.05**
Hindu	0.76 (0.02)	0.78 (0.02)	-0.02
Household members	5.01 (0.08)	4.97 (0.09)	0.04
Unemployed	0.33 (0.02)	0.32 (0.027)	0.01
No education	0.22 (0.02)	0.22 (0.02)	0.00
Completed primary education	0.20 (0.02)	0.21 (0.01)	0.00
Completed secondary education	0.36 (0.02)	0.39 (0.02)	-0.03
Completed tertiary education	0.21 (0.02)	0.19 (0.01)	0.02
Past-month income	17968.91 (3941.90)	13262.10 (1754.74)	4706.80
Past-month savings	7381.29 (1211.84)	7933.16 (1168.38)	-551.87
Past-month temptation spending (INR)	128.06 (23.11)	126.14 (16.84)	1.92
Temptation index	1.47 (0.03)	1.44 (0.02)	0.03
Self-efficacy index	6.12 (0.09)	6.15 (0.09)	-0.03
Female empowerment index	0.02 (0.04)	-0.02 (0.04)	0.04
Outstanding debt	8810.23 (1994.31)	10651.75 (2736.34)	-1841.52
Past-month household expenditures (for selected goods)	5088.60 (398.14)	5491.69 (435.67)	-403.09
Resilience index	0.27 (0.02)	0.28 (0.02)	-0.01
F-test of joint significance (F-stat)			0.765
F-test, number of observations			1525

Notes: Standard errors in parentheses.

3.2 Impacts on primary outcomes

Results for the study's primary outcomes, total savings and total temptation expenditures, are reported in Table 2. For each outcome, we first ran a regression specification that only included the stratifying variables as a predictor, then, in the second specification, we added the lagged outcome, and additional socioeconomic controls in the third specification. We show that magnitude and significance of the program effects were largely robust to all three specifications and therefore focus on the final regression model in the following paragraphs.

The ITT effect on total savings are shown in Table 2, Column (1). In the control group, the average endline amount of total savings was at 8,400.85 INR (approx. 118 USD), compared to 15,109.10 INR (approx. 211 USD) in the treatment group. The treatment effect estimated in specification III thereby corresponds to a 81% increase in total savings, which was significant at the 5%-level. The treatment effect remained significant at the 10%-level when we adjusted for multiple hypothesis testing.

In Table A2, we present disaggregated estimates for past-month deposits and past-month withdrawals. We show that the intervention effect on total savings was largely driven by significantly lower withdrawal rates among treatment group participants. Interestingly, the coefficient for deposit rates was also negative (but insignificant), suggesting that treatment group participants, when compared to control group participants, either put less frequent or smaller amounts of money into their saving devices. We could tentatively interpret this as indication for a stronger perceived constraint or binding force associated with the portable saving device. Against this backdrop, participants in the treatment group may ascribe stronger feelings of guilt or failure to potential withdrawals from their zip purse (or box) and thus, in anticipation thereof, decide to only deposit money, which they do not plan to withdraw.

In Table A3, we assess whether there is crowding-out from one saving type to another. Overall, the intervention did not affect participants' formal savings. We found similar levels of total savings held in a bank account, mobile money account or in a savings club among participants in the treatment and control group. The coefficient for informal savings held at home was negative, suggesting that participants in the treatment arm may have moved some of their previous home savings into their portable saving device. However, the coefficient was not significant and we can thus infer that the observed treatment effect not simply a reflection of savings being shifted from one place to another.

Moving to the other primary outcome, we found ITT effect estimates that were close to null. While we observed a general reduction (in the full study sample) in temptation spending between baseline and endline (corresponding to a 48 percentage-points decline), there were no significant differences in spending amounts between study arms at endline. This finding was corroborated by an alternative measurement of temptation spending: in Table A4, we provide ITT effect estimates for a self-reported temptation index[4] and, again, find coefficients very close to zero.

Table 2. ITT Estimates: Primary Outcomes

	(1) **Total Past-Month Savings**			(2) **Temptation Expenditures**		
	I	II	III	I	II	III
ITT: Received Program	6708.26** (3111.33) [0.072]	6555.50** (3085.91) [0.068]	6802.30** (3351.22) [0.086]	1.37 (14.91) [0.927]	1.34 (14.91) [0.929]	2.65 (15.25) [0.862]
Stratification variables	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.09* (0.05)	0.09* (0.05)	no	0.00 (0.01)	0.00 (0.01)
Controls	no	no	yes	no	no	yes
Observations	1421	1421	1379	1421	1421	1379
Mean Control		**8400.83 (20463.25)**			**82.27 (262.25)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

3.3 Impacts on secondary outcomes

In the next step, we examined whether the portable saving device had positive impacts on secondary outcomes. Results are presented in Table 3, Columns (1)-(6). First, as shown in Table 3, Column (1), we observe no significant intervention effect on participants' gender attitudes and the difference between study arms was effectively zero. However, albeit non-causal, we observed improvements in gender

[4]The index was composed of the following three items, rated on a 1-5 Likert scale: *"In the past month, I spent money on things that I did not really need", "...I bought something and later regret that I did", "... I found it difficult to really control how I spend my money".*

attitudes over time in both study groups equally. More specifically, there was a 32 percentage-point improvement on the principal-component-weighted gender attitudes index when comparing baseline to endline measures. To put this into more meaningful terms, at baseline, 88% of participants agreed to the statement that girls and boys should have equal inheritance rights, compared to 95% of participants agreeing at endline. Similarly, at baseline, 12% agreed that boys should be given more food relative to their sisters, compared with only 8% at endline. Of course, our study design does not allow us to establish whether these improvements are indeed a direct result of our savings intervention – and its active integration of women into household financial management – or whether these are linked to other co-occurring but unobserved factors.

We further examined whether receipt of the portable saving device led to any improvements in actual female empowerment, captured by women's self-reported autonomy rights in relation to their husband or partner. For this analysis, we focused exclusively on the sub-sample of female respondents. As documented in Table 3, Column (2) we found significant improvements on the female empowerment index among participants in the treatment group. The treatment effect corresponds to 17% increase in female empowerment. Put differently, 26% of women in the control group indicated that they would face negative consequences if they did not inform their husband/partner about leaving the house. In the treatment group, only 20% of women anticipated such consequences. Likewise, in the control group, 78% of women were allowed to leave their homes unescorted, compared to 82% in the treatment group. The overall treatment effect was significant on the 5%-level when using naïve p-values but lost significance (p=0.11) after FDR adjustments.

In Column (3) of Table 3, we present treatment effects on financial self-efficacy. We found no significant differences in financial self-efficacy between study arms and coefficients were again very small in magnitude. Interestingly, we also did not find any changes in self-efficacy levels in either of the groups from baseline to endline. Our null findings thus contradict our initial hypothesis – as well as evidence put forth in previous literature (e.g., Steinert et al., 2018) suggesting that financial self-efficacy is a central mediator for changes in financial behavior, such as improvements in savings behavior.

In Columns (4)- (6) of Table 3, we report ITT estimates for outcomes that are more distal and thus more reflective of potential changes in economic welfare. First, we assessed participants' financial resilience to unforeseen emergencies such as health

Table 3. ITT Estimates: Secondary Outcomes

	(1) **Gender Attitudes Index (full sample)**			(2) **Female Empowerment Index (women only)**			(3) **Self-Efficacy Index**		
	I	II	III	I	II	III	I	II	III
ITT: Received Program	-0.00 (0.12) [0.973]	-0.00 (0.12) [0.992]	-0.02 (0.12) [0.862]	0.05** (0.02) [0.144]	0.04** (0.02) [0.171]	0.04** (0.02) [0.114]	0.03 (0.13) [0.945]	0.04 (0.12) [0.936]	0.04 (0.12) [0.862]
Stratification Variables	yes	yes	yes	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.08*** (0.02)	0.07*** (0.03)	no	0.20*** (0.03)	0.17*** (0.03)	no	0.14*** (0.03)	0.17*** (0.03)
Controls	no	no	yes	no	no	yes	no	no	yes
Observations	1420	1417	1375	1186	1176	1134	1415	1409	1368
Mean Control		**-0.69** **(0.57)**			**-0.23** **(0.37)**			**6.00** **(2.41)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected sharpened q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

shocks. ITT effects were non-significant and effectively zero, pointing to identical resilience levels in both study arms. Similar to other outcomes, participants in both groups reported substantial improvements in economic resilience from baseline to endline, corresponding to a 59 percentage-point improvement on the resilience index. In focus group discussions and during some of the home visits, participants also qualitatively noted that they had used savings from their lockbox to cope with a given emergency. For instance, one participant stated: *"Last month, suddenly my daughter got sick. That time the money in the box proved to be very useful"* and another noted: *"This time it rained heavily so we had to spend money on repairing our house. We used the savings in the box"*. While, overall, there might have been improvements in financial resilience, the portable saving device did not have an add-on effect.

Apart from this, we report ITT program effects on participants' debt. We revealed significantly lower levels of debt among participants who had received both the lockbox and the zip purse. Precisely, treatment group participants reported a reduction in total outstanding debts by 30%. This suggests that the demand for loans may have been partly substituted by higher savings rates among treatment group participants and thus enabled them to make investments or respond to emergencies without reliance on external money lenders. The treatment effect was significant when using naïve p-values (p=0.038) but narrowly failed to reach significance after FDR adjustments (p=0.11).

Lastly, to assess the potential positive downstream impact of higher savings (see Dupas et al, 2018; Brune et al., 2015; Dupas & Robinson, 2013), we estimated program effects on past-month household expenditures (see Table 3, Column (6)). The ITT coefficient was positive, suggesting that past-month expenditures were 11% higher in treatment group households. However, the difference was insignificant and imprecisely estimated with relatively large confidence intervals. While we cannot confirm any substantial downstream impacts of higher savings, we can at least rule out that higher savings rates were realized through cutting down spending on essential consumption goods – and thus a potential harmful effect of our intervention.

Table 3. (ctd.) ITT Estimates: Secondary Outcomes

	(4) **Resilience Index**			(5) **Outstanding Debt**			(6) **Household Expenditures**		
	I	II	III	I	II	III	I	II	III
ITT: Received Program	-0.01 (0.02) [0.946]	-0.01 (0.02) [0.936]	-0.01 (0.02) [0.774]	-498.81* (270.40) [0.195]	-506.13* (266.18) [0.171]	-579.65** (278.87) [0.114]	499.91 (1631.67) [0.946]	458.72 (1653.83) [0.936]	635.62 (1688.83) [0.862]
Stratification Variables	yes	yes	yes	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.22*** (0.03)	0.19*** (0.03)	no	0.01 (0.01)	0.01 (0.01)	no	0.08 (0.08)	0.08 (0.08)
Controls	no	no	no	no	no	yes	no	no	yes
Observations	1421	1421	1379	1421	1421	1379	1421	1421	1379
Mean Control		**0.44** **(0.44)**			**1956.51** **(5319.36)**			**5767.05** **(29496.46)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected sharpened q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

3.4 Complier analyses

For estimation of the average treatment effect for compliers, we instrumented self-reported active usage of the portable saving device with being assigned to the treatment. Overall, 23% of respondents in the treatment arm indicated that they had regularly used the zip purse since its distribution. Usage barriers were identified based on our qualitative data and included restrictions for women to leave their homes (e.g., one participant noted: *"We women do not go out much. So I don't use the purse much."*), fears from theft, and not seeing any potential for saving due to very low income.

As we show in Tables A5-A6, the overall pattern of the instrumented estimates remains broadly similar to the ITT estimates from Tables 2-3.

3.5 Heterogeneity in treatment effects

This section examines heterogeneity in treatment effects based on four pre-specified observable characteristics, namely participants' sex, baseline income level, present bias, and women's involvement in household financial decision-making. We found no significant evidence for heterogeneity in treatment effects and are therefore only elaborating on a few tentative points below.

With regards to the trial's primary outcomes (see Table A7), treatment effects did not vary between men and women. There was tentative evidence that the treatment effect on total savings turned insignificant for participants in the lowest income quantile. It is possible that these individuals were simply "too poor" to save and did not have any surplus money to be stored in their portable devices. With regards to present bias, we can cautiously interpret our results as pointing to lower treatment effects on total savings for participants with a higher level of present bias. Again, this finding suggests that our portable saving device did not effectively work as a commitment tool to enhance self-control among more impatient participants. Lastly, there appears to be some heterogeneity in both primary outcomes with regards to women's involvement in financial decisions. Although effect estimates turn insignificant after our FDR-adjustments, the treatment effects are higher for households in which the female spouse is actively involved in financial decisions. Accordingly, total savings amounts are larger in magnitude for this subgroup and the total effect coefficient for temptation expenditures changes signs between subgroups, thus indicating that temptation spending is only reduced if women can influence financial decision-making processes.

As reported in Table A8, we do not find any substantial heterogeneity in treatment effects on secondary outcomes. The only noteworthy finding were subgroup variations in treatment effects on past-month debts (see Table A8, Column (5)). The treatment effect was more pronounced, and debt amounts more significantly reduced, for female participants. This could be a direct consequence of the (slightly) higher savings rates among female treatment group participants and indicate a lower need for loans as investment or emergency capital.

4 Discussion of potential mechanisms

Our results reveal that the portable saving device was effective in increasing participants' total savings – and thus one of our two primary outcomes. A more nuanced understanding of the channels through which the saving device has helped promote saving activities is crucial for future policy making. We therefore draw on additional quantitative data on the reported device usage and perception, as well as on qualitative data from four focus group discussions and enumerators' field observations. Motivated by prior literature, we explore three distinct channels through which the portable saving device could have worked.

Self-control channel

A first possible channel is **enhanced self-control** linked to the (soft) commitment function of the zip purse. Accordingly, we would expect that the portable saving device fosters participants' self-discipline, willpower, and intrinsic motivation, and thereby helped them to better resist their impulses and temptations (Soman & Cheema, 2011; Bryan, 2010; Ashraf et al., 2006; Benabou & Tirole, 2004). While numerous previous studies have presented evidence on the effectiveness of lockboxes as a commitment tool (Aker et al., 2020), the central question of our research design refers to the potential add-on effect of the portable device. We argue that the zip purse could have activated an additional commitment effect in two ways. Firstly, the portable saving device allowed for physical segregation of the liquid cash that participants carried with them and could thus have amplified mental accounting mechanisms in their day-to-day budgeting decisions. Nobel prize winner Richard Thaler (1990) argues, based on ethnographic evidence, that money is perceived as less fungible if it is mentally earmarked for a specific purpose (Benabou & Tirole, 2004). By distributing portable savings purses, we may have helped to materialize these mental rules: we allow participants to keep "money to spent" physically separate from "money earmarked for savings". In line with this logic, treatment group participants may have considered money in their zip purse as explicitly "reserved" for savings purposes and have thus abstained from spending it on other purposes (Karlan & Linden, 2014; Dupas & Robinson, 2013; Shafir & Thaler, 2006). Secondly, the portable saving device was designed to create temporal concurrency between the psychological commitment effect and actual spending decisions. Since the portable device was supposed to be carried during the day, it was likely physically present when most spending decisions occurred. Salience of its commitment function was thus more pronounced in comparison to the lockbox that was kept at home and thus more distant and abstract. We could therefore assume that any violations of saving intentions would have *instantly* induced negative emotions and feelings of guilt and thus a higher perceived

psychological costs for a person who had carried the portable saving device (Shefrin & Thaler, 1992).

We empirically assessed this first possible channel and found little supportive evidence for a self-control pathway. Most importantly, if the portable saving device had effectively increased participants' self-control, we would expect to see significant reductions in their temptation expenditures. However, we reported null-effects on this outcome across all regression specifications. We further exploited survey data on participants' views on the purpose of the portable device. In Table 4, Column (1), we show that the self-control purpose was not a significant predictor of higher savings amounts among treatment group participants. In the full regression specification (see Table 4, Column (4)), the coefficient even turned negative, corresponding to a decrease in savings amounts. Conversely, in our qualitative data, we revealed some narratives that pointed to the self-control function of the zip purse (e.g., *"When I go shopping or some other work, I keep the change in the purse instead of spending it here and there. So I don't buy unnecessary things because of the purse"*, see Table 6, Column (1)). However, it is possible that these accounts were more suggestive of a certain priming or social desirability effect, whereby participants described their usage and perception of the portable device closely in line with the instructions they were given on their intended usage during the intervention delivery. Furthermore, it is also possible that the zip purse did in fact represent an effective commitment device for some participants but that these were only few and thus became statistically irrelevant in the quantitative analyses.

Reminder channel

A second possible channel could have been materialized through a possible **reminder function** of the portable saving device. Participants were instructed to carry their portable device during the day in order to ensure that it was physically present (and possibly visible) whenever spending decisions were made. Accordingly, the zip purse could have served as a salient saving reminder and brought participants' saving intentions "to the top of their mind" (Karlan et al., 2016). Several previous studies have corroborated this argument, demonstrating how reminders have helped to increase savings rates (Kast, Meier & Pomeranz, 2018; Karlan et al. 2016; Akbas et al. 2016; Atkinson et al., 2013). 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 that inform participants about their own and their peers' saving performance almost tripled weekly deposit amounts

Table 4. Predictors of total savings amounts in the treatment arm

	Total Savings (1)	**Total Savings (2)**	**Total Savings (3)**	**Total Savings (4)**
Self-Control Purpose	**773.25**			**-9463.58***
	(1613.97)			**(5038.55)**
Reminder Purpose		**4645.75**		**7077.37**
		(3373.92)		**(4830.38)**
Hiding Purpose			**6876.99***	**8189.97****
			(3450.01)	**(3680.32)**
Female	6515.64	5527.11	3936.45	4801.37
	(5592.23)	(4893.84)	(4595.69)	(9599.92)
Baseline Savings	4126.74**	4283.84	3812.15	4242.19*
	(1609.07)	(1713.08)	(1817.17)	(2343.39)
Baseline present bias	-2408.11	-2318.34	-1900.43	-1961.51
	(3181.44)	(3009.79)	(2794.89)	(2358.03)
Married	3151.00	2867.61	2356.84	3034.55
	(3549.45)	(3244.11)	(3110.29)	(9524.23)
Age	339.71	444.06**	372.509*	353.98
	(212.69)	(224.61)	(197.44)	(273.91)
Education level	3927.05	3633.17	3704.90	3555.78*
	(2628.03)	(2333.74)	(2445.72)	(1923.43)
Employed	-4849.21	-7862.40	-4892.63	-4360.43
	(7835.76)	(8529.77)	(8037.81)	(7544.69)
Household size	1965.59	1734.07	1946.96	1722.35
	(2827.50)	(2865.16)	(2888.24)	(1247.09)
Baseline income	-0.02	-0.02	-0.02	-0.03
	(0.02)	(0.02)	(0.02)	(0.07)
N	641	645	639	638

*Notes: *p<0.1, **p<0.05, ***p<0.01. Robust standard errors in parentheses. The Self-control purpose is captured with the following item: "When I am tempted to buy something I do not really need, the purse helps me to resist my temptations", reminder purpose is captured with: "When I see the purse, it reminds me of the importance to save money", and hiding purpose is captured with: "This purse helps me to keep money for myself and not to give it to other people (my partner, children, friends...)". All three items are rated on a 1-5 Likert-scale with higher values indicating higher agreement.*

in the treatment arm (Kast, Meier & Pomeranz, 2018). Although our zip purse takes the form of a physical rather than a text-message-delivered electronic reminder, we contend that its effect could be similar and potentially even more cost-effective and sustainable.

However, our empirical findings did not corroborate the existence of a possible reminder channel. While we report on average higher savings amounts for treatment group participants who are more inclined to perceive their zip purse as a savings reminder, the coefficient remains insignificant (see Table 4, Columns (2) and (4)). Likewise, the qualitative participant accounts included very few indications of a distinct reminder effect of the portable saving device. Only two of 40 focus group participants made some reference to the device's reminder function and none of the ten

enumerators identified this channel in their field observations (see Table 5, Column (2)). This finding is in line with a recent field experiment in Niger, in which receipt of a lockbox significantly increased participants' savings amounts while there was no add-on effect for the treatment group who received SMS reminders on top (Aker et al., 2020). While take-up and usage of the box were high, only 20% of SMS recipients even remembered receiving a text message.

Hiding channel

A last possible channel is the conversion of the zip purse into a tool to **hide money from others**. This assumption was motivated based on our field observations (and hand count data) suggesting that treatment group participants had not always moved the savings from their zip purse into their lockbox – even though the program facilitator had instructed them to do so when delivering the devices. In contrast, we found that many participants had kept a relatively large share of their savings amounts in their purses. For some, the amount stored in the portable saving device even exceeded the amount held in the stationary saving device. This puzzle is closely linked to a body of literature that points to a high social demand on any untapped financial resources that an individual may hold, thus putting major constraints on a person's capacity to accumulate future-oriented savings (World Development Report, 2015). Accordingly, disposable income is often bound by social approbation mechanisms and webs of reciprocal social obligations (Dizon, Gong & Jones, 2016). Individuals experience social pressure to share disposable cash with their spouses, other family members or friends in need (Ambec & Treich, 2007; Noponen & Kantor, 2004; Platteau, 2000). The consequences are twofold: individuals may either meet their social obligations and simply undersave or decide to quickly spend all liquid cash so that demands from family members, friends or neighbours can be turned down (Ky et al., 2016; Brune et al., 2011).

Such social obligations are also a defining feature of financial management dynamics between spouses. For instance, Schaner (2015) conducts a field experiment in Kenya to evaluate a model of non-cooperative household savings behavior. She reveals that women are willing to accept economic utility losses (here in the form of lower interest rates) in exchange for holding an account of their own, rather than a joint account with a husband who "will simply withdraw all her savings and spend the funds on current consumption" (p. 136). Similarly, Anderson and Baland (2002) argue that women tend to join Rotating Savings and Credit Associations (ROSCAs) with the intention to protect their savings from their husbands' immediate consumption needs. In addition, several laboratory field experiments suggest that participants choose income-hiding over profit maximization when allocating experimental endowments

(Castilla, 2019; Jakiela & Ozier, 2016). While both men and women may follow such hiding motives, we can assume that demand for private saving devices is particularly high among women with low hierarchical status and thus substantially constrained decision-making power vis-a-vis a spouse (Dupas & Robinson, 2013).

Our empirical analysis corroborates the assumption that similar hiding motives are at play in our target population. In Table 4, Columns (3)-(4), we show significant increases in total savings amounts among participants who reported that the zip purse enabled them to keep money for themselves and conceal their savings from their spouse or someone else. Specifically, for participants who used their portable device as a means of hiding money, relative to those who reported to strongly disagree with this purpose of the device, total endline savings were 102% higher. We show further quantitative causal evidence in support of the hypothesized hiding mechanism. Essentially, additional analyses (see 5 6)reveal that participants in the treatment group reported significantly lower past-month transfers of money to other household members. More specifically, treatment group participants shared 35% less money compared to participants in the control group. While we did not observe the same reduction with regards to financial transfers to people outside participants' homes, this finding still suggests that the portable saving device has likely helped treatment group participants to keep their saved money for themselves, rather than giving it away to their husband or other household members.

Our qualitative findings further underline this mechanism. First, enumerator field observations document ample evidence on the social barriers that many participants had faced with regards to savings practices – and particularly female participants with regards to their spouse. For example, our field enumerators noted: *"She [female participant] said her husband would spend all her savings if he got information about it."*, or: *"When it was told that the savings box is for entire family, one female participant asked: 'what if I save the money by hook or by crook and my husband takes it away for drinking alcohol?'"* (see Table 6, Column (3)). In our focus group discussions, we also identified more direct narratives on how women had used their portable saving device as a means of hiding money from their husbands. Accordingly, one treatment group participant noted: *"I keep some money with me in the purse. If he [husband] wants money he takes it from the box or asks me. I give him the money from the box but he does not know that I have more money with me in my purse"*; and another one described: *"I always keep money in the purse so that I can keep it for myself"* (see Table 6, Column (3)).

Table 5. ITT Estimates: Responding to Social Demand

	(1) **Past-month transfer to to a household member**			(2) **Past-month transfer to a person outside the household**		
	I	II	III	I	II	III
ITT: Received Program	-395.27*** (151.36)	397.49*** (150.76)	-409.16** (162.34)	222.23 (202.28)	222.49 (202.56)	204.25 (204.59)
Stratification variables	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.00 (0.01)	0.00 (0.01)	no	-0.00 (0.00)	-0.00 (0.00)
Controls	no	no	yes	no	no	yes
Observations	1421	1421	1379	1421	1421	1379
Mean Control		**1185.53 (3030.57)**			**340.98 (2269.30)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

The described usage of the portable device for concealment purposes is also consistent with the main outcome analyses presented in here (3.2). In these, we reported substantial increases in total savings at endline, but it was unclear how these higher savings amounts were actually realized – considering that we found neither reductions in temptation expenditures nor reductions in other past-month expenditures or increases in levels of debt. However, it is still possible that, prior to the intervention, any untapped financial resources of participants, rather than being saved, were spent by spouses or other household members – potentially even on temptation goods. It follows that our participants have benefited from the intervention exactly because it provided them with a private saving device that enabled them to protect their financial resources from such social demands. These changes in spending patterns of participants' spouses or other network partners were not recorded as part of the survey, but based on this logic, we would assume that they had likely declined.

Table 6. Qualitative Evidence

Self-Control Channel (quoted by 11/50)	**Reminder Channel (quoted by 2/50)**	**Hiding Channel (quoted by 21/50)**
Improved self-control: *"When I go shopping or some other work, I keep the change in the purse instead of spending it here and there. So I don't buy unnecessary things because of the purse."* (female participant, FGD2) *"When I used to go in the market, I couldn't resist myself buying those things. Which were not much useful. But after you told me the importance of saving, I realized that the money I am going to save in the purse you had given us, would turn out to be useful in crunch times."* (female participant, FGD2) *Women used the purse and they benefitted from it also. They told us that they used it to put money in it. Some told us that they fought with their desires because they had decided to save money."* (enumerator field observation) *"I have learnt to resist myself. If I resist myself now, then I can use the same amount of money in the future."* (male participant, FGD4)	***Savings reminder:*** *"Earlier I used to buy vegetables and put the remaining money somewhere. I never saved it. But now I see the purse and put the remaining amount from the purse in the box."* (female participant, FGD4) *"Earlier I used to buy anything I saw. Now I don't because the purse is there."* (female participant, FGD4)	***Hiding tool:*** *"I keep some money with me in the purse. If he [husband] wants money he takes it from the box or asks me. I give him the money from the box but he does not know that I have more money with me in my purse."* (female participant, FGD2) *"I always keep money in the purse so that I can keep it for myself."* (female participant, FGD3) *"My husband can take money from the box. But not from the purse."* (female participant, FGD2) *"Many women saved money without telling it to their family and mostly their husband. They said that if their husbands got to know about their savings, they will spend it."* (enumerator field observation) ***Spousal control:*** *"They also have to take care of not disclosing the amount to their family members so that they won't face any trouble from their family members. That is why not disclosing their savings is one of their priorities."* (enumerator field observation)

Table continued on next page

Table 6 – continued from previous page

Self-Control Channel	Reminder Channel	Hiding Channel
Mental Accounting Effect: *"It was useful. What I do is, I put small purse inside the big one. When I go out, I put my remaining money in it and when I come back, I put it in the box."* (female participant, FGD4) *"I keep the money aside in it for medical expenses."* (male participant, FGD4) *"I kept the notes in the purse and the coins in the box. So both of them proved to be useful."* (female participant, FGD4)		*"While I was interviewing another household, she came and said she lied about her saving because of her husband's presence. She said her husband would spend all her savings if he got information about it."* (enumerator field observation) *"Some households have so much of patriarchy that women in those households are interested in savings but the son/husband is not ready for it."* (enumerator field observation) *"Lot of men tortures their wives. So when I asked them, they told me that they couldn't keep the boxes at their homes as they live in small place. Their husband might take money from the box."* (enumerator field observation) *"When it was told that the savings box is for entire family, one female participant asked: 'what if I save the money by hook or by crook and my husband takes it away for drinking alcohol?'"* (enumerator field observation)

Notes: Counts based on 40 focus group participants and ten enumerators. Quotes were translated from Marathi into English.

5 Conclusion

We set out to test the effectiveness of a portable saving device that was distributed to low-income slum dwellers in India's Maharashtra province. Based on a randomized controlled trial design, we reveal causal impacts on participants' total savings amounts while their temptation expenditures remained unchanged. We can therefore infer that the portable device worked – but that its purpose and usage diverged from what we had initially hypothesized. That is, rather than serving as a portable commitment device that helped increase participants' self-control, the device was primarily used with the intention to conceal individual savings and financial resources from others, most likely from a spouse, from in-laws, or other family members.

Some caveats are in order. A first limitation is the reliance on self-report data for measures other than savings held in the lockbox and in the zip purse. This might be particularly problematic for the outcome of temptation spending, which is likely prone to social desirability biases due to the cultural and social stigma attached to alcohol consumption, smoking and gambling. However, while our participants might well under-report their actual inclination to give in to temptations, we do not expect any systematic reporting differences between both study arms. Considering that our trial was set up with an active control group that was also given a saving device, we contend that participants in both study arms would have been equally exposed to potential interviewer demand or Hawthorne effects. More importantly, several recent studies have suggested that demand effects might not impair the validity of findings to a large extent. For instance, de Quidt and colleagues (2018) conducted a series of online experiments with 19,000 participants in which they manipulated information that was given on the researchers' hypothesis (weak treatment: "we expect that participants do X"; strong treatment: "you would do us a favour if you did X") and found very limited evidence for demand effects. Likewise, Mummolo and colleagues (2018) showed in another experiment that revealing the purpose of the experiment as well as the key hypotheses did not change the outcomes. Related to this, Jayachandran (2018) integrated a social desirability measure in baseline surveys but did not find heterogeneous treatment effects on gender attitudes for respondents with lower versus higher social desirability ratings.

Another shortcoming is that the final sample size of our study was lower than we had initially budgeted for. The primary reason for this lower turnout was that data collection efforts were thwarted by heavy flooding and temporary evacuations of several treatment locations. In consequence of this, our field experiment was – albeit

being decently powered for the main outcome analyses – possibly insufficiently powered to detect heterogeneity in treatment effects (see Porter, 2018; Karlan & Appel, 2016). This prevents us from drawing valuable policy guidance on possible targeting strategies or profiling candidates who may benefit the most (Frölich & Huber, 2015; Ravallion, 2009).

Lastly, our trial covered a follow-up timeframe of only six months. In view of this, a meta-analysis of existing saving promotion interventions has highlighted a negative association between treatment effect sizes and the length of follow-up, suggesting that program impact tends to "fade out" over time (Steinert et al., 2018). In line with this, it is unclear whether the treatment effects documented in this study as well as the usage of the portable saving device will last beyond the six months observed in here.

Overall, our results demonstrate positive effects of receiving a simple portable saving device in the form of a zip purse. Participants in the treatment group did not only report higher savings amounts but there was also tentative indication of increases in female empowerment – most likely activated by giving women ownership over a private saving device and thus individual decision-making power over how financial resources are used. In addition, we also reported decreases in levels of debt, which could have important positive future downstream impacts such as higher financial independence and substantial cost savings on high interest rates. A broader policy implication of our findings is that the distribution of *private* saving devices can likely help to boost savings rates among low-income individuals, and particularly among women who hold low financial bargaining power relative to their spouse. Our findings also motivate the conclusion that access restrictions to safe and private saving infrastructure appear to be a greater saving barrier than lack of self-control and vulnerability to temptations.

Acknowledgments: The authors wish to thank Kay Tuschen, Louisa Barzen and Aditi Kuber for their considerable conceptual contributions to the intervention and study design. We further wish to thank Henrike Sternberg, Lena Merkel and Regina Dworschak for substantial contributions to data collection and research design. The study would not have been possible without the dedication and local expertize of the research team in Pune, including Hrishikesh Patil, Aishwarya Joshi, Bhagyashree Patole, Sagar Ramdas Mankar, Dorthi Robin Joseph, Ravindra Dnyandeo Ambore, Aruna Hitendra Kamble, Sunita Zumbar Salunke, Sujata Sanjay Kale, Sangita Mahesh Kumbhar, Samina Abdul Shaikh, Shrutika Mungi, Pratibha Tenkale, Shilpa Tathe, Ashwini Dawre, Devika Ghule, Shewale Aishwarya Balkrishna, Umesh Ramdas Madane, Anita Dattu Bhalekar, Aditya Mahendra Chandanshive, Mugdha Vijay Narkar, Sarita Goraksh Murtadak, Daingade Vijay Pandurang, Abhijeet Kumar. The authors also wish to thank the SNEH Foundation and Shraddha Deo in Pune and Ashwini Kuber for their logistical support. We are also grateful for feedback from Nianbo Dong and Yulia Shenderovich with regard to the power calculations. Lastly, we thank the families participating in this study who contributed their time and data.

References

Aggarwal, S., Brailovskaya, V., & Robinson, J. (2020). Saving for Multiple Financial Needs: Evidence from Lockboxes and Mobile Money in Malawi. National Bureau of Economic Research.

Aggarwal, S., Francis, E., & Robinson, J. (2018). Grain today, gain tomorrow: Evidence from a storage experiment with savings clubs in Kenya. Journal of Development Economics, 134, 1–15. https://doi.org/10.1016/j.jdeveco.2018.04.001

Akbaş, M., Ariely, D., Robalino, D. A., & Weber, M. (2016). How to Help the Poor to Save a Bit: Evidence from a Field Experiment in Kenya (SSRN Scholarly Paper No. ID 2803856).

Aker, J. C., Sawyer, M., Goldstein, M., O'Sullivan, M., & McConnell, M. (2020). Just a bit of cushion: The role of a simple saving device in meeting planned and unplanned expenses in rural Niger. World Development, 128, 104772.

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

Ambec, S., & Treich, N. (2007). ROSCAs as financial agreements to cope with self-control problems. Journal of Development Economics, 82(1), 120–137.

Anderson, S., & Baland, J.-M. (2002). The Economics of Roscas and Intrahousehold Resource Allocation. The Quarterly Journal of Economics, 117(3), 963-95.

Anderson, C. L., Reynolds, T. W., & Gugerty, M. K. (2017). Husband and Wife Perspectives on Farm Household Decision-making Authority and Evidence on Intrahousehold Accord in Rural Tanzania. World Development, 90, 169–183.

Ashraf, N., Karlan, D., & Yin, W. (2010). Female Empowerment: Impact of a Commitment Savings Product in the Philippines. World Development, 38(3), 333–344.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. The Quarterly Journal of Economics, 121(2), 635–672.

Atkinson, J., de Janvry, A., McIntosh, C., & Sadoulet, E. (2013). Prompting Microfinance Borrowers to Save: A Field Experiment from Guatemala. Economic Development and Cultural Change, 62(1), 21–64.

Avdeenko, A., Bohne, A., Frölich, 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).

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), 141–167.

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), 848–86.

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), 289–300.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive Linear Step-up Procedures That Control the False Discovery Rate. Biometrika, 93(3), 491–507.

Bernhardt, A., Field, E., Pande, R., & Rigol, N. (2019). Household Matters: Revisiting the Returns to Capital among Female Microentrepreneurs. American Economic Review: Insights, 1(2), 141–160.

Bobonis, G. J. (2009). Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment. Journal of Political Economy, 117(3), 453–503.

Borooah, V. K. (2008). Caste, inequality, and poverty in India. Review of Development Economics, 9(3), 399–414.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77–101.

Brune, L., Gine, X., Goldberg, J., & Yang, D. (2011). Commitments to save: a field experiment in rural Malawi. Policy Research Working Paper Series No. 5748, The World Bank.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. Annual Review Economics, 2(1), 671–698.

Castilla, C. (2019). What's yours is mine, and what's mine is mine: Field experiment on income concealing between spouses in India. Journal of Development Economics, 137, 125–140.

Chowa, G. (2006). Savings performance among rural households in sub-Saharan Africa: the effect of gender. Social Development Issues, 28(2), 106.

Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2009). Portfolios of the Poor: How the World's Poor Live on USD 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), 61–96.

Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2018). The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution. Washington, DC: World Bank.

De Mel, S., McKenzie, D., & Woodruff, C. (2009). Are Women More Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns. American Economic Journal: Applied Economics, 1(3), 1–32.

Desai, S., Dubey, A., Joshi, B. L., Sen, M., Shariff, A., & Vanneman, R. (2010). Human Development in India: Challenges for a Society in Transition. New Delhi: Oxford University Press.

Dizon, F., Gong, E., & Jones, K. (2016). Does Financial Inclusion Exclude? The Effect of Access to Savings on Informal Risk-Sharing in Kenya. Mimeo.

Doss, C. (2013). Intrahousehold Bargaining and Resource Allocation in Developing Countries. The World Bank Research Observer, 28(1), 52–78.

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), 745–753.

Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. World Bank Economic Review, 17(1), 1–25.

Duflo, E., & Udry, C. (2004). Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices. National Bureau of Economic Research (NBER), Working Paper 10498.

Dupas, P., Karlan, D., Robinson, J., & Ubfal, D. (2018). Banking the Unbanked? Evidence from three countries. American Economic Journal: Applied Economics, 10(2), 257–297.

Dupas, P., & Robinson, J. (2013). Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. American Economic Journal:

Applied Economics, 5(1), 163–192.

Duvendack, M., Palmer-Jones, R., & Vaessen, J. (2014) Meta-analysis of the impact of microcredit on women's control over household decisions: methodological issues and substantive findings, Journal of Development Effectiveness, 6(2), 73–96.

Fiala, N. (2017). Business is Tough, but Family is Worse: Household Bargaining and Investment in Microenterprises in Uganda. University of Connecticut Department of Economics Working Paper Series, 2017–05.

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

Frölich, M., & Huber, M. (2015). Direct and indirect treatment effects: causal chains and mediation analysis with instrumental variables. Working Paper Series No. 2025:12, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Gang, I. N., Sen, K., & Yun, M. S. (2008). Poverty in rural India: caste and tribe. Review of Income and Wealth, 54, 50–70.

Glennerster, R., Walsh, C., & Diaz-Martin, L. (2020). A Practical Guide to Measuring Women's and Girls' Empowerment in Impact Evaluations. The Abdul Latif Jameel Poverty Action Lab.

Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. Science, 344(6186), 862–867.

Herskowitz, S. (2020). Gambling, Saving, and Lumpy Liquidity Needs. American Economic Journal: Applied Economics. Forthcoming.

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.

International Institute for Population Sciences (IIPS) (2018). National Family Health Survey (NFHS-4), India, 2015-16: Maharashtra. Mumbai: IIPS.

Iversen, V., Jackson, C., Kebede, B., Munro, A., & Verschoor, A. (2011). Do Spouses Realise Cooperative Gains? Experimental Evidence from Rural Uganda. World Development, 39(4), 569–578.

Jakiela, P., & Ozier, O. (2016). Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies. The Review of Economic Studies, 83(1), 231–268.

Jayashandran, S. (2019). Social norms as a barrier to women's employment in developing countries. WIDER Working Paper Series No. WP-2019-74.

Jayachandran, S. (2018). Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India. 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), 3393–3411.

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), 36–78.

Karlan, D., Savonitto, B., Thuysbaert, B., & Udry, C. (2017). Impact of savings groups on the lives of the poor. Proceedings of the National Academy of Sciences, 114(12), 3079–3084.

Karlan, D., & Zinman, J. (2018). 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, 275–294.

Ky, S., Rugemintwari, C., & Sauviat, A. (2016). Does Mobile Money Affect Saving Behavior? Evidence from a Developing Country (No. ID 2815090). Rochester, NY: Social Science Research Network.

Lastrapes, W. D., & Rajaram, R. (2016). Gender, caste and poverty in India: evidence from the National Family Health Survey. Eurasian Economic Review, 6(2), 153–171.

Levenson, H. (1973), Reliability of the I. P. and C. scales: a multidimensional view of locus of control. Proceedings of the American Psychological Association, Montreal.

Locke, E. A. & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. American Psychologist, 57, 705–717.

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

Mehta, A. K., & Shah, A. (2003). Chronic poverty in India: incidences, causes and policies. World Development, 31(3), 491–511.

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), 255–264.

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

Mummolo, J., & Peterson, E. (2018). Demand Effects in Survey Experiments: An Empirical Assessment (SSRN Scholarly Paper No. ID 2956147). Rochester, NY: Social Science Research Network.

Noponen, H., & Kantor, P. (2004), Crises, setbacks and chronic problems—the determinants of economic stress events among poor households in India. Journal of International Development, 16, 529–545.

Platteau, J.-P. (2000). Institutions, social norms, and economic development. Amsterdam: Harwood Academic Publishers.

Porter, K.I. (2018). Statistical Power in Evaluations That Investigate Effects on Multiple Outcomes: A Guide for Researchers. Journal of Research on Educational Effectiveness, 11(2), 267–295.

Quidt, J. de, & Haushofer, J. (2016). Depression for Economists. Working Paper No. 22973, National Bureau of Economic Research.

Ravallion, M. (2009). Should the Randomistas Rule? The Economists' Voice, 6(2), 1–5.

Rogers, T., Milkman, K. L., John, L. K., & Norton, M. I. (2016). Making the Best Laid Plans Better: How Plan-Making Increases Follow-Through. Mimeo.

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

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

Schaner, S. (2015). Do Opposites Detract? Intrahousehold Preference Heterogeneity and Inefficient Strategic Savings. American Economic Journal. Applied Economics, 7(2), 135–174.

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), 694–712.

Shefrin, H. M., & Thaler, R. H. (1992). Mental accounting, saving, and self-control. In G. Loewenstein & J. Elster (Eds.), Choice over time (p. 287–330).

Soman, D., & Cheema, A. (2011). Earmarking and Partitioning: Increasing Saving by Low Income Households. Journal of Marketing Research, 48, 14–22.

Soman, D., & Zhao, M. (2011). The Fewer, the Better: Number of Goals and Savings Behavior. Journal of Marketing Research, 48 (6), 944–957.

Steinert, J. I., Cluver, L. D., Meinck, F., Doubt, J.,& Vollmer, S. (2018). Household Economic Strengthening through Saving and Budgeting: Evidence from a Field Experiment in South Africa. Journal of Development Economics, 134, 443-466.

Steinert, J. I., Zenker, J., Filipiak, U., Movsisyan, A., Shenderovich, Y., & Cluver, L.D. (2018). Do Saving Promotion Interventions Increase Household Savings, Consumption, and Investments in Sub-Saharan Africa? A Systematic Review and Meta-Analysis. World Development, 104(C), 238-256.

Thaler, R. H. (1990). Anomalies: Saving, Fungibility, and Mental Accounts. The Journal of Economic Perspectives, 4(1), 193–205.

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 Bank (2017). Maharashtra - Poverty, growth, and inequality (English). India state briefs. Washington, D.C. : World Bank Group.

World Development Report (2015). Mind, Society, and Behaviour.Washington D.C.: World Bank.

World Development Report (2012). Gender Equality and Development. Washington D.C.: World Bank.

Wright, G., & Mutesasira, L. (2001). The Relative Risks to Poor People's Savings. Journal of Small Enterprise Development, 12(3), 33–45.

Appendix

Table A1. Sample Attrition

	Participants not completed endline
Treatment	0.01
	(0.01)
Female	0.10***
	(0.02)
Married	-0.03*
	(0.02)
Age	-0.00
	(0.00)
Education	0.00
	(0.00)
Employed	-0.03**
	(0.01)
Household Size	-0.00
	(0.00)
Income	0.00
	(0.00)
Baseline Savings	-0.01*
	(0.00)
Baseline Time Preference	0.00
	(0.00)
Observations	1482
R^2	0.048

*Note: *p<0.1, **p<0.05, ***p<0.01. Robust standard errors in parentheses.*

Table A2. ITT Estimates: Impact on Withdrawals and Deposits

	(1) **Total Past-Month Withdrawals**			(2) **Total Past-Month Deposits**		
	I	II	III	I	II	III
ITT: Received Program	-3445.25** (1444.96)	-3470.71** (1450.53)	-3542.79** (1525.09)	-482.42 (332.18)	-491.12 (332.48)	-512.00 (342.72)
Stratification Variables	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.02 (0.02)	0.01 (0.02)	no	0.06 (0.06)	0.04 (0.07)
Controls	no	no	yes	no	no	yes
Observations	1421	1421	1379	1421	1421	1379
Mean Control		**8765.12 (34986.16)**			**2315.65 (7344.15)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

Table A3. ITT Estimates: Impact on Withdrawals and Deposits

	Bank Savings	Mobile Money Savings	Savings Club	Home Savings	Savings held by Relatives
ITT: Received Program	1345.19 (1062.99)	60.31 (62.64)	-27.00 (83.21)	-218.19 (229.65)	3798.32 (2714.16)
Stratification Variables	yes	yes	yes	yes	yes
Lagged Outcome	0.05 (0.03)	-0.04 (0.03)	0.05 (0.06)	0.05 (0.03)	0.05 (0.06)
Controls	yes	yes	yes	yes	yes
Observations	1379	1379	1379	1379	1379

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected q-values in square brackets. Robust standard errors in parentheses. Estimates are for Model III, which includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

Table A4. ITT Estimates: Impact on Self-rated Temptations

	Self-rated Temptations		
	I	II	III
ITT: Received Program	-0.03 (0.02)	-0.03 (0.02)	-0.02 (0.02)
Stratification Variables	yes	yes	yes
Lagged Outcome	no	0.10*** (0.03)	0.09*** (0.03)
Controls	no	no	no
Observations	1417	1414	1414
Mean Control		**1.29** **(0.46)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income*

Table A5. CACE Estimates: Primary Outcomes

	(1) **Total Past-Month Savings**			(2) **Temptation Expenditures**		
	I	II	III	I	II	III
CACE: Used the portable device	13272.83** (6320.16) [0.072]	13439.03** (6306.20) [0.066]	13914.47** (6818.33) [0.082]	2.81 (30.58) [0.927]	2.746 (30.58) [0.928]	5.43 (31.16) [0.862]
Stratification Variables	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.09* (0.05)	0.10* (0.05)	no	0.00 (0.01)	0.00 (0.01)
Controls	no	no	yes	no	no	yes
Observations	1421	1421	1379	1421	1421	1379
Mean Control		**8400.83** **(20463.25)**			**82.27** **(262.25)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

Table A6. CACE Estimates: Secondary Outcomes

	(1) **Gender Attitudes Index (full sample)**			(2) **Female Empowerment Index (women only)**			(3) **Self-Efficacy Index**		
	I	II	III	I	II	III	I	II	III
CACE: Used the portable device	-0.01 (0.25) [0.973]	0.01 (0.25) [0.977]	-0.04 (0.25) [0.888]	0.09** (0.04) [0.144]	0.08** (0.04) [0.174]	0.08** (0.04) [0.114]	0.07 (0.26) [0.944]	0.07 (0.25) [0.937]	0.07 (0.25) [0.888]
Stratification Variables	yes	yes	yes	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.09*** (0.02)	0.07*** (0.03)	no	0.20*** (0.03)	0.17*** (0.03)	no	0.19*** (0.03)	0.17*** (0.03)
Controls	no	no	yes	no	no	yes	no	no	yes
Observations	1420	1417	1375	1186	1176	1134	1415	1409	1368
Mean Control		**-0.24** **(2.33)**			**-0.23** **(0.37)**			**6.00** **(2.41)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected sharpened q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

Table A6. (ctd.) CACE Estimates: Secondary Outcomes

	(4) **Resilience Index**			(5) **Outstanding Debt**			(6) **Household Expenditures**		
	I	II	III	I	II	III	I	II	III
CACE: Used the portable device	-0.02 (0.05) [0.944]	-0.03 (0.05) [0.937]	-0.02 (0.05) [0.888]	-1024.85* (556.43) [0.195]	-1040.04* (547.76) [0.174]	-1188.70** (571.79) [0.114]	1025.04 (3348.10) [0.944]	943.43 (3395.77) [0.937]	1304.79 (3453.89) [0.888]
Stratification Variables	yes	yes	yes	yes	yes	yes	yes	yes	yes
Lagged Outcome	no	0.22*** (0.03)	0.19*** (0.03)	no	0.01 (0.01)	0.01 (0.01)	no	0.08 (0.08)	0.08 (0.08)
Controls	no	no	yes	no	no	yes	no	no	yes
Observations	1420	1420	1378	1421	1421	1379	1421	1421	1379
Mean Control		**0.44** **(0.44)**			**1956.51** **(5319.36)**			**5767.05** **(29496.46)**	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected sharpened q-values in square brackets. Robust standard errors in parentheses. Model I includes trial arm and stratification variables, namely participant sex and baseline savings. Model II includes stratification variables (same as Model I) and the lagged outcomes. Model III includes additional controls: participants' age, marital status, educational status, employment, household size, household income.*

Table A7. Heterogeneity in Treatment Effects: Primary Outcomes

	(1) **Total Past-Month Savings**		(2) **Temptation Expenditures**	
	I **Main & Interaction Effect**	II **Total Effect**	I **Main & Interaction Effect**	II **Total Effect**
ITT Received Program	6414.72	6782.25*	43.17	-6.69
	(4673.90)	(3637.71)	(48.93)	(14.99)
x TRAIT	367.53		-49.86	
Female	(5922.70)		(51.18)	
	[0.951]		[0.660]	
ITT: Received Program	5484.95**	10904.61	6.03	-12.67
	(2313.22)	11654.22	(17.45)	(27.96)
x TRAIT	-2422.33*		-18.70	
Low Income	(1468.00)		(32.97)	
	[0.198]		[0.571]]	
ITT: Received Program	9796.39*	3211.85	11.30	-8.12
	(5536.30)	2302.87	(23.04)	(18.44)
x TRAIT	-6584.54		-19.42	
Present Bias	(5996.15)		(29.51)	
	[0.511]		[0.511]	
ITT: Received Program	596.46	8499.89*	23.01	-15.87
	(3132.58)	(4571.88)	(14.66)	(18.74)
x TRAIT	7903.44		38.88*	
Female Involvement	(5542.12)		(23.78)	
	[0.154]		[0.154]	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected sharpened q-values in square brackets. Robust standard errors in parentheses. Specification I presents results from OLS regressions with socioeconomic controls, stratification variables and the lagged outcome variable (equivalent to Specification III in the main outcome analyses). Specification II shows the total effect for participants exhibiting a given TRAIT, which corresponds to the sum of the main effect and the interaction effect as estimated in the previous column and includes corresponding significance levels. Low income is coded 1 for the lowest income quantile of the sample. Present bias is a composed index of four items ("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") and centered around zero. Present bias is coded 1 if the index score is greater than 0, thus indicating a higher level of present bias. Female involvement denotes whether the female spouse (partner) is actively involved in financial decision making processes within the household.*

Table A8. Heterogeneity in Treatment Effects: Secondary Outcomes

	(1) **Gender Attitudes**		(2) **Female Empowerment Index (women-only sample)**		(3) **Self-Efficacy Index**	
	I Main & Interaction Effect	II Total Effect	I Main & Interaction Effect	II Total Effect	I Main & Interaction Effect	II Total Effect
ITT: Received Program	0.33	-0.07	/	/	-0.03	0.07
	(0.30)	(0.14)			(0.26)	(0.14)
x TRAIT	-0.39				0.09	
Female	(0.33)				(0.30)	
	[0.369]				[0.760]	
ITT: Received Program	0.45	-0.21	0.02	0.09**	0.05	0.05
	(0.14)	(0.28)	(0.02)	(0.04)	(0.14)	(0.28)
x TRAIT	-0.25		0.07		0.00	
Low Income	(0.31)		(0.05)		(0.31)	
	[0.830]		[0.348]		[0.993]	
ITT: Received Program	0.04	-0.03	0.02	0.06**	0.15	-0.05
	(0.17)	(0.17)	(0.03)	(0.03)	(0.18)	(0.18)
x TRAIT	-0.08		0.05		-0.20	
Present Bias	(0.25)		(0.04)		(0.25)	
	[0.970]		[0.970]		[0.970]	
ITT: Received Program	0.02	-0.09	0.09*	0.03	-0.32	0.17
	(0.31)	(0.15)	(0.0)	(0.02)	(0.31)	(0.16)
x TRAIT	-0.11		-0.06		0.49	
Female Involvement	(0.35)		(0.05)		(0.35)	
	[0.889]		[0.889]		[0.448]	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected sharpened q-values in square brackets. Specification I presents results from OLS regressions with socioeconomic controls, stratification variables and the lagged outcome variable (equivalent to Specification III in the main outcome analyses). Specification II shows the total effect for participants exhibiting a given TRAIT, which corresponds to the sum of the main effect and the interaction effect as estimated in the previous column and includes corresponding significance levels. For variable specifications, see Table A6.*

Table A8. (ctd.) Heterogeneity in Treatment Effects: Secondary Outcomes

	(4) **Resilience Index**		(5) **Outstanding Debt**		(6) **Household Expenditures**	
	I Main & Interaction Effect	II Total Effect	I Main & Interaction Effect	II Total Effect	I Main & Interaction Effect	II Total Effect
ITT: Received Program	0.05	-0.03	1163.94	-837.90***	4536.09	-254.05
	(0.05)	(0.02)	(890.81)	(262.67)	(4103.31)	(1860.90)
x TRAIT	-0.08		-2001.84**		-4790.14	
Female	(0.06)		(925.12)		(4575.00)	
	[0.369]		[0.155]		[0.369]	
ITT: Received Program	-0.02	-0.00	-228.36	-1450.86***	722.26	-247.27
	(0.03)	(0.05)	(306.68)	(546.94)	(2167.81)	(451.68)
x TRAIT	0.01		-1222.50		-969.53	
Low Income	(0.06)		(627.98)		(2231.80)	
	[0.969]		[0.312]		[0.969]	
ITT: Received Program	-0.02	-0.01	-367.01	-639.40	817.17	272.87
	(0.03)	(0.03)	(304.90)	(423.49)	(2274.30)	(2418.19)
x TRAIT	0.00		-272.39		-544.29	
Present Bias	(0.05)		(500.30)		(3284.80)	
	[0.970]		[0.970]		[0.970]	
ITT: Received Program	0.03	-0.04	-573.79*	-923.77***	-338.78	-357.69
	(0.05)	(0.03)	(311.00)	(325.32)	(450.74)	(2420.46)
x TRAIT	-0.07		-349.99		-18.91	
Female Involvement	(0.06)		(450.11)		(2383.92)	
	[0.655]		[0.448]		[0.990]	

*Notes: *p<0.1, **p<0.05, ***p<0.01, based on naïve p-values. Multiple hypothesis corrected sharpened q-values in square brackets. Specification I presents results from OLS regressions with socioeconomic controls, stratification variables and the lagged outcome variable (equivalent to Specification III in the main outcome analyses). Specification II shows the total effect for participants exhibiting a given TRAIT, which corresponds to the sum of the main effect and the interaction effect as estimated in the previous column and includes corresponding significance levels. For variable specifications, see Table A6.*

Previous Munich Papers in Political Economy:

Betz, Timm and Amy Pond. “Political Ownership”, MPPE No. 1/2020, Munich. (ISSN)2701-3456

Chatziathanasiou, Konstantin, Hippel, Svenja and Michael Kurschilgen. “Property, Redistribution, and the Status Quo. A laboratory study”, MPPE No. 2/2020, Munich. (ISSN)2701-3456

Becker, Annette, Hottenrott, Hanna and Anwesha Mukherjee. “Division of Labor in R&D? Firm Size and Specialization in Corporate Research”, MPPE No. 3/2020, Munich. (ISSN)2701-3456

Steinert, Janina Isabel, Satish, Rucha Vasumati, Stips, Felix and Sebastian Vollmer. “Commitment or Concealment? Impacts and Use of a Portable Saving Device: Evidence from a Field Experiment in Urban India, MPPE No. 4/2020, Munich. (ISSN)2701-3456

Impressum:

ISSN: **2701-3456**
Editors: Tim Büthe, Hanna Hottenrott
Associate Editors: Timm Betz, Sebastian Goerg, Eugénia da Conceição Heldt, Michael Kurschilgen, Amy Pond, Sebastian Schwenen, Janina Steinert, Matthias Uhl
Managing Editor: Luca Messerschmidt

Contact: Technicial University of Munich, Arcisstraße 21, 80333 München
mppe@gov.tum.de, mppe@wi.tum.de
https://www.wi.tum.de/mppe/
Twitter: @MunichPapers