Saving for Multiple Financial Needs: Evidence from Malawi*

Shilpa Aggarwal† Valentina Brailovskaya‡ Jonathan Robinson§

December 27, 2018

Abstract

We test whether multiple labeled savings accounts affects savings decisions and downstream outcomes in a field experiment with 481 entrepreneurs in urban Malawi. Treatment respondents received either one or multiple accounts, while a control group received nothing. Multiple accounts increased savings in treatment accounts by about 30%. Savings accounts had sizeable effects on a number of outcomes, including labor supply, farming decisions, household expenditures, land purchases, credit extended to customers, and interpersonal transfers. However, we find no evidence that multiple accounts had larger downstream effects than single accounts.

JEL Codes: D14, L26, O12, O16

Keywords: savings, liquidity, labeling, mental accounting, lockboxes, goal-setting

*We are extremely grateful to Calvin Mhango and Gabriella Fleischman for overseeing the project and to Marble Karuu and David Park for research assistance. We thank Carly Farver for her continuous support, and IPA Malawi for implementing the research protocol in Malawi. We thank the Financial Inclusion Program at IPA, ISB, and UCSC for financial support. We are grateful to Sripad Devalkar for useful discussions. We thank seminar participants at the 2018 Williams CDE conference, UCSC, and IDinsight for valuable feedback. This research protocol was approved by the the National Commission for Science and Technology in Malawi and the IRBs of UCSC and the ISB. This trial is registered in the American Economic Association's registry for randomized controlled trials (AEARCTR–0002188). A pre-analysis plan is available on the AEA registry website. All errors are our own.

†Indian School of Business, email: shilpa_aggarwal@isb.edu

‡IDinsight, email: valentina.brailovskaya@idinsight.org

§University of California, Santa Cruz and NBER, email: jmrtwo@ucsc.edu

1 Introduction

Most people have multiple concurrent financial goals. For example, it is common for households to be saving up for large indivisible investments such as buying a house or paying for higher education, while also setting aside smaller amounts for day-to-day expenses or for dealing with unforeseen emergencies.[1] How does one effectively save towards multiple goals simultaneously? One potential strategy is to create separate labeled accounts. Research in behavioral economics suggests that once these accounts are created, withdrawals for any purpose other than the labeled one impose a utility cost on the account-holder (Ainslie 2001; Benabou and Tirole 2004; Koch and Nafziger 2016; Thaler 1990; 1999). Previous studies have shown that creating a single labeled account increases the probability of reaching the labeled goal by making money less fungible across uses (i.e. Brune et al., 2014; Dupas and Robinson 2013a; Karlan and Linden, 2014), and this finding likely generalizes to having more than one labeled account.

As a practical matter, however, it is not clear how one can accomplish the cognitively challenging task of keeping track of distinct sums of money that have been mentally allocated towards different purposes.[2] We conjecture that the effectiveness of mental accounts will likely be enhanced when accounts are accompanied by the physical separation of money. The practice of physically separating pots of money meant for distinct uses has precedent,[3] although it is not known if this method actually leads to an increase in deposits.

To test the efficacy of physically separated accounts in facilitating savings, we conduct an experiment with 481 micro-entrepreneurs in the city of Blantyre, Malawi. One treatment group was offered a single metal lockbox in which to save up for their goals, while a second group was offered multiple lockboxes (up to 3). A third group served as a control group. One of the main contributions of the paper is to carefully examine effects on a range of downstream outcomes. All respondents were given cell phones, and half of the sample was called once or twice per week to measure several outcomes at high frequency, including savings decisions, labor supply, income, expenditures, and transfers. In order to supplement the high frequency data and to collect information on the full sample, we also conducted two rounds of monitoring surveys with all respondents.

We have two main sets of results. First, we find clear evidence that respondents who were given multiple accounts saved more. While we find near-universal take-up in both treatments (92% of respondents used a project box at least once), people who were offered multiple boxes saved 27-34% more on average. Single account group reported depositing $117 in the first 5 months of having an account, while those in the multiple account group deposited about $40 more. This is substantial sum in this context given daily business profits are roughly $3.50 . In a debriefing survey at endline,

[1] In our study context, the average respondent reported having 2.4 savings goals at baseline.

[2] Research on the cognitive costs of scarcity (i.e., Carvalho, Meier and Wang, 2016; Mani et al., 2013; Shah, Mullainathan, and Shafir, 2012) suggests that this task might be even harder for the poor.

[3] For example, see this oft-quoted anecdote cited in previous work (i.e. Zelizer 1994; Soman and Cheema 2016), from Alice Bradley (1923): *"Take for instance Mrs. M's system as she told it to* Women's Home Companion *in the early 1920's: "I collected eight little cans, all the same size, and pasted on them the following words, in big letters: groceries, carfare, gas, laundry, rent, tithe, savings, miscellaneous.... [W]e speak of those cans now, as the grocery can, carfare can, etc."*

we found that the usefulness of multiple accounts was driven largely by the ability to set distinct goals for each account, followed by the fact that physical separation served as a deterrent to large withdrawals. A small minority of respondents reported that multiple accounts helped them save by diversifying the risk of theft and made it easier to hide money from others.[4]

Second, and in contrast to most of the existing literature which finds modest effects, we find strong evidence that the accounts had effects on a host of downstream outcomes. Treatment respondents reduced labor supply in their main business, and increased the time they spent working on their farms. They also invested in more farm inputs, and were more likely to purchase land. We find that treatment respondents substantially increase the amount of credit extended to customers in their main business, suggesting increased liquidity. We also find that treatment respondents increased expenditures on a number of categories.. Finally, we find that treatment respondents were more likely to give transfers to friends and neighbors and we find some evidence that they were more likely to receive them as well.

We make several contributions to the literature. We are one of the first papers to examine the effects of multiple savings accounts.[5] There have been many recent papers which have studied the effect of providing unbanked households with savings devices, including basic savings accounts,[6] commitment accounts which limit liquidity through external restrictions (i.e. Ashraf, Karlan, and Yin 2006; Beshears et al., 2015; Buehren et al. 2018; John 2018), or accounts with softer commitment such as labeling (i.e. Dupas and Robinson, 2013b; Karlan and Linden 2016; Habyarimana and Jack 2018). These studies typically compare a treatment group which receives an account to a control group which receives nothing, and thus the mechanism by which the accounts increase savings include several possible channels. Among other reasons, these include that savings accounts provide security, limit liquidity, or because they encourage the activation of mental accounts for particular goals. Our experimental design isolates the effect of an additional, physically separate savings location, since other pathways such as security are equalized across these groups. By documenting increases in savings for the multiple account group over and above those obtained by the single box group, we are able to show that physical separation of cash could be an important way to increase savings.[7]

A primary difference with much of the existing literature is that take-up and usage of the accounts in this study was much higher than in previous work on savings, including studies that have been set in Malawi.[8] The limited take-up in most of the literature, combined with heterogeneity in

[4]This is consistent with previous studies which document the presence of social pressure to share income with others (Jakiela and Ozier 2011, Platteau 2000, Dupas and Robinson 2013a).

[5]There is a small literature on a related issue about partitioning consumption items between physically separated places. Soman and Cheema (2011) report experiments in which experimental subjects were paid in different numbers of accounts (i.e. money split into multiple envelopes or chocolates split into separate packages). The authors find that immediate consumption is decreasing in the number of accounts.

[6]Among others, see Agarwal et al. (2017), Dupas and Robinson (2013a), Dupas, Keats, and Robinson (2018), Kast, Meier, and Pomeranz (2018), Kochar (2018), Prina (2015), and Schaner (2018).

[7]In addition to simplifying the process of keeping track of money meant for different purposes, physical separation could also work through risk diversification, i.e., by decreasing the likelihood that all of one's money could get stolen or claimed by others, or by imposing additional transaction costs on each withdrawal.

[8]For example, Dupas et al. (2018) find that only 41% of people in rural Malawi used experimental bank accounts

goals that people have, has made it hard to find downstream outcomes. A handful of previous studies have found statistically significant effects on downstream outcomes such as business investment and personal expenditures (Dupas and Robinson 2013a), health investment (Dupas and Robinson 2013b), educational expenditures (Prina, 2015), labor supply (Callen et al., 2014), self-reported financial well-being (Kast and Pomeranz, 2014; Prina, 2015) and debt (Kast and Pomeranz 2014). However, most of these studies find outcomes on only one of several potential outcomes, while other studies find no effects at all.[9] By contrast, we are unusually well-powered to find downstream effects because of high take-up and because our surveys measured a comprehensive set of outcomes at high frequency.

Because we find so many effects on downstream outcomes, it is difficult to discuss them all in detail here. In brief however, several of our findings have not (to our knowledge) been shown before, such as effects on land rentals and purchases[10] or on credit to customers.

Other results relate closely to previous recent work. We find that treatment respondents actually worked *less* in their main business, but worked more in farming, suggesting that on the margin the returns to farming exceed those to running a small business. This result is similar to several others in the literature which show that poor access to credit or savings may lead to suboptimal labor supply decisions. Improved access to credit, by relaxing the subsistence constraint, allows individuals to move away from occupations that may otherwise have low returns but pay during lean periods. This has been shown by Fink, Jack, and Masiye (2018), who find that providing credit to smallholder farmers decreases off-farm labor and increases own-farm labor, and by Pitt and Khandker (2002) , who show that agricultural households in Bangladesh take up secondary occupations to smooth consumption during the lean season and as insurance against weather shocks, and that credit may serve as an alternative source of funds for these purposes. In the same setting of rural Bangladesh, Bandiera et al. (2017) show that expanding access to credit for the poorest women enables them to reallocate labor towards higher-return occupations. On the savings side, improved access may work through relaxing subsistence constraints, in which case, one would observe a re-optimization of hours across occupations, as in Callen et al. (2014), who provide evidence from Sri Lanka that households given access to deposit collection increase their hours in wage work, and some exit self-employment altogether. Alternatively, savings accounts may affect labor supply through increasing the marginal benefit of working by allowing account-holders to save money for lean periods, when the marginal value of consumption is high. This is especially pertinent in the case of agricultural workers, and work by Brune, Chyn and Kerwin (in-progress) shows that agricultural wage-workers given access to savings via deferred wages increase their labor supply in that very occupation.

Our finding that treatment respondents gave more to social networks contributes to the literature on the relationship between access to savings and informal insurance networks. Ligon, Thomas

and the average amount saved was only $75, compared to 95% and $117 here.

[9] See Table 3 in Prina (2015) and Figure 5 in Dupas et al. (2018) for a summary of the effects found in these studies.

[10] While there is a large literature documenting the phenomenon of underinvestment in agriculture due to financial constraints (i.e., Croppenstedt et al. 2003, Flory 2018, Karlan et al. 2014), this strand of research is largely limited to inputs, and does not usually extend to factors of production like land. See Jack (2011) for a review.

and Worrall (2000) show theoretically that access to savings may reduce interpersonal risk sharing as the returns to autarky increase in the presence of savings. Chandrasekhar, Kinnan and Larreguy (2018) test this in a lab-in-the-field setting and find no impact of savings on risk-sharing, which they explain through the competing effects of social distance on the benefits from reneging in the presence of savings and the amount of risk-sharing in the absence of savings. Field experiments on savings provide mixed evidence in this regard, potentially because there is variation in the observability of saved amounts and the liquidity of savings vehicles provided in these studies, as well as in the structure of the underlying social network. Our results are in line with Dupas, Keats and Robinson (2018), who find that Kenyan households transfer more but receive no less when they receive accounts, and Flory (2018), who finds that rural Malawians who receive accounts make transfers to other households in the village during the hungry season, strengthening the local insurance network; though in contrast to Dizon, Gong and Jones (2018) who find that savings accounts given to vulnerable women in Kenya reduce their participation in risk-sharing networks.[11]

Finally, our finding that expenditures increased may sound counter-intuitive given that savings accounts would be expected to *reduce* spending, at least for some period of time (i.e., Bachas et al. 2018, Somville and Vandewalle 2018). However, a key difference between these papers and ours is the process through which increased savings come about. In the setting considered by Bachas et al. (2018), savings were specifically intended to serve as a precautionary buffer stock. In Somville and Vandewalle (2018), the additional savings came about simply as the outcome of a default option, similar to Carroll (2009) and Madrian and Shea (2001), and not due to the intention to purposely save towards a particular goal. In both the cases, therefore, savings are not intended to be used towards income-generating investments. In contrast, in our study, although we do not directly observe an increase in income, it is very likely that incomes went up given our findings regarding increased farming as well as credit to customers. It is possible, therefore, for savings and spending to comove due to an increase in income. In this sense, our result is similar to Dupas and Robinson (2013a) and Prina (2015) who also observe an increase in several categories of spending.

2 Experimental Design and Data

2.1 Context and sampling

The study described in this paper is part of a larger experiment which also included an evaluation of mobile money (see Aggarwal, Brailovskaya and Robinson 2018a for those results). The experiment took place with a representative sample of small entrepreneurs operating in Blantyre, the second largest city in Malawi. While Blantyre is an urban center with a population just over 1 million, the outskirts of the city contain farmland. Blantyre contains 26 wards and 392 enumeration areas (EAs). To construct a sample with coverage across the city, we aimed to randomly selected three EAs in each ward, ultimately selecting 77 (one EA did not have 3 wards).

[11] Related to this, Comola and Prina (2017) provide evidence from Nepal that access to formal savings alters the informal insurance networks leading to a greater diffusion of gains than is immediately apparent.

Market structure is heterogeneous across EAs – the number of business ranged from 0 to 1,649 (mean 104, median 48).[12] Because of the high number of businesses in some EAs, it was not logistically possible to census every business. We therefore decided to divide EAs between those with more than 100 and those with less than 100 businesses. In the smaller EAs, we censused all businesses; in the larger EAs, we counted all businesses but only censused a randomly selected subset of approximately 40% of businesses.[13] Please note however, that s We counted a total of 9,848 businesses and classified 8,078 (82.1%) of these as small businesses.[14] We attempted to conduct a census survey with 3,857 businesses and completed surveys with 2,842 (74%).[15]

After the census, we imposed additional exclusion criteria. First, we excluded any business that did not meet our definition of a small business, which we defined as having no more than 2 employees (6% of the census list). Second, we excluded businesses in which the business owner was operating as a mobile money agent (3%) as our companion study was about providing access to mobile money accounts. Third, we excluded businesses in which the owner was not actively involved in running operations (defined as working there at least 5 days per week) since such owners would not be able to reliably answer business-related questions (9%). Fourth, we excluded businesses that were planning to shut down within 6 months of the census, i.e., before the project was slated to end (16%).

Once we had a sample of businesses that met our criteria, we made two other exclusion criteria based on our planned data collection. First, we removed all polygamous households, which amounted to 5% of the sample. Second, since we initially planned to collect surveys with paper-and-pencil logbooks (we eventually changed to phone surveys), we excluded business owners who were illiterate (about 20% of the sample) and those whose eyesight prevented them from reading the printed page (about 10% of the sample).[16]

These exclusion criteria left us with approximately 1,640 eligible businesses from which we drew our final sample, stratified by financial access (defined by either having a mobile money or a bank account) and self-reported distance to the nearest mobile money agent (defined as above or below the sample median).[17] In drawing the sample, we chose to oversample businesses connected to the electricity grid: while 26% of eligible businesses were connected to the grid, we sampled 35%.[18] We replaced respondents who could not be found (about 6.5%) or refused to participate (another 6.5%) with randomly chosen backups, ultimately yielding a sample of 801 businesses for the two studies (the current paper and the mobile money evaluation). After removing the mobile money

[12]Two EAs contained no small businesses. One was an industrial area and the other was farmland.

[13]Since we counted all businesses, we have sampling weights for all EAs.

[14]We excluded several classes of businesses in this exercise since they were unlikely to qualify as a small business. This included gas stations, clinics, hospitals, banks, microfinance institutions, manufacturing plants, warehouses, wholesalers and supermarkets.

[15]Of the 1,012 businesses that were not censused, 552 (14%) refused to participate (either before or after we were able to explain the study), 346 (9%) were permanently closed, 114 (3%) were not reached (either because the shop was closed after 3 visits or the owner was under 18 years old).

[16]We identified these households either at the census or prior to the baseline survey.

[17]This last criteria was put in place for the mobile money evaluation.

[18]This decision was made to improve the power of the related paper, Brailovskaya (2018), which utilizes some of the data in this experiment to calculate the effect of power outages on business outcomes.

sample, our study includes 481 respondents, evenly split between a group offered one box, a group offered multiple boxes, and the control group.[19]

2.2 Experimental design

The experimental design is summarized in Web Appendix Figure A2 and the timeline of project activities is shown in Web Appendix Figure A3. As discussed below, the experimental design cross-cut the provision of savings accounts with the frequency of surveying.

2.2.1 Lockboxes

The savings device we offered were lockboxes, and were very similar to those offered in Dupas and Robinson (2013b) and subsequent projects.[20] These were metal boxes with a deposit slit in the top, and a latch that could be locked. The boxes were produced by a local artisan and cost about $3.40 (at wholesale prices). Respondents were also given a lock and key, worth about $1. These boxes were not commercially available in Malawi at the time. While a sizeable minority of people (22%) had lockboxes at baseline, these were considered to be of lower quality than the project boxes. Existing boxes were typically made of wood or cardboard, and either could not be locked or had to be broken to be opened. Finally, as in Dupas and Robinson (2013b), respondents were given a passbook to record withdrawals and deposits, and so that they could track the balance without having to open their box(es).

Those in the single box treatment group were offered only one box, while those in the multiple box group were offered up to 3 lockboxes. To differentiate the boxes, they were painted different colors – everyone received a silver box, while the second and third boxes were painted black and brown respectively. For those in the multiple box groups, the project passbooks allowed for separate tracking of deposits and withdrawals for each box.

During baseline, we asked all respondents about their current savings goals. At the time when boxes were handed out (which was about 2 months after baseline), those in the treatment groups were encouraged to use the project box(es) to save towards their savings goal. Specifically, we asked respondents to write their savings goal on a sticker, which was then attached to the lid on the inside of their project boxes. Summary statistics of goals are reported in Web Appendix Table A1. Column 1 shows goals elicited at baseline for the entire sample, while the next 2 columns show goals elicited at a lockbox check, which was conducted about halfway through the study and covered a random subset of the treatment group only. Therefore, the goals summarized in Columns 2 and 3 are for the treatment groups only (note that people could pick more than one goal for the account). Respondents had a variety of savings goals, of varying duration. At baseline 64% of business owners save for business reinvestment, 40% save up for emergencies, 42% save up for regular household

[19] See Web Appendix Figure A1 for the geographic distribution and spread of the various treatment arms across the city of Blantyre.

[20] These include Ashraf, Karlan and Yin (2006), Karlan and Linden (2014), Dupas, Keats and Robinson (forthcoming), Francis (2018) and Herskowitz (2018), among others.

expenses, 18% save up for buying land, and 12% each save up for buying durable goods and for children's education (with smaller percentages on other items). At the lockbox check, we asked box respondents for the goals exclusively on their project accounts and find similar heterogeneity. However, the prevalence of each goal is lower, since there could be other saving places where they accumulate savings besides the project box accounts. As expected, multiple box groups are more likely to save for each goal.

2.2.2 Phone surveys

A primary goal of this project was to measure the effect of savings accounts at high frequency. While this has been attempted before (i.e. Dupas and Robinson 2013a), obtaining high quality data is challenging. In this project, we opted to measure outcomes using high-frequency phone surveys (described in more detail in section 2.3).[21] To implement this effectively, we gave everyone in our sample a basic feature phone (worth about $12 in Malawi). We did this even though 95% of the sample had a phone at baseline, because we were concerned that the phones people already owned were of low quality and might break or need to be replaced prior to the end of the experiment.

Since it is possible that the high-frequency survey itself is a treatment, we randomly split the sample into two groups: one was administered a high-frequency phone survey (which we call the "HFPS") while the other was not. To measure the effect of surveying itself, we administered two monitoring surveys to the entire sample, so that we could compare responses between those given high frequency surveys and others who were not. A detailed analysis of the effect of high frequency surveying can be found in the companion paper Aggarwal, Brailovskaya and Robinson (2018b), which finds that the HFPS had effects on several outcomes (likely through focusing attention on those outcomes) but had no effect on the estimated treatment effect of the accounts. For the purposes of this paper, we show results in separate panels for the HFPS and the monitoring survey. We believe that the HFPS results are superior for many of the outcomes because the recall period in the HFPS is only a few days, whereas it is much longer in the monitoring surveys. However, many of the results are similar across the different surveying methods.

2.3 Data

We utilize information collected through 5 different types of surveys to conduct our analysis. First, the baseline survey contains a host of questions on household and demographic characteristics, business outcomes, savings, cash flows, and related measures.

Second, we have two main sources of data to measure treatment effects: the high-frequency phone surveys (HFPS) and two monitoring surveys. The HFPS measured business outcomes and labor supply at the daily level, and household expenditures, transfers, savings, credit, shocks and related outcomes at the weekly level. The HFPS was conducted in two waves, one in September-October 2017 (covering 8 weeks) and another in February-March 2018 (covering 6 weeks). In Wave

[21] We initially planned to ask respondents to keep pencil-and-paper logbooks (which is why we screened on literacy and on eyesight), but ultimately changed to phone surveys.

1, respondents were called twice per week, with the 2 calls being 3-4 days apart. In one of the weekly interviews, the respondent was administered a "short" survey which took about 15 minutes and which asked about business outcomes over the past 3 days (day by day). The survey took about 15 minutes to complete. The other "long" survey took about 40 minutes and included all the questions in the short call, but also added a recall module for other outcomes that were expected to be rarer or more memorable and thus could be reliably remembered over a week. These included shocks such as household illness and funerals, deposits and withdrawals from various saving source, and transfers given and received. The long survey also included questions about expenditures over the past 7 days. Due to budgetary constraints, respondents were called only once per week in Wave 2 and were administered only the "long" version of the survey from Wave 1.

Respondents were randomized into which days they were to be called, and this day was kept the same throughout the project.[22] Respondents were able to pick the time of day when they wished to be called. To encourage compliance, respondents were given US $2 in airtime for each week of the survey and were also enrolled into a lottery in which there was a 1/4 chance of winning an additional US $3. Airtime was sent directly to the respondents' phones after each phone call was completed.[23] We control for lottery payments in all HFPS regressions. If the respondent was not reached on the original day of the phone call after 3 attempts (which were spaced by at least an hour on the day of the call), a make-up call was scheduled for the following day. During this call, we asked for information for the preceding 4 days in order to reconstruct the lost days. However, we capped this at a recall of only up to 4 days since we expected that longer recall would be problematic.

Third, we conducted two monitoring surveys, after each round of the HFPS, in January 2018 and in March 2018. These were also done over the phone, but (as discussed in section 2.2) included the entire sample. These surveys took approximately 75 minutes to complete, and included a host of questions similar or identical to the HFPS, though over a longer recall period of up to 3 months for some outcomes. The surveys also included a number of other questions, such as questions about self-reported financial security, land purchases, and tuition payments. In addition, the second monitoring survey (the endline survey) included some debriefing questions about people's experience with the boxes, pressures to share money (if any) and experiences with the surveys themselves. Respondents were compensated $2 USD via airtime per survey round.

Fourth, as an independent verification of the information collected from respondents during the surveys, we conducted an unannounced in-person "lockbox check" visit with a random sub-sample of respondents in the box groups in December 2017. During this visit, we first asked a number of questions about usage of the project boxes, including cumulative deposits and withdrawals since July 2017 as well as the current balance. After the survey, enumerators requested to visually inspect the box (as in Dupas and Robinson 2013b) in order to verify the balance. Seventy-nine percent of

[22]Out of 240 HFPS respondents, only 1 respondent asked to be called on different days and the rest complied with the original day assignment.

[23]Respondents were enrolled in the lottery even if they were not reached for a specific call, but they did not receive the US $2 payment for survey completion.

respondents who we interviewed were willing and able to open at least one of the boxes.[24] In a regression of the verified amount in the box on self-reported deposits during the box-check survey, we find the coefficient very close to 1 (1.05), which we take as evidence that self-reporting of box usage was accurate. [25]

Finally, we administered a short intake survey at the time of account opening, i.e., when the boxes were handed out. This survey included many of the questions that would later constitute the HFPS, recalled over the 3 days prior to account opening. These pre-treatment values are used as control variables to improve power of the main regressions (Bruhn and McKenzie 2009), so that our specifications in HFPS regressions are always ANCOVA.

2.4 Attrition

Attrition for the different surveys is shown in Web Appendix Table A2. Columns 1-4 show attrition during the HFPS, with the odd columns showing whether a respondent appears at least once and the even columns showing the percentage of calls that were successfully completed. In round 1, 99% of respondents completed at least one survey and 89% of calls were made; in round 2, survey completion fell to 84% and 74%, respectively.[26]While not unexpected, lower compliance in round 2 points to the problem of conducting phone surveys in general, as people lose their phones, change phone numbers, or become fatigued with the surveys. We do not find any differences in attrition across the various box treatment groups and the coefficients on the treatment indicators are not significant for any of the surveys. However, HFPS respondents were 12 percentage points less likely to complete a first monitoring survey, which might be because HFPS respondents were more likely be fatigued by the surveying process. Survey completion for the second monitoring survey is balanced across HFPS and non-HFPS respondents.

2.5 Summary statistics and randomization check

Summary statistics and a check of randomization balance are presented in Table 1 (from the baseline survey) and Web Appendix Table A3 (from the intake survey). From Table 1 Panel A, 46% of the sample is male and the average business owner is 34 years old and has 9 years of education. Ninety-three percent of respondents have an iron roof on their homes (suggesting that these business-owners are less poor than the average Malawian – in the 2016 Integrated Household Survey for the country, only about half the households reported having iron sheets as the material of their roof) and the average value of assets owned by respondents households is $873. Panel B shows statistics on business outcomes. Sixty-eight percent of the businesses are in retail, with the remainder

[24]Of the remaining 21%, 5.4% could not open the box because they did not have the key with them, 9.4% refused to open the box or travel home to show the box to the field officer, 2.4% did not have access to the boxes at box checks, 3.3% respondents with boxes were not checked because the respondent could not travel home with the field officer on the day of the survey.

[25]

[26]In round 1, 1.25% of respondents were reached but refused to do the survey; in round 2, this figure was 6%. The remaining attriting business owners were not reached, which could be due to respondents changing numbers, moving, or because they were screening our phone calls.

predominantly in services, which includes occupations such as barbershops, tailoring, and welding. These businesses are very small: average weekly profits are about $19 per week and the average firm has only $293 in equipment and inventory.

Panel C shows statistics on savings. Average savings across all sources was $120 at baseline, split across an average of 2.5 savings places. We also observe that people already engage in physical separation of cash for different goals: the average respondent has 2.4 goals and saves up for these goals in different places. In particular, 78% of people have separate saving places for different goals, and only 35% save for more than 1 goal in a single savings place. The most common saving place is keeping cash at home, reported by 82% of respondents, suggesting that despite access to more formal financial services in an urban environment (such as banks), transaction costs and accessibility are a still major deterrent to usage. Saving groups (VSLAs and ROSCAs) are used by 52% of the business owners, and 47% report using mobile money accounts to save. [27]Thirty percent of the sample have access to a bank account and 22% save in saving boxes. Note that a sizable minority of the sample already saves in lockboxes; as we see by the take-up rates, however, the project lockboxes were seen as being of higher quality and were preferred to the original boxes.

In columns 2 and 3, we present regression coefficients from bivariate regressions of dependent variables on an indicator of each treatment arm for the full sample. In columns 5 and 6, the regressions are replicated for HFPS groups only, where each coefficient represents the effect of being in a particular treatment group relative to the HFPS control group. Reassuringly, we find minimal differences between treatment groups, with a very few characteristics displaying imbalance by chance. In our main regressions, we control for these covariates.

We also examine baseline imbalance on the variables measured in the intake survey in Web Appendix Table A3. The only dimension along which treatment and control do not appear to be balanced is the amount of net non-spousal transfers, which are lower for the multi-box group, which is likely by chance. The majority of the characteristics suggest comparability of groups.

3 Results

3.1 Take-up

Figure 1 shows a CDF of deposits into the single and multiple accounts measured in the HFPS (Panel A) and monitoring surveys (Panel B). Note that both of these figures only show deposits for periods covered by the surveys – we lack an comprehensive measure of total deposit activity throughout the study. Figure 1 shows substantial usage of the accounts across the distribution, and indicates usage was clearly higher in the multiple account group than in the single account group.[28]

In Table 2 we present take-up statistics to supplement the figures, using project records from account opening (Panel A), the lockbox check (Panel B), and the HFPS and monitoring surveys

[27]Sixty two percent of the sample have access to a mobile money account, defined as either owning one or using someone else's. About half of the the respondents have their own accounts at baseline.

[28]Web Appendix Figure A4 confirms this pattern, using data collected during the in-person lockbox check in December 2017.

(Panel C). From Panel A, we see that nearly all respondents accepted a box. Take-up of multiple boxes was not universal: only 76% of people took at least 2 boxes when offered, and only 42% took all 3 that were offered to them. This is unexpected since the boxes were of high quality, were offered free of cost and could have been sold.

Turning to usage, Panel B displays measurement of cumulative usage from the date of receiving the box until the in-person lockbox check in December 2017, about 5 months later. Usage of boxes was nearly universal: 95-97% of people used a box at least once, and 92% used the box at least 5 times. This level of take-up is far higher than in many prior studies, including several in Malawi with banks (Dupas et al. 2018; Brune et al. 2014) or VSLAs (i.e. Ksoll et al. 2016). Our preferred measure of usage, the value of deposits, is also substantial: mean deposits were \$117 in the single box group, a sizable amount since daily income is only \$2.50 a day in this sample. We also observe clear evidence that savings were higher in the multiple box group, by about \$40 (about 34%).[29]

Panel C shows usage as measured in the surveys. These surveys are likely less prone to measurement error than the lockbox check (which asked retrospectively about deposits over a 5 months period), but also only cover a small fraction of the study period and are not comprehensive measures of cumulative usage. Nevertheless, the pattern is very similar – 81-94% of people reported using the box at least once during these windows, and deposit values were meaningfully large. We continue to observe that savings in the multiple box group are substantially higher – the value of deposits was 65% higher in the HFPS and 28% higher in the monitoring surveys.

In Table 3, we replicate these results in a regression and report the correlates between usage and baseline characteristics treatment groups. We use three different measures of savings – the value of deposits recalled during the lockbox check, the HFPS, and the monitoring surveys. Across the three measures, we consistently observe that deposits in the multiple box group was statistically significantly higher, with a magnitude of 32% in the lockbox check, 74% in the HFPS and 27% in the monitoring surveys. While the specific effect size is not meaningful here (since deposits were measured over different windows and may reflect seasonal changes in savings behaviors), these results all point to higher savings from the introduction of an additional box.[30]

[29]In Web Appendix Table A4, we further analyze take-up and usage for the multiple box group by looking at these measures for the dominant (the one with the highest value of deposits) and other box(es) separately. While the likelihood of making a deposit in the other box was much lower than the dominant box, it was fairly high (between 60 and 70%). In fact, these rates of usage of the non-dominant account are significantly higher than the ones found in the rest of the savings literature in studies where only one account has been provided. We also observe that while the majority of the usage is in the dominant box, people did actively use other boxes as well. In general, the amount saved in the dominant box was about 2-4 times bigger than the other boxes.

[30]We also show some non-random correlations in this table. As we would expect, and consistent with prior work (see Dupas et al. 2018), we see that individuals who had more savings at baseline deposited more in project lockboxes, although the pattern of statistical significance varies across specifications. We also find that those who are more "taxed" by their networks, i.e., net givers of money in their informal social network have greater deposits, potentially due to a greater need to accumulate money. Interestingly, those who had a bank account at baseline also saved more in the boxes, suggesting that people need not one, but a portfolio of savings instruments to meet varying liquidity and temporal requirements for their savings.

3.2 Estimating treatment effects from survey data

For both the HFPS and monitoring surveys, we have 2 main intent-to-treat specifications. In our first specification, we pool the single box and multiple box treatments into a single dummy variable for having any box as follows:

$$Y_{ist} = \alpha_1 + \alpha_2 LB_i + \beta X_i + \gamma L_t + \mu_s + \delta_t + \epsilon_{ist} \tag{1}$$

We also report results from a second specification with disaggregated treatment groups, which is given by the following:

$$Y_{ist} = \theta_1 + \theta_2 LB_i^1 + \theta_3 LB_i^{mult} + \lambda X_i + \tau L_t + \mu_s + \delta_t + \epsilon_{ist} \tag{2}$$

In both Equations (1) and (2), Y_{ist} is an outcome for individual i at time t in strata s, μ_s is a strata fixed effect, δ_t is a fixed effect for the date of the interview, X_i are a limited set of individual controls from the baseline, and L_t is an indicator of airtime lottery wins measured either daily or weekly , depending on the measurement window of the outcome variable. LB_i is a dummy for being offered any lockbox, LB_i^1 is a dummy for being offered 1 lockbox, and LB_i^{mult} is a dummy for being offered multiple lockboxes. For the HFPS, we also include the mean of the dependent variable from the intake survey in X_i, making the specification ANCOVA. Standard errors are clustered at the individual level, and the regressions are estimated with population weights. In all tables, we present the high frequency surveys in Panel A and the monitoring surveys in Panel B, both of which display (1) the pooled and (2) disaggregated by group results . All monetary values in all tables are winsorized at 5% and expressed in USD.[31]

3.3 Savings

We present effects of accounts on savings in Table 4. Consistent with the prior literature, in this table we focus on the value of deposits. In the Table, Column 1 shows effects on savings in the project lockboxes, which are unsurprisingly statistically significant in both Panels since the control group had zero savings in those accounts by design. Columns 2-6 show effects on other savings sources, from which we observe that a substantial fraction of the savings in treatment accounts displaced other sources – there are statistically significant declines in , most notably savings at home and in other lockboxes (for those that use lockboxes on their own)[32] . Surprisingly, even though the effect on treatment box savings is substantially higher for the multiple box group, we see similar levels of crowd out for single and multiple boxes.

As is common in the savings literature, the effect on total savings (Column 7) is imprecise, but greatly suggestive of a positive treatment effect, particularly for multiple accounts. While the coefficient from the HFPS in Panel A is statistically insignificant, the monitoring surveys in Panel

[31] The exchange rate was about 700 MWK to $1 US during the sample period.

[32] Anecdotal accounts from our field staff suggest that non-project lockboxes were inferior in quality to what we provided - many were made of cardboard, and were almost universally not capable of being locked.

B show an increase in total savings of $26 over the two past month in the pooled box treatment. This is equivalent to about 28%, given the base of $92 in the control group. Disaggregated in Panel B2, we see an increase of $14.50 (16%) for single boxes and an increase of $38 (41%) for multiple boxes. Compared to the control mean of $93, these are large effects, equivalent to increases of 16% and 41% compared to the control group. [33]

3.4 Labor supply, business outcomes and productive investment

In Table 5, we examine labor supply, business outcomes, farming, and land purchases. Columns 1-5 show effects on the respondent's primary business. Interestingly, we observe some evidence that treatment respondents worked less in their main business in Panel A1, although an examination of coefficients from the disaggregated specification in Panel A2 reveals that this is driven entirely by the single box group – people in the single box group were 8 percentage points less likely to work in their main business, and worked 0.8 less hours per day (on a base of about 8 hours). The point estimate for the monitoring surveys in Panels B1 and B2 is the same sign, though not significant. We do not have a good explanation for why labor supply should be impacted negatively only for the single box group. From the remaining columns, we do not observe changes in profits or investment. However, we see a statistically significant increase in credit offered to customers (which was only measured in the monitoring surveys) for both groups. This result suggests that increased liquidity may have been passed on to customers, which may be a way of expanding business in a highly competitive retail sector. Casaburi and Reed (2017) show a similar finding, though in the different setting of traders buying cocoa from farmers in Sierra Leone. They show that when liquidity constraints are relaxed for the traders through a random allocation of subsidies, they are 14 percentage points more likely (more than a 100% increase over the control mean of 11%) to advance credit to farmers.

Columns 6-8 show farming outcomes, in particular hours spent on agriculture and investment in inputs like fertilizer and hybrid seeds. We find consistent evidence of an increase in farming activity. From the HFPS in Panels A1 and A2, we observe that farmers are 3-8 percentage points more likely to have worked on their farm in the past week (on a base of only 6%) and work 0.23-0.51 more hours (on a base of 0.12 hours). The point estimates in Panels B1 and B2 are also positive, though not statistically significant.[34] We also observe a large increase in spending on farm inputs of about 25%, significant at 10%.

Finally, Columns 9-10 show land transactions (measured only in the monitoring surveys). We show results for dummy variables for buying or renting land (rather than for amounts spent) because these variables are noisy and contain mostly zeroes. We observe large, statistically significant

[33] In Web Appendix Tables A5 and A6, we examine the effects on withdrawals and on balances for the project accounts as well as for other major savings places. The results are very similar – we find statistically significant effects for withdrawals in the monitoring surveys (though not the HFPS), and for balances (which were only measured only in the monitoring surveys). As with deposits, we see larger effects for the multiple box group than for the single box.

[34] Note that the control means differ for farming hours across the two sets of surveys, due to seasonal labor requirements for agriculture.

increases in both types of land acquisition – treatment respondents are about 4-7 percentage points more likely to buy land (base of 2%) and 9-10 percentage points more likely to rent (base of 5%).

For many of these outcomes, we observe larger point estimates for multiple boxes than for single boxes. However, this is not universally the case, and only a few of these differences are statistically significant. However, we do observe strong evidence that *both* the single and multiple account treatments had effects relative to control groups. We hypothesize that the incremental effects of an additional box are relatively small compared to effect of having one savings account, and we are likely underpowered to detect it in this case.

Taken together, the results in Table 5 suggest an increase in investment in farming, and a reallocation of labor away from the main business. These results contribute to a small recent literature on financial constraints of poor households which suggests that households re-optimize their labor supply decision once a subsistence constraint is relaxed via the provision of credit/capital (Bandiera et al. 2017; Brune, Chyn and Kerwin (in-progress); Callen et al. 2014; Fink, Jack, and Masiye 2018; Pitt and Khandker 2002). Similar to these papers, our results imply that entrepreneurs prefer to work less in their primary business and some would prefer to work more on farms and, possibly, in other occupations. This implies that entrepreneurial activity is partly designed to meet daily subsistence needs and if anything people work *too much*. Please note however that in our case, part of the increase in labor supply for farming may be driven by greater investments in land, which would have increased the marginal product of agricultural labor.

3.5 Expenditures

Table 6 shows effects on a variety of expenditure categories. In Columns 1-6, we show results on aggregate categories measured in both the high frequency and monitoring surveys. We find positive point estimates in almost every specification, though the pattern of statistical significance varies between the two surveys. Looking at total expenditures, we see an increase of $3.5 in the HFPS (on a base of $27 and significant at 10%) and $7 in the monitoring surveys (on a base of $34 and significant at 1%). While these are substantial effects, they are similar to some of the previous literature, in particularly Dupas and Robinson (2013a), who find large though imprecise increases in several expenditure categories. Among subcategories, we find an increase in school expenses in the HFPS, and effects on staple foods, personal and household expenses in the monitoring surveys.

In Columns 7-8, we show effects on some outcomes measured only in the monitoring surveys, including expenses on holidays and on school fees and tuition. For holidays, we observe an increase of $2.5, on a base of $8, and similar (and jointly significant) treatment effects for the disaggregated treatments. Holiday spending can be a large expense for households in many developing countries (i.e. Banerjee and Duflo 2007), and the Malawian context is similar with large expenses at mothers' day as well as at Christmas. We find that savings accounts allow people to spend more on these events, a finding that is consistent with Prina (2015). Whether this increase is ultimately money well-spent is, however, an open question. We do not observe an effect on school fees and tuition. We conjecture that the reason for the discrepancy between this and the school spending in the

HFPS is that the HFPS measured more common, smaller expenses like school supplies, which are easier to change on the intensive margin. Tuition expenses, on the other hand, require changes in enrollment on the extensive margin and are likely hard to induce in general, and especially in this setting as primary education is free in Malawi.

As in Table 5, we do not observe significant differences between the one box and multiple box groups, however, and the increases in expenditures across the various categories are similar in magnitude for the two treatment groups. While this could be for power reasons as noted above, it is also possible that the first account was often used for immediate expenses and so effects on outcomes like expenditures are similar between treatment groups. The second account may have allowed study participants to create a separate account for longer-term goals, which is why the effects in the previous table were somewhat larger for the multiple box group. This is purely speculative however.

3.6 Inter-personal transfers and loans

Table 7 shows effects of the accounts on inter-personal transfers and loans. Columns 1-4 show transfers to and from individuals outside the households, Columns 5-8 show effects on transfers to the spouse, and Columns 9-10 show effects on loans. In Panel A, transfers are measured over the previous week, while in Panel B they are measured over the past month. We conjecture that accuracy of recall is superior in HFPS as small transfers will likely be far more salient in the short run, and therefore, place somewhat more weight on the those results. In the pooled HFPS in Panel A1, we see that treatment respondents were more likely to give a transfer and also increased the value of their transfers, with no effect on receiving transfers. As with previous results, these effects are not statistically different between the one and multiple box groups. Our results are consistent with Dupas, Keats and Robinson (2018) and Flory (2018), but in contrast with Dizon, Gong and Jones (2018). An important caveat, however, is that we do not observe an effect on giving transfers in Panel B (though the point estimate on the value of transfers is positive). We conjecture that this may be due to recall issues in the monitoring survey. For instance, notice that the probability of giving transfer over the past 1 week among the control group in HFPS is 0.44, while the same probability in the monitoring survey measured over the past month is 0.41, suggesting of substantial recall bias.

In both panels, we observe no consistent pattern in regards to transfers to spouses. There are no effects in the HFPS and we observe some evidence of a decrease in transfers received from the spouse in the monitoring survey, though no change in transfers given. This could be consistent with respondents being less dependent on spouses. Finally, in Columns 9-10, we find no effects of accounts on probability of taking a loan and values of loans taken.[35]

[35] Another potential channel would be that the accounts made households less vulnerable to shocks. We test this in Web Appendix Tables A7 and A8, focusing on health shocks. Web Appendix Table A7 shows no differential probability of experiencing a health shock between the treatment and control groups, and Web Appendix Table A8 shows no differential ability to cope with shocks as a result of having a savings account.

4 Conclusion

People throughout the world save up simultaneously for multiple goals of varying in amounts and duration. A simple strategy for saving towards several goals might be to create multiple physically separated accounts, but this may be challenging in developing countries like Malawi where two thirds of the adult population lacks even a single bank account (see the 2017 Findex – Demirgüç-Kunt et al. 2018). In this paper, we show that entrepreneurs who were given multiple lockboxes saved about 30% more than those given only one. These results strongly suggest that a simple policy of providing multiple accounts with labeled goals may cost-effectively increase savings.

In addition, we find robust evidence that provision of savings account has strong impacts on downstream outcomes. In particular, we observe that entrepreneurs who received savings accounts invested more in farming, and substituted away from their small business. A possible explanation for this result which is consistent with recent literature is that small businesses might be operated in part to generate income for subsistence consumption, but on the margin farming is more profitable. The lockboxes also had statistically significant effects on a variety of other outcomes, including credit provision, expenditures, and interpersonal transfers. The magnitude of these effects may be surprising at first glance, since many prior studies show modest effects on downstream outcomes such as these. However, we are unusually well-powered to detect downstream effects of savings accounts due to almost universal takeup of accounts in compared to other studies in which this number rarely exceeds 20%. While we can only conjecture on why this may be, one possibility is that metal lockboxes are better suited to fit the savings needs of this population. As such, lockboxes may be a compelling temporary solution to financial exclusion in developing countries, where formal financial services continue to be inadequate.

References

[1] Agarwal, Sumit, Shashwat Alok, Pulak Ghosh, Soumya Ghosh, Tomasz Piskorski, and Amit Seru (2017). "Banking the Unbanked: What Do 255 Million New Bank Accounts Reveal about Financial Access?" Unpublished.

[2] Aggarwal, Shilpa, Valentina Brailovskaya and Jonathan Robinson (2018a). "The Effect of Mobile Money in Urban Malawi." Unpublished.

[3] Aggarwal, Shilpa, Valentina Brailovskaya and Jonathan Robinson (2018b). "Evaluating the Effect of High-Frequency Surveying: Evidence from Malawi." Unpublished.

[4] Ainslie, George (2001). *Breakdown of Will*. Cambridge: Cambridge University Press.

[5] Ashraf, Nava, Dean Karlan, and Wesley Yin (2006). "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics* 121 (2): 673–97.

[6] Bachas, Pierre, Paul Gertler, Sean Higgins, and Enrique Seira (2018). "How Debit Cards Enable the Poor to Save More." Unpublished.

[7] Banerjee, Abhijit V., and Esther Duflo (2007). "The economic lives of the poor." *Journal of Economic Perspectives* 21 (1): 141-168.

[8] Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman (2017). "Labor Markets and Poverty in Village Economies." *Quarterly Journal of Economics* 132 (2): 811–870.

[9] Benabou, Roland and Jean Tirole (2004). "Willpower and Personal Rules." *Journal of Political Economy* 112 (4): 848-886

[10] Beshears, John, James Choi, Christopher Harris, David Laibson, Brigitte C. Madrian, and Jung Sakong (2015). "Self Control and Commitment: Can Decreasing the Liquidity of a Savings Account Increase Deposits?" NBER Working Paper No. 21474.

[11] Blumenstock, Joshua, Eilin Francis and Jonathan Robinson (2017). "Digital Credit: A Snapshot of the Current Landscape and Open Research Questions." Unpublished.

[12] Bradley, Alice (1923). *Fifty family budgets*. New York: Woman's Home Companion.

[13] Brailovskaya, Valentina (2018). "Unreliable Electricity: Effects of Power Outages on Small Businesses in Urban Malawi." Unpublished.

[14] Bruhn, Miriam and David McKenzie (2009). "In pursuit of balance: Randomization in practice in development field experiments." *American Economic Journal: Applied Economics* 1 (4): 200-232.

[15] Brune, Lasse, Xavier Giné, and Jessica Goldberg, and Dean Yang (2014). "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi." *Economic Development and Cultural Change* 64 (2): 187-220.

[16] Brune, Lasse, Eric Chyn, and Jason Kerwin (in-progress). "Pay Me Later: A Simple Employer-based Saving Scheme." In-progress.

[17] Buehren, Niklas, Markus Goldstein, Leora Klapper, Tricia Koroknay-Palicz, and Simone Schaner (2018). "The Limits of Commitment: Who Benefits from Illiquid Savings Products?" Unpublished.

[18] Callen, Michael Suresh De Mel, Craig McIntosh and Christopher Woodruff (2014). "What are the Headwaters of Formal Savings? Experimental Evidence from Sri Lanka." NBER Working paper No. 20736.

[19] Carroll, Gabriel, James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick (2009). "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124 (4): 1639 - 1674.

[20] Carvalho, Leandro S., Stephan Meier, and Stephanie W. Wang (2016). "Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday." *American Economic Review*, 106 (2): 260-84.

[21] Casaburi, Lorenzo and Tristan Reed (2017). "Competition in Agricultural Markets: An Experimental Approach." Unpublished.

[22] Chandrasekhar, Arun, Cynthia Kinnan, and Horacio Larreguy (2018). "Social Networks as Contract Enforcement: Evidence from a Lab Experiment in the Field." *American Economic Journal: Applied Economics* 10 (4): 43-78.

[23] Comola, Margherita and Silvia Prina (2014). "Treatment Effect Accounting for Network Changes: Evidence from a Randomized Intervention." Working Paper

[24] Croppenstedt, Andre, Mulat Demeke, and Meloria M. Meschi (2003). "Technology Adoption in the Presence of Constraints: the Case of Fertilizer Demand in Ethiopia." *Review of Development Economics* 7 (1): 58-70.

[25] Demirgüç-Kunt, Asli, Leora Klapper, Dorothe Singer, Saniya Ansar, and Jake Hess (2018). *The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution.* https://globalfindex.worldbank.org.

[26] Dizon, Felipe, Erick Gong and Kelly Jones (2018). "Does Financial Inclusion Exclude? The Effect of Access to Savings on Informal Risk-Sharing in Kenya." Unpublished.

[27] Dupas, Pascaline and Jonathan Robinson (2013a). "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics* 5 (1): 163-192.

[28] Dupas, Pascaline and Jonathan Robinson (2013b). "Why Don't the Poor Save More? Evidence from Health Savings Experiments." *American Economic Review* 103 (4): 1138-1171

[29] Dupas, Pascaline, Anthony Keats and Jonathan Robinson (2017). "The Effect of Savings Accounts on Interpersonal Financial Relationships: Evidence from a Field Experiment in Rural Kenya." Forthcoming, *Economic Journal.*

[30] Dupas, Pascaline, Dean Karlan, Jonathan Robinson and Diego Ubfal (2018). "Banking the Unbanked? Evidence from three countries." *American Economic Journal: Applied Economics* 10 (2): 257-297.

[31] Fink, Gunther, B. Kelsey Jack and Felix Masiye (2018). "Seasonal Liquidity, Rural Labor Markets and Agricultural Production." NBER Working Paper 24564.

[32] Flory, Jeffrey (2018). "Formal Finance and Informal Safety Nets of the Poor: Evidence from a Savings Field Experiment." Forthcoming, *Journal of Development Economics*.

[33] Francis, Eilin (2018). "Paying to Repay? Experimental Evidence on Repayment Commitment." Unpublished.

[34] Habyarimana, James and William Jack (2018). "High Hopes: Experimental Evidence on Saving and the Transition to High School in Kenya." Unpublished.

[35] Herskowitz, Sylvan (2018). "Gambling, Saving, and Lumpy Expenditures: Sports Betting in Uganda." Unpublished.

[36] Jack, B. Kelsey (2011). "Market inefficiencies and the adoption of agricultural technologies in developing countries" White paper prepared for the Agricultural Technology Adoption Initiative, JPAL (MIT) / CEGA (Berkeley).

[37] Jakiela, Pamela, and Owen Ozier. 2011. "Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies." Unpublished

[38] Karlan, Dean and Leigh Linden (2016). "Loose Knots: Strong versus Weak Commitments to Save for Education in Uganda." NBER Working Paper 19863.

[39] Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher R. Udry (2014). "Agricultural Decisions after Relaxing Credit and Risk Constraints." *Quarterly Journal of Economics* 129 (2): 597–652.

[40] Kast, Felipe, and Dina Pomeranz (2014). "Saving More to Borrow Less: Experimental Evidence from Access to Formal Savings Accounts in Chile." NBER Working Paper 20239

[41] Koch, Alexander and Julia Nafziger (2016). "Goals and Bracketing under Mental Accounting." *Journal of Economic Theory* 162: 305-351.

[42] Kochar, Anjini (2018). "Branchless banking: Evaluating the Doorstep Delivery of Financial Services in Rural India." *Journal of Development Economics* 135: 160-175.

[43] Ksoll, Christopher, Helene Bie Lilleør, Jonas Helth Lønborg, and Ole Dahl Rasmussen (2016). "Impact of Village Savings and Loan Associations: Evidence from a cluster randomized trial." *Journal of Development Economics* 120: 70-85.

[44] Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall (2000). "Mutual Insurance, Individual Savings, and Limited Commitment." *Review of Economic Dynamics* 3 (2): 216-246.

[45] Madrian, Brigitte and Dennis Shea (2001). "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (4): 1149–87.

[46] Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiyaying Zhao (2017). "Poverty Impedes Cognitive Function," *Science* 341: 976–980.

[47] Pitt Mark M. and Shahidur Khandker (2002). "Credit Programs for the Poor and Seasonality in Rural Bangladesh." *Journal of Development Studies* 39 (2): 1-24.

[48] Platteau, Jean-Philippe. 2000. "Institutions, Social Norms, and Economic Development". Amsterdam: Harwood Academic Publishers

[49] Prina, Silvia (2015). "Banking the Poor via Savings Accounts: Evidence from a Field Experiment." *Journal of Development Economics* 115: 16-31.

[50] Schaner, Simone (2018). "The Persistent Power of Behavioral Change: Long-Run Impacts of Temporary Savings Subsidies for the Poor." *American Economic Journal: Applied Economics* 10 (3): 67-100.

[51] Shah, Anuj, Sendhil Mullainathan, and Eldar Shafir (2012). "Some Consequences of Having Too Little," *Science* 338: 682–685.

[52] Soman, Dilp and Amar Cheema (2011). "Earmarking and partitioning: increasing saving by low-income households." *Journal of Marketing Research* 48 (SPL): S14-S22.

[53] Somville, Vincent and Lore Vandewalle, "Saving by Default: Evidence from a Field Experiment in Rural india," *American Economic Journal: Applied Economics* 10(3): 39-66.

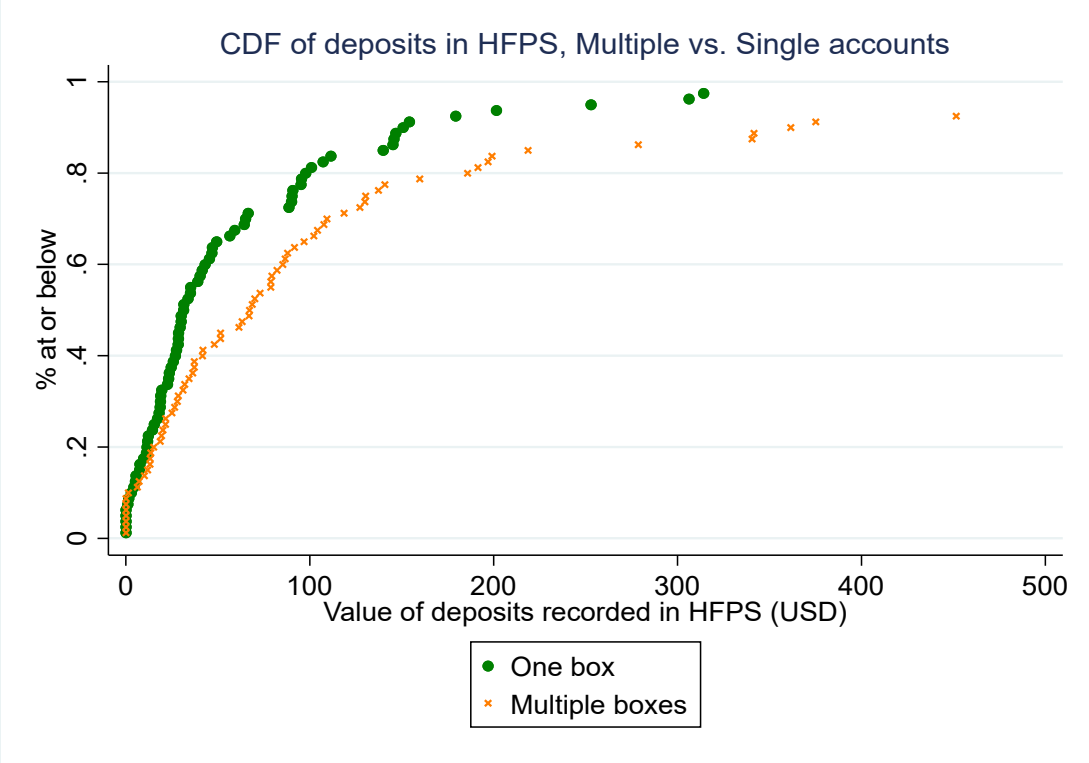
[54] Thaler, Richard (1990). "Anomalies: Saving, Fungibility, and Mental Accounts." *Journal of Economic Perspectives* 4 (1): 193-205.

[55] Thaler, Richard (1999). "Mental Accounting Matters." *Journal of Behavioral Decision Making* 12: 183-206.

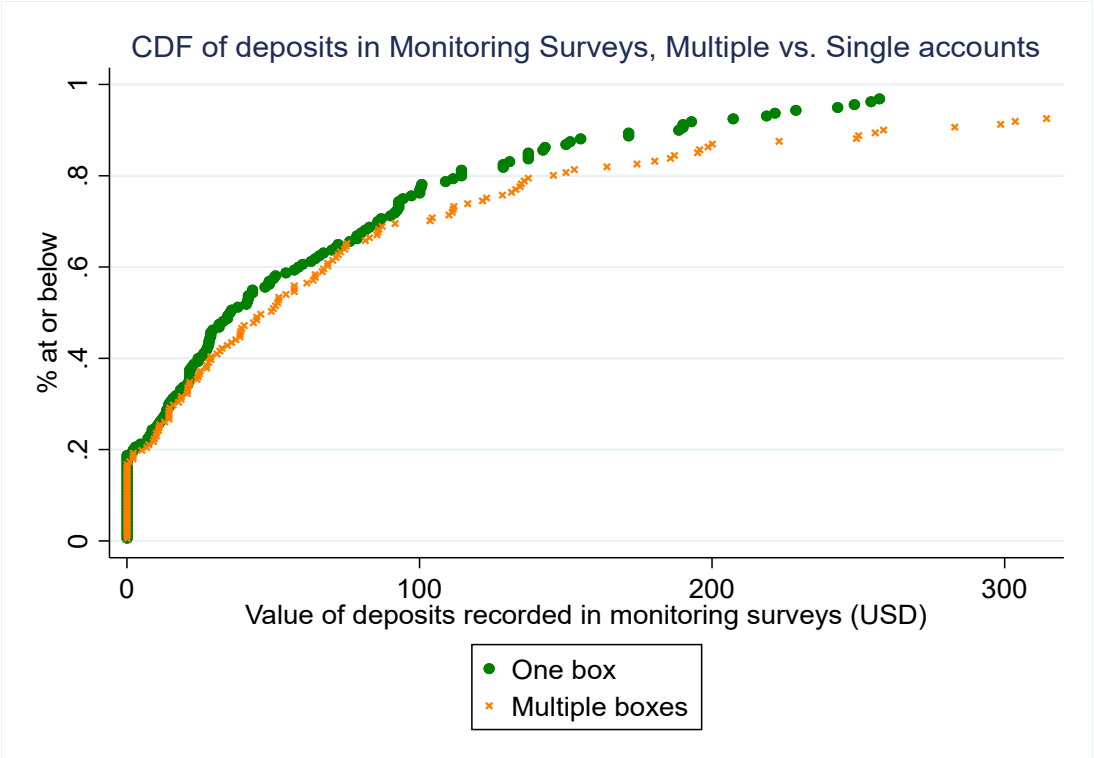
[56] Zelizer, Viviana A. (1994). "The Creation of Domestic Currencies," *American Economic Review (Papers and Proceedings)* 84 (2): 138-142.

Figure 1. CDFs by number of accounts

Panel A. High Frequency Phone Surveys



Panel B. Monitoring Surveys



Notes: In Panel A, data is from the high frequency phone surveys; in Panel B, data is from monitoring surveys. Monetary values are in USD and CDF shows only below 95th percentile (since there are several large values).

Table 1. Summary Statistics and Randomization Check

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample			HFPS sample only		
		Regression Coefficient for:			Regression Coefficient for:	
	Control Mean	One Box	Multiple Boxes	Control Mean	One Box	Multiple Boxes
Panel A. Demographic information and asset ownership						
Age	34.42	2.17	0.83	33.68	2.74	2.5
	(10.16)	(1.39)	(1.42)	(9.27)	(1.79)	(1.71)
Married	0.81	-0.08	-0.07	0.79	-0.11	-0.05
		(0.06)	(0.07)		(0.09)	(0.08)
Male	0.46	0.03	0.08	0.57	-0.09	-0.04
		(0.08)	(0.09)		(0.11)	(0.10)
Household Farms	0.71	0	-0.02	0.68	-0.07	0.1
	(0.46)	(0.07)	(0.09)	(0.47)	(0.11)	(0.09)
Years of Education	9.05	-0.09	0.01	9.15	-0.28	-0.15
	(2.83)	(0.40)	(0.41)	(2.97)	(0.65)	(0.61)
Land Owned (Acres)	0.66	0.03	-0.09	0.59	0.25	-0.11
	(1.18)	(0.19)	(0.15)	(0.93)	(0.30)	(0.13)
Value of Durable Assets and Livestock	872.60	-188.03	-12.85	976.10	-179.42	191.02
	(1547.00)	(158.98)	(199.56)	(1620.00)	(155.69)	(300.66)
House has iron roof	0.93	0.02	0.02	0.92	0.02	0.03*
		(0.01)	(0.01)		(0.02)	(0.02)
Owns a cell phone	0.94	0.01	-0.09	0.92	0.01	-0.04
	(0.24)	(0.02)	(0.06)	(0.27)	(0.04)	(0.05)
Has mobile money account	0.56	0.04	0.18**	0.53	0.1	0.16
	(0.50)	(0.08)	(0.08)	(0.50)	(0.10)	(0.10)
Pavel B: Business						
=1 if Retail	0.68	0.02	-0.04	0.66	0	0.1
	(0.47)	(0.07)	(0.09)	(0.48)	(0.10)	(0.08)
Average Weekly Revenue	66.37	4.45	-8.54	67.28	-4.98	12.26
	(99.00)	(14.00)	(12.96)	(107.30)	(12.91)	(14.60)
Average Weekly Profit	18.77	0.27	-0.69	17.98	-0.38	2.81
	(23.20)	(2.96)	(3.09)	(23.33)	(2.60)	(3.15)
Value of equipment and inventory	292.90	9.8	6.76	265.30	82.53	163.53
	(457.40)	(76.58)	(90.60)	(356.90)	(86.80)	(133.77)
Pavel C: Savings						
Number of savings places	2.54	0.06	0.02	2.54	0.08	-0.05
	(1.17)	(0.17)	(0.17)	(1.20)	(0.24)	(0.25)
Number of savings goals	2.40	0.15	0.16	2.29	0.26	0.08
	(1.27)	(0.15)	(0.22)	(1.13)	(0.20)	(0.21)
=1 if has separate savings	0.78	0.07	0.08	0.76	0.12	0.08
places for different purposes	(0.42)	(0.07)	(0.07)	(0.43)	(0.10)	(0.10)
=1 if saves for multiple goals in a	0.35	0.1	0.08	0.29	0.1	0.03
single savings place	(0.48)	(0.07)	(0.09)	(0.46)	(0.11)	(0.10)
Total Cash Savings (Balance)	119.80	-6.01	-6.53	120.40	18.28	35.9
	(196.70)	(25.43)	(27.24)	(222.00)	(24.47)	(34.45)
Saves in:						
Mobile money	0.47	0.06	0.13	0.48	-0.07	0.03
		(0.09)	(0.11)		(0.13)	(0.13)
Bank account	0.30	0.02	0.11	0.32	0.05	0.09
		(0.07)	(0.08)		(0.09)	(0.09)
VSLA / ROSCA	0.52	0.05	-0.04	0.48	0.1	0.01
		(0.07)	(0.09)		(0.11)	(0.10)
Secret place at home	0.82	-0.03	0.03	0.80	0.01	-0.01
		(0.06)	(0.06)		(0.09)	(0.09)
Savings box	0.22	0.01	-0.10*	0.29	-0.08	-0.16*
		(0.06)	(0.05)		(0.09)	(0.08)

Notes: There are 457 observations in columns 1-3 and 235 in columns 4-6. Means are population weighted. Randomization check is performed only for businesses that appear in the analysis sample (i.e. completed a monitoring survey or at least one phone survey). Monetary values are winsorized at 1% and expressed in USD. In Columns 1 and 4, standard deviations in parentheses; in the other columns, standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

Table 2: Take-up of project accounts

	(1) One box group	(2) Multiple box group
Panel A. Administrative data on take-up		
Took at least one box	0.99	0.99
If offered multiple boxes: took more than 1 box	-	0.76
If offered multiple boxes: took 3 boxes	-	0.42
Observations	160	161
Panel B. Lockbox check visit in December 2017		
Reported at least 1 deposit	0.95	0.97
Reported at least 2 deposits	0.95	0.95
Reported at least 5 deposits	0.92	0.92
Total value of deposits	116.93	156.49
	(152.89)	(180.91)
Number of deposits	49.94	71.01
	(53.67)	(67.23)
Total value of withdrawals	82.32	119.81
	(124.73)	(156.84)
Number of withdrawals	7.42	13.06
	(17.61)	(21.60)
Balance	15.87	14.70
	(34.26)	(30.97)
Observations	121	120
Panel C. Survey responses		
C1. HFPS		
Made at least 1 transaction	0.94	0.91
Total value of deposits	70.11	116.43
	(108.34)	(151.60)
Total value of withdrawals	14.80	49.15
	(37.79)	(137.90)
Observations	80	80
C2. Monitoring survey		
Made at least 1 transaction	0.81	0.83
Total value of deposits	66.16	84.60
	(74.95)	(99.09)
Total value of withdrawals	46.15	63.43
	(67.98)	(90.94)
Observations	160	161

Notes: See text for discussion of data sources. Means are presented, with standard deviations in parentheses. In Panel C, the HFPS covers transactions over approximately 14 weeks (from September-October 2017 and February-March 2018) while the monitoring survey records information on deposits recalled over a 2 month period before each monitoring survey (conducted in January and March 2018), so covering 4 months total. The value and number of deposits and withdrawals are winsorized at 5% level.

Table 3. Determinants of Account Usage

	(1)	(2)	(3)
	Total deposits recalled in lockbox check survey in Dec 17	Total deposits observed in HFPS	Total deposits observed in monitoring surveys
Multiple accounts	37.25*	52.15**	18.82*
	(20.54)	(21.11)	(9.62)
Other Covariates			
Female	-96.15**	-33.26	-24.36
	(45.57)	(46.36)	(21.24)
Married	-54.28	-28.47	-16.96
	(38.59)	(38.96)	(17.89)
Female * married	47.05	2.02	3.2
	(50.55)	(52.29)	(23.59)
Age (tens of years)	11.87	-6.19	2.06
	(12.01)	(12.41)	(5.63)
Years of education	6.53	-1.29	0.3
	(4.30)	(4.26)	(1.96)
Has kids under 18	28.14	25.5	9.7
	(36.68)	(37.30)	(17.38)
Had savings box at baseline	-1.84	-4.8	-12.21
	(29.08)	(28.07)	(12.78)
Had bank account at baseline	21.79	28.97	24.85**
	(25.93)	(25.84)	(11.96)
Had mobile money account	37.83*	28.64	26.81***
at baseline	(21.65)	(22.27)	(10.19)
Inverse hyperbolic sine of	6.33	13.39*	3.48
baseline monetary savings	(6.38)	(7.10)	(3.11)
Log assets	22.45**	7.51	4.36
	(8.75)	(8.23)	(4.11)
"Taxed" (i.e. gives money but does	56.06**	16.8	17.60*
not receive)	(22.37)	(22.35)	(10.27)
Observations	238	159	317
Mean (1 box group)	116.90	70.11	66.16
Std. dev. (1 box group)	152.90	108.30	74.95

Notes: Values are in USD and winsorized at 5%. The source of data is (1) the lockbox check in December 2017 in Column 1, (2) the high frequency phone surveys in Column 2, and (2) the monitoring surveys in Columns 3. Deposits are winsorized at 5% before logging. Standard errors in parentheses.
*, **, and *** indicate significance at 10%, 5%, and 1% respectively.

Table 4. Treatment Effects on Deposits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Deposits into project lockboxes	Deposits into other savings sources					Total Deposits[2]
		Other lockboxes	Cash at home	Bank accounts	Savings groups[1]	Mobile money	
Panel A. High frequency phone surveys							
A1. Pooled							
Any box	7.24***	-2.06***	-3.84***	0.07	-0.91**	-0.73**	0.78
	(1.07)	(0.39)	(1.00)	(0.61)	(0.41)	(0.35)	(2.10)
A2. Disaggregated							
One box	4.07***	-1.99***	-3.86***	0.06	-1.20**	-0.68*	-2.65
	(1.05)	(0.38)	(1.00)	(1.01)	(0.46)	(0.40)	(2.58)
Multiple boxes	10.11***	-2.13***	-3.82***	0.07	-0.63	-0.78*	3.87
	(1.58)	(0.41)	(1.20)	(0.74)	(0.52)	(0.40)	(2.65)
p -values for F-tests of:							
1 Box = Multiple Boxes	0.001***	0.45	0.97	0.99	0.29	0.79	0.04**
Joint significance	0.001***	0.001***	0.001***	0.99	0.035**	0.12	0.12
Observations	2670	2670	2670	2670	2670	2670	2670
Number of Businesses	234	234	234	234	234	234	234
Control Mean	0.00	2.32	7.32	2.03	3.21	1.67	17.22
Control SD	0.00	5.03	12.67	19.48	5.79	4.78	27.66
Panel B. Monitoring Surveys							
B1. Pooled							
Any box	52.34***	-9.98***	-13.76***	6.72	-5.63	-2.29	26.59**
	(4.51)	(2.26)	(2.84)	(4.90)	(5.82)	(2.10)	(11.11)
B2. Disaggregated							
One box	42.23***	-9.00***	-12.57***	0.12	-3.36	-2.47	14.59
	(4.72)	(2.19)	(3.01)	(3.93)	(6.06)	(2.22)	(10.58)
Multiple boxes	62.09***	-10.94***	-14.91***	13.10*	-7.83	-2.11	38.17**
	(6.09)	(2.48)	(3.04)	(7.41)	(6.77)	(2.40)	(14.97)
p -values for F-tests of:							
1 Box = Multiple Boxes	0.002***	0.11	0.26	0.058*	0.41	0.86	0.086*
Joint significance	0.001***	0.001***	0.001***	0.16	0.50	0.53	0.039**
Observations	786	786	786	786	786	786	786
Number of Businesses	429	429	429	429	429	429	429
Control Mean	0.00	14.83	18.31	14.04	32.43	10.64	92.52
Control SD	0.00	36.08	33.26	59.26	55.81	22.20	108.00

Notes: In Panel A, deposits are measured over the 7 days prior to the survey. In Panel B, deposits are cumulative deposits in the 2 months prior to the monitoring surveys (and thus cover November-December 2017 for monitoring survey 1 and February-March 2018 for monitoring survey 2). All regressions in Panel A include a measure of the dependent variable during the intake survey as a control - see text for details. All regressions control for strata, date fixed effects and baseline controls, and are probability weighted (see in the text for details). All monetary variables are expressed in USD and are winsorized at 5%. Standard errors clustered at individual level in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

[1]Savings groups include VSLAs and ROSCAs.

[2]Total deposits is the sum of the other columns, as well as other less common types of savings (such as safekeeping with shopkeepers or friends).

Table 5. Treatment effects on business outcomes, labor supply, and farming

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Main Business					Farming			Land[4]	
	=1 if worked[1]	Hours	Profits	Investment	Credit to customers[2]	=1 if farmed	Hours	Farm inputs[3]	=1 if bought land	=1 if rented land
Panel A. High frequency phone surveys										
A1. Pooled										
Any box	-0.04*	-0.33	-0.38	5.42		0.06***	0.38***			
	(0.02)	(0.30)	(0.23)	(7.39)		(0.02)	(0.12)			
A2. Disaggregated										
One box	-0.08***	-0.84**	-0.34	4.84		0.03	0.23**			
	(0.03)	(0.36)	(0.25)	(10.02)		(0.02)	(0.11)			
Multiple boxes	0.00	0.13	-0.42	5.98		0.08***	0.51***			
	(0.02)	(0.34)	(0.29)	(8.15)		(0.03)	(0.17)			
p -values for F-tests of:										
1 Box = Multiple Boxes	0.005***	0.01***	0.76	0.91		0.04**	0.12			
Joint significance	0.009***	0.02**	0.27	0.75		0.01**	0.01***			
Observations	18589	18588	15483	2674		2678	2678			
Number of Businesses	234	234	234	234		234	234			
Control Mean	0.82	8.27	3.42	48.08		0.06	0.21			
Control SD	-	4.64	3.59	73.12		-	1.21			
Panel B. Monitoring Surveys										
B1. Pooled										
Any box	-0.03	-0.35	-0.12	4.78	5.32***	0.06	1.04**	7.80*	0.05**	0.09**
	(0.03)	(0.39)	(0.29)	(6.51)	(1.74)	(0.04)	(0.50)	(4.40)	(0.02)	(0.04)
B2. Disaggregated										
One box	-0.03	-0.24	-0.28	6.72	5.13***	0.03	0.44	9.40**	0.04	0.09*
	(0.03)	(0.38)	(0.32)	(7.84)	(1.97)	(0.05)	(0.59)	(4.58)	(0.03)	(0.05)
Multiple boxes	-0.01	-0.38	-0.07	3.39	5.49**	0.07	1.55**	6.29	0.07**	0.10*
	(0.03)	(0.38)	(0.31)	(7.49)	(2.53)	(0.05)	(0.66)	(5.30)	(0.03)	(0.05)
p -values for F-tests of:										
1 Box = Multiple Boxes	0.61	0.73	0.56	0.69	0.90	0.46	0.14	0.50	0.29	0.99
Joint significance	0.68	0.60	0.68	0.69	0.008***	0.37	0.06*	0.12	0.06*	0.10
Observations	786	786	777	785	783	786	785	775	384	385
Number of Businesses	429	429	428	429	427	429	429	427	384	385
Control Mean	0.75	7.38	3.04	54.21	7.02	0.24	2.07	27.89	0.02	0.05
Control SD	0.31	3.90	3.37	84.33	16.25	-	5.78	40.99	0.13	0.22

Note: In Panel A, all outcomes are at the daily level, except for farming and investment (which are measured at the weekly level). Investment includes inventory and equipment. There are fewer observations for profits because this was only asked for the past 4 days in the second round of HFPS. In Panel B, all outcomes are measured over the 7 days prior to the survey, other than credit to customers (which is measured over the past month) and lan questions. Credit to customers, land transactions, and farm input expenditures were only measured during monitoring surveys. Land questions were asked only at endline. All regressions in Panel A include a measure of the dependent variable during the intake survey as a control - see text for details. All regressions control for strata, date fixed effects and baseline controls, and are probability weighted (see in the text for details). All monetary variables are expressed in USD and are winsorized at 5%. Standard errors clustered at individual level in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

[1]In Panel B, this variable is the proportion of days worked over the 7 days prior to the survey.

[2]In Panel B, credit to customers is measured over the month prior to the survey.

[3]Farm inputs are measured cumulatively since July 2017.

[4]Land purchases are measured over the entire sample period (from July 2017 to the date of the 2nd monitoring survey in March 2018). Conditional on being positive, the average amount spent on land is $386 (mean) and $250 (median); while the amount spent on rent per month is $35 (mean) and $2.3 (median).

Table 6. Treatment effects on expenditures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Common expenditures measured over the past 7 days						Lumpy expenditures measured in monitoring surveys	
	Staple foods	Personal expenses	Airtime	Household expenses	School expenses	Total	Holiday spending	School fees, tuition and school supplies
Panel A. High frequency phone surveys								
A1. Pooled								
Any box	0.24	0.33	0.1	0.61	0.82***	3.47*		
	(0.70)	(0.38)	(0.11)	(0.45)	(0.28)	(1.98)		
A2. Disaggregated								
One box	0.32	0.72	0.15	0.45	0.71**	3.99*		
	(0.82)	(0.45)	(0.12)	(0.53)	(0.33)	(2.37)		
Multiple boxes	0.18	-0.02	0.07	0.78	0.91**	3.04		
	(0.81)	(0.46)	(0.12)	(0.57)	(0.38)	(2.47)		
p -values for F-tests of:								
1 Box = Multiple Boxes	0.87	0.13	0.55	0.61	0.65	0.74		
Joint significance	0.93	0.19	0.47	0.36	0.02**	0.19		
Observations	2,667	2,678	2,674	2,678	2,678	2,678		
Number of Businesses	234	234	234	234	234	234		
Control Mean	10.30	3.45	0.84	3.75	1.59	27.38		
Control SD	6.31	4.08	1.05	6.36	4.09	20.01		
Panel B. Monitoring surveys								
B1. Pooled								
Any box	1.98**	1.45**	0.25	2.21**		7.34***	2.45**	-1.46
	(0.83)	(0.63)	(0.18)	(1.08)		(2.48)	(1.12)	(6.12)
B2. Disaggregated								
One box	1.52	1.58**	0.32	3.07***		8.44***	2.86**	-1.44
	(0.94)	(0.70)	(0.20)	(1.18)		(2.80)	(1.34)	(6.94)
Multiple boxes	1.86*	1.12	0.18	1.64		6.28**	2.06	-1.49
	(0.95)	(0.70)	(0.22)	(1.24)		(2.98)	(1.38)	(7.09)
p -values for F-tests of:								
1 Box = Multiple Boxes	0.74	0.52	0.50	0.22		0.47	0.60	0.99
Joint significance	0.10	0.07*	0.28	0.04**		0.008***	0.08*	0.97
Observations	784	786	786	786		786	786	786
Number of Businesses	429	429	429	429		429	429	429
Control Mean	13.21	5.04	1.82	6.80		34.14	8.31	49.77
Control SD	8.96	6.13	1.77	10.23		24.46	13.36	75.61

Note: In both Panels, expenditures in Columns 1-6 are measured over the 7 days prior to the survey. Total expenses in Column 6 includes the first 5 columns in addition to other categories not shown here. Holiday spending was only measured in the monitoring surveys, and include the total amount spent on Mother's Day, Christmas, Easter and any other holidays. Education expenses (Column 8) include all expenses since the start of the project (July 2017). All regressions in Panel A include a measure of the dependent variable during the intake survey as a control - see text for details. All regressions control for strata, date fixed effects and baseline controls, and are probability weighted (see in the text for details). All monetary variables are expressed in USD and are winsorized at 5%.Standard errors clustered at individual level in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are expressed in USD and winsorized at 5%.

Table 7. Treatment effects on loans and transfers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Transfers to / from friends & family				Transfers to / from spouse				Credit[1]	
	Received transfer	Value	Gave transfer	Value	Received transfer	Value	Gave transfer	Value	Took out loan	Value
Panel A. High frequency phone surveys										
A1. Pooled										
Any box	0.04	-0.13	0.15***	0.93*	0.02	-0.08	-0.01	-0.01	0.04	0.10
	(0.04)	(0.65)	(0.04)	(0.55)	(0.04)	(0.53)	(0.05)	(0.41)	(0.04)	(0.32)
A2. Disaggregated										
One box	0.08*	0.25	0.15***	1.09	0.01	-0.48	0.00	0.30	0.03	0.44
	(0.05)	(0.83)	(0.06)	(0.76)	(0.04)	(0.57)	(0.06)	(0.49)	(0.04)	(0.48)
Multiple boxes	0.00	-0.44	0.15***	0.79	0.03	0.26	-0.01	-0.27	0.06	-0.17
	(0.04)	(0.69)	(0.05)	(0.64)	(0.04)	(0.61)	(0.05)	(0.50)	(0.04)	(0.35)
p -values for F-tests of:										
1 Box = Multiple Boxes	0.06*	0.38	1.00	0.73	0.59	0.17	0.88	0.30	0.60	0.24
Joint significance	0.12	0.64	0.003***	0.24	0.73	0.37	0.98	0.58	0.44	0.50
Observations	2690	2678	2678	2678	2678	2672	2672	2672	2678	2739
Number of Businesses	234	234	234	234	234	234	234	234	234	234
Control Mean	0.23	3.00	0.44	2.92	0.21	2.42	0.33	2.10	0.32	0.98
Control SD	0.42	10.90	0.50	6.21	0.41	6.91	0.47	4.34	0.47	4.27
Panel B. Monitoring Surveys										
B1. Pooled										
Any box	-0.03	2.36**	-0.01	0.79	-0.08**	-0.95	0.00	0.84	0.00	1.38
	(0.05)	(1.17)	(0.05)	(0.75)	(0.04)	(1.11)	(0.05)	(0.76)	(0.05)	(1.09)
B2. Disaggregated										
One box	-0.01	1.57	-0.01	0.84	-0.06	-1.39	-0.03	0.77	0.03	1.25
	(0.06)	(1.37)	(0.05)	(0.83)	(0.04)	(1.09)	(0.05)	(0.85)	(0.05)	(1.17)
Multiple boxes	-0.06	1.87	-0.02	0.71	-0.09**	-0.90	0.00	0.35	-0.01	1.12
	(0.05)	(1.27)	(0.05)	(0.93)	(0.04)	(1.26)	(0.05)	(0.85)	(0.05)	(1.40)
p -values for F-tests of:										
1 Box = Multiple Boxes	0.44	0.83	0.83	0.89	0.48	0.66	0.39	0.64	0.37	0.92
Joint significance	0.54	0.29	0.95	0.56	0.07*	0.45	0.67	0.67	0.64	0.53
Observations	786	786	784	784	786	786	786	786	786	786
Number of Businesses	429	429	427	427	429	429	429	429	429	429
Control Mean	0.38	5.67	0.41	3.97	0.33	6.63	0.38	3.91	0.67	5.60
Control SD	0.49	13.40	0.49	7.52	0.47	15.12	0.49	7.40	0.47	10.75

Notes: Transfers include gifts and loans, and both cash and in-kind payments and do not include survey compensation. In Panel A, variables are measured over the 7 days prior to the survey. In Panel B, all values are over the past month, except for loans which are measured over the previous 3 months. All regressions in Panel A include a measure of the dependent variable during the intake survey as a control - see text for details. All regressions control for strata, date fixed effects and baseline controls, and are probability weighted (see in the text for details). All monetary variables are expressed in USD and are winsorized at 5%. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

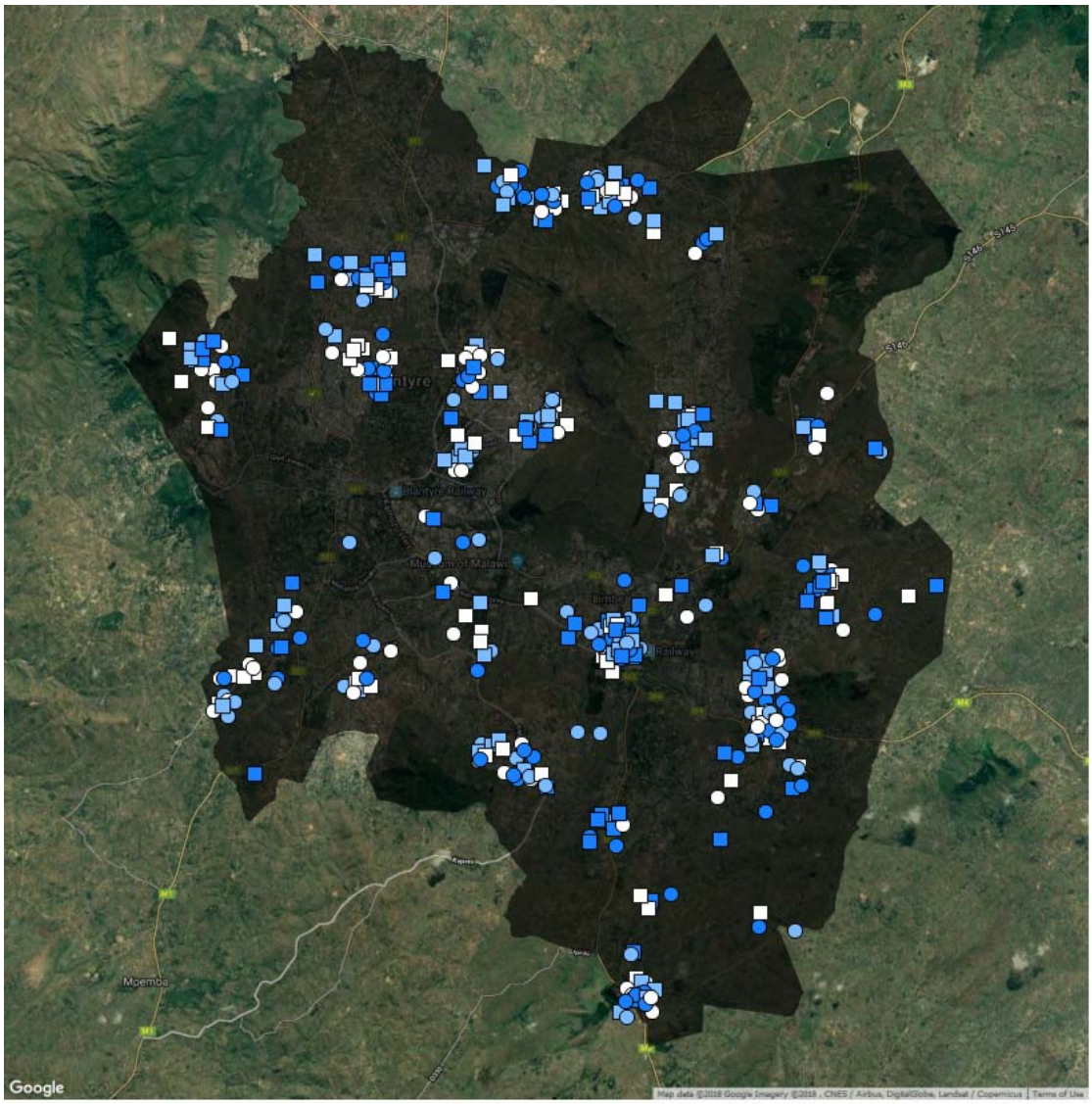
[1]Credit include digital loans and loans from VSLAs, ROSCAs, banks, microfinance institutions, and moneylenders.

Web Appendix

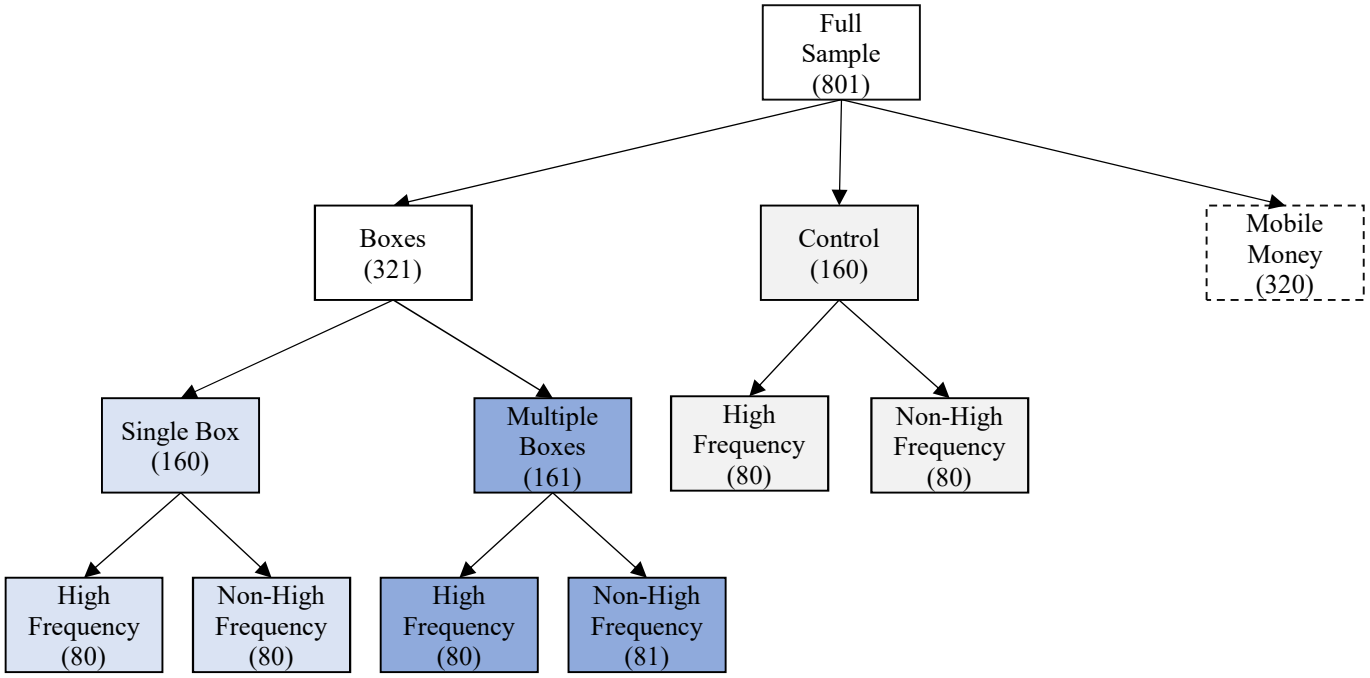
Saving for Multiple Financial Needs: Evidence from Malawi

Shilpa Aggarwal, Valentina Brailovskaya and Jonathan Robinson

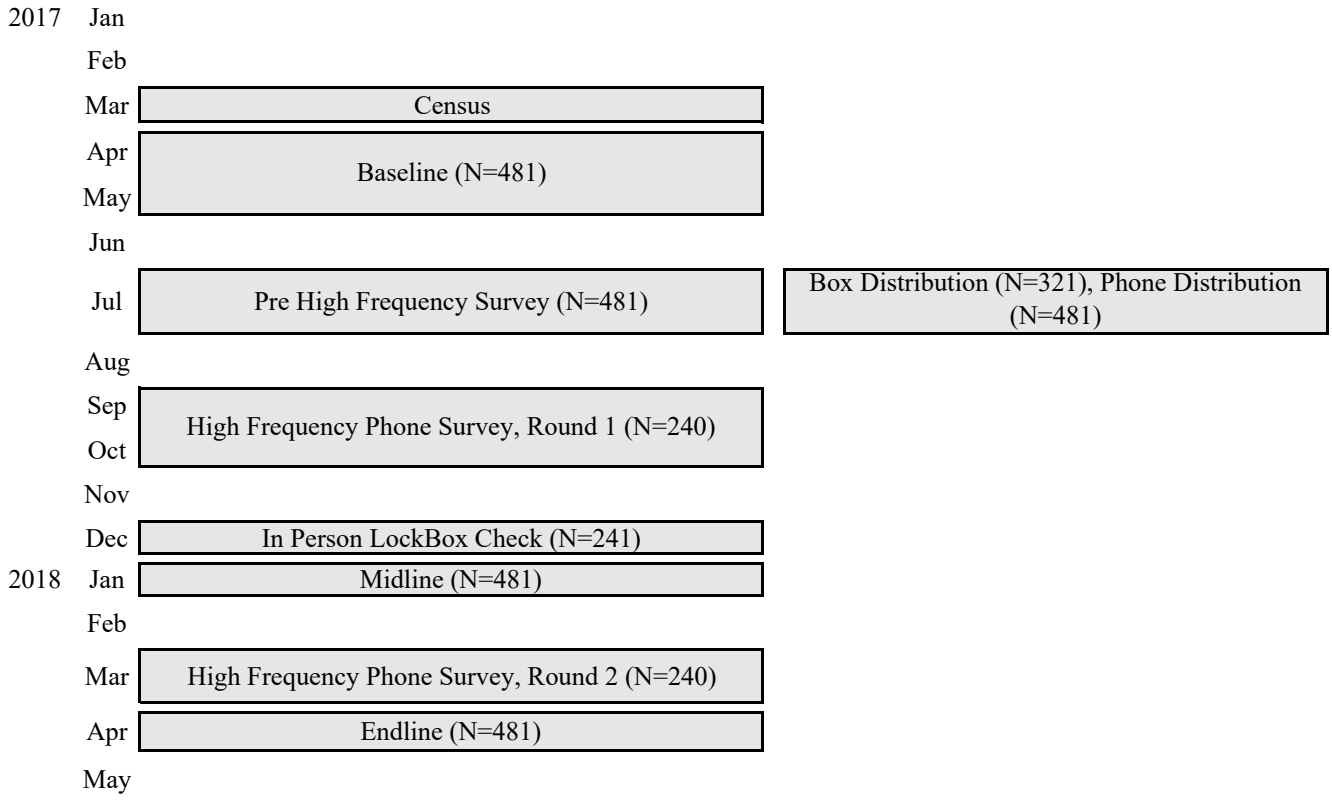
Web Appendix Figure A1. Map of Blantyre with Treatment Assignment



Web Appendix Figure A2. Experimental Design

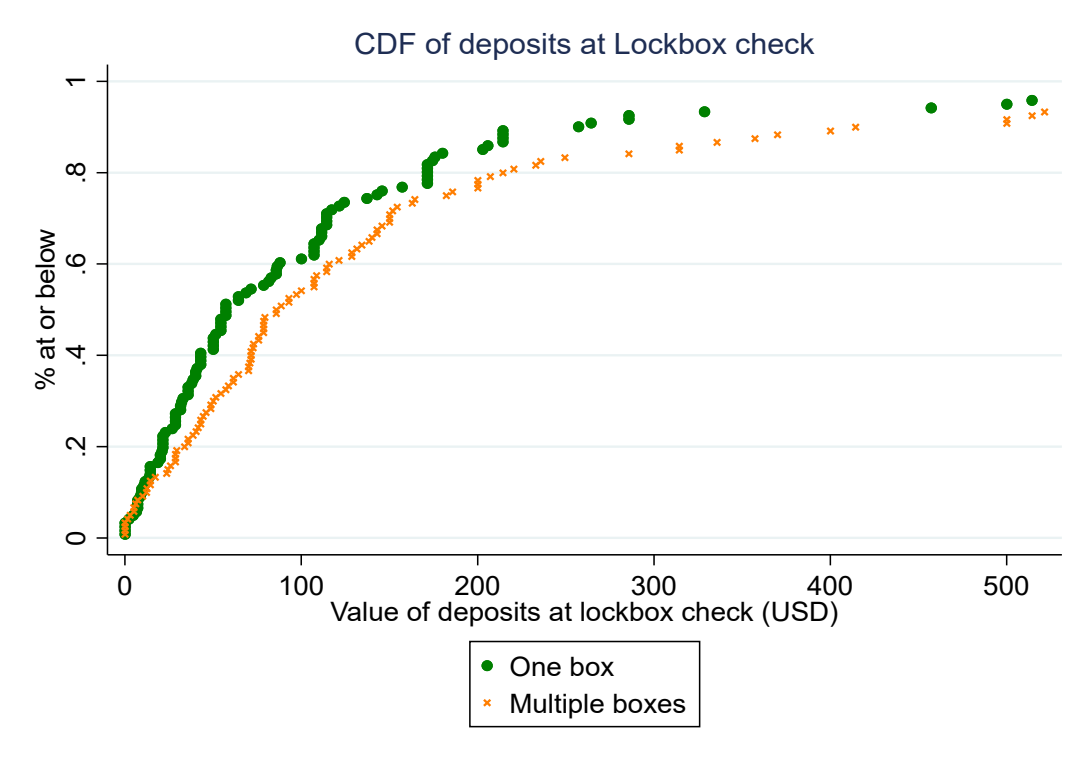


Web Appendix Figure A3. Project and Survey Timeline

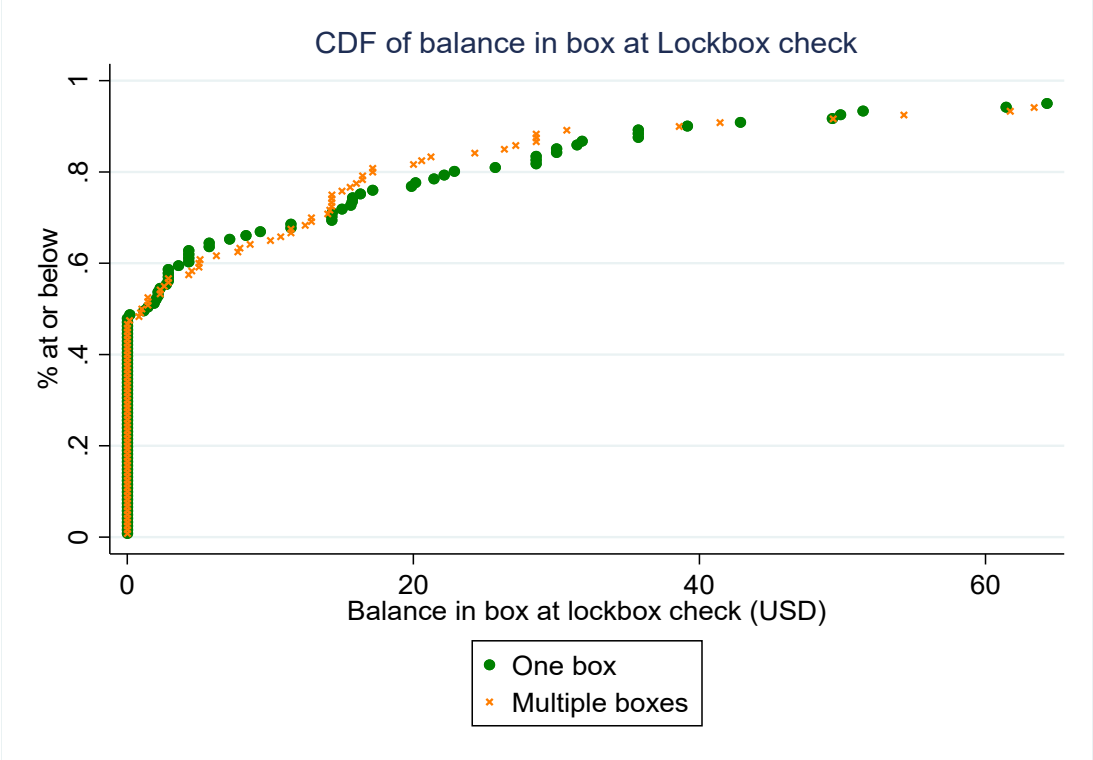


Web Appendix Figure A4. CDFs by number of accounts (Lockbox check data)

Panel A. Deposits



Panel B. Balance in December 2017



Notes: Data is from in-person lockbox checks in December 2017. In Panel A, Deposits are self-reported; in Panel B, balances were verified by asking respondents to open the box. Monetary values are in USD and CDF shows only below 95th percentile.

Web Appendix Table A1. Savings goals

	(1)	(2)	(3)
	Savings goals listed at baseline (all respondents)	Savings goals listed for treatment accounts	
		One box	Multiple boxes
Expand business, start new business, or invest in inventory	0.64	0.36	0.50
General consumption	0.42	0.02	0.19
Emergencies	0.40	0.15	0.44
Buying land	0.18	0.25	0.32
Durable goods	0.12	0.05	0.09
Children's education	0.12	0.10	0.28
Home improvement	0.08	0.06	0.09
Home expenses (rent, utilities, etc.)	0.07	0.00	0.03
Agriculture	0.05	0.06	0.06
Observations	481	114	116

Notes: The unit of observation is the individual. List is restricted to goals that were named by at least 5% of people at baseline. At baseline, goals were measured by asking respondents about their current savings in various sources. If the amount saved in a source was non-zero, respondents were asked about what the savings were for. Goals were elicited for the lockbox (columns 2-3) at the lockbox check visit in December 2017.

Web Appendix Table A2. Attrition

	(1)	(2)	(3)	(4)	(5)	(6)
	High Frequency Phone Surveys (HFPS)				Monitoring Surveys	
	Round 1 (September-October 2018)		Round 2 (March 2018)		Completed round 1 (January 2018)	Completed round 2 (April 2018)
	Completed at least 1 survey	Percentage completed	Completed at least 1 survey	Percentage completed		
One box	0.00	-0.02	0.00	-0.03	0.01	0.00
	(0.02)	(0.04)	(0.06)	(0.06)	(0.04)	(0.04)
Multiple boxes	-0.04	-0.05	0.03	-0.01	-0.01	0.05
	(0.02)	(0.04)	(0.06)	(0.06)	(0.04)	(0.04)
HFPS	-	-	-	-	-0.12***	-0.04
	-	-	-	-	(0.03)	(0.04)
Observations	240	240	240	240	481	481
Control Mean	0.99	0.89	0.84	0.74	0.86	0.78

Notes: See text for discussion of surveys and Figure A2 for a project timeline. Standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

Web Appendix Table A3. Summary Statistics and Randomization Check based on the Intake Survey

	(1)	(2)	(3)	(5)	(6)	(7)
	Full Sample			HFPS sample only		
	Control Mean	Regression Coefficient for:		Control Mean	Regression Coefficient for:	
		One Box	Multiple Boxes		One Box	Multiple Boxes
Average hours worked (daily)	7.70	-0.53	-0.53	8.13	-0.45	-0.74
	(3.88)	(0.53)	(0.60)	(3.82)	(0.68)	(0.79)
Average daily profits	4.27	-0.14	-0.81	4.57	-0.65	-0.23
		(0.84)	(0.76)		(0.85)	(0.97)
Hours farmed in week prior to survey	0.22	-0.02	0.24	0.09	0.06	-0.05
		(0.21)	(0.31)		(0.23)	(0.20)
Total deposits to savings (daily)	3.40	2.12	-0.68	3.37	1.57	-0.17
	(6.01)	(1.42)	(0.84)	(5.47)	(1.22)	(1.09)
Total withdrawals from savings accounts (daily)	2.14	1.44	-0.68	2.38	1.72	-0.46
	(5.09)	(1.23)	(0.50)	(6.30)	(1.92)	(0.69)
Average Food expenditures (daily)	1.56	-0.02	0.09	1.60	-0.01	0.03
	(1.16)	(0.17)	(0.17)	(1.11)	(0.23)	(0.24)
Average Total daily expenditures[1]	2.49	0.06	0	2.40	0.2	0.01
	(1.79)	(0.26)	(0.24)	(1.45)	(0.28)	(0.28)
Net transfers to friends and family	1.21	-2.23	-4.46*	2.37	-2.61	-3.17**
	(7.86)	(1.40)	(2.29)	(8.75)	(1.99)	(1.60)
Net transfers to spouse	(1.10)	0.35	1.32	(0.50)	-2.1	-1.82
	(9.98)	(1.67)	(2.13)	(8.04)	(2.71)	(1.82)
Observations	457			235		

Note: Variables were measured over the 4 days before the intake survey. Means are population weighted. Randomization check is performed only for businesses that appear in the analysis sample (i.e. completed a monitoring survey or at least one phone survey). Monetary values are winsorized at 1% and expressed in USD.

[1]Expenditures measured at the intake survey are a subset of those in later surveys, and include food, personal items, household items, and transportation.

Web Appendix Table A4. Usage of multiple boxes

	(1) Dominant box	(2) Other boxes
Panel A. High frequency phone surveys		
Made at least 1 deposit during survey period	0.91	0.68
Total value of deposits	78.45	34.44
	(104.43)	(53.97)
Total value of withdrawals	38.02	13.19
	(151.09)	(32.60)
Observations	80	80
Panel B. Monitoring surveys		
Made at least 1 deposit during survey period	0.83	0.57
Total value of deposits	126.67	33.74
	(418.57)	(98.76)
Total value of withdrawals	113.32	23.32
	(429.29)	(76.42)
Observations	161	161

Notes: Panel A is from the high frequency phone surveys and Panel B from the monitoring surveys. Statistics are for group offered multiple boxes only. Standard deviations in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

Web Appendix Table A5. Treatment Effects on Withdrawals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Withdrawals from other savings sources					
	Withdrawals from project lockboxes	Other lockboxes	Cash at home	Bank accounts	Savings groups[1]	Mobile money	Total Withdrawals[2]
Panel A. High frequency phone surveys							
A1. Pooled							
Any box	1.96***	-0.51**	-1.25	1.11	-0.38	-0.44	0.96
	(0.62)	(0.20)	(1.03)	(0.85)	(0.37)	(0.31)	(1.79)
A2. Disaggregated							
One box	0.30	-0.33	0.02	1.82*	-0.59	-0.58	1.25
	(0.88)	(0.21)	(1.27)	(1.02)	(0.38)	(0.37)	(2.02)
Multiple boxes	3.47***	-0.68***	-2.40**	0.46	-0.19	-0.32	0.71
	(0.94)	(0.22)	(1.10)	(1.19)	(0.43)	(0.31)	(2.24)
Observations	2670	2670	2670	2670	2670	2670	2670
Number of Businesses	234	234	234	234	234	234	234
Control Mean	0.00	0.91	5.39	2.75	1.51	1.15	11.97
Control SD	0.00	4.14	13.82	43.06	9.61	4.71	47.53
p -values for F-tests of:							
1 Box = Multiple Boxes	0.02**	0.02**	0.05**	0.35	0.25	0.37	0.82
Joint significance	0.001***	0.006***	0.04**	0.21	0.23	0.29	0.82
Panel B. Monitoring Surveys							
B1. Pooled							
Any box	39.29***	-4.94***	-7.60***	8.76*	-5.63	-0.14	29.13***
	(3.94)	(1.70)	(1.99)	(5.30)	(5.82)	(1.63)	(9.87)
B2. Disaggregated							
One box	29.45***	-4.16**	-7.15***	2.97	-3.36	-1.64	15.43
	(4.16)	(1.92)	(2.04)	(4.84)	(6.06)	(1.66)	(9.92)
Multiple boxes	48.79***	-5.69***	-8.04***	14.35*	-7.83	1.31	42.35***
	(5.61)	(1.65)	(2.13)	(8.10)	(6.77)	(1.97)	(13.52)
Observations	786	786	786	786	786	786	786
Number of Businesses	429	429	429	429	429	429	429
Control Mean	0.00	8.75	11.36	9.75	32.43	7.10	69.98
Control SD	0.00	27.92	24.95	74.86	55.81	17.04	113.20
p -values for F-tests of:							
1 Box = Multiple Boxes	0.002***	0.18	0.48	0.16	0.41	0.08*	0.05**
Joint significance	0.001***	0.002***	0.001***	0.21	0.50	0.18	0.007***

Note: In Panel A, withdrawals are measured over the 7 days prior to the survey. In Panel B, withdrawals are cumulative withdrawals in the 2 months prior to the monitoring surveys (and thus cover November-December 2017 for monitoring survey 1 and February-March 2018 for monitoring survey 2). All regressions in Panel A include a measure of the dependent variable during the intake survey as a control - see text for details. All regressions control for strata, date fixed effects and baseline controls and are population weighted (see in the text for details). All monetary variables are expressed in USD and are winsorized at 5%. Standard errors clustered at individual level in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

[1]Savings groups include VSLAs and ROSCAs.

[2]Total withdrawals is the sum of the other columns, as well as other less common types of savings (such as safekeeping with shopkeepers or friends).

Web Appendix Table A6. Treatment Effects on balances measured in monitoring surveys

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Balance in project lockboxes	Balances in other savings sources					Total Balance[2]
		Other lockboxes	Cash at home	Bank accounts	Savings groups[1]	Mobile money	
Pooled							
Any box	40.80***	0.52	-8.83***	-5.33	-1.43	-3.86*	20.49*
	(3.92)	(0.80)	(2.37)	(11.87)	(3.84)	(2.23)	(12.30)
Disaggregated							
One box	41.05***	1.05	-6.65**	-17.11	0.95	-3.03	16.48
	(5.54)	(1.21)	(2.82)	(11.14)	(4.44)	(2.44)	(13.27)
Multiple boxes	40.55***	0	-10.94***	6.56	-3.72	-4.66**	24.36
	(4.84)	(0.76)	(2.39)	(15.77)	(4.46)	(2.31)	(16.14)
Observations	785	784	782	769	786	780	786
Number of Businesses	429	427	427	424	429	427	429
Control Mean	0.00	1.71	14.46	36.59	17.04	9.00	80.77
Control SD	0.00	10.93	26.24	101.70	37.54	20.26	117.80
p -values for F-tests of:							
1 Box = Multiple Boxes	0.941	0.389	0.055*	0.079*	0.299	0.313	0.632
Joint significance	0.001***	0.660	0.001***	0.101	0.548	0.116	0.250

Note: Balances were measured as of the date of the monitoring surveys. All regressions control for strata, date fixed effects and baseline controls and are population weighted (see in the text for details). All monetary variables are expressed in USD and are winsorized at 5%. Standard errors clustered at individual level in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.

[1]Savings groups include VSLAs and ROSCAs.

[2]Total withdrawals is the sum of the other columns, as well as other less common types of savings (such as safekeeping with shopkeepers or friends).

Web Appendix Table A7. Treatment effects on prevalence of sickness

	(1) =1 if member of household sick in past week	(2) =1 if respondent sick in past week
One box	0.02	0.03
	(0.03)	(0.03)
Multiple Box	0.02	0.03
	(0.02)	(0.03)
Observations	2678	2678
Number of businesses	234	234
Control Mean	0.02	0.02
p -values for F-tests of:		
1 Box = Multiple Boxes	0.98	0.92
Joint significance	0.78	0.53

Notes: Standard errors clustered at individual level in parentheses. Regressions are population weighted.
*, **, and *** indicate significance at 10%, 5%, and 1% respectively.

Web Appendix Table A8. Coping with health shocks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Income		Expenditures		Net transfers		Net withdrawals (all savings sources)
	Total income	Total hours	Medical Expenses	Total	Outside household	Spouse	
Panel A. Household member sick							
Member sick	-0.6	-0.13	1.76***	3.45**	-1.06	-0.07	2.09
	(0.56)	(0.34)	(0.48)	(1.34)	(1.45)	(0.45)	(1.62)
1 Box*Member sick	1.28	0.22	-0.76	-3.04	5.79*	-0.12	-2.11
	(0.97)	(0.47)	(0.60)	(2.54)	(3.01)	(0.83)	(3.86)
MultipleBox*Member sick	0.36	-0.65	-0.32	0.5	-1.52	-0.22	-4.27
	(0.80)	(0.75)	(0.58)	(2.13)	(1.98)	(1.28)	(4.26)
Observations	2692	2692	2692	2692	2692	2489	2688
Number of businesses	235	235	235	235	235	234	235
Control Mean	2.94	1.80	0.74	27.52	0.12	0.55	-9.78
Control SD	9.70	7.27	1.93	20.10	11.50	8.92	30.81
p -values for F-tests of:							
1 Box = Multiple Boxes	0.348	0.246	0.369	0.195	0.014**	0.945	0.681
Joint significance	0.42	0.505	0.423	0.392	0.048**	0.979	0.562
Panel B. Respondent sick							
Respondent sick	1.02*	0.5	2.52***	6.06***	2.58	-0.96	-0.32
	(0.56)	(0.32)	(0.37)	(2.09)	(2.48)	(0.61)	(1.88)
1 Box * respondent sick	-1.29	-0.25	0.79	-2.01	-3.71	0.96	-0.37
	(1.16)	(0.62)	(0.56)	(2.55)	(3.07)	(0.92)	(3.05)
Multiple Box * respondent sick	0.54	0.39	-0.58	-2.16	-1.54	1.91	-1.03
	(1.19)	(0.49)	(0.49)	(3.03)	(2.75)	(1.31)	(3.59)
Observations	2692	2692	2692	2692	2692	2489	2688
Number of Businesses	235	235	235	235	235	234	235
Control Mean	2.94	1.80	0.74	27.52	0.12	0.55	-9.78
Control SD	9.70	7.27	1.93	20.10	11.50	8.92	30.81
p -values for F-tests of:							
1 Box = Multiple Boxes	0.212	0.307	0.012**	0.954	0.32	0.48	0.865
Joint significance	0.416	0.543	0.041**	0.696	0.436	0.284	0.96

Notes: Fixed effects regressions with date fixed effects. Regressions are population weighted. Transfers include gifts and loans, and both cash and in-kind payments. Net transfers and withdrawals are positive for inflows and negative for outflows. Standard errors clustered at individual level in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% respectively.