Linking savings behavior, confidence and individual feedback: A field experiment in Ethiopia

Alexandra Avdeenko [a], Albrecht Bohne [b,*], Markus Frölich [c]

[a] Center for Evaluation and Development (C4ED), O7, 3, Mannheim 68161, Germany
[b] ZEW Mannheim, L7, 1, Mannheim 68161, Germany
[c] Center for Evaluation and Development (C4ED) and University of Mannheim, Department of Economics, L7, 3-5, Mannheim 68131, Germany

ABSTRACT

In this paper we investigate behavioral constraints to savings among smallholder farmers in rural Ethiopia. Increasing savings by overcoming such behavioral constraints has been documented to have positive effects on various outcomes such as health, education, and agricultural investments. We causally identify a strong increase in savings to a soft commitment device in the form of a moneybox with a regular savings plan. In our randomized field experiment, we also provide personalized feedback consisting of recommendations to self-set saving goals. These recommendations trigger increases in savings of up to 36 percent. In a detailed analysis of the behavioral characteristics driving these results, we find a strong and robust link between financial confidence and savings behavior. In particular, the savings of underconfident individuals are less than 2/3 of the savings of overconfident individuals — an association stronger than other behavioral traits such as risk-lovingness and present-biasedness. Remarkably, the effect of our personalized feedback is particularly strong for underconfident individuals. We discuss possible underlying mechanisms, rule out a set of alternative behavioral explanations, and address crowding-out behavior into other forms of saving.[1]

1. Introduction

Recent research has highlighted the importance of savings for individuals in developing countries. People in these countries are often exposed to potentially large idiosyncratic shocks while facing seasonal income patterns and lacking access to social insurance schemes. Finding ways to increase savings among these households has attracted considerable attention

---

* Corresponding author.
E-mail addresses: avdeenko@c4ed.org (A. Avdeenko), albrecht.bohne@zew.de (A. Bohne), froelich@uni-mannheim.de (M. Frölich).

from economists. In fact, a broad literature has shown positive effects of increased savings on a number of development outcomes in areas ranging from agricultural investments (Brune et al., 2016) to health (Dupas and Robinson, 2013) and education (Karlan and Linden, 2017).[2]

In devising strategies to increase savings, previous research has put a strong focus on overcoming behavioral biases, especially through the introduction of innovative savings products serving as commitment devices (Ashraf et al., 2006; Dupas and Robinson, 2013; John, 2017).[3] We build on this literature and distribute a simple savings technology in the form of a moneybox together with a savings plan. Our results draw on a randomized control trial we conducted in 94 rural villages in Northern Ethiopia. Smallholder farmers in this region provide the ideal study population to analyze savings behavior due to the high variability in their income reflecting seasonal patterns in agriculture. We distribute moneyboxes coupled with savings plans to around 900 randomly selected farmers during peak income season. Through the earmarking and labeling of these regular savings installments towards a specific self-set savings goal, the moneyboxes exert a certain amount of commitment. However, since the money can be accessed at any time without transaction costs, they are a soft commitment device with sufficient flexibility for every individual user. They initiate large increases in savings attainment, with farmers saving about 22 percent more than those in a randomly chosen control group.

We randomly combine this savings technology with a behavioral treatment providing a subsample of participants with individualized recommendations on their savings goals. Thereby we test an innovative feedback loop in the form of individualized recommendations on respondents' self-set savings goals. Farmers receive individual feedback consisting of recommendations to save either more or less than their originally stated savings goal. We find that individuals who receive this individualized feedback to reflect and reconsider their savings goal save more (an increase of 181 Birr, or 36 percent). Interestingly, the direction of the recommendation we provided had no impact on actual savings accrual. To get a bit closer at understanding potential underlying mechanisms, we conduct a detailed analysis of the behavioral characteristics of those individuals particularly affected by the recommendations. We begin by studying general correlation patterns between saving levels and observable behavioral characteristics. We put particular emphasis on confidence types by classifying individuals as underconfident if they underestimate their ability to correctly answer a standard World Bank (1998) financial literacy module. In a detailed analysis of baseline (pre-treatment) savings levels, we find underconfidence to be an independent driver of low savings, over and beyond the effects of present-biasedness, risk preferences, and a range of socio-demographic characteristics. The interaction between receiving recommendations and exhibiting underconfidence is strongly positive, irrespective of the direction the recommendation takes. We interpret these findings as lending credibility to two potential mechanisms underlying the effect of the recommendations, in particular additional attention and re-assessment. The feedback may make the participants feel that they are taken more seriously and may help them set their savings goals. Likewise, the re-assessment associated with reconsidering and revising the savings goal might be particularly helpful for indecisive participants and the associated cognitive process might increase the importance respondents attach to the savings process. To support these hypotheses we present suggestive evidence that our recommendation treatments could be effective in real-life interactions and problems. In detailed robustness analyses we further rule out a number of alternative mechanisms and the crowding out of other savings vehicles.

Our paper relates to the growing literature on savings, especially in developing countries. Even though individuals in these countries may be poor, they still have considerable scope in deciding what to spend their money on (Banerjee and Duflo, 2006). In settings without access to formal savings mechanisms, people often save in informal mechanisms ("under-the-mattress"), informal saving groups or investments into livestock (Karlan et al., 2014). Numerous studies have shown the importance of savings for a range of development outcomes for individuals in all income ranges (Brune et al., 2016; Dupas and Robinson, 2013; Karlan and Linden, 2017). Therefore, the literature has focused on identifying a set of constraints on savings, such as lack of financial knowledge (Perry and Morris, 2005; Berg and Zia, 2013; Bayer et al., 2009; Karlan et al., 2014), market frictions and reduced access to financial services (Dupas et al., 2017; Karlan et al., 2014). Moreover, numerous behavioral frictions have been put forward such as time-inconsistent preferences (Ashraf et al., 2006; Dupas and Robinson, 2013; John, 2017), inattention problems (Karlan et al., 2016; Kast et al., 2018), intra-household barriers (Dupas and Robinson, 2013), and procrastination in financial decision-making (Becchetti et al., 2015; Bisin and Hyndman, 2014; Brown and Previtero, 2014; Duflo et al., 2009; Thaler and Benartzi, 2004; Linardi and Tanaka, 2013). Within this large literature, our most direct contribution is to a number of studies showing the effectiveness of lockboxes as a simple savings technology (Dupas and Robinson, 2013; Shipton et al., 1990). Moreover, we particularly contribute to recent studies highlighting the importance of reminders and feedback in savings behavior. Karlan et al. (2016) show that reminders to save increased total bank savings and savings goal attainment by 6%. Kast et al. (2018) highlight the relevance of feedback and follow-up text messages in self-help saving groups. Carvalho et al. (2016) show that bank accounts and weekly visits by deposit collectors lead to a higher accumulation of wealth. In our study, we find strong effects of a relatively weak feedback mechanism and provide evidence on the possible behavioral channels underlying our findings.

The second broad literature our paper connects to is one linking confidence and financial decision-making. Overconfident individuals are known to take higher risks in financial markets (Kirchler and Maciejovsky, 2002; Caliendo and Huang, 2007;

---

[2] Please refer to Karlan et al. (2014) and Kremer et al. (2019) for comprehensive overviews.

[3] The literature has also proposed various other barriers to savings, including transaction costs, information asymmetries, lack of access to financial services, and social constraints, to name just a few.

Doerr et al., 2011), trade too much, too aggressive and earn lower net returns (Barber and Odean, 2001; Gervais and Odean, 2001; Barberis and Thaler, 2003). Studies by Malmendier and Tate (2005) and Garrard and Robinson (2015) argue that over-confident managers over-invest and take lower quality decisions when firms have abundant internal funds.[4] However, within the literature on confidence and financial decision-making there are only few studies relating underconfidence to economic outcomes, especially savings behavior. Perhaps closest to this question is a study in the field of psychology by Tang and Baker (2016), who state that financial actions might be "intimidating, and short-term failures or distractions can undermine responsible long-term financial behaviors". They suggest that the related concept of self-esteem relates to financial behavior both directly and indirectly through subjective financial knowledge. In the same vein, (Campbell, 2016) hypothesizes that financial ignorance could be a driver of wealth inequality, and that approaches to nudge behaviorally biased households, in particular those with low confidence, could help to avoid costly mistakes on the financial market. Likewise, a range of studies relate measurable personality traits such as locus of control and perceived control to financial decisions. In fact, high levels of perceived control seem to be key to savings decisions (Rotter, 1966; Ajzen, 2002; Perry and Morris, 2005; Cobb-Clark et al., 2013; Fouarge et al., 2013). Importantly, (Cobb-Clark et al., 2013) find that households with internal locus of control (believing to be in full control of their lives) save more than households with an external locus of control (believing lives are controlled by external factors). These findings are particularly relevant since self-confidence is one of the defining elements of the locus of control. Our most direct contribution to this broad literature lies in analyzing the effects of confidence for individuals in developing countries. More importantly, we complement the prior focus on overconfidence and take a closer look at underconfident individuals and their saving behavior. We establish a clear link between the two and implement a simple individualized feedback mechanism helping individuals overcome this constraint.

We can draw several policy implications from this paper. First, in our setting, a soft commitment device is sufficient to increase savings. This provides further evidence to the growing consensus that soft commitment devices are more effective than hard commitment devices (Kremer et al., 2019; Karlan and Linden, 2017). Second, our main findings underline the importance of personalized encouragement in the attainment of saving goals, especially for underconfident individuals. Here our findings are in line with recent literature establishing the link between confidence and financial decision-making (Kramer, 2016; Campbell, 2016; Tang and Baker, 2016). The mechanisms we explore suggest scope for improving consultation processes in financial interactions in general. In these interactions, particularly underconfident individuals might need to be addressed with more attention, encouragement, and individual feedback. Extrapolating to even more general settings, our results imply that incorporating individualized feedback might be an avenue to help increase the effectiveness of policy interventions in other domains. Lastly, the fact that individuals in the moneybox treatment group were told that they would be visited again in the future (to check on their progress towards reaching the savings goal), might have further encouraged the usage of the moneybox. While these experimenter demand effects are common in such settings (de Quidt et al., 2018), this aspect is important for the external validity of our findings and scale-up considerations for this particular treatment.

The rest of the paper is organized as follows: In Section 2 we describe the background and experimental design. Section 3 presents our data and key descriptive statistics. In Section 4 we present our results on the moneyboxes, recommendations and underconfidence. Section 5 provides robustness checks and Section 6 concludes.

2. Background and experimental design

We conducted the study in Northern Ethiopia (Tigray region). The experiment took place in December 2014 together with smallholder farmers in 94 rural villages. Participants were randomly selected out of lists of current and former microfinance clients.

2.1. Background

Economic activity in Northern rural Ethiopia is almost entirely dependent on agriculture. More than 90% of the households in our study directly engage in agriculture as their main source of income, and almost all of this is generated through crop production. Agricultural activity in Northern Ethiopia is heavily dependent on rainfall patterns, creating three seasons that are of relevance to smallholder farmers: *Belg* (March to June), associated with little agricultural activity, *Kiremti* (July to September), where cultivation and heavy rainfalls take place, and *Kewie* (October to February), which is the season in which farmers harvest, sell or store their goods. Naturally, seasonal patterns in household income and cultural activities are a direct consequence of these agricultural seasons. During the harvest period *Kewie* (October to February), households typically generate relatively high income streams due to irregular sales of their goods at local markets, wage work as harvest helpers and a general uptake in all other economic activities. The subsequent season associated with little agricultural activities (*Belg*, March to June) typically entails several religious festivities and weddings, for which households are socially expected to make non-negligible expenditures. Agricultural investments, such as fertilizer or seed acquisition, usually tend to take place in *Kiremti* (July to September), the planting and cultivation season. Obviously, the lag between income generation and investment causes difficulties for our study population.

---

[4] Further studies have focused on the behavior of consumers and have shown that overconfident consumers overpay due to overestimation of the benefits of the product or service (Grubb, 2009, Grubb, 2015 and Li et al., 2016).

Our partner organization, a local microfinance institution (MFI), has long tried to devise strategies to increase savings and investment among their client population. Smallholder farmers in our regions of Tigray usually have access to several savings technologies: cash savings in their home, savings with the MFI, or savings in informal savings arrangements, most importantly *Equb* (savings society) and *Iddir* (funeral society). Only roughly 5% of the study population has a formal bank account, and in the subsequent paper we will refer to "bank savings" as savings in the local MFI. Savings at the local microfinance institution, however, also incur sizable transaction costs, since clients usually need to travel to the next larger-sized village. In our sample, mean travel time to the next MFI branch is slightly above 60 min. Especially during harvest season, it may be difficult for farmers to find this additional time.

Apart from savings, farmers in Northern Ethiopia have considerable demand for further microfinance products. Almost 90% of our sample has taken out a loan from the MFI at some point in time. The most frequent months for these loans are June and July, right before cultivation begins and agricultural investments need to be made. Due to this, the MFI has frequently promoted savings behavior among the local population. We chose the introduction of our intervention and provision of savings technology to coincide with the seasonality patterns of agriculture and income generation. Therefore, we implemented the intervention in December, the middle of the harvest season *Kewie*, where most households are selling crops or otherwise in the middle of their income generating activities.

2.2. Experimental design

We visited 940 households and asked to talk to the household head along with a further adult household member (generally the spouse). After a few survey questions, we started with a discussion on the general importance of savings leading towards the setting of an individualized savings goal. To do so, we asked respondents to formulate and talk about the most important savings goal they would like to save for. Examples for these goals include livestock (cows, goats), school books, and fertilizer. We proceeded by asking respondents how much they want to save for this goal and in how much time they want to accumulate this amount. In order to guarantee comparability, we pre-set the timeline to be between 8 and 24 week.[5] For those 300 households randomly selected to the pure control group not receiving a moneybox, the savings discussion ended at this point and was followed by a few more survey questions.[6]

Our treatment was composed of two main elements: First, we offered smallholder farmers a new savings technology, namely moneyboxes with individual savings plans. Second, we provided individualized feedback on the chosen savings amount and randomized the type of feedback.

Moneyboxes with Savings Plans

An emerging consensus in the literature states that savers have a demand for commitment and that softer commitments may be more effective in inducing behavioral change with respect to savings than harder commitments (Kremer et al., 2019; Karlan and Linden, 2017; John, 2017). Thus, we choose a soft commitment device – a moneybox provided to a subset of randomly chosen farmers – to leverage savings. Moneyboxes induce a certain amount of commitment since the cash inside the box is earmarked for a specific goal and using it for something else may induce unease for the owner. However, as opposed to savings in banks, MFIs, or commitment savings accounts, the money held inside moneyboxes is available at all times in case of emergency and does not entail any constraints on the individual's future choice sets.[7] Moreover, the households in our sample dwell in remote areas. Traveling to the next microfinance branch is time-consuming and visits in the village by the branches' savings officers are infrequent. Using a moneybox allows for saving at high, even daily, frequencies at virtually no transaction costs. This is especially important for the detailed saving plans designed together with the participants.

Together with every moneybox, we provided an individualized savings plan. Given the goal amount participants chose, the enumerator calculated regular (daily or weekly) savings installments necessary to reach the goal by the self-set deadline.[8] We asked participants to write down the following information on the label of the box: (1) the savings goal (written or drawn, e.g. cow), (2) the chosen goal amount in *Birr*, (3) the regular savings installment in *Birr*, and (4) the savings end date. At the end of the interview, we put 30 *Birr* into the moneybox.[9] Additionally, the interviewers let households know that they would be visited again at an unspecified future date to check on their progress towards reaching the savings goal.

Feedback to Reconsider Original Savings Goal

Among those farmers receiving the moneybox and savings plan, a random subset received individual recommendations to reconsider their original savings goal. Directly after stating their initial goal amount, the participants were asked to

---

[5] If the costs of the goal (say, a cow) exceed the amount an individual believes to be able to accumulate in the chosen time frame, we asked that person to state the amount of money they would like to save for this same goal during the chosen time period.

[6] We wanted to ensure that also households not receiving the moneybox were given the same general discussion about savings as treatment households.

[7] The moneyboxes we employed are round cylindrical plastic boxes with a slit at the top to facilitate the introduction of cash. The bottom of the moneybox has a hole with a small lock. Two keys were handed out to each household.

[8] Whether the installments were daily or weekly was also randomized. This treatment was cross-randomized, i.e. it did not have an impact on the results presented in this study. Moreover, we also cross-randomized whether or not the moneybox was transparent. These results are part of a separate study and available upon request.

[9] 30 *Birr* ≈ 1.5 USD. The pure control group without moneybox also received the payment.

reconsider the amount of their planned savings. The enumerator asked them if they would like to *increase* their initial goal amount by 20 percent (treatment arm I) or 40 percent (treatment arm II). A further set of farmers were given the recommendation to *decrease* their initial goal amount by 20 percent (treatment arm III) or 40 percent (treatment arm IV).[10] These treatment arms were randomized. Our core goal was to choose thresholds that are easy to implement in practice. We did not follow a theoretical justification for this choice. They had to be realistic and allow for a simple and quick calculation in the field, reflect varying intensities, and leaving sufficient number of observations to detect an effect. Following these recommendations, we asked participants whether they would like to revise their originally stated goal amount and what this new goal amount should be. Given this possibly revised savings amount, the enumerators then proceeded as described above by calculating the regular savings installments and labeling the box accordingly. A fifth treatment arm received a confirmation of their initially stated goal without any recommended change.

These recommendation treatments draw on individual information – the initially stated savings amount – and adjust this according to a pre-determined rule. In other words, they provide participants with individualized feedback on their saving goals.

Sampling and Randomization

We sampled a total of 940 households in 94 village clusters. These village clusters basically represent a census of the zone in which the study was conducted. Within each village cluster, 10 households were randomly chosen from lists of current and former clients at our partner MFI. Enumerator teams from the University of Mekelle conducted the interviews with the respondents. These enumerators were members of the local population and received extensive training in administering the survey questionnaires.

The sample of 940 households was split into random subsamples as follows: 640 households received a moneybox and 300 households belong to the pure control group without moneybox. The 640 households receiving the moneybox were randomized by equal parts (128 each) into the five recommendation treatments, including the one receiving no recommended change. Randomization was conducted at the individual household level.

For the randomization, we stratified the different blocks according to a range of baseline characteristics. These included important financial and economic measures, specifically savings in cash (amount in *Birr*), current savings with the local MFI (amount in *Birr*), total outstanding debt (amount in *Birr*),[11] whether the household is a member of an *Iddir*, an informal funeral insurance, or *Equb*, an informal savings group (both binary indicators), the total land area under cultivation (in hectare), the total revenues from crop sales (in *Birr*) and the per-capita consumption expenditures (in *Birr*). Furthermore, we considered the demographic composition of the household (measured by the number of household members between 0–5, 6–11, 11–17, 18–65 and more than 65 years of age), whether the household head is female or married (both binary indicators) and the years of education of the household head. We chose this combination of economic and socio-demographic variables to reflect those aspects of our baseline data which we believe are most important in determining savings outcomes and for which we sought balance.

Randomization was conducted in December 2014 before treatment implementation and after collection of the first baseline survey. Household-level treatment status was randomly assigned while balancing the Mahalanobis distances in the above-mentioned baseline variables used for stratification. The 940 households were assigned to either pure control group or one of the treatment arms. Thereafter the Mahalanobis distance with respect to the baseline covariates given above was calculated. Mahalanobis distance was calculated pairwise between any two treatment arms and also for each treatment arm relative to control. The random assignment was only accepted if the maximum distance to control was smaller than the 0.01 percentile of its empirical distribution (0.605) and the maximum distance between any of the treatment arms was smaller than the 0.01 percentile of its empirical distribution (1.322). The treatments in the study were assigned according to the first random assignment vector that passed the balance thresholds.

3. Data and descriptives

3.1. Timeline, data, and attrition

We use data from three rounds of comprehensive household-level surveys. Fig. 1 provides a timeline of the data collection and treatment implementation process. An extensive baseline survey was administered in November/December 2013 on all sample households ("1st baseline"). We distributed the moneyboxes and implemented the recommendations in December 2014. This implementation was combined with a short survey (administered before providing the moneyboxes) to

---

[10] While our experiment encourages all participants to increase savings, these individual recommendations towards their saving goals can be seen as an additional variation in the intensity with which we encourage increased savings. We made a number of efforts to contextualize this research design in the local context and make sure that it would be well understood by the participants. The exact phrasing in the questionnaire was decided upon after a long process including discussions with local experts from the University of Mekelle and two rounds of pilots in Tigray. This final version involved the enumerators referring to their experience when introducing the additional encouragement. See Appendix A.1 for experimental details, including the original text read out to the participants and a further discussion on potential ethical implications.

[11] The variable debt is the value of total outstanding debt. As such, we consider the sum of all possible sources of debt.

Treatment Implementation

1st Baseline   2nd Baseline   Endline
                                   Distribution of Moneyboxes and Recommendations

[saving goals reached with

ongoing saving (all participants)   individually varying timing]

AEA Registry
November/December 2013   *December* 2014   *January/February* 2015
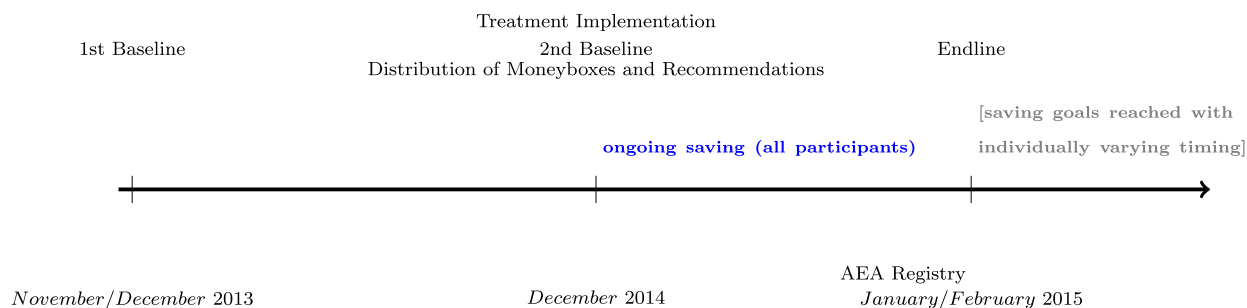
Fig. 1. Timeline.

enable gathering further data on key baseline variables, among them the financial literacy module at the heart of our measure of confidence. We call this survey the "2nd" baseline. We timed the savings intervention in December 2014 to coincide with the middle of the harvest season, when our study population enjoys high levels of income and has the opportunity to save (refer to Section 2.1 for details). This setup effectively provides us two baseline surveys. For important variables available at both points in time, we average them across both waves to create cleaner measures less prone to measurement error.

We administered the endline survey in January/February 2015, right at the end of the harvest season and the beginning of the period with little agricultural activity for two main reasons. First, this point marks the end of the high-income period for agricultural activity and is the ideal timing to measure the stock of savings households were able to accumulate during this period of relative prosperity. Second, in the design of the experiment, we asked households to decide on the endpoint of their savings plan (length restricted to two to six months). This flexibility creates an issue for the data collection process: At any point during the savings process, some individuals may not have reached the end of their savings plan. We conducted the unannounced endline survey almost two full months after treatment implementation, thereby guaranteeing that all participants were still within their savings timeline. This allows us to check whether households are on track towards reaching their savings goal right before the first households are scheduled to reach their goal date.

Our baseline and endline surveys encompass a large range of economically and behaviorally relevant variables and measures. We put a special focus into gathering comprehensive savings data. Given that savings (especially non-bank savings) are typically difficult to observe and measure correctly, we exploit the fact that the moneyboxes let us accurately observe how much cash is inside them at a given point in time, thereby reducing measurement error compared to conventional survey questions considerably. We consider total monetary savings, which are savings in the moneybox (opened and directly counted by our enumerators), *plus* savings in cash the household head holds during the interview. In most cases, these further cash holdings outside of the moneybox were also physically counted by the enumerators. Therefore, our measure of monetary savings is based on observations with little measurement error and does not, as is often the case, represent only self-reported savings.[12]

Moreover, to assess possible crowing-out behavior, we cover a range of additional saving measures, including savings at the local MFI ("bank") and contributions to informal savings arrangements, specifically *Iddir* (funeral society) and *Equb* (savings society). These measures will be introduced and discussed in more detail in Section 5 on the robustness of our results.

The study draws on three waves of extensive surveys with relatively low levels of attrition. Between the first baseline at the end of 2013 and the implementation of the treatments in 2014, we have a modest amount of attrition of 3.7% (the enumerators relocated 905 out of 940 individuals). However, the attrition between treatment implementation and the endline survey is considerably lower and amounts to only 0.7% — we were able to find 899 of the 905 household that received the treatment. Since this second level of attrition would be a stronger threat to identification, this reinforces our findings. After deleting households with missing information in key outcome variables, the final sample in our analysis amounts to 881 households (599 in the moneybox sample).[13] While the problem of attrition and/or missing data is limited in its scope, we analyze it in more detail in Section 3.3 below.

---

[12] We cannot, however, fully exclude the possibility that the households still have some hidden savings they are not showing to the enumerators. However, this is a general issue in the savings literature. We go to great length to reduce measurement error. We ask respondents to show us their MFI savings passbooks so that the enumerators can copy the last entries. Moreover, we elicit cash savings by asking respondents to report the composition of their further cash holdings in bills and coins, which typically led to enumerators counting the money together with respondents and thereby considerably reducing measurement error and ensure truthful reporting.

[13] Even when applying a more stringent definition of attrition including these cases of missing information, attrition between treatment and endline would still only amount to 2.7%. Note that in the subsample of households receiving the moneybox, attrition is even smaller: While between the first baseline (640 households) and the treatment implementation (614 households) attrition amounts to 4%, in the endline we relocated all individuals that received a moneybox (0% attrition).

3.2. Confidence

This subsection introduces and discusses the central behavioral aspect analyzed in this paper, namely confidence regarding financial decision-making. We provide information on how we construct our measure of confidence, discuss the related literature, and validate our measure. Lastly, we analyze the relation between confidence and pre-treatment saving levels and provide evidence for how underconfidence can be seen as an independent driver for low savings.

Concept of Confidence

Given the aim of our study in analyzing household savings behavior in developing countries, our core focus will be on measuring confidence regarding financial decision-making. However, eliciting such financial literacy scores and the corresponding perceptions of ability in the developing country context poses its own set of difficulties which, we believe, can only be overcome through simple, clear, and direct survey questions. As described above, assessing confidence (according to our preferred concept of estimation) consists of two parts: Creating a measure of true ability and a measure of perceived ability. As we are interested in assessing financial confidence, our measure of true ability needs to capture ability related to aspects of financial decision-making which are relevant to our study population. We therefore limit the scope of our questions to those dealing with basic financial numeracy and refrain from asking additional questions related to financial instruments not accessible to our study population.[14]

The second part needed to construct our measure of confidence relates to the perceptions about own ability. The standard approach in more developed settings is to ask individuals for a general self-assessment of their knowledge on financial matters, usually in the form of a Likert-scale (Bannier and Schwarz, 2018; Allgood and Walstad, 2016; Lusardi and Mitchell, 2014; Disney and Gathergood, 2013; van Rooij et al., 2012). These general assessments usually consist of questions like "How do you rate your knowledge about financial matters". To make these answers comparable to measures of true ability, these studies determine confidence levels depending on whether individuals rank higher in their perceived knowledge or their true knowledge. In a more sophisticated approach, (Kramer, 2016) regresses a measure of self-assessed literacy on an index of actual financial literacy and uses the residual to construct a measure of confidence. In our context, however, all of these approaches would require participants to understand what is meant by "financial knowledge" or "financial literacy". And while the development literature has acknowledged that psychological traits such as confidence influence financial literacy and low savings (Kremer et al., 2019), previous studies have taken different approaches to measure related concepts. They all have in common that the measured concepts are easily understandable. A number of papers draw on generalized self-efficacy scores to construct a proxy for the broader concept of self-confidence (McKelway, 2018; Roy et al., 2018; Krishnan and Krutikova, 2013). In alternative approaches to measuring perceptions, instead of asking for perceptions of general concepts such as financial knowledge, researchers have asked about knowledge on specific, but generalizable, concepts. Steinert et al. (2018), for example, measure financial self-efficacy by asking respondents about smoothing consumption over the month and effectively planning a monthly budget. Bedi and Marshall (2002) asked teachers about self-assessed teacher quality and (McKenzie and Sansone, 2019) ask entrepreneurs about confidence in business plans. However, these studies were not primarily interested in creating measures of confidence and used their measure of perceptions per-se as measures of financial self-efficacy.

In order to establish a specific measure of ability which is directly comparable to our measure of true ability, we kept our measurement of financial confidence as straightforward as possible. In the original sense of focusing our measure of confidence on over- or under-*estimation* (Moore and Healy, 2008), we elicit the number answers individuals believe to have answered correctly and compare this to their actual number of correct answers.[15] One notably very similar approach is used by Suzuki et al. (2018) who compare individual predictions on the number of correct answers on a maths test to the actual number of correct answers in Ethiopia. Similarly, we apply this approach because it is in essence comparable to other well-established, often more advanced approaches discussed above yet still relatively simple and clean. By making a comparison of beliefs with actual test outcomes it goes further than simple survey questions. At the same time it avoids complex terms and instructions in rural Ethiopia. Moreover, we follow the standard approach in the literature and use a non-incentivized, domain-specific measure. To the best of our knowledge, we are the first to apply this approach to measure financial confidence in developing countries. For simplicity, we will refer to our measure of financial confidence as confidence throughout the rest of the paper.

Operationalization

Specifically, we measure confidence using financial literacy questions. We apply a standard World Bank financial literacy questionnaire and compare the correctness of the answers to an individual's *perceptions* of their correctness (details and exact wording below). To design our measure of confidence, we assess whether individuals are (1) overconfident, (2) underconfident or (3) correctly confident using the following standard (World Bank, 1998) financial literacy questions:

---

[14] Established papers in the literature often ask a number of questions related to more or less complex financial instruments (Bannier and Schwarz, 2018; Allgood and Walstad, 2016; Lusardi and Mitchell, 2014; Disney and Gathergood, 2013; van Rooij et al., 2012). Our list of three questions is limited though not unusual in the literature. Kramer (2016) uses five question for basic numeracy and Ortoleva and Snowberg (2015) look at four factual questions.

[15] The example Moore and Healy (2008) make is: "If a student who took a 10-item quiz believes that he answered five of the questions correctly when, in fact, he got only three correct, then he has overestimated his score."

   I. Let's assume that you deposited 1000 *Birr* in a bank account at 5 percent annual interest rate. How much money will you have in your account in 2 years if you do not withdraw from or add to this account any money?

   II. Let's assume that in 2015 your income is twice as now, and the food prices also grow twofold. Do you think that in 2015 you will be able to buy more, less, or the same amount of goods and services as today?

   III. Let's assume that you saw a mobile phone of the same model offered from two different sellers. The initial retail price of it was *Birr* 1000. One seller offered a discount of *Birr* 150, while the other one offered a 10 percent discount. Which one is a better bargain a discount of *Birr* 150 or 10 percent?

Every question has three possible answers but only one of them is correct. After individuals answer these questions we ask how many questions they believe they have answered correctly.[16] We classify as underconfident those individuals who believe to answer less of the financial literacy questions correctly than they actually did. Conversely, an individual that overestimates the amount of correct answers is classified as overconfident. Individuals that neither over- nor underestimate their performance are classified as correctly confident.

Validity

To give more credibility to our chosen approach we correlate our preferred measure with (1) alternative, yet related measures of financial confidence referred to as underplacement and (2) with the individuals' assessments of their abilities to reach their financial goals.

In our survey, we additionally asked respondents how many questions they believe other people in their village answered correctly. This gives rise to an alternative, but related concept, namely that of *relative placement*.[17] Specifically, we calculate the difference between an individual's perception about their own knowledge and her perception about the knowledge of other individuals. In this approach, we stipulate that individuals exhibit "underplacement" if they believe to be less knowledgeable than others in their village about the financial literacy questions. Additionally, we design a further measure of "unjustified underplacement" by drawing on the actual responses of other village members in our sample. To this end, we construct a double difference between our previous measure of "underplacement" (the difference between the perceptions about own ability and that of others) and the difference between the respondent's true number of correct answers and the leave-out number of correct answers by other village members in our sample. As a result of this double difference, if someone believes to know less than others and this turns out to be true, they are not deemed as underplaced. Conversely, if a respondent believes to know as much as others in the village while in fact exhibiting higher knowledge, they are classified as "unjustifiably underplaced".[18] We find strong correlations between our preferred measure of underconfidence and both of these alternative measures regarding relative placement to other village members (see Table A.1 in the Appendix).

Apart from these alternative behavioral measures, we also correlate our measure of underconfidence with elicited perceptions about individual's confidence in reaching saving goals. To do so, we analyze responses from a set of questions on general savings behavior and also information related to the new savings goals that we set in the course of our experiment. Our measure of underconfidence is highly correlated with statements such as *I find working towards my savings goals to be very stressful* and *I have too many savings goals*, which can characterize underconfident individuals as doubtful and insecure. Moreover, following our experiment motivating individuals to save, we asked about their specific expectations in reaching the savings goal. Both a binary indicator of individuals expecting some type of problem in reaching their savings goal and a binary indicator for individuals with a subjective probability of reaching their goal of less than 100 percent are correlated with our chosen measure of underconfidence. Even when including baseline controls and controls for treatment status, individuals who expect problems in reaching their goal are seven percentage points more likely to be deemed as underconfident at baseline according to our preferred measure (please refer to Table A.1 in the Appendix for details).

Lastly, one potential additional issue regarding our measure of confidence are floor and ceiling effects: Individuals answering all questions correctly cannot be overconfident and individuals answering everything wrong cannot be underconfident. In robustness tests in Section 5 we exclude these individuals and find even stronger results.

Confidence and Baseline Savings

In order to better understand the driving forces behind savings accumulation, we begin by analyzing which behavioral factors are important in determining savings levels. To do so, we analyze correlations of baseline saving levels *before* the implementation of any experimental treatments with observable behavioral characteristics. We provide evidence of underconfidence as an independent driver of low savings levels.

Specifically, we construct a comprehensive measure of baseline savings as the sum of cash holdings, bank savings, and money in informal saving arrangements. We run a simple OLS regression with our measure of baseline savings as the

---

[16] The exact wording of the question is "What do you think: how many of the last 3 questions did you answer correctly?"

[17] As (Moore and Healy, 2008) show, this concept is often confounded with our preferred measure, while in fact they argue that it would be drawing on different underlying psychological factors.

[18] As in any double difference, it is also possible to rearrange the differences the other way around. By doing so, we would look at the difference between the confidence regarding an individuals own decisions (perceptions - number of correct answers) and the confidence regarding the actions of others in the village (perceptions about others - leave-out correct answers of others).
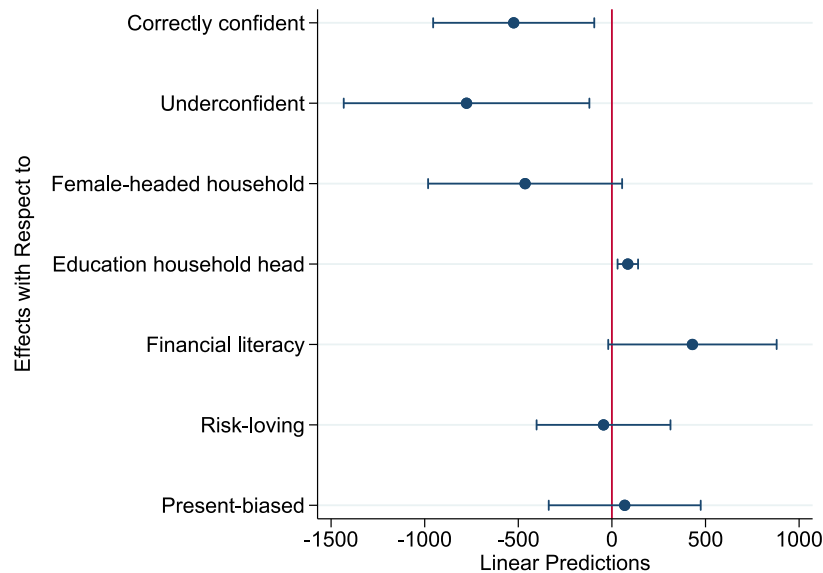
Fig. 2. Savings Behavior Before Treatment *Note*: N=599. This figure depicts the coefficients from a simple OLS regression with a broad measure of savings at baseline (cash + bank + informal arrangements) as the dependent variable. Additional control variables include all baseline variables used for stratification during randomization excluding savings outcomes (socio-demographics, land and consumption variables) as well as enumerator indicators. The savings measure is taken *before* treatment implementation. The variables depicted in the figure are measured as indicator variables with the exception of education, which is measured in years of schooling. 95% confidence intervals are depicted around the point estimates.

outcome and indicators for our central behavioral measure of underconfident and correctly confident individuals. We control for important socio-demographic observables and behavioral traits, including whether a household is female-headed, the years of education of the household head, our score for financial literacy, and indicators for risk-loving and present-biased individuals. Additionally, we also control for a range of socio-demographic and economically relevant variables: indicators for the age structure of the household, marital status of the household head, land ownings, revenues, outstanding debt, per capita consumption expenditures, and lastly enumerator indicators. We are interested in quantifying the relative importance of various individual characteristics and behavioral traits in explaining savings behavior. Thus, we compare confidence to other key characteristics introduced above. Fig. 2 shows the coefficients from this OLS regression along with 95% confidence intervals comparing confidence to an indicator for female-headed households, the education of the household head, financial literacy, as well as risk and time preferences.[19]

The figure clearly depicts the strong association between the different confidence types and savings behavior. Being underconfident is a statistically significant predictor for holding less savings at baseline, even when controlling for a range of socio-demographic and economic variables. Underconfident individuals save more than one third less than overconfident individuals. Interestingly, the confidence levels are more important in predicting savings behavior than other behavioral traits frequently associated with financial outcomes such as risk-lovingness and present-biasedness. The coefficients of the indicator for financial literacy and our measure for years of education of the household head are positively associated with saving levels, as would be expected. In Fig. A.1 in the Appendix, we show these same coefficients for both of our alternative measures of confidence with regard to others and find extremely similar results. We interpret these results as underlining the importance of confidence measures in being strongly associated with financial outcomes and will investigate the effectiveness of our behavioral treatments especially with respect to our measure of confidence.

---

[19] Risk and time preferences were incentivized measures. In more detail, in our baseline survey we elicit time and risk preferences of household heads. Risk preferences are elicited with a standard incentivized framework going back to Holt and Laury (2002). We present 7 distinct lotteries to participants, each consisting of two alternatives which will materialize with a 50% probability. The first lottery has a payout of 3 *Birr* in both states of the world. The following lotteries incrementally increase the expected payout but also the variance. The respondents are asked to choose which of these lotteries they would like to participate in. In a further section of the baseline survey, we elicit time preferences. To this end, we ask participants whether they would like to receive a certain amount of money tomorrow or within one month. By setting the time for receiving the first payment to the day after the survey instead of the day of the survey, we guarantee that individuals make the decision purely based on their time preferences and are not affected by other considerations like the perceived credibility of actually receiving the payment. In five subsequent questions, the payoff occurring in one month is incrementally increased. The switching point provides us with a measure for the underlying discount rate that an individual applies. In a later part of the survey, we repeat the elicitation with questions comparing payoffs in one year to those in one year and one month. If an individuals discounts values faster in the present than in one year, they are classified as present-biased. Individuals were informed at the beginning of the survey that one of the questions would be randomly chosen and the payouts provided. Future payouts were provided through the branch office of the local MFI.

3.3. Descriptive statistics

This section presents balance statistics for our two main randomized treatments: the distribution of moneyboxes and the recommendations. We continue the section by comparing individuals according to their confidence levels.

Balance Statistics: Moneyboxes and Recommendations

Table 1 reports detailed balance statistics for the moneybox and the recommendation treatments. The first three columns focus on the full sample and are divided into individuals that received the moneybox in column (1), those that did not in column (2), and a p-value of their differences in column (3). Additionally, to address potential selective attrition throughout our survey process (refer to Section 3 for details), we include the p-value for the difference between treatment and control group in the original sample surveyed in the first baseline in column (4). The remaining columns focus on the subsample receiving the moneybox and are categorized by whether or not an individual received a recommendation. While columns (5)-(8) follow the same logic as columns (1)-(4), we additionally include column (9) which reports the p-value of a joint orthogonality test (JOT) for the individual recommendations (+/−40%, +/−20%, and 0%). In Panel I., we report the differences for the baseline variables on which we randomized. All variables, which will also be the control variables in all our regressions, are well balanced with the difference in means being insignificant in all but one case.

Panel II. reports means and differences for a number of additional measures of personality, including our central measure of confidence, the expected financial literacy score, actual financial literacy score (including dummies for each number of possibly correct answers), present-biasedness, and risk-lovingness. Even though these variables were not included in the randomization process, overall the balance is also very good, with the exception of our measure for present-biasedness in the moneybox treatment vs. control comparison.[20] Lastly, Panel III. conducts the same comparison for the initially set goal amount, the saving duration, and analyzes the propensity for attrition or missing data with no notable differences. In both the moneybox and the recommendation treatments, attrition and/or missing data between the first and the last wave is not systematically related to treatment status.[21]

Finally, in a test for joint significance of all baseline variables (all those reported in Table 1 with the expectation of the confidence levels), we cannot reject the null that these baseline values do not explain treatment status (p-values of 0.6 or 0.8 respectively, last line of Table 1). We interpret these results in a way that our main treatments are well balanced along key baseline indicators and along our important personality measures.[22]

Detailed Summary Statistics: Moneybox Sample.

Given the focus of this paper on different degrees of confidence, we further split the descriptive statistics using this central behavioral measure and focus on the sample of individuals receiving the moneybox. Table 2 reports these summary statistics (means and standard deviations) for the moneybox sample, split up according to their confidence levels. The last column displays the p-value for the null hypothesis that there is no difference between the average characteristics of underconfident individuals and others.[23]

Out of 599 individuals in the moneybox sample, 16 percent are characterized as underconfident (94 individuals), 47 percent as overconfident (283 individuals) and the remaining 37 percent are correctly confident (222 individuals). The majority of household heads are married (74 percent) and 22 of households are led by women. Household members are on average relatively young, with the average household having more members below 18 years of age (3.05 individuals on average) than above 18 years (2.6). The household head has on average 3 years of education. Comparing the characteristics across the three types of confidence, we find that underconfident individuals are as likely to live in a female-headed household as others. Their household head is also as likely to be married. Yet, at baseline underconfident individuals do seem to differ to overconfident individuals in a number of characteristics: The household head has about one year less in education, land size is slightly bigger, debt is considerably smaller, and revenues from business activities are smaller.

The variation in baseline savings is big. However, this is partly due to possible measurement error in our reported savings variables, which is eliminated in our endline survey due to our improved measurement process based on direct observations. The baseline average of cash savings is about 4430 *Birr*. However, our baseline savings measures show striking differences according to an individual's confidence level: Underconfident individuals save 2816 *Birr* in cash, while overconfident ones save twice as much (5633 *Birr*). Bank savings are at considerably lower overall levels (2290 *Birr* on average) and are again smaller for underconfident individuals (1150 *Birr*) than for overconfident individuals (3175 *Birr*). Furthermore, underconfident

---

[20] To address this disbalance, we re-estimate all our results controlling for present-biasedness and find that this disbalance does not drive our results. Please find the results in the Online Appendix for your reference.

[21] Please note that due to our multiple survey waves, it is possible to define attrition differently according to the wave in which the data was lost. The results presented here are the most conservative ones, that is we look at attrition between the first baseline and the endline. For the same analysis with alternative attrition definitions, please refer to the Online Appendix. Please also note, we do not include all variables from Table 2 into Table 1 (such as revised saving goal, investment goal, lack of expected problems, and subjected probability to reach the goal), because differences in means are likely to be an outcome of the recommendation treatment. In fact, this is what we find when we analyze the variables in the Section 4.2.

[22] We also compare balance across different recommendation types and find no evidence of systematic imbalance (refer to Online Appendix).

[23] Additional p-values comparing means of underconfident vs. overconfident, underconfident vs. correctly confident, and overconfident vs. correctly confident individuals are provided in the Online Appendix.

Table 1
Balance Statistics.

Treatment	Moneybox?				Recommendation?				JOT
	Yes	No	Difference		Yes	No	Difference		
	[est.] (1)	[est.] (2)	p-value [est.] (3)	p-value [orig.] (4)	[est.] (5)	[est.] (6)	p-value [est.] (7)	p-value [orig.] (8)	p-value [est.] (9)
I. Baseline Control Variables:									
Cash savings	4429.933 (506.937)	3845.156 (475.062)	0.469	0.843	4171.175 (530.764)	5495.923 (1399.312)	0.301	0.391	0.420
Bank savings	2289.851 (384.676)	3301.262 (1451.681)	0.376	0.393	2160.774 (371.362)	2821.607 (1243.637)	0.496	0.559	0.591
Iddir member	0.439 (0.020)	0.401 (0.029)	0.283	0.291	0.434 (0.023)	0.462 (0.046)	0.586	0.524	0.950
Equb member	0.297 (0.019)	0.301 (0.027)	0.898	0.717	0.280 (0.020)	0.368 (0.045)	0.064	0.069	0.056
Household members aged 0–5	0.850 (0.032)	0.826 (0.045)	0.673	0.591	0.873 (0.035)	0.752 (0.071)	0.130	0.133	0.161
Household members aged 6–11	1.078 (0.036)	1.124 (0.054)	0.478	0.262	1.089 (0.041)	1.034 (0.077)	0.544	0.512	0.911
Household members aged 11–17	1.120 (0.038)	1.011 (0.053)	0.098	0.189	1.098 (0.042)	1.214 (0.088)	0.224	0.277	0.081
Household members aged 18–64	2.548 (0.043)	2.571 (0.074)	0.771	0.688	2.541 (0.048)	2.573 (0.096)	0.772	0.798	0.648
Household members aged 65+	0.048 (0.009)	0.028 (0.011)	0.190	0.169	0.046 (0.010)	0.060 (0.022)	0.537	0.652	0.919
Female household head	0.215 (0.017)	0.238 (0.025)	0.460	0.538	0.216 (0.019)	0.214 (0.038)	0.961	0.925	0.801
Married household head	0.736 (0.018)	0.727 (0.027)	0.772	0.880	0.730 (0.020)	0.761 (0.040)	0.504	0.597	0.716
Education household head	3.280 (0.153)	3.418 (0.228)	0.613	0.782	3.259 (0.171)	3.368 (0.342)	0.780	0.703	0.305
Land size	4.577 (0.224)	4.201 (0.213)	0.294	0.402	4.495 (0.226)	4.917 (0.673)	0.457	0.592	0.559
Revenue	3951.387 (530.294)	18249.379 (12535.627)	0.098	0.093	4023.166 (641.143)	3655.684 (631.707)	0.784	0.705	0.144
Debt	2974.743 (182.291)	3709.057 (632.438)	0.149	0.192	2932.512 (202.619)	3148.718 (418.805)	0.639	0.901	0.892
Consumption expenditures	736.633 (40.050)	722.567 (53.714)	0.839	0.658	747.178 (44.341)	693.190 (93.427)	0.593	0.572	0.694
II. Personality:									
Underconfident	0.157 (0.015)	0.142 (0.021)	0.561	0.558	0.151 (0.016)	0.179 (0.036)	0.455	0.416	0.789
Overconfident	0.472 (0.020)	0.475 (0.030)	0.940	0.916	0.463 (0.023)	0.513 (0.046)	0.330	0.314	0.703
Expected financial literacy score	1.993 (0.037)	2.032 (0.054)	0.554	0.491	1.981 (0.042)	2.043 (0.076)	0.509	0.488	0.454
Financial literacy score	1.566 (0.038)	1.532 (0.055)	0.610	0.657	1.560 (0.042)	1.590 (0.082)	0.756	0.611	0.513
- 0 answered correctly	0.147 (0.014)	0.156 (0.022)	0.724	0.941	0.156 (0.017)	0.111 (0.029)	0.223	0.226	0.536
- 1 answered correctly	0.294 (0.019)	0.305 (0.027)	0.736	0.599	0.282 (0.021)	0.342 (0.044)	0.204	0.348	0.273
- 2 answered correctly	0.406 (0.020)	0.390 (0.029)	0.660	0.757	0.409 (0.022)	0.393 (0.045)	0.759	0.955	0.611
- 3 answered correctly	0.154 (0.015)	0.149 (0.021)	0.858	0.746	0.154 (0.016)	0.154 (0.033)	0.993	0.928	0.435
Present-biased	0.282 (0.018)	0.355 (0.029)	0.029	0.028	0.278 (0.020)	0.299 (0.043)	0.649	0.661	0.621
Risk-loving	0.489 (0.020)	0.457 (0.030)	0.380	0.432	0.485 (0.023)	0.504 (0.046)	0.716	0.969	0.439
III. Other:									
Initial goal amount (*Birr*)	1975.745 (214.239)	1884.675 (191.931)	0.789	0.750	1940.686 (196.779)	2121.121 (744.192)	0.739	0.796	0.475

(*continued on next page*)

Table 1 (*continued*)

Treatment	Moneybox?				Recommendation?				JOT
	Yes	No	Difference		Yes	No	Difference		
	[est.] (1)	[est.] (2)	p-value [est.] (3)	p-value [orig.] (4)	[est.] (5)	[est.] (6)	p-value [est.] (7)	p-value [orig.] (8)	p-value [est.] (9)
Saving duration (in weeks)	22.956 (0.114)	22.884 (0.182)	0.731	0.696	22.944 (0.125)	23.009 (0.272)	0.822	0.673	0.746
Attrited/ missing (waves 1–3)				0.536				0.259	
N	599	282	881	938	482	117	599	640	599
Proportion	0.680	0.320			0.805	0.195			
F-Test			0.859				0.700		
p-value			0.623				0.804		

Note: Standard deviations in parenthesis. Columns (1)/(5) and (2)/(6) denote the average value of the relevant variable depending on treatment status. Column (3)/(7) denotes the p-value of the difference between treatment and control for the estimation sample and column (4)/(8) denotes the p-value of the difference between treatment and control for the original sample (without attrition). In the last column (Column (9)) we report the p-value of the joint orthogonality test (JOT) for recommendations (+/−40%, +/−20%, and 0%). The significance is given as follows: * p < 0.1, ** p < 0.05, *** p < 0.01. Part I of the table includes baseline variables (measured at baseline before randomization) for which we control in all estimations. Part II displays key personality measures. For a test of joint significance, we estimate a regression with a binary indicator for treatment status on the left-hand side and all the balance variables used for stratification during randomization in the table on the right-hand side. The F-statistic and corresponding p-value for the null of joint zero effects are denoted in the last two rows.

individuals seem to save slightly more in traditional saving groups (*Iddir* membership 12 percentage points above sample average). Our baseline measures for time and risk preferences (present-biasedness and risk-lovingness) show no systematic differences between individuals with varying levels of confidence.

Table 2 further reports summary statistics of key factors related to the introduction of the treatments and the simultaneous measurement of our confidence measures. The self-set savings plans and individual responses to the recommendations strongly differ between under- and overconfident individuals. Underconfident individuals set lower savings goals and are more likely to revise them. While the average original goal amount in the sample is 1976 *Birr*, the corresponding amount for underconfident individuals is about 19 percent lower. After being encouraged by interviewers to reconsider their savings goal, underconfident individuals change their goals upwards by 56 *Birr* on average. This is especially remarkable considering the fact that respondents received recommendations to increase and decrease saving goals in equal proportions. Overconfident individuals, on the contrary, decreased their saving goals by 23 *Birr* on average. Most respondents report saving towards some form of investment goal (73 percent).

Additionally, we collected information on individuals' beliefs about their ability to save and their expected obstacles. 66 percent of the underconfident respondents name specific obstacles in reaching their goals, whereas almost half of the overconfident individuals do not expect any problems in reaching their goal. Likewise, underconfident individuals report slightly lower probabilities in reaching their goal (on average 81 percent compared to 83 percent for overconfident individuals). It is remarkable that while underconfident individuals believe to have lower financial literacy scores than overconfident respondents, in reality their scores are higher than those of individuals with high confidence levels.

In our endline survey, we document that moneyboxes have been a useful and frequently employed savings technology. Only 3 percent of our sample did not possess the box at the endline survey, mostly because they were given away as presents or were damaged. Nearly all individuals used the moneybox for saving purposes, and only 4 percent of the respondents had no money in the box. Every sixth person had taken out money at least once. On average individuals reported to have taken 462 *Birr* out of their savings device. Interestingly, underconfident individuals, although their probability of having taken out money is comparable to overconfident individuals, reported taking out lower amounts (only 184 *Birr*). Given the small amount of time respondents had access to this new savings device, they used it considerably. The average amount in the moneybox was 239 *Birr*, with underconfident respondents having lower amounts (just like with overall savings) than overconfident individuals.

Note, while we did not randomize with respect to the individuals confidence levels, in Table 1 we observed that the share of those who received a recommendation and those who did not is comparable across both groups. 15 percent of those who received an encouragement were underconfident, while for the group of those who did not receive one the share is 18 percent. The difference of 2.8 percent is statistically insignificant. Similarly, 46 percent of those who received a recommendation were overconfident. Among those who did not receive a recommendation, 51 percent were classified as overconfident. The difference between these two groups is statistically insignificant as well.

4. Results

In this section we discuss our empirical approach, present the main results of the paper, and provide a detailed discussion. The first subsection documents the positive effects of the moneyboxes on savings attainment and explores how the recommendations increase savings and how this increase is related to the key behavioral characteristic confidence. The

Table 2
Summary Statistics.

	(I) All		(II) Underconfident		(III) Overconfident		(IV) Correctly Confident		(V) Diff. II vs. III+IV
	(1) Mean	(2) SD	(3) Mean	(4) SD	(5) Mean	(6) SD	(7) Mean	(8) SD	(9) p-value
I. Baseline Control Variables:									
Cash savings	4429.93	12407.02	2816.49	3605.52	5632.93	15975.10	3579.55	9040.21	0.232
Bank savings	2289.85	9414.75	1149.94	1707.37	3174.63	12581.11	1644.63	5888.07	0.409
Iddir member	0.44	0.50	0.56	0.50	0.39	0.49	0.45	0.50	0.025
Equb member	0.30	0.46	0.29	0.45	0.32	0.47	0.27	0.45	1.000
Household members aged 0–5	0.85	0.78	0.83	0.84	0.82	0.72	0.90	0.81	0.821
Household members aged 6–11	1.08	0.88	1.11	0.92	1.12	0.85	1.01	0.89	0.560
Household members aged 11–17	1.12	0.93	1.28	0.88	1.16	0.96	1.00	0.88	0.057
Household members aged 18–64	2.55	1.04	2.67	1.19	2.58	1.03	2.45	0.99	0.160
Household members aged 65+	0.05	0.22	0.10	0.30	0.04	0.19	0.04	0.22	0.135
Female household head	0.22	0.41	0.23	0.43	0.20	0.40	0.22	0.42	0.789
Married household head	0.74	0.44	0.73	0.44	0.77	0.42	0.70	0.46	0.875
Education household head	3.28	3.75	2.35	3.37	3.57	3.63	3.30	3.99	0.001
Land size	4.58	5.49	5.45	8.31	4.37	4.88	4.47	4.67	0.113
Revenue	3951.39	12978.67	2938.40	4113.87	5186.33	18163.36	2806.04	4898.77	0.602
Debt	2974.74	4461.47	1923.40	3110.10	3479.88	4863.12	2775.97	4333.01	0.064
Consumption expenditures	736.63	980.21	649.28	704.98	762.39	1138.63	740.79	855.71	0.208
II. Personality:									
Underconfident	0.16	0.36	1.00	0.00	0.00	0.00	0.00	0.00	
Overconfident	0.47	0.50	0.00	0.00	1.00	0.00	0.00	0.00	0.000
Expected financial literacy score	1.99	0.90	1.32	0.59	2.46	0.70	1.68	0.93	0.000
Financial literacy score	1.57	0.92	2.47	0.62	1.17	0.75	1.68	0.93	0.000
- 0 answered correctly	0.15	0.35	0.00	0.00	0.21	0.41	0.13	0.34	0.000
- 1 answered correctly	0.29	0.46	0.06	0.25	0.41	0.49	0.24	0.43	0.000
- 2 answered correctly	0.41	0.49	0.40	0.49	0.38	0.49	0.44	0.50	0.607
- 3 answered correctly	0.15	0.36	0.53	0.50	0.00	0.00	0.19	0.39	0.000
Present-biased	0.28	0.45	0.32	0.47	0.29	0.45	0.26	0.44	0.825
Risk-loving	0.49	0.50	0.55	0.50	0.46	0.50	0.50	0.50	0.143
III. Additional Treatment Information:									
Positive recommendation	0.41	0.49	0.39	0.49	0.39	0.49	0.44	0.50	0.993
Negative recommendation	0.40	0.49	0.38	0.49	0.40	0.49	0.40	0.49	0.941
Measured at Baseline:									
Initial goal amount (*Birr*)	1975.75	5234.63	1593.62	3778.63	1799.22	4520.26	2366.09	6476.24	0.375
Saving duration (in weeks)	22.96	2.77	22.71	3.21	23.22	2.43	22.72	2.95	0.435
Revised goal amount (*Birr*)	1988.40	5251.24	1650.04	3893.76	1776.68	4466.38	2403.43	6524.62	0.497
Investment goal	0.73	0.45	0.71	0.45	0.71	0.46	0.76	0.43	0.851
No expected problem in reaching goal	0.45	0.50	0.34	0.48	0.46	0.50	0.47	0.50	0.022
Subjective probability to reach goal	82.89	20.59	80.85	21.49	82.17	20.45	84.67	20.33	0.169
Measured at Endline:									
Does not possess box anymore	0.03	0.17	0.02	0.15	0.04	0.19	0.02	0.13	0.392
Total amount in box	239.01	468.03	175.17	240.91	270.59	567.78	226.54	395.49	0.146
No money in box	0.04	0.19	0.03	0.18	0.03	0.17	0.05	0.23	0.770
Opened the box to take money out	0.16	0.37	0.15	0.36	0.13	0.34	0.20	0.40	0.745
Amount taken out last time	462.31	1810.20	138.44	184.41	365.52	1159.98	651.02	2448.84	0.440
N	599		94		283		222		94 vs. 505

Note: Sample of participants with moneyboxes (N=599). Column (1) reports means for all individuals in the sample, followed by the standard deviations (SD) in column (2). The remaining columns follow the same pattern for underconfident, overconfident, and correctly confident individuals respectively. The last column (9) displays the p-value for the null hypothesis that there's no difference between the means in column (3) vs. columns (5) and (7), i.e. between underconfident individuals and others. All monetary variables are given in local currency *Birr*, exchange rate at time of treatment (Dec 2014) was 1 USD ≈ 20 *Birr*.

second subsection provides an in-depth discussion of the results by analyzing the direction of the recommendations, proposing two key potential mechanisms and ruling out a number of alternative explanations.

To estimate the effects of the treatments on our outcomes of interest, we use standard methods from the analysis of randomized control trials and estimate OLS regressions with treatment indicators and baseline controls. Our preferred regression specification is

$$Y_i = \alpha + \mathbf{T}_i \beta + \mathbf{X}_{i0} \gamma + \delta_1 Y_{i0} + \delta_2 lit_i + \mathbf{I}_i \zeta + \epsilon_{it}, \tag{1}$$

with Y_i being the outcome of interest of household i at time t and $\mathbf{T}_i$ the vector of treatment dummies. $\mathbf{X}_{i0}$ represents the vector of baseline control variables: savings in cash and at the local MFI, membership in the informal savings societies *Iddir* and *Equb*, land area, revenue from crop sales, per-capita consumption expenditures, total outstanding debt, number of

Table 3

Effect of Moneybox and Recommendations on Savings.

	(1) Savings	(2) Savings	(3) Savings	(4) Savings
Moneybox	111.759*	117.643*		
	(63.694)	(61.370)		
Recommendation			162.573*	181.463**
			(84.055)	(80.733)
N	881	881	599	599
Baseline Control Variables[1]		X		X
Mean	527.596	527.596	496.009	496.009

Note: This table reports results from OLS regressions with savings as the outcome variable. *Savings* is measured as money held in the moneybox (0 for respondents without moneybox) plus cash holdings. Column (1) provides results for individuals who received a moneybox to those who did not receive a moneybox, only controlling for enumerator fixed effects. Column (2) adds all baseline covariates. [1]Control variables include all baseline variables used during randomization (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator indicators and our measure of financial literacy (see Table 1, Panel I, for a full list of variables). Column (3) reports results from an OLS regression comparing individuals who received a recommendation to those who did not receive a recommendation, only controlling for enumerator fixed effects. Column (4) includes baseline covariates. The sample is smaller since only those individuals receiving a moneybox were assigned to one of the recommendation treatments. All variables are winsorized at the 95% percentile. In columns (1) and (2) the last row gives the mean value of the outcome variable in the control group (no moneybox). In column (3) and (4) the last row gives the mean value of the outcome variable in the control group (moneybox, but no recommendation). Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

household members by age, gender, marital status and education of the household head (these were also used for stratification, refer to Section 2.2 for details). We further control for the baseline values of the outcome variables Y_{i0} in order to improve precision. We explicitly control for lit_i, an indicator for above average financial literacy score, in order to rule out any possible mechanical effects of financial literacy on confidence. Although the enumerators followed a detailed protocol and received intensive training, we include a vector of enumerator dummies $\mathbf{I}_i$ at treatment to control for any possible enumerator-specific effects in administrating the treatments. The remaining error term is denoted by ϵ_{it}. Throughout the rest of this section, we keep this estimation framework and vary the definition of the treatment vector $\mathbf{T}_i$. We also introduce several interactions along behavioral dimensions, especially comparing individuals by confidence levels.

We would like to point out that while the variability in the baseline values of our key variables is significant, our specific measurement approach during the endline data collection leads to a strong decrease in the variance and thereby an increase in the precision of our estimates. Specifically, the enumerators opened the moneybox and counted the money inside and asked respondents to detail their additional cash savings by naming the exact composition in bills and coins, leading to a de-facto counting of bills and coins in most cases.[24]

Finally, the study is registered at the AEA RCT Registry under # *AEARCTR-0000613* with the title "Overconfidence, underconfidence and the use of persuasive messages in the attainment of savings goals". We have additionally uploaded a pre-analysis plan in which we outline in detail all estimations and define the variables used. We have closely followed the plan, with two noticeable changes: First, we have included financial literacy as a control variable to our main specification. None of the main results changed, yet given the construction of our confidence measures, we decided that it represents a more robust approach. Results are available upon request. Second, part of the pre-analysis plan (with an additional treatment arm) is preserved for a separate investigation. The AEA RCT registration and the upload of the pre-analysis plan took place prior to the endline data collection (January 26, 2015). Our study obtained IRB approval from the University of Mannheim.

4.1. Main results

Moneyboxes

We begin our analysis by estimating the effect of our soft commitment device on savings attainment. Our results show that the moneyboxes coupled with savings plans with regular savings installments are a very effective method to increase savings among smallholder farmers in Northern Ethiopia. Table 3, Panel A, compares those (randomly selected) individuals who received a moneybox to those in the pure control group without a moneybox, and documents substantial increases in savings. Cash savings plus savings in the moneybox increased by about 117 *Birr* (22 percent increase compared to the control group). These results stem from simple OLS regressions following our standard estimation approach set out in Eq. (1). They are in line with a large literature documenting the effectiveness of moneyboxes in increasing savings (among others Dupas and Robinson, 2013). The moneyboxes thus helped to initiate larger savings and further permit us to measure cash savings with little measurement error (through direct counting of the money in the box).

---

[24] In essence, our approach is similar to Dupas et al. (2018) who ask for "cash at home or in a secret place" to reduce reporting bias. We additionally ask respondents for the composition of bills, making misreporting even more difficult.

Table 4

Effect of Recommendations on Savings of Underconfident Participants.

	Savings
Recommendation	106.141
	(92.405)
Recom x Underconfident	409.802**
	(162.633)
Underconfident	-440.383***
	(138.329)
Financial literacy	23.315
	(83.985)
N	599

Note: N=599. The dependent variable savings is calculated as the value of savings in the moneybox plus cash holdings. In this table we introduce the effect of underconfidence, measured by an indicator equal to one for individuals believing to answer fewer questions on financial literacy correctly than they actually do. We interact underconfidence with the recommendations. All variables are winsorized at the 95% percentile. Additional control variables include all baseline variables used for stratification during randomization (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator indicators. Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

One feature of our findings is the fact that by providing the moneybox and explicitly telling participants that we would come back we might be inducing an experimental demand effect and nudge participants into saving more. The literature has extensively analyzed these demand effects and while there have been recent advancements in gauging its size (de Quidt et al., 2018), it remains difficult to control for these effects and we cannot rule out such a demand effect. This does not pose a threat to the internal validity of our study, but should be kept in mind when reaching policy conclusions. However, we made it very clear in our survey instructions that the enumerators were not affiliated with the microfinance institution in any way. Therefore, the participants should not have the feeling that accumulating higher savings would increase their standing with the local MFI.

Individualized Recommendations and Confidence

We continue by analyzing the effect of the individualized recommendations on the actual savings accumulation of the smallholder farmers. For this second aspect of our randomized experiment, we focus on the subsample of individuals receiving the moneybox.[25] Among these households, we compare those who received a recommendation to reconsider the originally stated savings goal to those who did not receive such a recommendation.[26] Table 3 reports the main results. The outcome variable consists of savings in the moneybox and further cash savings. The estimation strategy follows the exact setting laid out in Eq. (1). Columns (3) and (4) of Table 3 show that receiving a recommendation leads to a significant increase in savings of 181.5 *Birr* (specification with controls), statistically significant at the 5 percent level. This corresponds to a sizable increase of 36 percent compared to the control group (receiving no recommendation).[27] Thus, a simple message motivating smallholder farmers to reconsider their original savings goal leads these individuals to save more. We would like to stress that this happens irrespective of the direction the recommendations take, i.e irrespective of whether we encourage the individual to save more or less.

In Section 3.3 we have shown that in our context low degrees of confidence are an independent predictor for low baseline savings levels. We now analyze whether underconfident individuals, who save less at baseline, may be differentially affected by the recommendations. Table 4 introduces the indicator variable underconfidence and its interaction with the

---

[25] Our analysis in this section draws on additional experimental variation in the individualized recommendations and their interactions with behavioral characteristics. While the design of our intervention does not permit drawing on the same experimental variation within the control group not receiving moneyboxes, there could also be an interaction between the behavioral characteristics and the overall moneybox effect. In unreported results, we interacted the moneybox effect with all of the behavioral characteristics proposed and find no significant differences in the effects.

[26] As individuals who received a recommendation we combine all those with the treatments I–IV presented in Section 2.2.

[27] As discussed in Section 3, our endline measurement approach relies on counting the money in the box, asking for the composition of bills of other cash savings, and copying entries from the MFI's savings booklet. This led to a large and significant decrease in standard errors from our original baseline to the endline measure – a reduction by more than half. Nevertheless, to address any further variability in savings measures, in Table A.3 in the Appendix we look at log savings and estimate quantile regressions for both our main results accruing to moneyboxes and recommendations. The coefficients for the moneyboxes from standard OLS and quantile regressions of the first 4 quintiles are positive and highly significant. Likewise, the effect of the recommendation treatments is positive, though not significant in standard OLS, but positive and significant in all but the lowest quintile of the outcome distribution.

recommendation treatment. Just as at baseline, underconfidence leads individuals to save significantly less (on average 440 *Birr* less). We find that the recommendation treatment is especially helpful in increasing savings for these underconfident individuals. The estimate of the interaction term between both corresponds to an increase of 409 *Birr*.

4.2. Discussion

The previous section has demonstrated that recommendations are particularly helpful for underconfident individuals. This section attempts to provide more depth to these results. We start with a detailed analysis of whether our nudging to change the savings amount up- or downwards worked as expected, i.e. we look at intermediate outcomes of our experiment. In continuation, we explore potential underlying mechanisms behind the observed results and analyze additional, pre-treatment survey evidence on differential patterns in administering the savings process and setting savings goals. While there could be a multitude of behavioral and cognitive factors at play that explain why underconfident individuals may be differentially affected by the recommendations, we provide suggestive evidence for two potential mechanisms: Attention and re-assessment.

Direction of Recommendation

Effectively, the recommendations provide instantaneous and individually-tailored feedback to the respondents in the form of a nudge to save more or less than originally planned. Here, we first explore the effect of the recommendations on the savings goal formation process. In particular, we look at a set of intermediate outcomes measuring how the initial saving goals changed with the individualized recommendations. The following analysis allows us to test whether the messages were well delivered by the enumerators, understood by the farmers, and consequently actually triggered a change in intended saving goals in the recommended direction. We then analyze whether the direction of the recommendation has an effect on the observed results.

Table 5 presents results for intermediate outcomes which capture the savings *goal-setting* instead of actual savings *attainment* behavior. We construct a number of outcome measures of the extent to which individuals revised their self-set

Table 5
Intermediate Results.

	(1) Revised - Initial Goal	(2) Revised - Initial Goal (% of Initial)	(3) Share of HH Changed Goal	(4) Share of HH with Recommended Change
Panel A: Direction of Recommendation				
Positive recommendation	218.969***	0.102***	0.286***	0.346***
	(53.338)	(0.034)	(0.038)	(0.033)
Negative recommendation	-220.270**	-0.037	0.180***	0.142***
	(100.622)	(0.032)	(0.038)	(0.028)
N	597	597	597	597
Panel B: Intensity of Recommendation				
Positive recommendation	248.721***	0.105***	0.257***	0.318***
+ 40%	(81.470)	(0.029)	(0.049)	(0.045)
Positive recommendation	195.715***	0.100**	0.312***	0.372***
+ 20%	(67.020)	(0.051)	(0.047)	(0.042)
Negative recommendation	-116.177	-0.006	0.189***	0.137***
– 20%	(129.461)	(0.041)	(0.049)	(0.036)
Negative recommendation	-324.837**	-0.067**	0.169***	0.145***
– 40%	(144.722)	(0.034)	(0.044)	(0.034)
N	597	597	597	597
Panel C: Underconfident Individuals Only				
Positive recommendation	110.735*	0.067**	0.323***	0.323***
	(63.462)	(0.028)	(0.100)	(0.100)
Negative recommendation	-71.466	-0.034	0.135	0.135
	(75.656)	(0.028)	(0.112)	(0.112)
N	94	94	94	94

Note: This table shows the effect of the recommendations on intermediate outcomes regarding the formation of the savings goal (not actual savings attainment). The outcome variable in column (1) is the difference between the revised and initial savings goal. In column (2) the outcome is this difference as a fraction of the initial goal amount, more specifically: (revised - initial goal)/initial goal. The dependent variable in column (3) is an indicator for changing the goal at all and in column (4) we take an indicator for changing the goal in the recommended direction. Panel A depicts results for recommendations grouped by their direction, Panel B disaggregates the recommendations by their intensity and Panel C focuses on the subsample of underconfident indivinduals. Two observations are missing: one without data on the initial goal level, another without data on any goal level. In 24 cases we had missing data on the *revised goal* and assumed the *revised goal* to be equal to the *initial goal*. Further control variables include all baseline variables used for stratification in the randomization process (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator controls. Robust standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

goals.[28] In column (1), we look at the difference between the revised and initial savings goal. For example, an individual planned to save 1000 *Birr* for a cow, but we recommended saving 1400. The respondent could then change the goal in any way and the revised amount was noted down on the moneybox. Column (2) reports this change as a percentage of the original amount.[29] The outcome in column (3) is an indicator whether an individual has changed the goal at all and column (4) is an indicator whether an individual has changed the goal in the recommended direction.[30]

For Panel A of Table 5, we group the four different types of recommendations (please refer to Section 2.2 for details) into "positive" and "negative" ones, i.e. messages that encourage to save more or less, respectively. As we see from the first column of Table 5, the direction of the recommendation clearly has an effect on how the goals are revised. While positive recommendations increase savings goals, negative recommendations induce respondents to decrease their goals. These differences are statistically significant and the magnitudes of positive and negative recommendations are almost identical and correspond to roughly an 11 percent change in the goal amount on average. Columns (2) to (4) support these findings. Both individuals receiving positive and negative recommendations change their goals, but it seems that the probability of revising the goal is slightly stronger for recommendations to save more. Panel B analyzes the intensity of the recommended change, whereby we distinguish between recommended revisions of plus/minus 40 percent and plus/minus 20 percent. In general, stronger recommendations are accompanied by higher changes in goal amounts and positive recommendations seem to be more effective at influencing the process of savings goal formulation.[31] Overall, these results on intermediate outcomes relate to a large literature on goal revision behavior and the updating of beliefs. In line with the literature, we also find evidence for the conservative bias, which refers to the finding that when receiving new information, economic actors update their beliefs, but to a lesser degree than suggested by the signal (Coutts, 2019; Ertac, 2011; Mobius et al., 2011). In our setting, this corresponds to individuals changing their goal in the intended direction, but not as much as proposed by the recommendation.[32]

Given the focus of our analysis on underconfident individuals, Panel C focuses on the subsample of underconfident individuals. At the stage of formulating their savings goals, underconfident individuals also respond to the recommendations in the suggested directions. Their increase in saving goals following a positive recommendation is comparable to that of the whole sample in relative terms, as the coefficient of 111 *Birr* corresponds to a 7 percent increase in their goal amount.[33] However, underconfident individuals respond significantly more strongly to positive and encouraging recommendations as compared to negative recommendations. While the effect of positive recommendations is significant in all specifications, the effect of the negative recommendations is insignificant throughout. This is in line with our expectations and descriptive evidence that farmers with low degrees of confidence set low goals and might be more likely to revise these upwards when prompted.

Going back to our main estimation framework, we additionally distinguish explicitly between positive and negative recommendations. Table 6 provides similar results as before: The effect of the recommendations on savings behavior is driven by underconfident individuals. In line with the previous findings, both positive as well as negative recommendations increase savings for this group. Moreover, the coefficients of the interactions between underconfidence and both positive and negative recommendations are almost identical. We take this as evidence that the direction of the recommendation is in our experiment irrelevant for the actual savings behavior.

Mechanisms

Why are the recommendations so successful in increasing savings and why do underconfident individuals react so strongly to them? Here we provide suggestive evidence on potential mechanisms and connect our measure of confidence to established survey measures in the goal-setting literature.

Specifically, we propose two potential mechanisms driving the results: Attention and re-assessment. Attention refers to the positive motivational effects of being taken seriously and receiving recognition from peers or others. In our context this

---

[28] Note, we lose two observations due to missing information on intermediate outcomes in at least one of the four regressions.

[29] The variable is calculated as: (revised amount - initial amount)/ initial amount.

[30] In additional descriptive analyses in the Online Appendix, we break up the original goal amount by positive vs. negative recommendations and find no statistical differences. The revised *Birr* amounts after the negative and positive recommendation is 1617.76 and 2284.27, respectively. This difference is statistically significant at the 10 percent significance level (p-value 0.092), which is in line with the finding that the recommendations were successful in changing the goal amounts. When we further compare other aspects of the treatments, we observe that in the case of the positive recommendation participants were more likely to open the box and take money out (13.0 vs. 18.9 percent, p-value for difference is 0.074). We also see that in the case of a positive revision, respondents are more likely to expect problems in reaching their goals (53.8 vs. 43.4 percent, p-value for difference is 0.010). Individuals that were discouraged from setting an originally higher goal were less likely to not possess the moneybox at endline (1.7 vs. 5.1 percent p-value for difference is 0.074).

[31] For example, being recommended to save 40 percent more leads to an upward revision of the initial savings goal by 249 *Birr* compared to individuals who did not receive any recommendation. On the other hand, the recommendation to increase the savings goal by only 20 percent results in a slightly lower change of only 196 *Birr*. We see a similar pattern for the intensity of the negative recommendations.

[32] In this same literature on Bayesian updating, important contributions have been made to distinguish between "positive" and "negative" news, with a frequent (albeit contested) result being that individuals respond more strongly to positive signals than to negative ones (Mobius et al., 2011). Ertac (2011) even disaggregate this finding by degrees of confidence. Even though in our analysis we also find stronger responses to positive recommendations than to negative ones, it is difficult to establish a direct link to this literature, as in our setting there is no "good" or "bad" news as is typical in the signals on individual performance.

[33] The relatively low absolute value of the coefficient can be explained by the fact that underconfident individuals have lower initial goal amounts.

Table 6
Direction of Recommendations and Underconfidence.

	(1) Savings
Positive Recommendation	98.610
	(102.346)
Negative Recommendation	114.151
	(102.251)
Positive Recommendation x Underconfident	419.866**
	(187.637)
Negative Recommendation x Underconfident	400.443*
	(205.225)
Underconfident	-440.564***
	(138.597)
Financial literacy	23.459
	(84.606)
N	599
Mean	496.009

Note: The dependent variable savings is calculated as the value of savings in the moneybox plus cash holdings. This table disaggregates the effect of recommendations on savings by their direction (positive or negative) and interacts these with our measure of underconfidence. Further control variables include all baseline variables used for stratification during randomization (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator controls. The last row gives the mean value of the outcome variable in the control group (no recommendation). Robust standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

refers to the individual, non-generic component of the recommendation: The enumerator taking a respondent's individual savings goal and providing an individually tailored recommendation in the form of a proposed percentage change. The second mechanism of re-assessment refers to the respondent's cognitive process behind re-assessing and revising the savings goal. After individuals set a savings goal, they are provided with one of the randomized recommendations, and then asked if they would like to change their goal and what they would like to change it towards. This process of reassessment could lead to improved focusing of efforts on a given goal, away from a set of possibly conflicting objectives. Moreover, the reassessment could be beneficial to individuals that are indecisive about financial decisions. They might also attach more importance to the savings process after additional reflection. As highlighted above, we cannot perfectly pin down these mechanisms and cleanly disentangling them from each other is beyond the scope of this paper. However, in this section we hypothesize that these underlying behavioral forces are particularly relevant for underconfident individuals and that this offers one potential explanation for our results.

In the baseline survey, we elicited a battery of questions related to respondents' perceptions of their savings goals and savings process. Our survey instruments are an adaption of the established goal setting scale used in the psychology literature. Originally designed to measure goal setting behavior at the workplace, the seminal studies by Latham and Locke (1979) and Lee et al. (1991) provide a framework for measuring attributes of goal setting and perceptions about goals. The items capture various components of goal setting behavior like "performance feedback", "supervisor support", and "goal stress". We selected a subset of their questions and adapted them to the context of financial decision-making in rural Ethiopia.[34] In particular, we administered two sets of survey instruments designed to measure stress associated with saving goals and the support participants receive from the local savings officer at their MFI. Table A.7 in the Appendix provides details on all questions and displays T-tests comparing the responses of individuals classified as underconfident to those classified as overconfident or correctly confident. We focus on items with particularly high baseline difference between underconfident individuals and others, which to some extent also capture our hypothesized mechanisms of attention and re-assessment.

In Table 7 we replicate our main results from Tables 4 and 6. In Panels A we report results substituting our measure of underconfidence with information on the interactions with savings officers and the perceived support the respondents receive when interacting with their financial institutions. Panel A.1 uses an indicator for whether the savings officer listens openly to a respondent's problems and Panel A.2 an indicator for whether the respondent reports coming to an agreement with the savings offer on steps to be taken by each of them to solve any savings performance problem. Both of these indicators serve as good substitutes for our measure of confidence and provide estimates similar to our main results: Individuals that feel taken less seriously by their savings officer are more likely to benefit from the recommendations. In Panel B we use an indicator that the individual had described herself as someone who has "too many savings goals" at baseline. Similar

---

[34] A number of the questions and concepts we omitted are very specific to working in white-collar jobs in advanced economies and are therefore unsuited to financial decision-making in general, and particularly in developing countries.

Table 7
Real-life Measures.

	(1) Savings	(2) Savings
Panel A: Suggestive Evidence for the Attention Channel		
A1: Savings Goal Setting and Performance Appraisal (I)		
Savings offers listens openly to problems	382.470** (158.211)	378.941** (158.530)
Recommendation	409.041*** (134.423)	
Positive recommendation		340.758** (155.122)
Negative recommendation		468.684*** (151.308)
Recom. X Savings offers listens openly to problems	-317.487* (170.457)	
Positive recom. X Savings offers listens openly to problems		-227.859 (195.204)
Negative recom. X Savings offers listens openly to problems		-401.573** (188.789)
N	578	578
A2: Savings Goal Setting and Performance Appraisal (II)		
Savings comes to agreement with me on steps to be taken	391.486** (156.175)	388.320** (156.609)
Recommendation	391.385*** (132.169)	
Positive recommendation		327.311** (152.738)
Negative recommendation		437.458*** (147.909)
Recom. X Savings comes to agreement with me on steps to be taken	-300.555* (170.345)	
Positive recom. X Savings comes to agreement with me on steps to be taken		-222.753 (194.464)
Negative recom. X Savings comes to agreement with me on steps to be taken		-361.661* (188.077)
N	578	578
Panel B: Suggestive Evidence for the Re-Assessment Channel		
B: Savings Goal Stress		
Too many savings goals	-219.416 (173.587)	-205.777 (174.122)
Recommendation	-54.576 (120.896)	
Positive recommendation		24.594 (142.924)
Negative recommendation		-122.578 (125.692)
Recom. X Too many savings goals	373.955** (156.734)	
Positive recom. X Too many savings goals		257.049 (180.824)
Negative recom. X Too many savings goals		481.935*** (171.904)
N	585	585

Note: The dependent variable savings is calculated as the value of savings in the moneybox plus cash holdings. Evidence on self-reported behavior with savings officers is presented in Panel A and self-assessed (general) savings behavior in Panel B. Column (1) disaggregates the effect of recommendations on savings by their direction (positive or negative) and interacts these with three types of additional measures. This analysis is along the lines to the one presented in Table 4. Column (2) disaggregates the effect of recommendations on savings by their direction (positive or negative) and interacts these with our real-life measure. This analysis is along the lines to the one presented in Table 6. Further control variables include all baseline variables used for stratification during randomization (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator controls. Table A.7 in the Appendix provides the means for the three types of measures. Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

to our results analyzing the interaction between underconfidence and recommendations, we find that the recommendations are particularly helpful for individuals perceiving to have too many saving goals. Moreover, when differentiating by positive and negative recommendations in column (2), the interaction coefficients with both directions are positive, although significant only when interacted with the negative recommendation. We interpret the evidence from these "real-life" measures as suggestive evidence for the validity of our proposed mechanisms. The results related to interactions with the savings officer and the perceived support speak towards the mechanism of attention. Underconfident individuals are more likely to feel

that they are not being taken seriously and this may drive some of the differential results we observed above. We associate the second mechanism of re-assessment to the real-life measure related to goal stress and the quantity of goals. Individuals who are underconfident in their financial decisions are more likely to have too many goals and this indecisiveness may be one of the reasons why specific recommendation may help these people increase their savings.

To summarize, this section has provided suggestive evidence for two possible mechanisms: Attention and re-assessment. We show that real-life measures associated to these mechanisms can lead to the same results as our principal measure of confidence. The next section addresses a number of alternative plausible mechanisms that could also be driving our main results. In particular, we look at further baseline differences in savings behavior, alternative behavioral traits, and financial literacy.

5. Robustness

This section presents a range of robustness checks. A number of behavioral factors are possibly related to confidence levels and could partly explain some of the observed results. Here we address alternative mechanisms and also discuss further robustness checks for our main results.

Alternative Mechanisms

We conduct a thorough analysis, replacing our underconfidence measure with alternative, related measures. The results are presented in Table A.2 in the Appendix. To start with, an open question is whether the differential recommendation results for underconfident individuals are due to different risk and time preferences. These personality traits have been shown to be highly correlated with financial decisions, including savings behavior. In our sample we have already shown that underconfident individuals do not display significantly higher or lower levels in these traits (see Table 2). These additional results, in which we interact recommendations with these traits, show that in our setting present-biasedness and risk-lovingness cannot explain differences in the reaction to our recommendations.

We also take a closer look at financial literacy. Throughout our regressions we always control for financial literacy to eliminate any mechanical effects on confidence levels as well as any direct effects on savings behavior. A remaining issue might be that the ability to properly apply the recommendations also differs by financial literacy. To test this hypothesis, we interact an indicator for having above average financial literacy with the recommendation treatment. We find no statistically significant effect and conclude that financial literacy seems not to be an independent driver of our results. Additionally, there could be an independent effect of individuals with exactly zero correct answers. By construction of our measure of confidence, these individuals cannot be classified as underconfident. To rule out any possible direct effects of these individuals, we also replicate our interaction with an indicator for having exactly zero correct answers and find no significant effects.

Respondents in our sample might differ in their ability or willingness to anticipate obstacles hindering them from achieving their savings goal. This may in turn have implications for the effectiveness of our recommendation treatments and might be an alternative mechanism to general confidence levels. In our detailed survey data, we ask respondents to explicitly state any problems they anticipate in reaching their savings goal. About 55 percent of the sample stated one or more potential problems, the remaining 45 percent did not see any problem hindering them from reaching their savings goal. We include an indicator for expecting at least one problem and interact this with our recommendation treatment. The recommendations did not work differentially for those individuals. We conclude that the ability or willingness to anticipate obstacles does not drive the observed results.

In the descriptive analysis we have also observe that the debt levels of underconfident people are 36 percent lower (2172.20 vs. 3395.92 *Birr*, p-value for difference is 0.06) than of others. At the same time, while underconfident individuals take out money at the same rate as overconfident people, the quantities taken out are lower relative to overconfident people and substantially lower than correctly confident people.[35] Thus, a potential alternative explanation for the results could be that overconfident individuals used the money in the moneybox to service their higher debt levels. To address this alternative mechanism, we also mimic our main results with an interaction of whether individuals ever opened the box to take money out instead of our measure for underconfidence and find no significant effect.

An additional alternative mechanism associated with the process of savings may be the savings duration, which respondents could freely choose within certain limits. Even though the savings duration is very similar between individuals with varying degrees of confidence (refer to Table 2), the savings duration might be an additional proxy for the importance farmers attach to the savings process. However, this is not the case.

Robustness of Main Results

Finally, we address further concerns with regards to the robustness our main findings. The results are presented in an Online Appendix, if not otherwise indicated. Our results are robust to (1) the inclusion of the present-biasedness variable,

---

[35] These differences are not statistically significant. The share of underconfident people to have opened the box to take money out is 0.15, while it is 0.16 for the others (p-value of the difference is 0.75).

(2) savings duration, (3) omitting baseline control variables, (4) the inclusion of an indicator for having taken out money,[36] and (5) using categorical variables of our confidence score.

Furthermore, our central results show that the moneyboxes and the recommendations increase savings, measured as the sum of savings in the moneybox and cash savings. A key concern in this measure of savings is crowding out of other forms of savings. An increase in one form of savings could lead to a reduction in other saving vehicles.[37] To address this issue, we draw on a large range of alternative savings measures in Table A.4 in the Appendix. We find no evidence of crowding out behavior. Both the moneybox treatment and our recommendation treatments do not increase bank savings (measured at the local MFI), holdings in the informal savings arrangements *Equb* (savings group), or *Iddir* (funeral society). We also look into two alternative measures of savings: an indicator for storing any crops and the total expenditure for building materials in the last four months. Our recommendation treatment has no effect on all of these alternative saving measures. We conclude that we find no evidence for any crowding-out into alternative savings vehicles.

One issue in the construction of our confidence measure is the fact that individuals which answered the financial literacy questions either all correctly or all incorrectly cannot be classified as either over- or underconfident. This could potentially introduce a bias into our results. To address this issue, we exclude all individuals with zero or three correct questions in the financial literacy score. These additional sample restrictions reduce the sample by about one third. The sample loss is equally divided between respondents with 0 correct answers and 3 correct answers (each represent 15%). Tables A.5 and A.6 in the Appendix replicate our main findings and find very similar results. If anything, the overall effect increased in this more robust subsample.

6. Conclusion

In this paper, we measure the effectiveness of a soft commitment device in increasing savings among banked small-holder farmers in Ethiopia. Our moneyboxes coupled with regular saving plans lead to a substantial increase in savings attainment of about 22 percent. In a detailed analysis of the behavioral factors associated with savings outcomes, we observe a strong relationship between confidence and savings behavior. In particular, we find evidence that underconfident individuals have significantly lower savings levels, conditional on well-known behavioral factors like present-biasedness. Accordingly, we provide savers with individualized feedback on their saving goals, generating random variation to increase or decrease the originally intended savings amount. We find that the recommendations cause strong increases in savings, especially for underconfident individuals.

Why is this individual feedback in the form of recommendations so effective at increasing savings, particularly for underconfident individuals? To answer this question our analysis first focuses on the confidence measure per se. We show a range of correlations related to measures of confidence and real-life behavior. We continue by analyzing intermediate outcomes of the experiment, presenting evidence that it was well-administered and that the farmers changed their goals in the experimentally expected (upward or downward) direction. Interestingly, we find that the recommendations increase savings per se, independently of their direction. This motivates us to discuss two likely channels. First, the individualized recommendations address the personal needs and wishes of respondents in a way that generic advice cannot. This might make savers feel as being taken more seriously and this effect might be particularly strong for underconfident individuals. Second, the farmers are asked to reassess the amount they would like to save for, which might help them to focus on a given goal. Moreover, the additional reflection itself might increase the importance respondents attach to the savings process and their goals. In a detailed analysis of the farmers' baseline behavioral characteristics we present additional real-life evidence that underconfident individuals indeed feel stressed when setting their goals. Underconfident individuals are more likely to report having too many savings goals and not being taken seriously when trying to solve savings problems with their microfinance institution. We replicate our main results using these additional indicators of challenges in financial decision-making.

The findings in our paper have clear implications for both research and policy on savings behavior. To start with, our strong positive results on savings attainment after implementing a soft commitment device are in line with previous literature contributing to a rising consent that soft commitment devices are better than hard commitment devices (Kremer et al., 2019; Karlan and Linden, 2017). Moreover, our results stress the importance of interventions that influence behavior by setting goals and providing reminders (see e.g. Dupas et al., 2018). This is supported by evidence that focusing on behavioral processes such as self-control (Kaur et al., 2015), aspirations (Beaman et al., 2012), and even self-efficacy, social support, and optimism (Steinert et al., 2018) can increase savings in developing countries. In this paper, we provide related evidence, in particular that behavioral channels matter strongly for saving outcomes and that "one size fits all" financial literacy approaches are unlikely to motivate the poor to reach higher savings goals. Relying on a well-established literature on the impacts of confidence on financial decision-making (Kramer, 2016; Malmendier and Tate, 2005; Tang and Baker, 2016), we are among the first to systematically analyze this relationship in a developing context. In essence, we find evidence for a hypothesis suggested in the work by Campbell (2016). He argues that households lack the knowledge to manage their financial affairs effectively, and that these difficulties are exacerbated by behavioral constraints. His analysis identifies a specific

---

[36] We also ran our main estimations on a subsample with only those individuals who never opened the moneybox. The results are very similar to our main specification (refer to Online Appendix).

[37] Note, however, that (Brune et al., 2016) observe the opposite: an increase in one type of saving increases the demand for other forms of savings.

subset of individuals we should be particularly concerned about: "Financially illiterate people with low self-confidence are likely to have a strong demand for assistance with financial problems, and to be relatively responsive to defaults and other nudges." Although his setting is one of consumer finance in a developed country, our findings for smallholder farmers in Ethiopia are very similar.

We know that details matter and that the poor face many choices and biases (Duflo, 2017). Underconfident savers could make better financial choices if provided with more individualized feedback as well as increased attention, re-assessment, and encouragement – a cost-effective way to increase savings. Saving officers could be trained to improve their communication skills, in particular taking into account the different personality traits of their customers. Importantly, the assessment of these traits should not only be based on self-reporting but needs to be thoroughly designed considering also financial literacy (Kramer, 2016). In general, incorporating individual feedback or similar individualized components into interventions might help underconfident individuals also in other domains. This in turn might be an overall avenue to help increase the effectiveness of policy interventions.

Declaration of Competing Interest

None.

Appendix

A.1. Experimental details

Recommendations. As described in the main text, we provided participants with individualized recommendations effectively varying the intensity with which we encourage participants to increase their savings. The exact wording of the randomized recommendations read out by the enumerators was as follows:

   I. "Our experience shows that people are more likely to reach their savings goal if they have higher goal amounts. Do you want to increase the amount of your savings goal to... [*initial goal amount × 1.4*]?"
  II. "Our experience shows that people are more likely to reach their savings goal if they have higher goal amounts. Do you want to increase the amount of your savings goal to... [*initial goal amount × 1.2*]?"
 III. "Our experience shows that people are more likely to reach their savings goal if they have lower goal amounts. Do you want to decrease the amount of your savings goal to... [*initial goal amount × 0.8*]?"
  IV. "Our experience shows that people are more likely to reach their savings goal if they have lower goal amounts. Do you want to decrease the amount of your savings goal to... [*initial goal amount × 0.6*]?"

In choosing these specific values for our recommendations, we focused on thresholds that are easy to implement in practice. We did not follow a theoretical justification for this choice. They had to allow for a simple and quick calculation in the field, reflect varying intensities, and leaving sufficient number of observations to detect an effect with from the different treatment arms.

The exact phrasing of the recommendation treatments was agreed upon after two pilots and several rounds of discussions with local experts from the University of Mekelle. This turned out to be the best way to contextualize our research design within the local institutional setting. However, there may be some ethical concerns due to the fact that the enumerators referred to their experience. Within the local context, this phrasing guaranteed that the participants would fully understand the recommendations and that these would have the intended effect. Moreover, our intervention encouraged all participants to increase their savings, albeit to varying degrees. All participants received 50 Birr (baseline), 30 Birr during the implementation of the treatments (both the treatment and control group) and 30 Birr (endline) to compensate them for participating in the interviews. Therefore, receiving these recommendations did in no way provide participants with any risk or harm whatsoever. Lastly, the study was carried out in extremely remote and sparsely populated areas in rural Tigray (Northern Ethiopia), where participants had little to no interaction between each other and any so-called "contamination" of the experimental pool is not an issue. Most importantly, the research received full IRB clearance from the University of Mannheim Ethical Committee and was conducted in collaboration with the local University of Mekelle and the microfinance institution Dedebit Microfinance (DECSI).

A.2. More results

Fig. A.1 and Tables A.1–A.7
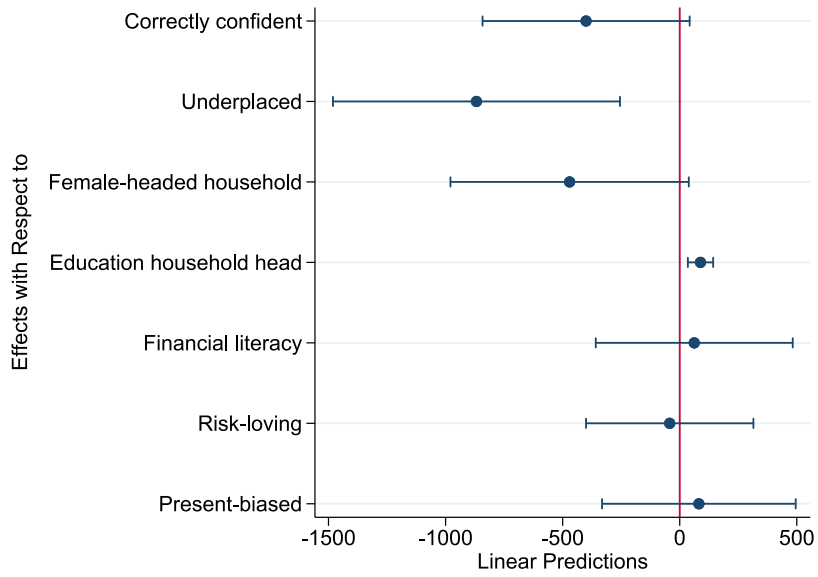
(a) Underplacement
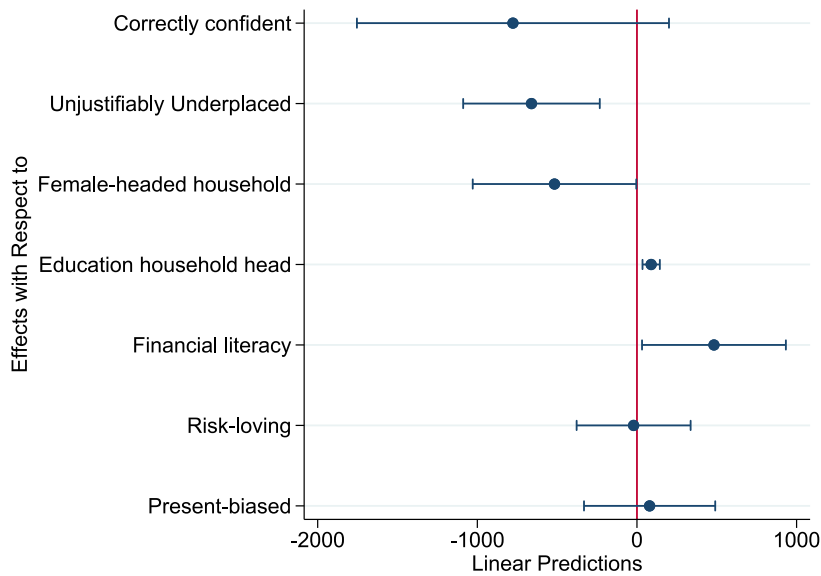
(b) Unjustified Underplacement

Fig. A.1. Savings Before Treatment – Alternative Confidence Measures *Note*: N=599. This figure depicts the coefficients from a simple OLS regression with a broad measure of savings at baseline (cash + bank + informal arrangements) as the dependent variable. The underlying regression is the same as in Fig. 2 but with differing measures of underconfidence. Panel (a) draws on the measure of "Underplacement" and Panel (b) on the measure of "Unjustified Underplacement". Additional control variables include all baseline variables used for stratification during randomization excluding savings outcomes (socio-demographics, land and consumption variables) as well as enumerator indicators. The savings measure is taken *before* treatment implementation. The variables depicted in the figure are measured as indicator variables with the exception of education, which is measured in years of schooling. 95% confidence intervals are depicted around the point estimates.

Table A.1
Alternative Confidence Measures.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Dep. Var: Underconfident													
Underplaced	0.118*** (0.0389)	0.208*** (0.0351)												
Unjustifiably underplaced			0.264*** (0.0269)	0.218*** (0.0272)										
Savings goals are stressfull					0.0710*** (0.0237)	0.105*** (0.0301)								
Too many savings goals							0.112*** (0.0226)	0.0703* (0.0363)						
Often fail to reach goals									-0.0485** (0.0245)	0.0174 (0.0279)				
Expect problems reaching goal											0.0669** (0.0292)	0.0702** (0.0335)		
Subjective probability <100													0.0452 (0.0297)	0.0611* (0.0325)
R^2	0.0144	0.2404	0.1280	0.2685	0.0097	0.2125	0.0216	0.2086	0.0042	0.2012	0.0084	0.2250	0.0039	0.2232
Controls		X		X		X		X		X		X		X
Treatment Controls												X		X
Observations	881	881	881	881	858	858	857	857	856	856	599	599	599	599

Notes: This table provides correlations for our primary measure of confidence with a number of alternative measures. The reported coefficients are from an OLS regression with an indicator for underconfident individuals as the dependent variable. In columns (1) and (2), the main explanatory variable is an indicator for a respondent believing to be less knowledgeable about the financial literacy questions than others in the same village (Underplacement). Columns (3) and (4) focus on an indicator for "unjustified underplacement". Specifically, the indicator is equal to one if the difference between own ability *minus* others' ability in the same village is *larger* than the difference between a respondent's perceptions about their own ability *minus* their perceptions about others' ability. Columns (5) and (6) look at baseline responses to the question "I find working towards my savings goals to be very stressful". Columns (7) and (8) analyze the baseline question "I have too many savings goals". Columns (9) and (10) analyze whether underconfidence is correlated with a perception of failing savings goals. Columns (11) and (12) report results for whether the respondents expected some type of problem in reaching their self-set saving goals. The last columns (13) and (14) use an indicator for an individual having a subjective, self-reported probability of reaching their savings goal of less than 100 percent. The controls in columns (2), (4), (6), (8), (10), (12) and (14) include all control variables used for stratification during randomization process (socio-demographics, land and consumption variables), enumerator indicators, and the measure of financial literacy. In columns (12) and (14), as these measures were reported after treatment implementation, we additionally control for treatment allocation (recommendation or not). The samples used in the estimations in columns (11)-(14) are only of those individuals which received a moneybox, since perceived probabilities in reaching goals are only comparable among those individuals. Robust standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2
Alternative Mechanisms.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dep. Var. throughout: Savings						
	Behavioral Measures		Financial Literacy		Savings Process		
Recommendation	168.343	245.207**	150.271	199.874**	209.927**	-177.805	211.147*
	(112.288)	(97.593)	(107.137)	(90.404)	(87.759)	(808.880)	(124.178)
Financial literacy	-6.832	-5.671	-52.161	-36.341	-2.289	-0.723	-12.491
	(76.997)	(76.619)	(146.091)	(84.665)	(76.074)	(77.223)	(76.610)
Risk-loving	1.683						
	(141.168)						
Recom × Risk-loving	27.819						
	(163.991)						
Present-biased		147.703					
		(157.771)					
Recom × Present-biased		-216.725					
		(176.381)					
Recom × Financial literacy			56.946				
			(158.793)				
0 correct financial literacy				-26.363			
				(124.143)			
Recom × 0 correct				-104.646			
				(130.992)			
Taken out money					79.077		
					(181.002)		
Recom × Taken out money					-181.449		
					(201.468)		
Savings duration						-21.315	
						(32.343)	
Recom × Savings duration						15.753	
						(34.515)	
Expect problems reaching goal							-40.156
							(141.636)
Recom × Expect problem							-39.020
							(160.883)
N	599	599	599	599	599	593	599
Control Group Mean	496.009	496.009	496.009	496.009	496.009	496.009	496.009

Note: This table tests a range of alternative behavioral mechanisms. Each column presents results from an OLS regression of an indicator for receiving a recommendation, an indicator for being financial literate, and an interaction with one of the alternative behavioral channels: being present-biased in column (1), being risk-loving in column (2), financial literacy (above average) in column (3), having zero correct answers on the financial literacy test in column (4), having taken money out of the box in column (5), savings duration in column (6) and expecting problems in reaching the goal in column (7). The dependent variable *Savings* is calculated as the value of savings in the moneybox plus cash holdings. Further control variables include all variables used for stratification at baseline plus enumerator controls. Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.3
Quantile Regressions: Effect of Moneybox and Recommendations on Log(Savings).

	(1) OLS	(2) Q0.20	(3) Q0.40	(4) Q0.60	(5) Q080
Panel A: Moneybox					
Moneybox	1.756***	4.133***	2.980***	0.516***	0.194
	(0.191)	(0.171)	(0.291)	(0.167)	(0.170)
N	881	881	881	881	881
Panel B: Recommendation					
Recommendation	0.258	0.134	0.334*	0.388**	0.330*
	(0.166)	(0.270)	(0.177)	(0.185)	(0.187)
N	599	599	599	599	599

Note: Panel A: N=881. Panel B: N=599. *Log(Savings)* measures money held in the moneybox (0 for respondents without moneybox) plus cash holdings. OLS regression are presented in column (1) (baseline specification), columns (2)-(5) report results from quantile regression, e.g. in column (2) we report the results for the 20th quantile, in column (3) for the 40th quantile, in column (4) for the 60th quantile, and in column (4) for the 80th quantile, respectively. Panel A of this table reports results from comparing individuals who received a moneybox to those who did not receive a moneybox. Panel B of this table reports results from comparing individuals who received a recommendation to those who did not receive a recommendation. All variables are winsorized at the 95% percentile. Control variables include all baseline variables used during randomization (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator indicators and our measure of financial literacy. Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.4

Alternative Savings Measures.

	(1) Bank Savings	(2)	(3) Equb	(4)	(5) Iddir	(6)	(7) Crop Storage	(8)	(9) Building Materials	(10)
Person received a moneybox	74.027		-36.367		-3.076		0.004		5.289	
	(114.187)		(149.200)		(19.261)		(0.030)		(4.656)	
Recommendation		50.238		-40.957		13.746		0.029		-0.793
		(157.002)		(221.486)		(26.793)		(0.044)		(7.288)
Financial literacy	105.735	61.484	61.365	7.299	-19.897	-19.288	0.037	0.035	7.346	6.171
	(121.969)	(158.480)	(158.164)	(207.573)	(19.343)	(24.363)	(0.031)	(0.038)	(4.741)	(5.698)
N	881	599	881	599	881	599	881	599	881	599
Control	1224.243	1188.416	975.206	1172.650	143.475	136.778	0.780	0.769	12.766	18.803

Note: This table provides evidence for alternative savings measures and potential crowding-out effects. The table reports OLS estimates of a given savings measure on the moneybox treatment in Columns (1), (3), (5), (7), (9) and on an indicator for receiving a recommendation in columns (2), (4), (6), (8), (10). Columns (1)-(2) focus on *Bank savings* as the current balance of an individual's savings accounts (copied from savings booklet). Columns (3)-(4) report estimates for *Equb*, the amount of money people currently hold with *Equb*, an informal savings arrangement. *Iddir* is the amount of money that a person would receive from the funeral society *Iddir* in case of death (column (5)-(6)). In columns (7)-(8) the dependent variable is an indicator for storing any crop and in columns (9)-(10) we take the total expenditure for building materials in the last four months. All variables are winsorized at the 95% percentile. Control variables include all baseline variables used for stratification and the baseline values of the respective outcome variable. The last row gives the mean value of the outcome variable in the control group (no recommendation). Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.5

Effect of Moneybox and Recommendations on Savings – Robustness.

	(1) Savings	(2) Savings
Moneybox	170.939**	
	(72.539)	
Recommendation		263.001***
		(95.646)
Financial literacy	-6.014	-0.939
	(73.710)	(96.054)
N	615	419
- Control	488.226	473.509

Note: This table reports results from OLS regressions with savings as the outcome variable. The results here draw on a subsample excluding individuals with either all financial literacy questions correct or all wrong, thereby abstracting from floor and ceiling effects in the measure for underconfidence. *Savings* is measured as money held in the moneybox (0 for respondents without moneybox) plus cash holdings. Column (1) provides results for individuals who received a moneybox to those who did not receive a moneybox. Column (2) reports results from an OLS regression comparing individuals who received a recommendation to those who did not receive a recommendation. The sample is smaller since only those individuals receiving a moneybox were assigned to one of the recommendation treatments. All variables are winsorized at the 95% percentile. Control variables include all baseline variables used during randomization (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator indicators and our measure of financial literacy. In column (1) the last row gives the mean value of the outcome variable in the control group (no moneybox). In column (2) the last row gives the mean value of the outcome variable in the control group (moneybox, but no recommendation). Robust standard errors in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.6
Effect of Recommendations on Savings of Underconfident Participants – Robustness.

	Savings
Recommendation	216.940**
	(104.363)
Recom x Underconfident	435.150**
	(199.865)
Underconfident	-504.920***
	(169.322)
Financial literacy	25.080
	(99.547)
N	419

Note: N=599. The dependent variable savings is calculated as the value of savings in the moneybox plus cash holdings. The results here draw on a subsample excluding individuals with either all financial literacy questions correct or all wrong, thereby abstracting from floor and ceiling effects in the measure for underconfidence. In this table we introduce the effect of underconfidence, measured by an indicator equal to one for individuals believing to answer fewer questions on financial literacy correctly than they actually do. We interact underconfidence with the recommendations. Mean Control Group refers to the average value of the outcome variable in the control group (no recommendation). All variables are winsorized at the 95% percentile. Additional control variables include all baseline variables used for stratification during randomization (savings in cash, bank and informal arrangements, socio-demographics, land and consumption variables, outstanding debt) as well as enumerator indicators. Robust standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7
Real-Life Measures: Ttests of Underconfident vs. Others.

	(1) Underconfident	(2) Others	(3) p-value
Panel A: Savings Goal Stress			
I find working towards my savings goals to be very stressful.	0.695	0.559	0.004
My savings goals are very much too difficult.	0.323	0.380	0.217
I often fail to attain my savings goals.	0.268	0.354	0.059
The savings officer of the local microfinance instiution acts nonsupportively when I fail to reach my savings goals.	0.291	0.337	0.313
I have too many savings goals.	0.837	0.644	0.000
Panel B: Saving Goal Setting and Performance Appraisal with the Savings Offer			
During meetings with savings officers from the local microfinance institution, the savings officerÉ			
... explains the purpose of the meeting to me...	0.750	0.775	0.527
... asks me what I have done that deserves recognition...	0.680	0.703	0.596
... asks me whether there are any areas on which he or she can assist me...	0.630	0.692	0.169
... tells me what he or she thinks I have done that deserves recognition...	0.606	0.645	0.398
... if there are problems with my savings performance, never brings up more than two of them at once...	0.609	0.696	0.053
... listens openly to my explanations and concerns regarding any savings performance problems...	0.656	0.733	0.074
... comes to agreement with me on steps to be taken by each of us to solve any savings performance problems...	0.602	0.740	0.001
... makes sure that at the end of the meeting I have a specific savings goal or goals in mind that I am to achieve in the future.	0.734	0.790	0.163

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2019.09.024.

References

Ajzen, I., 2002. Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior 1. J. Appl. Soc. Psychol. 32 (4), 665–683.

Allgood, S., Walstad, W.B., 2016. The effects of perceived and actual financial literacy on financial behaviors. Econ. Inq. 54 (1), 675–697.

Ashraf, N., Karlan, D., Yin, W., 2006. Tying odysseus to the mast: evidence from a commitment savings product in the Philippines. Q. J. Econ. 121 (2), 635–672.

Banerjee, A., Duflo, E., 2006. The economic lives of the poor. J. Econ. Perspect. 21 (1), 141–167.

Bannier, C.E., Schwarz, M., 2018. Gender- and education-related effects of financial literacy and confidence on financial wealth. J. Econ. Psychol. 67, 66–86.

Barber, B.M., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Q. J. Econ. 261–292.

Barberis, N., Thaler, R., 2003. A survey of behavioral finance. Handb. Econ. Finance 1, 1053–1128.

Bayer, P.J., Bernheim, B.D., Scholz, J.K., 2009. The effects of financial education in the workplace: evidence from a survey of employers. Econ. Inq. 47 (4), 605–624.

Beaman, L., Duflo, E., Pande, R., Topalova, P., 2012. Female leadership raises aspirations and educational attainment for girls: a policy experiment in india. Science 335 (6068), 582–586.

Becchetti, L., Solferino, N., Tessitore, M.E., et al., 2015. Some insights on procrastination: a curse or a productive art? Rev. Behav. Econ. 2 (4), 331–351.

Bedi, A., Marshall, J.H., 2002. Primary school attendance in honduras. J. Develop. Econ. 69 (1), 129–153.

Berg, G., Zia, B., 2013. Harnessing emotional connections to improve financial decisions: evaluating the impact of financial education in mainstream media. World Bank Policy Res. Work. Paper (6407).

Bisin, A., Hyndman, K., 2014. Present-bias, procrastination and deadlines in a field experiment. Technical Report. National Bureau of Economic Research.

Brown, J.R., Previtero, A., 2014. Procrastination, present-biased preferences, and financial behaviors. Technical Report. National Bureau of Economic Research.

Brune, L., Giné, X., Goldberg, J., Yang, D., 2016. Facilitating savings for agriculture: field experimental evidence from malawi. Econ. Develop. Cult. Change 64 (2), 187–220.

Caliendo, F., Huang, K.X.D., 2007. Overconfidence in financial markets and consumption over the life cycle. Working Papers 07-3. Federal Reserve Bank of Philadelphia.

Campbell, J.Y., 2016. Restoring rational choice: the challenge of consumer financial regulation. Am. Econ. Rev. 106 (5), 1–30.

Carvalho, L.S., Prina, S., Sydnor, J., 2016. The effect of saving on risk attitudes and intertemporal choices. J. Develop. Econ. 120, 41–52.

Cobb-Clark, D.A., Kassenboehmer, S.C., Sinning, M., 2013. Locus of control and savings. Ruhr Econ. Paper (455).

Coutts, A., 2019. Good news and bad news are still news: experimental evidence on belief updating. Exper. Econ. 22 (2), 369–395.

Disney, R., Gathergood, J., 2013. Financial literacy and consumer credit portfolios. J. Bank. Finance 37 (7), 2246–2254.

Doerr, U., Toman, O.M., Schmidt, U., 2011. Overconfidence and Risk Management of Ethiopian Farmers. Technical Report. Working Paper.

Duflo, E., 2017. Richard t. ely lecture: the economist as plumber. Am. Econ. Rev. 107 (5), 1–26.

Duflo, E., Kremer, M., Robinson, J., 2009. Nudging Farmers to use Fertilizer: Theory and Experimental Evidence from Kenya. Technical Report. National Bureau of Economic Research.

Dupas, P., Karlan, D., Robinson, J., Ubfal, D., 2018. Banking the unbanked? evidence from three countries. Am. Econ. J. Appl. Econ. 10 (2). 257–97.

Dupas, P., Keats, A., Robinson, J., 2017. The effect of savings accounts on interpersonal financial relationships: evidence from a field experiment in rural Kenya. Econ. J. 129 (617), 273–310.

Dupas, P., Robinson, J., 2013. Why don't the poor save more? evidence from health savings experiments. Am. Econ. Rev. 103 (4), 1138–1171.

Ertac, S., 2011. Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. J. Econ. Behav. Organ. 80 (3), 532–545.

Fouarge, D., Salamanca, N., Montizaan, R., et al., 2013. Locus of Control and Investment in Risky Assets. Technical Report. Maastricht University, Research Centre for Education and the Labour Market (ROA).

Garrard, P., Robinson, G., 2015. The Intoxication of Power: Interdisciplinary Insights. Palgrave Macmillan.

Gervais, S., Odean, T., 2001. Learning to be overconfident. Review of Financial Studies 14 (1), 1–27.

Grubb, M.D., 2009. Selling to overconfident consumers. The Am. Econ. Rev. 99 (5), 1770–1807.

Grubb, M.D., 2015. Overconfident consumers in the marketplace. The Journal of Economic Perspectives 29 (4), 9–35.

Holt, C., Laury, S., 2002. Risk aversion and incentive effects. Am. Econ. Rev. 92 (5), 1644–1655.

John, A., 2017. When Commitment Fails - Evidence from a Field Experiment. Technical Report. CREST.

Karlan, D., Linden, L.L., 2017. Loose knots: strong versus weak commitments to save for education in uganda. Working Paper.

Karlan, D., McConnell, M., Mullainathan, S., Zinman, J., 2016. Getting to the top of mind: how reminders increase saving. Management Science 62 (12), 3393–3411.

Karlan, D., Ratan, A.L., Zinman, J., 2014. Savings by and for the poor: a research review and agenda. Review of Income and Wealth 60 (1), 36–78.

Kast, F., Meier, S., Pomeranz, D., 2018. Saving more in groups: field experimental evidence from chile. Journals of Development Economics 133, 275–294.

Kaur, S., Kremer, M., Mullainathan, S., 2015. Self-control at work. Journal of Political Economy 123 (6), 1227–1277.

Kirchler, E., Maciejovsky, B., 2002. Simultaneous over-and underconfidence: evidence from experimental asset markets. Journal of Risk and Uncertainty 25 (1), 65–85.

Kramer, M.M., 2016. Financial literacy, confidence, and financial advice seeking. Journal of Economic Behavior and Organization 131, 198–217.

Kremer, M., Rao, G., Schilbach, F., 2019. Behavioral development economics. Handbook of Behavioral Economics, Douglas Bernheim, Stefano DellaVigna, and David Laibson (eds.).

Krishnan, P., Krutikova, S., 2013. Non-cognitive skill formation in poor neighbourhoods of urban india. Labour Economics 24 (C), 68–85.

Latham, G.P., Locke, E.A., 1979. Goal setting—A motivational technique that works. Organizational Dynamics 8 (2), 68–80. doi:10.1016/0090-2616(79)90032-9.

Lee, C., Bobko, P., Christopher Earley, P., Locke, E.A., 1991. An empirical analysis of a goal setting questionnaire. Journal of Organizational Behavior 12 (6), 467–482.

Li, Y., Shan, M., Li, M.Z., 2016. Advance selling decisions with overconfident consumers. Management 12 (3), 891–905.

Linardi, S., Tanaka, T., 2013. Competition as a savings incentive: a field experiment at a homeless shelter. J. Econ. Behav. Organ. 95, 240–251.

Lusardi, A., Mitchell, O.S., 2014. The economic importance of financial literacy: theory and evidence. J. Econ. Lit. 52 (1), 5–44.

Malmendier, U., Tate, G., 2005. Ceo overconfidence and corporate investment. J. Finance 60 (6), 2661–2700.

McKelway, M., 2018. Women'S self-efficacy and women's employment: experimental evidence from india. Unpublished Working Paper.

McKenzie, D., Sansone, D., 2019. Predicting entrepreneurial success is hard: evidence from a business plan competition in Nigeria. J. Develop. Econ.

Mobius, M., Niederle, M., Niehaus, P., Rosenblat, T., 2011. Managing Self-Confidence: Theory and Experimental Evidence. NBER Working Papers 17014. National Bureau of Economic Research, Inc.

Moore, D.A., Healy, P.J., 2008. The trouble with overconfidence. Psychol. Rev. 115 (2), 502–517.

Ortoleva, P., Snowberg, E., 2015. Overconfidence in political behavior. Am. Econ. Rev. 105 (2). 504–35.

Perry, V.G., Morris, M.D., 2005. Who is in control? the role of self-perception, knowledge, and income in explaining consumer financial behavior. J. Consum. Aff. 39 (2), 299–313.

de Quidt, J., Haushofer, J., Roth, C., 2018. Measuring and bounding experimenter demand. Am. Econ. Rev. 108 (11), 3266–3302.

van Rooij, M.C., Lusardi, A., Alessie, R.J., 2012. Financial literacy, retirement planning and household wealth*. Econ. J. 122 (560), 449–478.

Rotter, J.B., 1966. Generalized expectancies for internal versus external control of reinforcement.. Psychol/ Monogr. Gener. Appl. 80 (1), 1.

Roy, S., Morton, M., Bhattacharya, S., 2018. Hidden human capital: self-efficacy, aspirations and achievements of adolescent and young women in India. World Develop. 111 (C), 161–180.

Shipton, P., et al., 1990. How Gambians Save and what their Strategies Imply for International Aid. Technical Report. The World Bank.

Steinert, J., Cluver, L., Meinck, F., Doubt, J., Vollmer, S., 2018. Household economic strengthening through financial and psychosocial programming: evidence from a field experiment in south africa. J. Develop. Econ. 134.

Suzuki, A., Mano, Y., Abebe, G., 2018. Earnings, savings, and job satisfaction in a labor-intensive export sector: evidence from the cut flower industry in ethiopia. World Develop. 110 (C), 176–191.

Tang, N., Baker, A., 2016. Self-esteem, financial knowledge and financial behavior. J. Econ. Psychol. 54, 164–176.

Thaler, R.H., Benartzi, S., 2004. Save more tomorrow: using behavioral economics to increase employee saving. J. Political Econ. 112 (S1), S164–S187.

World Bank, 1998. Financial Literacy Questionnaire. Technical Report. The World Bank.