Mental Accounting and Cash Transfers: Experimental Evidence from a Humanitarian Setting*

Till Wicker[1], Patricio Dalton[1], and Daan van Soest[1]

[1]Tilburg University

March 18, 2025

Abstract

Can mental accounting help economically disadvantaged people accumulate capital and grow their income? To answer this question, we run a field experiment with 861 refugee households in Uganda who received an unconditional cash transfer over seven months. Households in the treatment group were given the option to divide their monthly cash transfers across four labeled envelopes (*Education*, *Health*, *Investments*, *Others*). In contrast, households in the control group received their monthly cash in one unlabeled envelope. One year after the cash transfer program ended, the treatment group, relative to the control group, significantly increased their investments, particularly in lumpy assets, leading to a higher monthly income and savings. These results are primarily driven by households that continued using the labeled envelopes one year after the program ended, who were also more likely to report that the envelopes help with budgeting, planning, and spending discipline.

*AEA RCT identification number: AEARCTR-0010472. This project received IRB approval from Tilburg University (IRB FUL 2022-004) and Mildmay Institute of Health Sciences (MUREC-2022-144). We gratefully acknowledge funding from the Dutch Research Council (NWO) Open Competition grant 406.21.E8.004. The project has successfully undergone a Stage 1 Pre-Results Acceptance at the *Journal of Development Economics*, and we are grateful to the guidance of Co-Editor Dean Yang and the constructive feedback of two anonymous referees. We also thank seminar participants at Maastricht, Tilburg, Torino, Lund, Field Days Conference, LISER, NNH, Tilburg Development and Economic History reading group, Georgetown Qatar, UC Berkeley, and PACDEV and employees at the Danish Refugee Council, IMPACT Initiatives, UNHCR, WFP, 100Weeks, ZOA, Alight, LWF, GiveDirectly, and DG ECHO for helpful discussions.

1 Introduction

Despite progress in poverty reduction, impoverished individuals still struggle to accumulate assets and increase their income (Banerjee and Duflo, 2007). While external barriers, such as limited access to capital, play a role, psychological or "internal" barriers may prevent those living in poverty from pursuing investment opportunities, even when the potential returns are high (Kremer et al., 2019). One behavioral constraint discussed in the literature is the difficulty poor people can face in committing to savings and investment plans (Collins et al., 2009), which can explain the strong demand for commitment devices among the poor, as documented by Ashraf et al. (2006), Dupas and Robinson (2013), and Kaur et al. (2015).

Commitment challenges are likely to be exacerbated in humanitarian settings. In addition to their heightened vulnerability, refugees are more likely to experience mental health problems (Pozuelo et al., 2023), which can worsen self-control problems (Kim and Park, 2015). Ercolano and Woolfenden (2021) document that refugees in Uganda planned for future expenses but failed to save adequately. They also report direct demand for commitment devices, noting that "[some refugee households suggested] it might make sense for organizations to pay school fees directly to prevent [them] from being tempted." As put by Banerjee et al. (2023) "even the most destitute households often look for ways to accumulate sums of money large enough to make larger, lumpier purchases. Designing [cash] schemes in ways that respond to this need could make them a more compelling strategy for addressing extreme poverty over time."

To address this challenge, we conducted a field experiment among 861 refugee households in Rhino Camp and Imvepi refugee settlements in Uganda, all of whom are recipients of a seven-month-long unconditional cash transfer program, receiving $22.89 PPP, per household member per month. The intervention we designed only marginally altered the usual cash transfer modality by introducing a soft commitment device that leverages mental accounting. Instead of receiving their monthly cash transfer in one unlabeled envelope (the status quo), households in the treatment arm were offered the opportunity to receive their cash transfers across four envelopes labeled "Education", "Health", "Investments", and "Other".

The first key insight from our field experiment stems from the exceptionally high demand for the soft-commitment device we offered: 93% of treatment households chose to divide their cash transfers among the four labeled envelopes.[1] Of these households,

[1]The take up was higher than the typical uptake of commitment devices in low-income settings (for

83% stated that the four labeled envelopes would help them improve their financial discipline, savings, and to resist purchasing temptation goods. This is in line with the theory of change we pre-specified, which posited that the four labeled envelopes would act as a soft-commitment device and help address commitment challenges.[2]

One year after the cash transfer program ended, households in the treatment group invested 25% more in income-generating activities compared to the control group, driven by larger lumpy investments. These investments led to an 18% increase in monthly income, and a 22% increase in savings. The larger investments were financed primarily through the households' own savings, supplemented by loans taken out during the cash transfer program: both savings and loans were 70% higher immediately after the cash transfer program ended compared to households in the control group. One year later, these loans had been repaid. Interestingly, we do not find significant effects of the intervention on education or health expenditures, which may suggest that the need for a commitment device was lower in these domains, at least during the period of this study.

The results we observe are primarily driven by the 35% of households that continue using the labeled envelopes one year after the cash transfer program ended - hereafter referred to as the 'persistent users'. Compared to households that stopped using the labeled envelopes after the end of the cash transfer program, the 'persistent users' have larger outstanding loans at baseline (suggesting greater financial strain), were younger, and expressed a stronger desire for higher future income. Furthermore, 'persistent users' were also more likely to report at baseline that the partitioning and labeling of the money would help them with budgeting, planning, and spending discipline.

Beyond assessing the effectiveness of the soft commitment device, we were also interested in whether its impact would differ depending on whether recipients could decide how to distribute the money in each of the four envelopes, or received a suggested default allocation per envelope. This question is important not only for policy, but also for research, as Prelec and Herrnstein (1991) distinguishes between behavior-governing rules set by "agents who have our interests in mind" and those "constructed by ourselves as we see the need for them". To investigate this, we randomly assigned

an overview, see Table 1 of Schilbach (2019)). A possible explanation may be that the commitment device we offered was arguably softer than others evaluated, such us such as lockboxes or blocked savings accounts (Ashraf et al., 2006; Dupas and Robinson, 2013).

[2]The theoretical frameworks of Bernheim et al. (2015) and Laajaj (2017) also suggest cash transfers present an excellent opportunity to introduce mental accounts that encourage future-oriented saving and spending, as the cash transfer increases the recipient's self-control and prolongs their time horizon.

the treatment group into two sub-groups: one in which households could freely decide their allocation across the four envelopes (Mental Accounting - hereafter ***MA***), and another where households were first presented with an NGO-calculated default recommended allocation, which they could either accept or adjust (Mental Accounting with Default - hereafter ***MAD***). While the default provided guidance on dividing the cash transfer, it was not tailored to each household's specific needs. Interestingly, uptake of the default was high at 96%, and allocations differed significantly between the two sub-treatments.

One year after the cash transfer ended, ***MAD*** households made larger investments, financed through loans and savings, but only ***MA*** households experienced positive effects on income and savings. This difference seems to be attributed to different investment patterns across sub-treatments: ***MAD*** households focused on livestock and agriculture, while ***MA*** households diversified into enterprises. Notably, the likelihood of continuing to use the labeled envelopes did not differ between the sub-treatments.

This paper contributes to several strands of literature. First, it adds to the emerging field of behavioral development economics by addressing a behavioral constraint to saving and investing among the poor (Kremer et al., 2019). While previous studies have examined the impact of role models and aspirational workshops (Bernard and Taffesse, 2014; Orkin et al., 2024), planning interventions (Augenblick et al., 2024), defaults (Banerjee et al., 2025), pharmacotherapy (Angelucci and Bennett, 2024), or cognitive behavioral therapy (Blattman et al., 2017), our paper proposes a different approach based on mental accounting. Compared with other studies that have used commitment devices to promote savings, such as savings groups (Karlan et al., 2017), separate savings accounts (Ashraf et al., 2006; Brune et al., 2017, 2021), and lockboxes (Dupas and Robinson, 2013; Aggarwal et al., 2023), the commitment device we study is softer, cheaper, and thus more scalable.

Second, our paper contributes to the literature on mental accounting (Thaler and Shefrin, 1981; Thaler, 1985; Heath and Soll, 1996; Thaler and Benartzi, 2004). A related study by Soman and Cheema (2011) provided Indian workers with the opportunity to set aside a portion of their weekly income for their children's education by storing it in a labeled envelope, leading to higher savings for education. By offering multiple labeled envelopes – rather than a single one as in Soman and Cheema (2011) – we can study trade-offs between different mental accounts. Furthermore, by distinguishing between treatment arms with and without a default, we contribute to the discussion

on self-chosen versus externally recommended mental accounts (Prelec and Herrnstein, 1991). Finally, to the best of our knowledge, our paper is the first to integrate mental accounts within a cash transfer program.

The third strand of literature to which our paper contributes concerns the effectiveness of cash transfer programs as a social protection policy. Meta-analyses have documented lasting positive effects beyond the duration of cash transfer programs (Bastagli et al., 2016; Crosta et al., 2024). Although several studies have examined the effects of varying the frequency, amount, and duration of cash transfers (Haushofer and Shapiro, 2016; Banerjee et al., 2023), others have combined cash transfers with interventions designed to alleviate additional (behavioral) constraints to enhance their impact (Ahmed et al., 2025). Examples include psychological counseling (Haushofer et al., 2020), asset transfers (Bossuroy et al., 2022), and aspiration workshops (Orkin et al., 2024). In contrast, our study simply modifies the transfer modality by offering the cash in four labeled envelopes. As such, our intervention has several advantages: it requires negligible upfront fixed costs, seamlessly integrates into existing NGO operations, is highly scalable, and can be easily adaptable to new settings, including digital payment systems.

Finally, our paper contributes to policy discussions on humanitarian aid. The number of people relying on humanitarian assistance continues to rise, with 35 million refugees, 108 million displaced individuals, and over 400 million in need of humanitarian aid by the end of 2022 (Development Initiatives, 2023; UNHCR, 2023). Notably, 78% of humanitarian aid recipients live in protracted displacement settings, prompting humanitarian organizations to shift their focus from addressing only short-term basic needs to incorporating longer-term development objectives.[3] As a result, cash transfers have emerged as a widely favored humanitarian policy due to their scalability, flexibility, cost-effectiveness, and the greater autonomy they afford recipients. In 2022, the value of humanitarian cash transfers exceeded $7.9 billion (Development Initiatives, 2023).[4] Our mental accounting intervention has the potential to enhance the effectiveness of humanitarian cash transfers, as it is highly scalable and low-cost ($1.70 per household) and has demonstrated positive effects on households' financial resilience one year after the program's conclusion.

[3]Protracted refugee situations are "those in which at least 25,000 refugees from the same country have been living in exile for more than five consecutive years" (UNHCR, 2025).

[4]Several studies have conducted evaluations of cash transfer programs in humanitarian settings include Hidrobo et al. (2014); Aker (2017); Ozler et al. (2021); Altındağ and O'Connell (2023).

The remainder of this paper is structured as follows. Section 2 outlines the context and experimental design, while Section 3 presents the results. Sections 4 and 5 discuss heterogeneity and potential mechanisms. Section 6 examines cost-effectiveness and Section 7 concludes.

2 Context and Experimental Design

2.1 Context

Uganda experienced a significant influx of refugees from 2016 to 2018, with over 900,000 South Sudanese nationals fleeing a civil war. Since then, the number of refugees has continued to rise and exceeded 1.7 million by the end of 2024 (UNHCR, 2024).[5] Upon arrival at a refugee settlement, each refugee household is allocated a 30-by-30-meter plot of land for shelter construction and small-scale agriculture. Additionally, refugees in Uganda have freedom of movement and the right to work, and can rent additional agricultural land from Ugandan landlords. However, 91.5% of refugees continue to reside within designated refugee settlements. Within these settlements, the World Food Programme (WFP) provides food assistance, and health centers offer free medical services.[6] Schools in the settlements are operated by NGOs under the directive of the Ugandan Ministry of Education. Despite this, attending school remains costly, as families must cover expenses for supplies, uniforms, and school fees each term.[7]

In two of Uganda's refugee settlements, Rhino Camp and Imvepi, the Danish Refugee Council (DRC) provides unconditional cash transfers to the most vulnerable households.[8]

[5]Refugees do not believe the conflict will end soon, and hence do not have the desire to return to South Sudan: at baseline, only 7% of households said they would want to return to South Sudan in the next two years, the remaining households intended on staying in Uganda.

[6]Larger treatments (e.g., amputations) are also covered, however referrals need to be made to regional hospitals with the appropriate facilities. Health centers within settlements typically provide basic medical services.

[7]Tuition fees are paid per term, costing 2,000 UGX ($1.53 PPP) for primary school children, and 50,000-100,000 UGX ($38.30-76.60 PPP) for secondary school children. There are three terms per academic year. Furthermore, national examination fees cost 34,000, 179,000, and 201,000 UGX ($26.04, $137.12, $153.97 PPP) for primary, lower secondary, and upper secondary school exams, respectively. Scholastic materials cost around 15,000 UGX ($11.49 PPP) and 120,000 UGX ($91.92 PPP) per primary and secondary school child, respectively.

[8]Vulnerability is calculated using a 27-item Vulnerability Scoring Model, covering three broad categories: Household Demographics, Socio-Economic Situation and Food Security, and Sectoral. The individual questions, answers, and cut-off scores for vulnerability are confidential and hence cannot

These transfers are distributed in seven monthly installments of $22.89 PPP ($8.36) per household member per month.[9] These transfers aim to help recipient households meet their basic needs and work toward self-reliance through savings and investments. Recipients can choose their preferred transfer modality, either physical cash or mobile money. However, over 90% opt for physical cash due to limited mobile phone ownership and poor cellular connectivity within the settlements.

2.2 Experimental Design

2.2.1 Description of the sample

We enrolled 861 refugee households (HHs) eligible for DRC's seven-month-long unconditional cash transfer program in our RCT. Of these, 90% originate from South Sudan, while the remaining 10% are from the Democratic Republic of the Congo. Among household heads, 81.6% are female, with an average age of 38 years.[10] Additionally, 85% exhibit symptoms of moderate or severe depression. The mean (and median) monthly income of households — excluding cash transfers — is $44.25 PPP ($11.45 PPP), resulting in an average daily income of $0.23 PPP per household member.[11] Households primarily earn income from livestock rearing and crop cultivation, in addition to receiving a monthly food ration from the WFP. For 85% of households, the monthly cash transfer exceeds their baseline monthly income.

The 861 households were randomly assigned to one of three experimental arms: Control (***CO***), Mental Accounting (***MA***), and Mental Accounting with Default (***MAD***). Randomization was stratified based on the household head's median age, gender, household size, country of origin, geographic zone, timing of the cash transfer, year of arrival, and vulnerability score.[12] Treatment arms are balanced, as shown in Appendix Tables

be shared.

[9]The size of the transfer is based on the Minimum Expenditure Basket (MEB), a calculation done by the Uganda Cash Working Group that captures the costs of a refugee household meeting its basic needs. The MEB is divided into *food* and *non-food* items (see Appendix D), with DRC's cash transfers covering the MEB value for *non-food* items. The World Food Programme's food aid covers the food component of the MEB. The total value of the cash transfer is smaller than those typically given by GiveDirectly. Given taht Egger et al. (2022) document cash transfer-induced inflation of less than 1%, inflationary concerns as a result of the cash transfers are low.

[10]The majority of households are female-headed because the husbands typically stay in their native country, and send the wife and children to Uganda in search of safety. Given both South Sudan and the Democratic Republic of the Congo are patriarchic societies, for many women this is the first time they are responsible for the household, and the finances.

[11]The World Bank's extreme poverty line lies at $2.15 PPP per person per day.

[12]A median split was used for stratify by the household head's age, household size, year of arrival,

16 and 17.

2.2.2 Treatment implementation

During the baseline survey, all households were encouraged to consider their future spending and investment plans.[13] They also received an *Investment Opportunities* sheet, which outlined productive investment options identified through focus group discussions, with costs based on median market prices in the refugee settlements. This sheet aimed to reduce information constraints on productive investments.

For ***CO*** households, the baseline survey ended after receiving the *Investment Opportunities* sheet, after which they continued to receive their monthly cash transfer in a single unlabeled envelope, as per the NGO's status quo. Households in ***MA*** and ***MAD*** were given the opportunity to allocate their future monthly cash transfers among four smaller envelopes, labeled "Education", "Health", "Investments", and "Other" (see Figure 1), instead of a single unlabeled envelope.[14]

In ***MA***, if the household head opted-in for the four labeled envelopes, they were asked to allocate their monthly cash transfer across them. Future cash distributions followed this chosen allocation. In the ***MAD*** treatment, if the household head opted-in for the labeled envelopes, they were shown a recommended allocation across the four envelopes, based on the Minimum Expenditure Basket. The household head could choose to either accept or reject this recommendation. If rejected, they determined their own allocation, similar to the ***MA*** setup.

Households that opted-in for the four labeled envelopes (either in ***MA*** or ***MAD***) further received an *Envelope Allocation* sheet. This sheet displayed the monetary amounts allocated to each envelope category, allowing households to verify that their cash transfer was accurately distributed.[15]

---

and vulnerability score.

[13] We conducted the baseline survey between the first and second cash transfers due to logistical constraints, as DRC identified eligible households only shortly before the program's start, leaving insufficient time for a pre-transfer survey. As a result, all households had received their first transfer before the baseline survey.

[14] The envelope categories and labels (in the form of stickers) were piloted and refined through focus group discussions with past recipients of the unconditional cash transfer. Follow-up groups discussions conducted 1.5 years after program ended revealed that most households would not have chosen different categories. Only two households mentioned that a food envelope would have been helpful.

[15] Appendix C provides further details on the *Investment Opportunities* and *Envelope Allocation* sheets.



Figure 1. Four Labeled Envelopes ("Education", "Health", "Investments", "Other")

Demand for receiving the cash transfer in the four labeled envelopes was high: 93% of households in the ***MA*** and ***MAD*** treatments opted in ($p = 0.55$), suggesting a strong demand for commitment devices. Furthermore, 96% of households in ***MAD*** who opted in for the four labeled envelopes accepted the default allocation, indicating a preference for guidance or lack of strong prior preferences. Table 1 presents differences in uptake and between-envelope allocations between the ***MA*** and ***MAD*** treatments. Allocations across the four envelope categories differ substantially between the ***MA*** and ***MAD*** treatments, with ***MA*** households allocating more to education and health on average, while allocating less to investments and other expenses.

Logistics of Cash Transfers and Envelopes

Cash transfers were distributed monthly on a pre-specified date. A money van from a Ugandan bank arrived at designated locations in the refugee settlements at a pre-announced time. DRC staff first verified the identity of the household head, who then collected the cash transfer from the money van (see Figure 2).[16] After receiving their cash, the household head proceeded to the *Envelopes Stand* (see Figure 3).[17]

At the *Envelopes Stand*, DRC workers verified whether the household received the four labeled envelopes. If not, the money was placed in a single large unlabeled envelope.

If the household head had opted-in to receive their cash in the labeled envelopes,

[16]Bank tellers were unaware of households' treatment assignments. As such, we can rule out that denominations differed between treatment groups. This is important as denominations have been shown to impact spending patterns (Raghubir and Srivastava, 2009).

[17]Household heads waited in a queue standing three meters from the stand, and arrived one at a time. Order and safety were maintained by two armed security guards.

Table 1: Allocations Across Envelopes: ***MA*** vs. ***MAD***.

Variable	N	(1) ***MA*** Mean/(SD)	N	(2) ***MAD*** Mean/(SD)	(3) Pairwise t-test Difference
Uptake	288	0.938	281	0.925	0.013
		(0.242)		(0.263)	
Default Accepted			260	0.962	
				(0.193)	
Education Share	270	0.268	260	0.168	0.100***
		(0.149)		(0.021)	
Health Share	270	0.198	260	0.173	0.025***
		(0.112)		(0.017)	
Investment Share	270	0.288	260	0.330	-0.042***
		(0.148)		(0.023)	
Other Share	270	0.246	260	0.330	-0.084***
		(0.163)		(0.023)	
Joint distribution test					$\chi^2(2, 8) = 40.24$***

Notes: Columns (1) and (2) show the average value (and standard deviation) for respondents in the two intervention treatments: Mental Accounting and Mental Accounting with Default. Differences in shares are reported in column (3), with statistical significance as determined using standard pairwise t-tests. The Chi-squared test checks for the equality of the distributions over the four envelope categories between ***MA*** and ***MAD***. Appendix Figure 4 displays histograms of the allocation shares across the four envelopes for ***MA*** and ***MAD***. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

the amount of cash they were entitled to was divided between the four envelopes based on the allocations obtained during the initial survey. Once the money was divided between the four envelopes, the four envelopes were placed inside one large unlabeled envelope. This was done to avoid signaling a difference in treatments, as all households left the cash distribution site with their cash transfer in one large unlabeled envelope. Hence, we reduced the chance that households in the control group (***CO***) were aware of the other two treatments, minimizing spillovers.[18]

The cash distribution process had a Complaints Desk, where recipient households could lodge complaints to DRC staff. The staff members responsible for running the complaints desk were trained by the research team on how to document and respond to complaints regarding the RCT and its treatments. No complaint related to the field experiment was lodged.

[18]Focus discussions conducted 1.5 years after the cash transfer program ended indicated that households in the control group were unaware of the four labeled envelopes.



Figure 2. Cash Distribution.



Figure 3. Envelopes Stand.

2.3 Econometric Specification

Households were surveyed two weeks after the first cash transfer (baseline), one month after the program ended (midline), and again one year later (endline) to document both the immediate and longer-term effects of the intervention (see a detailed timeline of the project in Appendix F). Attrition was low at 5.92% and 16.8% for midline and endline, respectively. Appendix H shows that there was no differential attrition between experimental arms.

To estimate the effects of the four-envelope intervention on the outcomes of interest, we run the following model:

$$Y_{ht} = \beta_0 + \beta_1 4Envelopes_h + \delta_e + \gamma_z + X_h + Y_{h0} + Early_h + \varepsilon_h, \quad t = \{1, 2\}. \tag{1}$$

where Y_{ht} represents outcome variable Y for household h measured at midline ($t = 1$) and endline ($t = 2$). ***4Envelopes*** is an indicator capturing whether household h was randomized into the treatment group (combining both the ***MA*** and ***MAD*** treatments), hence β_1 is our key parameter of interest. X_h is a vector of pre-registered baseline covariates, consisting of stratification variables and those unbalanced at baseline (Bruhn and McKenzie, 2009). We also include fixed effects for the Settlement Zone in which the household lives (γ_z) and for the enumerator (δ_e), following Maio and Fiala (2020). Following McKenzie (2012), we control for the outcome variables measured at baseline, Y_{h0}, whenever available. The indicator $Early_h$ equals one for households that received their first cash transfer in August 2022, and zero for households that received their first cash transfer in September 2022, capturing any time-specific effects. Finally,

ε_h is a heteroskedasticity-robust error term.[19] In a second, pre-registered specification, we evaluate ***MA*** and ***MAD*** separately. Given we report treatment effects on several outcome variables, we report sharpened q-values based on Anderson (2008), as pre-registered (Wicker et al., 2023).

3 Results

We first present the treatment effects of the intervention at endline to document the longer-term changes in outcomes. Subsequently, we make use of the midline data, to explore the mechanisms driving the treatment effects after one year. We present results for the ***4Envelopes*** intervention (pooling ***MA*** and ***MAD***), as well as for each treatment separately.

3.1 Effects One Year Post-Cash Transfers

Table 2 presents the estimated treatment effects of the four-envelope intervention on economic outcomes one year after the end of the cash transfer program. Columns (1) and (2) report effects on total investment and lumpy investments.[20] Compared to households in the control group, households that were offered the four envelopes spent 25.3% more on investments (0.23 s.d.) in the year since the end of the cash transfer program, and 32.2% more on lumpy investments (0.20 s.d.). Columns (3) and (4) suggest that these larger investments translate into a 18.6% higher monthly income and 22.6% higher savings (0.16 and 0.14 s.d., respectively). Small and statistically insignificant reductions are observed in the household's level of outstanding debt, as well as the value of household durable goods purchased in the last year, as shown in columns (5) and (6). These results suggest that earmarking an envelope for "Investments" may help households allocate funds toward investments after the cash transfer program ends, potentially leading to higher returns, increased monthly income, and larger savings.

[19] Following Abadie et al. (2022), we cluster standard errors at the household level, as randomization was implemented at the household level.

[20] To measure total investment, respondents were presented a series of investment items and asked how many of those items they had purchased in the last year. These were then multiplied by the median market price taken from three vendors in the refugee settlement. To measure lumpy investments, instead, respondents were asked about the five largest investments made in the last year: what they were and what they cost. Examples of lumpy investments include pigs, machinery, and market stalls. Note that 'total' and 'lumpy' may overlap.

Table 2: Endline Outcomes (USD PPP)

	(1) Total Investment	(2) Lumpy Investment	(3) Monthly Income	(4) Savings	(5) Loans	(6) Durable Goods	(7) Educ. Exp.	(8) Health Exp.
Envelopes	60.05*	16.35**	4.65*	8.75*	-5.15	-5.52	-16.35	-47.10**
	(32.73)	(6.39)	(2.40)	(5.11)	(4.26)	(4.44)	(15.57)	(19.19)
Sharp. q-val	0.076	0.060	0.060	0.076	0.163	0.163	0.196	0.060
Control Group Mean	235.12	50.61	24.94	38.63	39.69	33.21	250.56	330.31
Control Group S.D.	264.59	82.60	29.28	62.07	68.81	59.08	230.35	291.17
N	737	737	737	737	737	737	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Total investments refer to the acquisition of a pre-specified list of productive assets/investments (based on pre-RCT focus group discussions) since the previous survey, multiplied by the median market price obtained from three vendors in the refugee settlements. Lumpy investments are self-reported large investments made since the previous survey in physical and human capital. Monthly income is self-reported monthly income in the last month, while Savings and Loans are total savings and loans held by the household at the time of the survey. Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets since the previous survey, multiplied by the median market price obtained from three vendors in the refugee settlements. Education-related expenses refer to expenses in pre-specified list of education-related goods since the previous survey, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported school fees paid. Health-related expenses refer to expenses in pre-specified list of preventive health-related items since the previous survey, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported medicine expenses. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Columns (7) and (8) report annual expenditures on education and health. Treatment assignment leads to a reduction in spending on both categories in the year following the cash transfer program, with a statistically significant decrease of 14.3% (0.16 standard deviations) in health-related expenses. Online Appendix A further decomposes these treatment effects on health expenditures, showing that the observed negative effects are entirely driven by relatively lower spending on latrines in the treatment group after the cash transfers ended. However, decomposing the treatment effect on health expenditures at midline (immediately after the cash transfer ended) reveals that treatment households increased their spending on during the cash transfer period. Hence, while total spending on health-related expenses, particularly on latrines, does not differ statistically significantly between treatment and control households over the 18 months between baseline and endline ($p = 0.26$), the timing of these purchases does. Treatment households upgrade the quality of their latrines (a lumpy health investment) *sooner* than those in the control group.[21]

Turning to the negative though statistically insignificant coefficient on education-related spending, an interesting pattern emerges. In the year between the midline

[21] Appendix A discusses health outcomes in more detail, including treatment effects on health-related indicators, such as the number of health needs, and household's ability to meet their health needs. No statistically significant differences are documented.

and endline survey, the DRC implemented the *Education in Emergencies* cash transfer program, targeting households whose children were not enrolled in school. Treatment households were 12% less likely to receive this additional conditional cash transfer (p=0.08), see Online Appendix B. As a result, the reduction in education-related spending we observe may be due to the treatment households being less likely to receive additional education-related funding.[22] This raises broader insight into evaluating the effectiveness of humanitarian interventions: when interventions significantly improve the living conditions of its recipients, these recipients are less likely to receive additional humanitarian assistance in the future. As a result, general equilibrium treatment effect estimates likely underestimate the true partial equilibrium treatment effect. This will be discussed further in Section 6.

3.2 Post-Cash Transfer Effects: Immediate Outcomes

To understand how households in the treatment arms financed the larger (lumpy) investments after the end of the cash transfer, Table 3 reports treatment effects of the intervention on economic outcomes immediately after the end of the cash transfer program. Columns (1) and (2) report large effects on households' savings (71.6%, 0.52 standard deviations) and on the value of loans pending to repay (70.8%, 0.34 standard deviations). Larger savings were expected, as the main aim of the four envelopes was to act as a commitment device provided. We note that the increased savings may have come at the expense of lower spending on durable goods, although not statistically significant (Column (3)). One possible interpretation of the decline in durable goods spending during the cash transfer period is that the soft commitment device provided by the four envelopes may have partially crowded out a more costly commitment mechanism of durable goods.[23]

While the increase in savings was expected, the rise in loans was not. Although the reason for this increase is not fully clear, one possible explanation is that these loans were needed to complement savings and pay for the lumpy investment we observe at endline.[24] Focus group discussions seven months after the endline survey provide

[22]Results from the focus group discussions conducted seven months after the endline survey are also in line with this explanation, as treatment households are more likely to agree that they were better able to pay school fees by saving money over time.

[23]Durable goods are often seen as commitment devices as they can be liquidated (at a cost) in case of emergencies (Kang and Kang, 2022).

[24]Savings could have in principle served as collateral for larger loans. However, we do not find evidence for such channel. Lenders interviewed in the settlements stated that they do not accept

suggestive evidence for this, as is discussed in greater detail in Section 5.

Column (4) of Table 3 shows a (statistically insignificant) reduction in total investment during the cash transfer period. Together with the increase in savings and loans, this may suggest that, relative to households in the control group, treated households postponed some investments and chose to save money and supplement savings with loans to finance lumpy investments.[25] Column (5) of Table 3 documents no treatment effect on monthly self-reported income, which aligns with expectations as treatment households have not yet made additional investments. Finally, Columns (6) and (7) report negative and positive effects on education and health-related expenses, respectively, but both are statistically insignificant.

Table 3: Midline Outcomes (USD PPP)

	(1) Savings	(2) Loans	(3) Durable Goods	(4) Total Investment	(5) Monthly Income	(6) Educ. Exp.	(7) Health Exp.
Envelopes	29.67***	16.27***	-11.15	-24.78	-0.59	-8.57	16.09
	(5.25)	(5.08)	(53.85)	(30.70)	(3.38)	(9.51)	(14.58)
Sharp. q-val	0.001	0.005	1.000	0.724	1.000	0.724	0.724
Control Group Mean	41.44	22.97	392.95	244.60	35.15	154.54	149.81
Control Group S.D.	56.88	47.10	694.58	428.83	44.99	152.88	237.07
N	810	810	810	810	810	810	810

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Savings and Loans are total savings and loans held by the household. Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Productive investments refer to investments in a pre-specified list of productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Monthly income is self-reported monthly income. Education-related expenses refer to expenses in pre-specified list of education-related goods, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported school fees paid. Health-related expenses refer to expenses in pre-specified list of preventive health-related items, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported medicine expenses. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

savings as collateral, instead preferring WFP food aid, and focus group discussions conducted seven months after the endline survey indicated that households did not use savings as collateral.

[25] Cumulative investments over the 18 months between the baseline and endline were 19.36% higher among treatment households ($p = 0.09$).

3.3 Decomposing MA vs. MAD

The analysis thus far has compared the mental accounting intervention—using the four labeled envelopes—to the control group. However, the treatment effects may vary depending on whether participants were shown the default allocation. On one hand, the default allocation can offer households valuable guidance on how to manage the large cash transfer. On the other hand, since the default is not personalized, it may not align with each household's unique needs and circumstances. To examine the differential effects of providing the default allocation recommendation, we decompose treatment effects reported in Tables 2 and 3 separately for ***MA*** and ***MAD***.

3.3.1 MA vs MAD: One Year Post-Cash Transfers

Table 4 reports treatment effects for ***MA*** and ***MAD*** one year after the end of the cash transfer program. Column (1) shows that households in ***MAD*** invested significantly more compared to those in the control group (by 40.2%, 0.36 standard deviations). While ***MA*** invested less than ***MAD***, the difference is not statistically significant ($p = 0.31$). In Column (2), both ***MA*** and ***MAD*** households spent significantly more on lumpy investments than households in the control group. While ***MAD*** households spend more on lumpy investments compared with ***MA*** households, this difference is again not statistically significant ($p = 0.56$).

Contrary to the investment patterns, Columns (3) and (4) show that households in ***MA*** earn a higher monthly income, and have greater savings than households both in the control group and ***MAD*** treatment. Despite larger investments, ***MAD*** households do not have higher income or savings than households in the control group. Section 5 discusses possible explanations, with suggestive evidence indicating that while ***MAD*** households primarily invested in more common income-generating activities such as agriculture and livestock, ***MA*** households are more likely to diversify by starting businesses.

Columns (5)-(8) show little variation in treatment effects between ***MA*** and ***MAD*** households with respect to outstanding debt, expenditure on durable goods, and education- and health-related expenses.

Table 4: MA vs MAD: Endline Outcomes (USD PPP)

	(1) Total Investment	(2) Lumpy Investment	(3) Monthly Income	(4) Savings	(5) Loans	(6) Durable Goods	(7) Educ. Exp.	(8) Health Exp.
MA	25.11	13.78*	7.71**	15.66**	-6.22	-3.36	-13.34	-47.01**
	(36.97)	(7.78)	(3.02)	(6.32)	(4.79)	(4.88)	(17.99)	(21.98)
MAD	94.54*	18.90**	1.59	1.92	-4.09	-7.66	-19.33	-47.19**
	(51.91)	(7.59)	(2.74)	(5.51)	(4.78)	(4.97)	(18.48)	(22.29)
Sharp. q-val MA	0.474	0.126	0.051	0.051	0.287	0.474	0.474	0.101
Sharp. q-val MAD	0.167	0.074	0.335	0.513	0.513	0.293	0.484	0.154
Control Group Mean	235.12	50.61	24.94	38.63	39.69	33.21	250.56	330.31
Control Group S.D.	264.59	82.60	29.28	62.07	68.81	59.08	230.35	291.17
t-test MA vs. MAD	0.31	0.56	0.11	0.07	0.61	1.00	0.89	0.98
F-test	0.16	0.03	0.04	0.03	0.43	0.29	0.55	0.05
N	737	737	737	737	737	737	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Productive investments refer to investments available in local markets, with respondents indicating how many of them they had purchased. The quantities are multiplied by the median market price from three vendors in local markets. Lumpy investments are self-reported large investments made in the previous year in physical and human capital. Monthly income is self-reported monthly income, while Savings and Loans are total savings and loans held by the household. Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Education-related expenses refer to expenses in pre-specified list of education-related goods, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported school fees paid. Health-related expenses refer to expenses in pre-specified list of preventive health-related items, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported medicine expenses. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

3.3.2 MA vs MAD Immediately Post-Cash Transfers

Table 5 reproduces Table 3, decomposing the combined treatment into ***MA*** and ***MAD***. While households in both treatments have substantially larger savings at the end of the cash transfer program (Column (1)), ***MAD*** households additionally take out larger loans, equaling 120.6% of the outstanding loan value of households in the control group (0.59 standard deviations, Column (2)). The larger loans taken out by ***MAD*** households are reflected in the subsequent larger investments made by ***MAD*** households in between midline and endline (see Table 4). These larger loans are paid back, as the treatment effect of ***MAD*** on the value of outstanding loans in Table 4 (Column (5)) is negative and statistically insignificant.

Columns (3) and (4) of Table 5 indicate that ***MAD*** households spend less on durable goods and invested less than households in the control group, while ***MA*** households spend significantly more on durable goods than ***MAD*** households ($p = 0.02$). However, both measures are very noisy. Columns (5)-(7) indicate that the treatment effects on monthly income, educational expenses, and health-related expenses are statistically indistinguishable between households in the three experimental arms.

Table 5: MA vs MAD: Midline Outcomes (USD PPP)

	(1) Savings	(2) Loans	(3) Durable Good	(4) Total Investment	(5) Monthly Income	(6) Educ. Exp.	(7) Health Exp.
MA	27.23***	5.82	57.46	-10.43	-1.40	-3.86	23.42
	(6.08)	(4.52)	(67.97)	(37.62)	(3.70)	(11.10)	(17.11)
MAD	32.34***	27.71***	-86.21	-40.45	0.31	-13.71	8.06
	(6.93)	(8.20)	(55.88)	(29.95)	(4.23)	(10.76)	(17.54)
Sharp. q-val MA	0.001	0.656	0.859	0.859	0.859	0.859	0.656
Sharp. q-val MAD	0.001	0.003	0.255	0.255	0.398	0.255	0.398
Control Group Mean	41.44	22.97	392.95	244.60	35.15	154.54	149.81
Control Group S.D.	56.88	47.10	694.58	428.83	44.99	152.88	237.07
t-test MA vs. MAD	0.45	0.02	0.06	0.28	0.86	0.32	0.58
F-test	0.00	0.00	0.06	0.31	0.90	0.42	0.39
N	810	810	810	810	810	810	810

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Savings and Loans are total savings and loans held by the household. Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Productive investments refer to investments in a pre-specified list of productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Monthly income is self-reported monthly income. Education-related expenses refer to expenses in pre-specified list of education-related goods, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported school fees paid. Health-related expenses refer to expenses in pre-specified list of preventive health-related items, multiplied by the median market price obtained from three vendors in the refugee settlements, and added to self-reported medicine expenses. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

3.4 Other Outcomes

While the focus is on investment, savings, and income, further outcome variables were pre-registered. Online Appendix C documents negative but statistically insignificant treatment effects on total monthly consumption and consumption patterns at midline and endline for treatment households, in contrast to Lee et al. (2024), who show that earmarking reduces the willingness to spend/consume among rural residents. Treatment effects on outcomes related to mental health, education, health, and welfare-related outcomes (such as self-reliance and subjective well-being) are reported in Online Appendix D, again with statistically insignificant effects. Given the intervention was incorporated within a large cash transfer program (the cash transfer value exceeded monthly baseline income for 85% of households) and cost less than 0.46% of the cash transfer value, it is not surprising that no effects are found on downstream outcome variables a year after the end of the intervention.

In summary, the results indicate that offering households the opportunity to divide

their monthly cash transfer across four labeled envelopes increased savings and loans during the cash transfer period. These, in turn, were used to finance larger investments after the cash transfer ended. These larger investments resulted in a higher monthly income and greater savings. Although the effects on income and savings are primarily driven by households in the ***MA*** treatment, households in ***MAD*** took out larger loans and spent more on investments after the end of the cash transfer program. In the next section, we further explore the heterogeneity of these results, while Section 5 discusses potential mechanisms.

4 Heterogeneity

We pre-registered several dimensions along which we expected heterogeneous treatment effects, including baseline levels of self-control, the gender of the household head, and mental health, among others. However, we do not observe consistent heterogeneity across these variables.[26] The only dimension along which we observe heterogeneity is one we have not prespecified, as we had not anticipated that the intervention would remain so popular after the end of the cash transfer program ended.

4.1 Persistent Envelope Users

Although the final set of four envelopes was distributed with the last cash transfer, 35% of households that opted in for the four envelopes at baseline were still using them one year after the intervention ended. We refer to these households as *Persistent* users. Compared to households that also opted in for the four envelopes at baseline but discontinued their use after the cash transfer program ended (*Non-Persistent*), the heads of *Persistent* households tend to be younger, have arrived in Uganda six months later on average, have greater outstanding loans at baseline, and exhibit a stronger desire for self-sufficiency (see Appendix Table 6). At endline, the proportion of *Persistent* users did not differ significantly between ***MA*** and ***MAD*** ($p = 0.23$), and *Persistent* users were not statistically more or less likely than *Non-Persistent* users to accept the default allocation in the ***MAD*** treatment ($p = 0.3532$).

An important question is whether the four-labeled envelopes benefited *Persistent* users more than *Non-Persistent* users. Clearly, being a *Persistent* user is not randomly assigned —households self-select into persistence, which can be correlated with

[26] Results are available from the authors upon request.

unobserved characteristics that can explain savings and investment decisions (e.g., motivation, financial discipline). To address this concern, we use treatment assignment as an instrument and isolate variation in persistence that is exogenously determined by random assignment, mitigating concerns about confounding factors.[27] Specifically, we estimate a Two-Stage Least Squares (2SLS) Instrumental Variable regression, using treatment assignment as an instrument for being a *Persistent* user. The model includes the same set of controls as Equation 1. Appendix Table 7 presents Treatment-on-the-Treated (TOT) regression results at endline, showing that the positive treatment effects on total and lumpy investments, income, and savings (as documented in Table 2) are primarily driven by *Persistent* users. Online Appendix Table 41 reports 2SLS regression results for midline data, confirming the same pattern: the positive treatment effects documented in Table 3 are driven by *Persistent* users.

The validity of the 2SLS results relies on the exclusion restriction, which assumes that treatment assignment affects the outcome variables only through its effect on persistence and not through any other direct channel. If treatment assignment independently influences financial behavior, this assumption would be violated, and the IV estimates could be biased. While we cannot directly test the exclusion restriction, we address concerns about selection bias by performing Propensity Score Matching (PSM) using control variables selected via a LASSO Machine Learning algorithm. PSM ensures that *Persistent* and *Non-Persistent* users are comparable in terms of observable characteristics, reducing concerns that baseline differences drive our results. Although PSM does not address the exclusion restriction, the fact that results remain qualitatively similar across both methods strengthens the robustness of our findings (see Online Appendix Tables 42 and 43).

5 Mechanisms

In this section, we discuss potential mechanisms that could drive the treatment effects on savings and investments documented in Sections 3 and 4. Overall, we find suggestive evidence that the four envelopes intervention functioned as a commitment device,

[27] Being assigned to treatment increases the probability of becoming a *Persistent* user of the four-labelled envelopes because households in the treatment group were directly offered to use the envelopes. Households in the control group did not receive the intervention, making it less likely that they would be able to use the four envelopes at endline. No households in the control group used the four-labeled envelopes at midline or endline.

helping households adhere to their savings and investment plans. The intervention helped households accumulate savings by facilitating ex-ante planning and reinforcing commitment through the partitioning and labeling of the cash transfers. Additionally, it influenced investments decisions by enabling households to delay investments, thereby facilitating those with higher upfront costs. In what follows, we present evidence supporting these mechanisms and rule out alternative explanations.

5.1 Accumulation of Savings

The treatment effects presented in Section 3 indicate substantial positive impacts on accumulated savings during the cash transfer period, as well as sustained savings one year later. These effects are particularly pronounced among recipients who continued using the four labeled envelopes even after the cash transfer program ended. In the following sub-sections, we examine which aspects of the intervention contributed to the increase in savings.

Planning

An essential requirement for accumulating savings is having a plan for their future use. This appears to be a key feature of our intervention, as reported by recipients themselves. At baseline, household heads who opted into the four-labeled envelopes were asked an open-ended question about what they considered to be the most valuable aspect of the envelopes. Their responses were categorized into four main themes: planning and committing, safety, resisting temptations, and savings (or other). As shown in Appendix Table 8, *Persistent* users - who benefit the most from the intervention - were significantly more likely than *Non-Persistent* users to report that the envelopes would help them with planning and committing to their financial goals. However, they were no more likely than others to cite safety, resisting temptation, or increasing savings as key benefits.

Interestingly, when asked immediately after the end of the cash transfer program (at midline), *Persistent* users were no more likely than *Non-Persistent* users to state that the best part of the four-labeled envelopes was planning (see Online Appendix Table 29). This is because the share of *Non-Persistent* users who stated that the four-labeled envelopes helped with planning increased compared to at baseline, suggesting that they may have learned that envelopes could help with planning throughout the

duration of the cash transfer.[28] Nevertheless, *Persistent* users more quickly recognized how the intervention could help them plan for future expenditures.

Partitioning and Labeling

A key feature of the four-labeled envelopes is the ability to partition the cash transfer into four designated accounts and discipline spending. We find suggestive evidence that this mechanism was at play. During the midline survey, households who opted into the four-labeled envelopes were asked about their perceptions of the impact of the envelopes on their decision-making. Appendix Table 9 presents OLS regression results comparing *Persistent* envelope users to *Non-Persistent* users. While the intervention did not differentially affect household heads' perceptions of their ability to reject remittance requests, resist temptation, or save for schooling and health emergencies (Columns (3)-(6)), *Persistent* envelope users were more likely to agree that dividing the cash transfer into multiple accounts, and labeling these accounts for specific categories helped discipline their spending (Columns (1)-(2)). The OLS regression does not allow for a causal interpretation, however it provides suggestive evidence that the users that benefited most from the intervention valued the division and labeling more.

Combining these insights suggests that those who benefited most from the intervention used the commitment device to plan their future expenditures. The partitioning and labeling of the cash transfer served as a monthly reminder, helping them discipline their spending in line with their initial plans. This aligns closely with the principles of mental accounting (Thaler, 1985) and supports our pre-specified hypothesis.

5.2 Investment Decisions

Tables 2 and 3 highlight differences in investment decisions between treated and control households. While control households tend to make more investments when the cash transfer is disbursed - albeit not statistically significantly - treatment households delay their investments until after the cash transfer ends, at which point they make larger, lumpier investments. Additionally, we find that the treatment postpones investments in enterprises and leads to shifts in the types of investments made.

[28] 39% of those who were offered the intervention at baseline but turned it down, regretted doing so at midline, with 64% stating it would have helped them with their planning and budgeting.

Lumpy Investments

The four-envelopes intervention may have played a crucial role in facilitating lumpier investments by allowing households to accumulate larger sums within their *Investment* envelope. Given that the cash transfer amounted to $23 PPP per household member per month, larger households naturally received a greater total transfer. If all households allocated the same proportion of their transfer to the *Investment* envelope, larger households would have had more funds available for investment than smaller households.[29]

Accumulating more money in the *Investment* envelope not only enables larger investments but is particularly important for lumpier investments, which require a substantial upfront payment. Appendix Table 10 provides empirical support for this mechanism, showing heterogeneous treatment effects on lumpy investment based on a median split of household size. The results indicate that while the intervention did not lead to greater total investments among larger households, it did result in more lumpier investments — likely because these households had more funds set aside in their *Investment* envelope.

These findings provide suggestive evidence that the four-labeled envelopes helped facilitate lumpier investments for households with sufficient funds, without necessarily increasing total investments.

Timing and Type of Investments

At baseline, most households derive their income from agriculture or livestock, consistent with the broader economic landscape of Uganda's refugee settlements (UNHCR, 2024). Investing in agriculture or livestock represents a continuation of the status quo, whereas investing in businesses—such as market stalls, kiosks, or restaurants—diversifies income sources. Unlike agriculture and livestock, which are subject to seasonal fluctuations, businesses provide a more stable income stream and are less vulnerable to climate shocks, such as droughts.

Appendix Table 11 decomposes investments across agriculture, livestock, and enterprises, reporting average treatment effects at midline and endline. During the cash transfer period, treatment households invested more in agriculture but less in enterprises. However, one year after the cash transfer ended, this pattern reversed: treat-

[29] Households above and below the median size allocated equal proportions of their cash transfer to the *Investment* envelope (30.9% vs. 30.8%, $p = 0.89$).

ment households made significantly larger investments in enterprises, more than doubling the enterprise investments of the control group (115% increase).

By combining midline and endline investment flows, we estimate the total value of the investment activities since the start of the cash transfer program. Panel C shows that the value of cumulative investments on business creation was 49% larger in the treatment group compared to the control group. Online Appendix Table 40 further decomposes these results across ***MA*** vs. ***MAD*** revealing distinct investment patterns: ***MA*** households invested more in enterprises, while ***MAD*** households invested more in livestock. This difference in investment choices may explain why positive effects on endline savings and monthly income are observed for ***MA*** households but not for ***MAD***.

Furthermore, examining the types of enterprises established one year after the cash transfer ended reveals notable differences. Although the number of observations is small—limiting statistical significance—Appendix Table 12 provides suggestive evidence that treatment households shifted away from food stalls (e.g., selling breakfast or snacks, Column (2)) and toward market stalls (Column (5)). Market stalls typically require larger upfront investments, including supplies and rent, whereas food stalls are generally informal and require minimal startup costs. This pattern suggests that the treatment enabled households to invest in larger, more capital-intensive enterprises, potentially offering greater long-term revenue opportunities

5.3 Loan Purpose

Table 3 shows that, compared to control households, treatment households had significantly more loans immediately after the cash transfer ended. But what were these loans used for? To explore this (and other) question(s), we conducted focus group discussions with households in ***CO***, ***MA***, and ***MAD*** treatment arms seven months after the endline survey. These discussions revealed distinct motivations for taking out loans across groups. While households in ***CO*** primarily used loans to cope with shocks and smooth consumption, households that received the four-labeled envelopes took out loans for investments instead. This suggests that the intervention shifted the purpose of borrowing - from merely bridging short-term budget gaps to financing forward-looking capital investments. Additionally, focus group discussions indicated that treatment households often placed borrowed funds inside the corresponding envelope category, effectively labeling the loan's purpose and reinforcing disciplined financial planning.

Moreover, loan patterns differed substantially between ***MA*** and ***MAD*** households, which might explain their investment choices. Appendix Table 5 shows that, in addition to accumulating savings, households in the ***MAD*** treatment took out significantly more loans than those in the ***MA*** group ($p = 0.02$).[30] Focus Group Discussions revealed that interest rates on loans in Rhino Camp and Imvepi are extremely high, ranging from 10–20% per month. Given these high borrowing costs, ***MAD*** households probably needed to invest in activities such as livestock ($p = 0.01$, Appendix Table 11, Panel B, Column 2) that generated immediate returns (e.g. milk, eggs) and could be sold when debt payments are due.

Why did ***MAD*** households take out larger loans? Focus Group Discussions provided some potential answer to this question: ***MAD*** households mentioned that they frequently moved money from the *Investment* envelopes to the *Education* envelope, as the amount was insufficient for those with secondary school children. This is because the fraction of the cash transfer allocated to each envelope category in ***MAD*** did not differ by family composition, while school fees are substantially higher for secondary school children.[31] As a conscequence, ***MAD*** households subsequently supplemented the investment-labeled savings with loans for investment purposes, in order to have sufficient liquidity for their investments. This highlights the strength of the *Investment* envelope, as households took out loans (with high interest rates) in order to make the investments they had initially planned. This may also explain why ***MAD*** households were less likely to accumulate durable goods, such as furniture, during the cash transfer program ($p = 0.06$, see Appendix Table 5, Column 3). Unlike the investments made by ***MAD*** households, durable goods are likely to be more easily liquidated in markets if needed, providing a form of financial flexibility. Moreover, Column 5 of Appendix Table 12 provides suggestive evidence ($p = 0.11$) that ***MA*** households invested more in (lumpy) enterprises. These enterprises do not offer immediate returns, making them less attractive to ***MAD*** households that need to repay loans, while ***MA*** households can use their savings and liquidate durable goods until the lumpy investment starts earning income.

[30] Collins et al. (2009) documented that nearly every household in their survey held both low-interest savings and high-interest loans simultaneously, highlighting the prevalence of this practice.

[31] This highlights a limitation of the ***MAD*** intervention, as the amount of money allocated to the *Education* envelope depended on the family size, but not its composition. Primary school fees are \$1.53 PPP per academic term, with three terms per year. Furthermore, scholastic materials cost \$11.45 PPP per term, per child. For secondary schools, the respective costs are \$38.15 PPP and \$91.56 PPP.

5.4 Different Strength of Commitment

While the nature of the commitment device - the four-labeled envelopes - was identical for households in ***MA*** and ***MAD***, the strength of the commitment may have differed. In the ***MAD*** treatment, the allocation was recommended by humanitarian NGOs, which could have reinforced adherence to the suggested budgeting framework. In contrast, this additional external potential influence was not present in the ***MA*** treatment. But was this consequential?

Appendix Table 13 compares ***MAD*** to ***MA*** households on responses regarding how they used the four envelopes during the cash transfer program. Column (1) shows that nearly all households moved money across envelopes and used funds for unrelated expenses (91% in ***MA***), highlighting the flexibility of the intervention, compared to harder commitments. Despite this, most households did not express a desire to change their allocation across categories, with no significant difference between ***MA*** and ***MAD*** (Column (2)). This finding is somewhat surprising given the differing degrees of personalization in the two treatments, suggesting that even when allocation recommendations were externally set, households largely accepted and adhered to them.

Column (3) illustrates that while households spent money outside of their intended categories, many made the commitment device *harder* by sealing the envelopes. Envelopes were not sealed once the money was placed inside them at the cash distribution point, and hence the household sealing the envelope reveals a preference for a harder commitment. Columns (4)-(10) provide further insights into the mechanisms through which the commitment device had differential effects on ***MA*** and ***MAD*** households, with the latter self-reporting improvements in their ability to avoid unnecessary expenditures, save for schooling, and for health needs.

5.5 Investments vs. Education vs. Health

The three labeled categories —education, health, and investment—— were selected based on Focus Group Discussions conducted before the start of the RCT. These discussions identified them as the most important yet unmet needs of refugee households. The three categories differ along several key dimensions, including their degree of lumpiness, certainty, and flexibility, which may influence the effectiveness of commitment devices. In contrast to education and health-related expenses, investments are more

flexible and the immediate consequences of inaction (in the form of not spending on investments) are less noticeable. These insights are based on Focus Group Discussions after the endline survey, however remain suggestive.

School-related costs, particularly for primary school children, are relatively low but must be paid each academic term by a specific deadline. This deadline is communicated in advance, meaning that household heads have certainty about when the fees are due. Moreover, there is no flexibility in this expense — if school fees are not paid, the child cannot attend school. As a result, this type of expenditure is well-suited for structured savings plans, as households know the exact amount required, the due date, and the consequences of non-payment.

Health shocks, in contrast, can be very expensive and unpredictable, as households do not know when these costs will arise. Focus Group Discussions conducted seven months after the endline survey revealed that households typically take out loans in response to health emergencies rather than relying on savings. This may explain why no significant treatment effects are found on health-related outcomes or on the likelihood of households using savings to cope with adverse shocks. Similarly, preventive health investments, such as latrine upgrades, are both lumpy and highly flexible — if the upgrade is not made this month, it can still be done next month. In the absence of dedicated savings, households are likely to procrastinate these investments until they become urgent. This pattern is reflected in the results from Section 3, which show that households in the treatment arms upgraded their latrines earlier than those in the control group.

Investments vary in their degree of lumpiness. While small-scale agriculture and poultry farming require minimal upfront costs, enterprises and larger livestock purchases demand substantial lump-sum investments. Investments are inherently uncertain, particularly in agriculture and livestock, which are weather and season dependent. However, they also offer high flexibility — households can adjust their investment plans based on the capital available. For instance, a household can buy fewer poultry than originally planned or rent a smaller plot of land for cultivation if funds are limited. The *Investment* envelope serves as a tool to help households save for large, lumpy investments while also increasing overall investments (as documented in Table 2). By reinforcing commitment to initial investment plans, it helps households allocate resources more effectively.

The different types of expenses — in particular their lumpiness, certainty, and

flexibility — can affect the effectiveness of the commitment devices, with Tables 2 and 3 suggesting that the labeled envelopes were particularly effective for investments. This remains an interesting area for future research, and can help explain why some studies find positive effects of commitment devices, while other studies find null or negative effects (John, 2019; Bai et al., 2021).

5.6 Alternative Explanations

We rule out several alternative explanations, with the supporting regression analyses documented in Online Appendix E. First, the treatment effects are not driven by greater self-control, as we find no improvement in the household head's self-control at either midline or endline. Additionally, we do not observe heterogeneous treatment effects based on baseline self-control levels (Online Appendix Table 30). Second, focusing on the self-control index components related to the ability to resist temptations, we find no evidence that the mental accounting intervention improved the household head's ability to resist temptations (Online Appendix Table 33). Third, we find no statistically significant treatment effects on the experience of theft incidents at midline or endline, despite 6% of households citing safety concerns as a motivation for adopting the four-labeled envelopes (**Online Appendix Table 34).[32]

In addition to influencing *how much* households save, the intervention could also affect *where* they save. One concern was that the four-labeled envelopes might discourage households from adopting more advanced savings mechanisms, such as mobile money or savings groups.[33] However, we find no evidence of this effect, with precisely estimated null treatment effects (**Online Appendix Table 35).

Another concern is that respondents decided their allocation in front of the enumerator, potentially making the commitment public rather than private. While this does not affect the treatment effect estimates, it could influence the underlying mechanisms (Bryan et al., 2010). However, we can rule out this explanation. Treatment effect estimates remain unchanged when including enumerator fixed effects at baseline. Additionally, enumerators were effectively randomly assigned, with only 2.79% of endline surveys conducted by the same enumerator-respondent pair as at baseline.

[32]The underlying logic behind potential safety effects is that if a cash transfer is divided across multiple envelopes, the money can be hidden in different places. Thus, if a thief steals one envelope, they do not take the entire amount.

[33]Bank penetration is extremely low in this context; no households had a bank account at baseline, midline, or endline.

Furthermore, treatment effects do not differ based on whether the midline or endline enumerator also conducted the baseline survey (results available upon request).

In addition, We find no evidence that the information provided by the default treatment benefits households. If the default allocation were useful, we would expect differential treatment effects within the ***MA*** group, depending on how far a household's allocation was from the default. Households closer to the default would learn less new information than those further away. However, we do not find support for this hypothesis (Online Appendix Tables 31 and 32).

Nevertheless, the documented differential treatment effects between ***MA*** and ***MAD*** allow us to rule out habit formation as a mechanism. If the intervention were only effective through habit formation, the specific allocation to each account would not matter. Furthermore, at midline, 15.5% of households who received the four-labeled envelopes reported that they would have liked to change their allocation during the cash transfer program — again, inconsistent with habit formation as the primary mechanism.

One of our initial hypotheses was that the commitment device, particularly ***MAD***, would help households turn down remittance requests from family, friends, and neighbors. Alternatively, De Arcangelis et al. (2015) and Batista et al. (2015) suggest that migrants are more likely to give remittances when the intended use becomes more binding. However, we find no evidence supporting either explanation, as remittances received and given remain unchanged at both midline and endline (Online Appendix Tables 38 and 39).

Lastly, we rule out experimenter demand effects in different ways. First, following Dhar et al. (2022), we measure social desirability bias and find no heterogeneous treatment effects (Online Appendix Tables 32 and 31). Second, treatment effects on education- and health-related expenses at endline are negative, contrary to what would be expected if experimenter demand effects were driving the results. Additionally, we find no treatment effect on households' self-reported expectations about receiving an additional cash transfer from DRC in the future (Online Appendix Table 37). Third, total investments and lumpy investments were measured using two different methodologies, providing further reassurance that treatment effects reflect truthful reporting rather than experimenter demand.

A related concern is that households may have feared retribution from the NGO—such as having future assistance withheld—if they did not allocate funds according to the envelopes' intended categories. However, focus group discussions 1.5 years after the

intervention indicate that this was not a concern, as households did not perceive any risk of retribution. This aligns with the emphasis placed during baseline surveys that the research was conducted independently of the NGO. Moreover, 40.88% of households who opted into the four-labeled envelopes reported moving money across envelopes, further suggesting that they did not fear noncompliance would affect future aid.[34]

As pre-registered, we perform robustness checks by winsorizing at the 5% level, and not winsorizing at all. Furthermore, we winsorized separately per treatment, and also winsorize the whole sample, as discussed by Wicker (2025). Finally, we use the PD-LASSO algorithm of Belloni et al. (2014) to select control variables, and randomized inference. Results are robust to all specification, and are available upon request from the authors.

6 Cost-Effectiveness

The total costs of the four labeled envelopes per household were $1.70 ($4.64 PPP, see Appendix F for detailed calculations). One year after the end of the cash transfer program, the intervention increased monthly income by 0.16 standard deviations. Similarly, it increased savings by 0.14 standard deviations. This equals a 0.09 and 0.08 standard deviation increase in monthly income and savings per dollar spent.

As a reference point, Aggarwal et al. (2023) offered multiple lockboxes (with locks and keys) to micro-entrepreneurs in urban Malawi as a form of commitment device. The intervention led to a 0.21-0.27 standard deviation increase in savings. As the lockboxes on average cost $9.50 per micro-entrepreneur, this results in a 0.02-0.03 standard deviation increase in savings per dollar spent. Therefore, our intervention is highly cost-effective, despite being a much softer commitment device, far cheaper, and easy to integrate within an NGO's ongoing operations.

Furthermore, our estimated treatment effects are likely to be a lower bound: eligibility for humanitarian aid is based on vulnerability: the more vulnerable a household, the more likely they are to receive humanitarian assistance. Given this, causal estimations of the treatment effects of successful humanitarian interventions are likely to

[34]During the baseline survey, we emphasized that households were free to spend the money as they wished ("As before, you are free to spend your money as you want") and that their individual data would not be shared with NGOs ("Your answers will be used for research purposes only and will be confidential. Your name and answers will not be given to anyone other than the researchers"). If household heads had feared that noncompliance would reduce their likelihood of receiving additional humanitarian assistance, this number should have been 0%.

understate the true treatment effect, as these households become less vulnerable and will thus be less likely to receive further assistance from other NGOs. This is less likely to be a concern in development settings (where the majority of behavioral interventions and cash transfer programs are evaluated), as less assistance is available, and refugees are typically more dependent on additional assistance due to their greater vulnerability. We therefore likely underestimate the true treatment effect, as suggested by the evidence we provide that treatment households are less likely to receive an additional conditional cash transfer earmarked for education - indicating they are also perceived as less vulnerable by independent NGO metrics.

7 Conclusion

We conducted a Randomized Controlled Trial among 861 refugee households who were eligible for a seven-month-long unconditional cash transfer equaling $22.89 PPP per household member per month. Households were randomized across a control group and two treatment arms, who were both offered the chance to receive their monthly cash transfer divided across four envelopes labeled "Education", "Health", "Investments", and "Others" - rather than the NGO's status quo of placing the money inside one unlabeled envelope.

Demand for the intervention was high, with 93% of households opting for the labeled envelopes. Compared with households in the control group, households in the treatment groups have larger savings and outstanding loans shortly after the cash transfer ends. These savings and loans are spent on (lumpy) productive investments, which result in a higher monthly income one year later. Furthermore, treatment households have larger savings, and no longer have larger outstanding loans compared with households in the control group.

The positive treatment effects are primarily driven by those households who still use the envelopes one year after the end of the intervention. Compared with households who stopped using the envelopes after the end of the intervention, these households are more likely to believe that the labeled envelopes will help them with their planning and committing to their plans, and after the end of the intervention were more likely to agree that partitioning the cash transfer, and having labels, helped discipline spending. The limited effectiveness of the intervention on non-persistent users is in line with John (2019), who argue that the limited effectiveness of commitment devices can be due to

individual's inability to follow through with their initial commitment. An alternative explanation is that the commitment device was too weak to effectively overcome self-control problems (Kremer et al., 2019).

At a cost of merely $1.70 per recipient, the intervention represents an additional cost of 0.46% on top of the total cash transfer value. The documented positive treatment effects result in an increase savings of 0.08 standard deviations per dollar spent. As such, the intervention is highly cost-effective, and easy to implement at scale. The treatment effects suggest that the "Investment" envelope was more effective than the "Education" and "Health" envelopes. As such, future programs may consider to only have an investment-related commitment device, which further reduces costs. This presents an interesting avenue for future research and policy, for both development and humanitarian actors.

This study demonstrates that the effectiveness of cash transfers can be improved by incorporating low-cost interventions grounded in behavioral economics. Our results suggest that commitment constraints that prevent lumpy, productive investments can be addressed through the use of commitment devices grounded in mental accounting. As such, this paper provides a further motivation why households under-invest in high-return investments (Kremer et al., 2019), but open questions remain whether these findings will replicate in other contexts, when integrated within digital payments, or when incorporated within different cash transfer payment structures, for example, lump sum transfers.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2022). When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Aggarwal, S., Brailovskaya, V., and Robinson, J. (2023). Saving for Multiple Financial Needs: Evidence from Lockboxes and Mobile Money in Malawi. *The Review of Economics and Statistics*, pages 1–19.

Ahmed, A., Hidrobo, M., Hoddinott, J., Kolt, B., Roy, S., and Tauseef, S. (2025). Sustainable poverty reduction through social assistance: Modality, context, and complementary programming in bangladesh. *American Economic Journal: Applied Economics*. Forthcoming.

Aker, J. C. (2017). Comparing cash and voucher transfers in a humanitarian context: Evidence from the democratic republic of congo. *The World Bank Economic Review*, 31(1):44–70.

Altındağ, O. and O'Connell, S. D. (2023). The short-lived effects of unconditional cash transfers to refugees. *Journal of Development Economics*, 160(C).

Anderson, M. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103:1481–1495.

Angelucci, M. and Bennett, D. (2024). The economic impact of depression treatment in india: Evidence from community-based provision of pharmacotherapy. *American Economic Review*, 114(1):169–98.

Ashraf, N., Karlan, D., and Yin, W. (2006). Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines. *The Quarterly Journal of Economics*, 121(2):635–672.

Augenblick, N., Jack, K., Masiye, F., Swanson, N., and Kaur, S. (2024). Retrieval failures and consumption smoothing: A field experiment on seasonal poverty. Revise and resubmit, Quarterly Journal of Economics.

Bai, L., Handel, B., Miguel, E., and Rao, G. (2021). Self-control and demand for preventive health: Evidence from hypertension in india. *The Review of Economics and Statistics*, 103(5):835–856.

Banerjee, A., Claudia, M. A., and Puentes, E. (2025). Better strategies for saving more: Evidence from three interventions in chile. *Journal of Development Economics*, 173:103405.

Banerjee, A. and Duflo, E. (2007). The economic lives of the poor. *Journal of Economic Perspectives*, 21(1):141–168.

Banerjee, A., Faye, M., Krueger, A., Niehaus, P., and Suri, T. (2023). The (many) million dollar question: Experimental evidence from kenya on universal basic income. Working Paper.

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., and Pellerano, L. (2016). Cash transfers: What does the evidence say? a rigorous review of programme impact and of the role of design and implementation features. Technical report, World Bank.

Batista, C., Silverman, D., and Yang, D. (2015). Directed giving: Evidence from an inter-household transfer experiment. *Journal of Economic Behavior & Organization*, 118:2–21. Economic Experiments in Developing Countries.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50.

Bernard, T. and Taffesse, A. (2014). Aspirations: An approach to measurement with validation using ethiopian data. *Journal of African Economies*, 23:189–224.

Bernheim, B. D., Ray, D., and Yeltekin, S. (2015). Poverty and self-control. *Econometrica*, 83(5):1877–1911.

Blattman, C., Jamison, J. C., and Sheridan, M. (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *American Economic Review*, 107(4):1165–1206.

Bossuroy, T., Goldstein, M., Karimou, B., Karlan, D., Kazianga, H., Parienté, W., Premand, P., Thomas, C. C., Udry, C., Vaillant, J., and Wright, K. A. (2022). Tackling psychosocial and capital constraints to alleviate poverty. *Nature*, 605(7909):291–297.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Brune, L., Chyn, E., and Kerwin, J. (2021). Pay me later: Savings constraints and the demand for deferred payments. *American Economic Review*, 111(7):2179–2212.

Brune, L., Giné, X., Goldberg, J., and Yang, D. (2017). Savings defaults and payment delays for cash transfers: Field experimental evidence from malawi. *Journal of Development Economics*, 129:1–13.

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

Collins, D., Morduch, J., Rutherford, S., and Ruthven, O. (2009). *Portfolios of the Poor: How the World's Poor Live on $2 a Day*. Princeton University Press.

Crosta, T., Karlan, D., Ong, F., Rüschenpöhler, J., and Udry, C. R. (2024). Unconditional cash transfers: A bayesian meta-analysis of randomized evaluations in low and middle income countries. Working Paper 32779, National Bureau of Economic Research.

De Arcangelis, G., Joxhe, M., McKenzie, D., Tiongson, E., and Yang, D. (2015). Directing remittances to education with soft and hard commitments: Evidence from a lab-in-the-field experiment and new product take-up among filipino migrants in rome. *Journal of Economic Behavior & Organization*, 111:197–208.

Development Initiatives (2023). Global humanitarian assistance report 2023. Technical report, Development Initiatives.

Dhar, D., Jain, T., and Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American Economic Review*, 112(3):899–927.

Dupas, P. and Robinson, J. (2013). Why don't the poor save more? evidence from health savings experiments. *American Economic Review*, 103(4):1138–71.

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., and Walker, M. (2022). General equilibrium effects of cash transfers: Experimental evidence from kenya. *Econometrica*, 90(6):2603–2643.

Ercolano, F. and Woolfenden, J. (2021). Endline evaluation of the dg echo-funded action for the uganda cash consortium *(Internal report, available upon request)*. Technical report, Action Against Hunger.

Haushofer, J., Mudida, R., and Shapiro, J. (2020). The comparative impact of cash transfers and a psychotherapy program on psychological and economic well-being. *SSRN Electronic Journal*.

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from kenya. *The Quarterly Journal of Economics*, 131(4):1973–2042.

Heath, C. and Soll, J. B. (1996). Mental budgeting and consumer decisions. *Journal of Consumer Research*, 23(1):40–52.

Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., and Moreira, V. (2014). Cash, food, or vouchers? evidence from a randomized experiment in northern ecuador. *Journal of Development Economics*, 107(C):144–156.

John, A. (2019). When commitment fails: Evidence from a field experiment. *Management Science*, 66(2):805–822.

Kaboski, J. P., Lipscomb, M., Midrigan, V., and Pelnik, C. (2024). How important are investment indivisibilities for development? experimental evidence from uganda. Revise and resubmit, Journal of Political Economy.

Kang, J. and Kang, M. (2022). Durable goods as commitment devices under quasi-hyperbolic discounting. *Journal of Mathematical Economics*, 99:102561.

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

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

Kim, J.-H. and Park, E. (2015). Mediating effect of self-control in relation to depression, stress, and activities of daily living in community residents with stroke. *Journal of Physical Therapy Science*, 27:2585 – 2589.

Kremer, M., Rao, G., and Schilbach, F. (2019). Chapter 5 - behavioral development economics. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics - Foundations and Applications 2*, volume 2 of *Handbook of Behavioral Economics: Applications and Foundations 1*, pages 345–458. North-Holland.

Laajaj, R. (2017). Endogenous time horizon and behavioral poverty trap: Theory and evidence from mozambique. *Journal of Development Economics*, 127(C):187–208.

Lee, J. N., Morduch, J., Ravindran, S., and Shonchoy, A. S. (2024). The social meaning of mobile money: Earmarking reduces the willingness to spend in migrant households. *Journal of Economic Behavior & Organization*, 221:675–688.

Maio, M. and Fiala, N. (2020). Be wary of those who ask: A randomized experiment on the size and determinants of the enumerator effect. *The World Bank Economic Review*, 34:654–669.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of Development Economics*, 99(2):210–221.

Orkin, K., Garlick, R., Mahmud, M., Sedlmayr, R., Haushofer, J., and Dercon, S. (2024). Aspiring to a better future: Can a simple psychological intervention reduce poverty? Working Paper.

Ozler, B., Celik, C., Cunningham, S., Cuevas, P. F., and Parisotto, L. (2021). Children on the move: Progressive redistribution of humanitarian cash transfers among refugees. *Journal of Development Economics*, 153:102733.

Pozuelo, J. R., Bradenbrink, R., Stierna, M. F., and Sterck, O. (2023). Depression, violence and socioeconomic outcomes among refugees in east africa: evidence from a multicountry representative survey. *BMJ Ment Health*, 26(1).

Prelec, D. and Herrnstein, R. J. (1991). Preferences or principles: Alternative guidelines for choice. In Zeckhauser, R. J., editor, *Strategy and Choice*. MIT Press, Cambridge, MA.

Raghubir, P. and Srivastava, J. (2009). The denomination effect. *Journal of Consumer Research*, 36(4):701–713.

Schilbach, F. (2019). Alcohol and self-control: A field experiment in india. *American Economic Review*, 109(4):1290–1322.

Soman, D. and Cheema, A. (2011). Earmarking and partitioning: Increasing saving by low-income households. *Journal of Marketing Research*, 48(SPL):S14–S22.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3):199–214.

Thaler, R. and Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1):S164–S187.

Thaler, R. and Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy*, 89(2):392–406.

UNHCR (2023). Global trends: Forced displacement in 2022. Technical report, United Nations High Commissioner on Refugees.

UNHCR (2024). Uganda - refugee statistics october 2024. Accessed: 2024-11-30.

UNHCR (2025). Protracted refugee situations explained. Accessed: 2025-01-03.

Wicker, T. (2025). Winsorizing and trimming with subgroups. Working paper, Tilburg University.

Wicker, T., Dalton, P., and van Soest, D. (2023). Pre-analysis plan: Helping cash transfer recipients prosper: Experimental evidence from a humanitarian setting. Working Paper, Tilburg University.

Appendix to:

Mental Accounting and Cash Transfers: Experimental Evidence from a Humanitarian Setting

by Till Wicker, Patricio Dalton, and Daan van Soest

A Additional Tables

Table 6: Baseline imbalances: Persistent vs. Non-Persistent Users

	(1) ***Persistent***		(2) ***Non-Persistent***		(3) Pairwise t-test
Variable	N	Mean/(SD)	N	Mean/(SD)	Difference
Loan Amount	170	12.85	316	8.57	0.043**
		(27.34)		(18.85)	
Share of (hyp.) CT invested	163	0.32	289	0.26	0.015**
		(0.27)		(0.26)	
HoHH Age	170	35.98	316	39.43	0.010***
		(12.79)		(14.63)	
Arrival Year	170	2018.37	316	2017.72	0.072*
		(3.79)		(3.79)	
Desire for Suff. Income	170	0.58	316	0.48	0.038**
		(0.50)		(0.50)	
Desires No Income Source	170	0.01	316	0.03	0.095*
		(0.08)		(0.17)	

Notes: Columns (1) and (2) show the average value (and standard deviation) for households that opted-in for the four labeled envelopes and are still using them at endline (Persistent), and households that opted-in for the four labeled envelopes and are not using them at endline anymore (Non-Persistent). The significance levels based on t-tests is reported in column (3). This table only reports variables with statistically significant differences between Persistent and Non-Persistent households. All other variables listed in Tables 16 and 17 are not statistically significantly different between Persistent and Non-Persistent households. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 7: 2SLS Regression: Persistent Envelope Users, Endline

	(1) Productive Investments	(2) Lumpy Investments	(3) Monthly Income	(4) Savings	(5) Loans	(6) Durable Goods
Persistent	165.97*	45.04***	12.70**	24.11*	-14.03	-24.16
	(87.76)	(16.89)	(6.30)	(13.37)	(11.24)	(16.04)
Sharp. q-val	0.099	0.050	0.099	0.099	0.121	0.121
Control Group Mean	235.12	50.61	24.94	38.63	39.69	33.21
Control Group S.D.	264.59	82.60	29.28	62.07	68.81	59.08
N	737	737	737	737	737	737

Notes: 2SLS IV estimates for 'Persistent'. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD. Persistent is coded as a dummy variable equal to one if the household is still using the envelopes at endline, and zero otherwise. Productive investments refer to investments in a pre-specified list of productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Lumpy investments are self-reported large investments made in the previous year in physical and human capital. Savings are total savings held by the household. Monthly income is self-reported monthly income, while Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 8: Balance Table for Best Aspect of Envelopes.

	(1)		(2)		(3)
	Persistent		***Not Persistent***		Pairwise t-test
Variable	N	Mean/(SD)	N	Mean/(SD)	Difference
Envelope Advantage: Planning	170	0.812	316	0.715	0.019**
		(0.392)		(0.452)	
Envelope Advantage: Safety	170	0.041	316	0.073	0.168
		(0.199)		(0.260)	
Envelope Advantage: Resist Temptation	170	0.065	316	0.063	0.952
		(0.247)		(0.244)	
Envelope Advantage: Savings	170	0.041	316	0.041	0.998
		(0.199)		(0.199)	

Notes: Columns (1) and (2) show the average value (and standard deviation) for households that opted-in for the four labeled envelopes and are still using them at endline (Persistent), and households that opted-in for the four labeled envelopes and are not using them at endline anymore (Not Persistent). The significance levels based on t-tests is reported in column (3). ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 9: Behavioral Responses to Labeled Envelopes.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dividing Cash	Labeling Categories	Using 4 labeled envelopes made it easier to:			
	Helped Discipline	Helped Discipline	Reject Requests	Avoid Temptation	Save for School	Save for Health
Persistent	0.14**	0.11*	0.11	0.06	0.05	0.08
	(0.07)	(0.07)	(0.08)	(0.07)	(0.06)	(0.06)
Control Group Mean	4.01	3.97	3.86	4.00	4.11	4.06
N	440	440	440	440	440	440

Notes: Intention to Treat estimates, based on PD Lasso regressions (Belloni et al., 2014). Persistent is a dummy variable equal to 1 if the household is still using the labeled envelope at endline. Respondents were asked to indicate the degree to which they agree with the following statements on a scale from 1 to 5: (1) "Dividing the money across four envelopes instead of one helped discipline spending and increase savings and investment"; (2) "Labeling the four envelopes for explicit categories helped discipline spending and increase savings and investment"; (3) "Using the four labeled envelopes made it easier to reject people's request to borrow money"; (4) "Using the four labeled envelopes made it easier to not spend money on unnecessary consumption (e.g. alcohol, snacks, etc.)"; (5) "Using the four labeled envelopes made it easier to save for school fees"; (6) "Using the four labeled envelopes made it easier to save for health expenses". Control mean refers to the meal value of the outcome among non-persistent envelope users at midline. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 10: HTE by Household Size

	(1) Productive Investments	(2) Lumpy Investments
Envelopes	78.76	1.14
	(57.11)	(9.77)
Large Family	36.72	-4.34
	(57.17)	(12.47)
Envelopes * Large Family	-31.89	26.10*
	(69.64)	(13.32)
Control Group Mean	235.12	50.61
Control Group S.D.	264.59	82.60
N	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. 'Large Family' is an indicator equal to one if the household's family size is larger than the median household size. Lumpy investments are self-reported large investments made in the previous year in physical and human capital. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 11: Decomposing Timing and Investment Type

	(1) Agriculture	(2) Livestock	(3) Enterprise
Panel A. Midline			
Envelopes	3.71**	7.02	-45.59**
	(1.81)	(14.54)	(22.04)
t-test MA vs. MAD	0.28	0.72	0.82
Control Group Mean	7.50	134.65	96.79
N	810	810	810
Panel B. Endline			
Envelopes	-3.27	-1.59	59.37***
	(2.97)	(17.20)	(20.59)
t-test MA vs. MAD	0.88	0.01	0.63
Control Group Mean	18.59	126.42	51.85
N	737	737	737
Panel C. Combined			
Envelopes	-0.21	8.93	58.55*
	(3.89)	(28.25)	(30.08)
t-test MA vs. MAD	0.93	0.14	0.51
Control Group Mean	26.26	249.97	119.28
N	707	707	707

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Agriculture, Livestock, and Enterprise refer to pre-specified investments in each of the three categories. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 12: Decomposing Enterprise Types

	(1) Crafts	(2) Food	(3) Transportation	(4) Construction	(5) Market Stall
Envelopes	-0.08	-0.13	0.04	0.02	0.15
	(0.06)	(0.18)	(0.05)	(0.03)	(0.18)
t-test MA vs. MAD	0.33	0.30	0.34	0.33	0.11
Control Group Mean	0.11	0.61	0.00	0.00	0.28
N	43	43	43	43	43

Notes: Intention to Treat estimates. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Crafts - Market Stall refer to pre-specified types of enterprises. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 13: Behavioral Responses to Labeled Envelopes

	(1) Spent Money on Items Outside of Env. Category	(2) Liked to Change Allocation per Env.	(3) I sealed the Envelopes	(4) Dividing Money Helped Discipline Spending	(5) Labeling Env. Helped Discipline Spending
MAD	-0.09	-0.04	0.07	0.11	0.16**
	(0.07)	(0.03)	(0.04)	(0.07)	(0.07)
MA Mean	0.91	0.17	0.50	0.88	0.85
MA SD	0.91	1.04	0.50	0.88	0.85
N	499.00	499.00	499.00	499.00	499.00
	(6) Reject Money Requests	(7) Using Labeled Envelopes Avoid Unnec. Spending	(8) Made it Easier to Save for School	(9) Save for Health	(10) Felt Obligation to Only Spend on Env. Category
MAD	0.05	0.11*	0.11*	0.10*	0.13
	(0.07)	(0.06)	(0.06)	(0.05)	(0.09)
MA Mean	3.85	3.98	4.07	4.05	1.04
MA SD	0.96	0.80	0.73	0.69	0.38
N	499.00	499.00	499.00	499.00	499.00

Notes: Intention to Treat estimates. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Question (1) was asked on a four point scale, with the answer options ranging from 'Rarely or none of the time' to 'Most or all of the time', while Questions (2) and (3) were yes/no questions. Questions (4)-(10) were answered on a five-point Likert Agree-ability scale . MA mean and standard deviation refer to the mean value and standard deviation of the outcome in the MA group at midline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

B Histograms

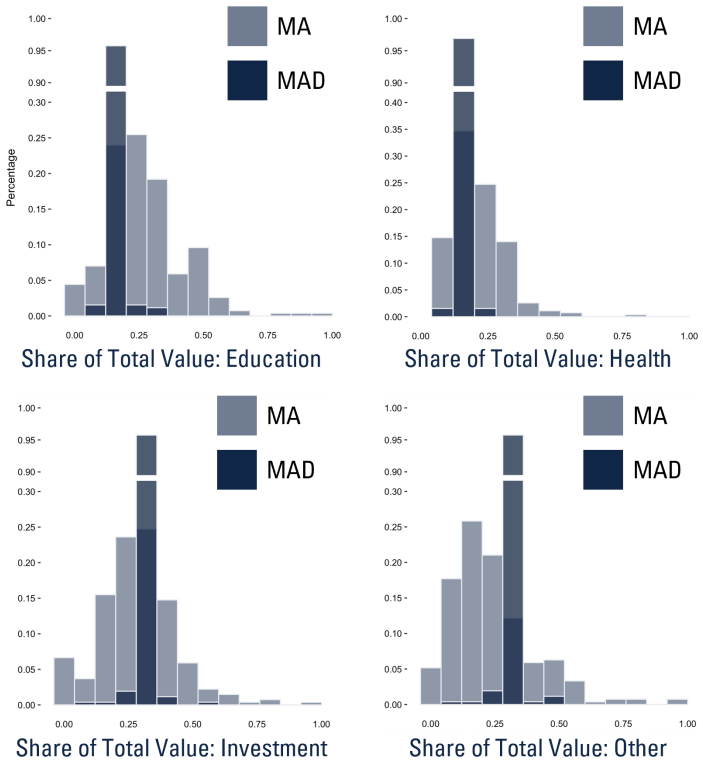


Figure 4. Histogram of Allocations across the Four Envelope Categories.

C Investment Opportunity and Envelopes Sheet

Investment Opportunities Sheet

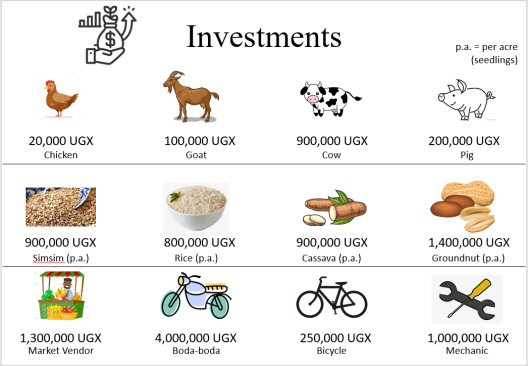


Figure 5. Investment Opportunities page 1.

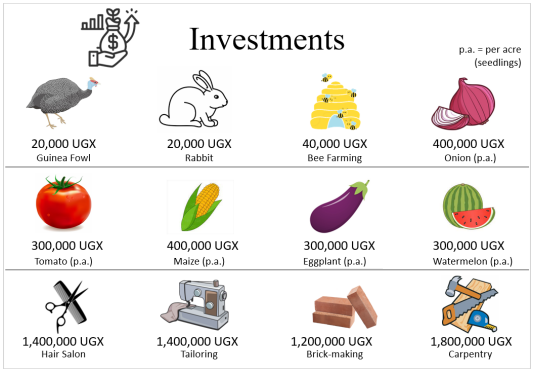


Figure 6. Investment Opportunities page 2.

At baseline, the *Investment Opportunities* sheet was given to households in all three treatments, to provide information about available investment opportunities and associated prices. Market prices are the median price after obtaining prices from three randomly chosen vendors from different markets across the refugee settlements. The prices were further confirmed by both DRC staff, the enumerators, and households that participated in the pilot.

Envelopes Overview Sheet

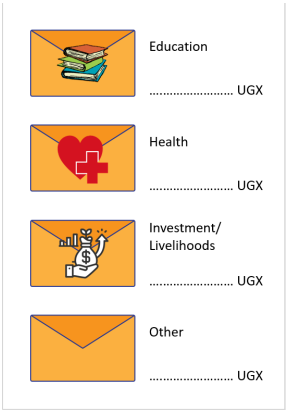


Figure 7. Envelopes Overview Sheet.

This *Envelopes Overview Sheet* was given to households in the ***MA*** and ***MAD*** treatments at the end of the baseline survey that opted-in to receive future cash transfers across the four

envelopes instead of the status quo. The enumerator wrote the monetary values allocated to each of the four envelopes, as a reminder for the households.

D Minimum Expenditure Basket

Table 14: Minimum Expenditure Basket

MEB Component	2021 (UGX)
Food	276,904
Hygiene	16,069
Water	3,750
Education	28,667
Energy	49,495
Transport	11,001
Communication	4,256
Clothing	3,806
Health	2,669
Personal Expenditure	6,080
Livelihood	37,705
Total	440,342

The Minimum Expenditure Basket (MEB) consists of eleven categories, divided into *food* and *non-food* items that are all deemed basic needs, and is specific to the setting of refugee settlements in Uganda. The United Nations and NGO partners in the Cash Working Group base the allocations per category on household surveys conducted with refugees across all settlements in Uganda (including Rhino Camp and Imvepi), and also consider local prices. In 2019, a harmonization of the MEB was conducted, during which each sub Working Group (e.g. the Health Working Group) identified basic needs within their domain - and hence the composition of each category is the same across all refugee settlements in Uganda. The cost of meeting these basic needs can vary per settlement based on local prices and is updated on a quarterly basis based on the prices per refugee settlement. The process of the MEB is used in most humanitarian settings, for example Ethiopia/Somalia, Jordan, Turkey, Bangladesh, etc.

The default allocation for ***MAD*** is: Education (16.6%), Health (16.6%), Investments (33.3%), and Others (33.3%). Percentages are in terms of the household's total cash transfer value.

E Deviation from Pre-Analysis Plan

We submitted a Pre-Analysis Plan to the Journal of Development Economics on February 24th 2022, and successfully underwent a Stage 1 review on July 21st, 2023. Below we outline how we deviate from our Pre-Analysis Plan, and why:

- Lumpy Investments were not pre-registered as an outcome variable. This is because the Pre-Analysis Plan placed a greater emphasis on consumption patterns, rather than investment patterns.
- The Focus Group Discussion after the endline survey was not pre-registered, but introduced to help understand some of the underlying mechanisms.
- Heterogeneity based on **Persistent** users was not pre-registered, as we had not anticipated the treatment to have such a persistent effect on households.
- Income is reported based on monthly income, rather than the average across the last quarter. This is because focus group discussions indicated that households thought about their income on a monthly (or shorter) basis, and struggled to recall income over the last three months.
- Savings and Durable Goods are reported separately, however are both reported.
- Marginal Propensity to Consume, and other consumption-related outcomes are not reported as primary outcomes, due to the noisy data collection.
- Winsorizing is done separately by treatment arm, as discussed in Wicker (2025). Results are robust to winsorizing the whole sample, including at the 5% level, and not winsorizing.

F Timeline

Table 15: Timeline of RCT

Event	**Timing**
Focus Group Discussion	July 2022
First Cash Transfer: Early Group	Third Week of August 2022
Baseline Survey: Early Group	First Week of September 2022
First Cash Transfer: Late Group	Third Week of September 2022
Baseline Survey: Late Group	First Week of October 2022
Last Cash Transfer: Early Group	Third Week of February 2023
Midline Survey: Early Group	First Week of March 2023
Last Cash Transfer: Late Group	Third Week of March 2023
Midline Survey: Late Group	First Week of April 2023
Endline Survey	April 2024
Focus Group Discussion	November 2024

G Balance Tables

Table 16: Balance Table for Stratified Variables.

		(1)		(2)		(3)		F-test	(1)-(2)	(1)-(3)	(2)-(3)
		CO		***MA***		***MAD***			Pairwise t-test		
Variable	N	Mean/(SD)	N	Mean/(SD)	N	Mean/(SD)	N	F-stat/P-value	P-value	P-value	P-value
Stratified Variables											
Age of HH Head	292	38.110	288	38.056	280	37.004	860	0.542	0.964	0.354	0.371
		(14.764)		(14.302)		(13.706)		0.582			
HH Head is Female	292	0.815	288	0.806	280	0.829	860	0.267	0.771	0.660	0.467
		(0.389)		(0.396)		(0.377)		0.766			
HH size	292	6.455	288	6.375	281	6.228	861	0.501	0.730	0.316	0.524
		(2.767)		(2.838)		(2.662)		0.606			
Arrival Year	292	2018.240	288	2018.201	281	2018.242	861	0.011	0.901	0.994	0.898
		(3.675)		(3.737)		(3.829)		0.989			
Country of Origin: South Sudan	292	0.901	288	0.910	281	0.900	861	0.093	0.711	0.990	0.704
		(0.300)		(0.287)		(0.300)		0.911			
DRC Vulnerability Score	119	60.057	112	59.479	111	59.442	342	0.616	0.362	0.327	0.953
		(4.888)		(4.731)		(4.585)		0.541			
Share of Protection Referrals	292	0.592	288	0.611	281	0.605	861	0.109	0.647	0.760	0.881
		(0.492)		(0.488)		(0.490)		0.897			

Notes: Columns (1), (2), and (3) show the average value (and standard deviation) for respondents in each of the three treatments: Cash Only, Mental Accounting, and Mental Accounting with Default. The F-test reports the joint test for orthogonality, including both the F-statistic and associated p-value. The normalized difference between means is reported, together with significance levels based on t-tests. 861 households were surveyed, of which one did not have an adult head of household and another did not answer whether the adults were working (hence Percentage Adults working has 859 observations). 342 households had Vulnerability Scores from DRC. Randomization was further stratified on the Zone of Residence, however as this is a categorical variable, it is not included in the balance table. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 17: Balance Table for Non-Stratified Variables.

		(1)		(2)		(3)		F-test	(1)-(2)	(1)-(3)	(2)-(3)
		CO		***MA***		***MAD***			Pairwise t-test		
Variable	N	Mean/(SD)	N	Mean/(SD)	N	Mean/(SD)	N	F-stat/P-value	P-value	P-value	P-value
Non-Stratified Variables											
Depressed	292	0.880	288	0.837	281	0.836	861	1.448	0.135	0.133	0.987
		(0.325)		(0.370)		(0.371)		0.236			
Outstanding loan amount ($)	292	12.330	288	10.279	281	9.326	861	1.226	0.313	0.139	0.601
		(26.624)		(22.015)		(21.461)		0.294			
Monthly Income ($ PPP)	292	32.956	288	34.863	281	43.102	861	1.723	0.718	0.086*	0.182
		(60.288)		(66.688)		(79.834)		0.179			
Savings ($ PPP)	292	25.495	288	26.817	281	25.291	861	0.080	0.753	0.961	0.708
		(51.872)		(49.011)		(48.078)		0.923			
Seasonal Migration	292	0.027	288	0.052	281	0.053	861	1.470	0.128	0.114	0.945
		(0.164)		(0.223)		(0.225)		0.231			
Self-Control	292	36.760	288	36.455	281	37.384	861	1.825	0.544	0.199	-0.059*
		(6.009)		(6.108)		(5.587)		0.162			
Self-Reliance Index	292	1.950	288	2.016	281	2.019	861	1.106	0.195	0.198	0.965
		(0.614)		(0.617)		(0.660)		0.331			
Aspirations	292	0.005	288	0.069	281	-0.017	861	1.173	0.253	0.726	0.137
		(0.717)		(0.640)		(0.735)		0.331			
Highest Schooling Attained	292	5.233	288	5.149	281	4.801	861	0.886	0.807	0.211	0.309
		(4.160)		(4.061)		(4.102)		0.413			
Fraction of Adults Working	291	0.375	288	0.404	280	0.412	859	0.563	0.432	0.310	0.820
		(0.427)		(0.439)		(0.435)		0.570			
Fraction of Kids in School	267	0.952	264	0.957	253	0.964	784	0.312	0.748	0.426	0.634
		(0.178)		(0.166)		(0.141)		0.732			
Livestock ($ PPP)	292	67.051	288	88.037	281	86.633	861	1.096	0.183	0.203	0.933
		(174.024)		(203.936)		(193.419)		0.335			
Acres of Land	56	1.304	54	1.734	60	1.350	170	0.203	0.543	0.945	0.642
		(2.619)		(4.663)		(4.307)		0.816			
Locus of Control	292	28.462	288	28.500	281	28.238	861	0.155	0.939	0.660	0.613
		(5.859)		(5.995)		(6.321)		0.857			
Poverty Likelihood	292	60.395	288	58.582	281	58.041	861	0.926	0.314	0.204	0.764
		(22.340)		(20.917)		(21.969)		0.396			
Experienced Shock	292	0.418	288	0.455	281	0.488	861	1.408	0.369	0.094*	0.436
		(0.494)		(0.499)		(0.501)		0.245			
Response to Hyp. Shock: Savings	292	0.479	288	0.517	281	0.480	861	0.536	0.362	0.981	0.379
		(0.500)		(0.501)		(0.501)		0.585			
Risk Preferences	292	4.305	288	4.003	281	4.064	861	0.639	0.278	0.402	0.832
		(3.364)		(3.315)		(3.510)		0.528			
Time Preferences	292	5.267	288	5.163	281	5.125	861	0.109	0.743	0.650	0.904
		(3.755)		(3.867)		(3.761)		0.897			
Hyperbolic Discounters	292	0.086	288	0.122	281	0.125	861	1.382	0.156	0.128	0.913
		(0.280)		(0.327)		(0.331)		0.252			
Remittances Given ($ PPP)	292	3.305	288	3.542	281	2.191	861	1.273	0.804	0.196	0.124
		(11.335)		(11.647)		(9.087)		0.281			
Remittances Received ($ PPP)	292	4.505	288	3.637	281	2.787	861	1.331	0.439	0.112	0.369
		(14.838)		(11.965)		(10.525)		0.265			
1st CT: Share on Educ.	277	0.224	268	0.221	261	0.220	806	0.046	0.845	0.768	0.916
		(0.181)		(0.169)		(0.170)		0.955			
1st CT: Share on Health	277	0.115	268	0.120	261	0.128	806	0.660	0.661	0.257	0.491
		(0.124)		(0.133)		(0.145)		0.517			
1st CT: Share on Inv.	277	0.284	268	0.272	261	0.285	806	0.194	0.601	0.960	0.575
		(0.264)		(0.257)		(0.270)		0.824			

Notes: Columns (1), (2), and (3) show the average value (and standard deviation) for respondents in each of the three treatments: Cash Only, Mental Accounting, and Mental Accounting with Default. The F-test reports the joint test for orthogonality, including both the F-statistic and associated p-value. The p-value between means is reported, together with significance levels based on t-tests. All monetary values are reported in 2021 USD PPP. 861 households were surveyed, of which one did not have an adult head of household and another did not answer whether the adults were working (hence Percentage Adults working has 859 observations). 170 had additional land, and 784 households had children in a school-going age. 55 households did not know how they intended to spend their first cash transfer (CT). Variables winsorized at the 1% level include: Outstanding Loan Value, Monthly Income, Savings Amount, Livestock, Acres of Land, Remittances Given, Remittances Received, and Aspirations. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

H Attrition

H.1 Midline

Table 18: Midline: Attrition

	(1)	(2)
	Attrition	
Envelopes	0.001	
	(0.017)	
MA		-0.006
		(0.019)
MAD		0.008
		(0.020)
Age of HoHH	-0.001	-0.001
	(0.001)	(0.001)
HH size	-0.002	-0.002
	(0.003)	(0.003)
Female	0.009	0.009
	(0.019)	(0.019)
Origin: SSD	-0.025	-0.025
	(0.033)	(0.033)
Arrival Year	0.002	0.002
	(0.002)	(0.002)
Protection Referral	-0.034	-0.034
	(0.022)	(0.022)
BL Monthly Income	-0.000	-0.000
	(0.000)	(0.000)
BL Self-Control	0.000	0.000
	(0.001)	(0.001)
BL Exp. Neg Shock	-0.003	-0.003
	(0.016)	(0.016)
Early	-0.099***	-0.098***
	(0.034)	(0.034)
Constant	-3.147	-3.181
	(3.398)	(3.376)
Observations	861	861
R-squared	0.062	0.062

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

Notes: Intention to Treat estimates. Attrition is a dummy variable equal to one if the household was surveyed at baseline, but not at midline. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Control mean refers to the meal value of the outcome in the control group at midline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

H.2 Endline

Table 19: Endline: Attrition

	(1)	(2)
	Attrition	
Envelopes	0.005	
	(0.025)	
MA		0.028
		(0.029)
MAD		-0.019
		(0.028)
Age of HoHH	-0.002**	-0.002**
	(0.001)	(0.001)
HH size	-0.003	-0.003
	(0.004)	(0.004)
Female	-0.005	-0.004
	(0.030)	(0.029)
Origin: SSD	-0.083	-0.082
	(0.078)	(0.078)
Arrival Year	0.011***	0.011***
	(0.004)	(0.004)
Protection Referral	-0.094***	-0.095***
	(0.035)	(0.035)
BL Monthly Income	0.000	0.000
	(0.001)	(0.001)
BL Self-Control	0.001	0.001
	(0.002)	(0.002)
BL Exp. Neg Shock	-0.004	-0.003
	(0.024)	(0.024)
Early	-0.119**	-0.120**
	(0.050)	(0.050)
Constant	-21.046***	-20.926***
	(7.231)	(7.164)
Observations	861	861
R-squared	0.077	0.080

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

Notes: Intention to Treat estimates. Attrition is a dummy variable equal to one if the household was surveyed at baseline, but not at endline. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Online Appendix to:

Mental Accounting and Cash Transfers: Experimental Evidence from a Humanitarian Setting

by Till Wicker, Patricio Dalton, and Daan van Soest

A Decomposing Health Expenses

A.1 Endline

Table 20: Endline: Health Expenditures Decomposed

	(1) Preventive Health	(2) Medicine
Envelopes	-43.90**	-6.14
	(18.10)	(7.09)
MA	-45.87**	-4.41
	(20.90)	(7.85)
MAD	-41.95**	-7.85
	(20.78)	(8.05)
Control Group Mean	276.57	54.68
N	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Preventive health refers to expenses on Latrines, Water Filters, ORS, Mosquito Nets, and WaterGuard. Respondents were asked how often they had purchased these items in the last 6 months, and the frequency was multiplied by the median price of three market vendors in the refugee settlements. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 21: Endline: Preventive Health Expenditures Decomposed

	(1) WaterFilter	(2) ORS	(3) Latrine	(4) Mosquito	(5) WaterGuard
Envelopes	0.18	0.03	-42.01**	-2.51	0.79
	(0.14)	(0.04)	(17.59)	(3.02)	(0.57)
MA	0.09	0.08	-41.69**	-4.46	0.34
	(0.11)	(0.05)	(20.25)	(3.37)	(0.68)
MAD	0.27	-0.02	-42.33**	-0.58	1.24
	(0.22)	(0.04)	(20.19)	(3.60)	(0.80)
Control Group Mean	0.26	0.08	240.45	34.90	0.91
N	737	737	737	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Respondents were asked how often they had purchased these items in the last 6 months, and the frequency was multiplied by the median price of three market vendors in the refugee settlements. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

A.2 Midline

Table 22: Midline: Health Expenditures Decomposed

	(1) Preventive Health	(2) Medicine
Envelopes	17.52	-0.11
	(14.54)	(0.82)
MA	25.41	-0.86
	(17.08)	(0.76)
MAD	8.88	0.71
	(17.52)	(1.09)
Control Group Mean	146.43	1.97
N	810	810

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Preventive health refers to expenses on Latrines, Water Filters, ORS, Mosquito Nets, and WaterGuard. Respondents were asked how often they had purchased these items in the last 6 months, and the frequency was multiplied by the median price of three market vendors in the refugee settlements. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 23: Midline: Preventive Health Expenditures Decomposed

	(1) WaterFilter	(2) ORS	(3) Latrine	(4) Mosquito	(5) WaterGuard
Envelopes	0.03	-0.01	45.81**	-3.04	-0.45
	(0.04)	(0.01)	(22.47)	(2.50)	(0.44)
MA	0.04	-0.01	57.80**	-1.93	-0.62
	(0.05)	(0.01)	(28.48)	(2.88)	(0.43)
MAD	0.01	-0.01	32.68	-4.26	-0.27
	(0.05)	(0.01)	(25.65)	(2.78)	(0.53)
Control Group Mean	0.07	0.01	120.44	26.66	0.71
N	810	810	810	810	810

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Respondents were asked how often they had purchased these items in the last 6 months, and the frequency was multiplied by the median price of three market vendors in the refugee settlements. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

B Decomposing Education Expenses

Table 24: Education In Emergencies (EiE) Cash Transfer

	(1) EiE Cash Transfer
Envelopes	-0.04*
	(0.03)
Control Group Mean	0.33
N	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. EiE refers to an Education in Emergencies cash transfer, given by DRC. It's selection is based on at-risk or out-of-school children, and hence the evaluation criteria is separate from the cash program we are evaluating. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

C Consumption

C.1 Midline

Table 25: Midline: Consumption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Monthly	Fraction of Monthly Spending on							
	Consumption ($PPP)	Education	Health	Invest.	Food	Clothes	HH Items	Tempt.	Loans
Envelopes	6.01	-0.80	-0.71	0.74	0.93	-0.08	0.03	-0.00	1.13**
	(13.06)	(1.33)	(1.09)	(1.11)	(1.31)	(0.59)	(0.57)	(0.04)	(0.44)
MA	2.74	0.25	-0.03	0.57	0.57	0.09	-0.21	-0.01	0.97**
	(15.39)	(1.58)	(1.29)	(1.31)	(1.50)	(0.67)	(0.63)	(0.06)	(0.49)
MAD	9.58	-1.95	-1.45	0.93	1.33	-0.27	0.29	0.00	1.30**
	(15.46)	(1.52)	(1.25)	(1.27)	(1.54)	(0.73)	(0.70)	(0.05)	(0.55)
t-test MA vs. MAD	0.81	0.17	0.25	0.85	0.71	0.92	0.31	0.87	0.22
Control Group Mean	216.28	26.47	16.04	8.69	28.63	2.66	6.84	0.07	1.32
Control Group S.D.	183.76	19.95	14.96	14.60	18.26	7.49	7.75	0.67	5.39
N	810	810	810	810	810	810	810	810	810

Notes: Intention to Treat estimates, based on PD Lasso regressions (Belloni et al., 2014). ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

C.2 Endline

Table 26: Endline: Consumption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Monthly	Fraction of Monthly Spending on							
	Consumption ($PPP)	Education	Health	Invest.	Food	Clothes	HH Items	Tempt.	Loans
Envelopes	-18.03	-1.61	-0.12	0.65	-2.06	0.60	-1.06	0.08	1.35*
	(13.62)	(1.57)	(1.22)	(0.90)	(1.52)	(0.51)	(0.83)	(0.12)	(0.70)
MA	-7.22	-2.14	0.14	0.31	-2.59	0.73	-1.10	0.25	1.28
	(16.52)	(1.80)	(1.38)	(1.00)	(1.70)	(0.57)	(0.96)	(0.21)	(0.86)
MAD	-28.68*	-1.09	-0.38	0.98	-1.54	0.47	-1.01	-0.09	1.43*
	(14.62)	(1.82)	(1.45)	(1.06)	(1.79)	(0.60)	(0.98)	(0.09)	(0.79)
t-test MA vs. MAD	0.15	0.87	0.97	0.48	0.59	0.78	0.95	0.09	0.67
Control Group Mean	176.76	20.91	12.67	4.43	30.53	1.93	9.65	0.14	2.70
Control Group S.D.	201.43	19.83	14.92	11.51	21.39	6.54	11.80	1.10	8.60
N	737	737	737	737	737	737	737	737	737

Notes: Intention to Treat estimates, based on PD Lasso regressions (Belloni et al., 2014). ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Consumption was measured in the last month, and hence we are unable to comment on the treatment's overall effects on consumption - as we only document it at one point in time. Ex-ante, it is unclear whether the treatment would impact the total value of consumption, and consumption expenditures. We document a decrease in monthly consumption - albeit statistically insignificant - for treatment households one year after the intervention. This is primarily driven by ***MAD*** households, where the reduction in consumption is statistically significant (not correcting for MHT). This can be post-rationalized using an Euler equation: for a given C_{t+1}, higher returns to investments imply a higher marginal utility of consumption

in time t ($u'(C_t)$). Given the non-linearities in returns to lumpy investments (see Kaboski et al. (2024)), we can assume that $f'(K_t)$ is higher, on average, for treatment households than control households, as they made more lumpy investments. The Euler equation – $u'(C_t) = f'(K_t)\delta u'(C_{t+1})$ – can hence rationalize why households who have high returns to investments, are willing to forgo consumption in the short run.

D Other Outcome Variables

D.1 Midline

Table 27: Midline: Consumption-Related Outcomes

	(1) School Attendance	(2) Health Needs Met	(3) Self-Reliance	(4) Food Security
Envelopes	0.02	0.00	0.15**	-0.02
	(0.06)	(0.01)	(0.06)	(0.06)
MA	0.05	0.01	0.13*	0.04
	(0.07)	(0.01)	(0.07)	(0.07)
MAD	-0.02	-0.00	0.17**	-0.08
	(0.08)	(0.02)	(0.08)	(0.07)
Control Group Mean	4.49	0.91	0.00	0.00
N	810	810	810	810

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. School attendance is the average school attendance of all children in a household across the last week. Health Needs Met refers to the fraction of cases in the last 90 days in which a household member had a healthcare need that they were able to meet. This value is averaged across all household members. Self-Reliance is based on a 13-item Self-Reliance Index, while Food Security is based on a WFP-created 5-item measure. More details on both can be found at Wicker et al. (2023). Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

D.2 Endline

Table 28: Endline: Consumption-Related Outcomes

	(1) School Attendance	(2) Health Needs Met	(3) Self-Reliance	(4) Food Security
Envelopes	0.00	-0.00	0.01	-0.01
	(0.08)	(0.01)	(0.07)	(0.06)
MA	0.04	0.01	-0.00	-0.01
	(0.10)	(0.02)	(0.08)	(0.07)
MAD	-0.04	-0.02	0.03	-0.00
	(0.09)	(0.02)	(0.08)	(0.07)
Control Group Mean	4.22	0.88	0.00	0.00
N	737	737	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. School attendance is the average school attendance of all children in a household across the last week. Health Needs Met refers to the fraction of cases in the last 90 days in which a household member had a healthcare need that they were able to meet. This value is averaged across all household members. Self-Reliance is based on a 13-item Self-Reliance Index, while Food Security is based on a WFP-created 5-item measure. More details on both can be found at Wicker et al. (2023). Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E Other Mechanisms

E.1 Partitioning at Midline

Table 29: Midline: Best Part of the labeled Envelopes

	(1)	(2)	(3)	(4)
	Best Part of the 4 labeled Envelopes			
	Easy to Hide	Safety	Saving	Planning
Persistent	0.00	-0.01	-0.01	0.01
	(0.01)	(0.02)	(0.03)	(0.04)
Control Group Mean	0.01	0.07	0.11	0.74
N	440.00	440.00	440.00	440.00

Notes: Intention to Treat estimates, based on PD Lasso regressions (Belloni et al., 2014). Persistent is a dummy variable equal to 1 if the household is still using the labeled envelope at endline. Respondents were asked an open-ended question of "What was the biggest added value of the envelopes?". Control mean refers to the meal value of the outcome among non-persistent envelope users at midline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.2 Self-Control

Table 30: Mechanism: Self-Control

	(1)	(2)
	Self-Control Index	
	Average Score	Anderson (2008)
Panel A. Midline		
Envelopes	0.15	-0.03
	(0.28)	(0.05)
Control Group Mean	39.01	0.02
N	810	810
Panel B. Endline		
Envelopes	0.05	-0.06
	(0.34)	(0.05)
Control Group Mean	40.40	0.03
N	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. The Self-Control measure is described in more detail in Wicker et al. (2023). Column (1) creates an average index, while Column (2) creates an Anderson (2008) index. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.3 Information from Default

We don't document differential levels of savings at midline between ***MA*** and ***MAD***. Furthermore, at endline, we document higher levels of savings among ***MA*** households. This suggests, if anything, that the information provided by the default allocation did not foster savings.

Tables 31 and 32 below reports heterogeneous treatment effects for ***MA*** (compared against ***MAD***) based on the *Sum of Squared Distance* between the household's allocation across the four envelopes, and the default.[35] The interaction term between ***MA*** and the Sum of Squared Distance is not statistically significant across any of the outcome variables, suggesting that the treatment effect does not vary depending on the distance of a household's allocation to the default - and hence the additional information obtained from the default.

Table 31: Distance from The MAD Default Recommendation: Endline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Productive Investments		Lumpy Investments		Monthly Income		Savings		Loans	
MA	-85.12	-83.76	-4.34	-6.34	5.22	4.39	12.38*	13.23**	-0.89	-1.24
	(73.29)	(73.95)	(9.11)	(9.05)	(3.26)	(3.34)	(6.41)	(6.63)	(4.43)	(4.46)
MA_SSD		-4.12		6.07		2.47		-2.59		1.09
		(17.94)		(6.63)		(1.65)		(2.14)		(1.70)
Control Group Mean	340.10	340.10	70.67	70.67	26.48	26.48	42.67	42.67	35.97	35.97
Control Group S.D.	785.04	785.04	103.28	103.28	35.38	35.38	67.25	67.25	54.70	54.70
N	456	456	456	456	456	456	456	456	456	456

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

[35] For the detailed specification, see the Pre-Analysis Plan (Wicker et al., 2023).

Table 32: Distance from The MAD Default Recommendation: Midline

	(1)	(2)	(3)	(4)	(5)	(6)
	Savings		Loans		Monthly Income	
MA	0.41	-0.19	-23.86**	-24.17**	0.50	1.15
	(7.98)	(7.88)	(9.94)	(9.76)	(4.05)	(4.10)
MA_SSD		2.03		1.03		-2.18
		(3.56)		(3.76)		(1.55)
Control Group Mean	76.81	76.81	46.29	46.29	33.39	33.39
Control Group S.D.	100.75	100.75	116.62	116.62	46.61	46.61
N	499	499	499	499	499	499

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.4 Reduced Temptation

Table 33: Mechanism: Reduced Temptation

	(1) SC-4	(2) SC-5	(3) SC-6
Panel A. Midline			
Envelopes	0.03	-0.01	-0.12*
	(0.07)	(0.07)	(0.07)
Control Group Mean	-0.01	0.01	0.06
N	810	810	810
Panel B. Endline			
Envelopes	-0.00	-0.07	-0.11
	(0.08)	(0.07)	(0.07)
Control Group Mean	-0.00	0.02	0.06
N	737	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. SC-4 asked "I refuse things that are bad for me, even if they are fun", SC-5 asked "I'm good at resisting temptation", and SC-6 asked "People would say that I have very strong self-discipline". Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.5 Greater Safety

Table 34: Mechanism: Greater Safety

	(1) Experienced Theft in Last 6 Months
Midline	
Envelopes	0.04
	(0.06)
Control Group Mean	0.32
N	810
Endline	
Envelopes	0.13
	(0.10)
Control Group Mean	0.80
N	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. The outcome variable is an indicator variable if the household experienced a theft incident in their house in the last six months. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.6 Formalized Savings Sources

Table 35: Mechanism: Formalized Savings Sources

	(1) Has Savings	(2) SACCO	(3) MoMo	(4) Less Formal Group	(5) Friend and Family	(6) Home	(7) VLSA	(8) ROSCA
Envelopes	0.02	-0.02	-0.00	0.06	0.00	-0.00	0.02	0.03
	(0.04)	(0.02)	(0.01)	(0.04)	(0.01)	(0.04)	(0.04)	(0.03)
Control Group Mean	0.58	0.09	0.02	0.73	0.01	0.18	0.45	0.28
N	737	438	438	438	438	438	438	438

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Has Savings is a dummy variable equal to 1 if the household has savings at endline. The other variables refer to savings locations. MoMo is mobile money, SACCO is a formally registered savings group, VSLA is a village savings and loans association, an informal savings group that does not have a revolving fund, which is how they differ from ROSCAs, who have a revolving fund that pays out every time. 'Less Formal Group' combines VSLA and ROSCAs. Friends and Family, and Home, are indicators for savings with friends and family, or saving at home. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.7 Social Desirability Bias - Experimenter Demand Effect

Table 36: Mechanism: Social Desirability Bias

	(1) Social Desirability Bias
Envelopes	0.106
	(0.130)
MA	0.115
	(0.147)
MAD	0.096
	(0.153)
Observations	810
R-squared	0.477

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Social Desirability Bias is based on a 14-item Marlowe-Crowne measure. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Furthermore, Experimenter Demand Effects can be ruled out by the patterns observed in the data: if respondents - particularly in the Treatment arms - wanted to appease the researchers, we would have expected to find positive, not negative, effects of the intervention on health- and education-related expenditures. Similarly, we document no treatment effects on mental health or downstream outcome variables (e.g., health needs met). Lastly, productive and lumpy investments were asked in two separate ways: for productive investments, households were read out a list of productive investments, and they indicated how many they had purchased in the last year. The combined value was computed based on median market prices. For lumpy investments, households were asked to name the type, date, and size of their five largest, lumpy investments made in the last year. The consistency of results across these two forms of elicitation further suggests experimenter demand effects are not driving the treatment effects.

Table 37: Expectation to Receive Additional Cash Transfer

	(1) Expect to Receive Add. Round of Cash Transfer
Envelopes	0.00
	(0.03)
Control Group Mean	0.61
Observations	737

Notes: Intention to Treat estimates. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Respondents were asked whether they expected to receive an additional cash transfer from DRC in the future. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.8 Remittances

Table 38: Mechanism: Remittances, Midline

	(1)	(2)
	Remittances (USD PPP)	
	Received	Given
Envelopes	1.12	-0.45
	(0.88)	(0.71)
Control Group Mean	5.05	4.65
Observations	810	810

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Remittances Received and Given are self-reported remittances received and given to/from neighbors, friends, and family in the last 30 days. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 39: Mechanism: Remittances, Endline

	(1)	(2)
	Remittances (USD PPP)	
	Received	Given
Envelopes	1.23	0.62
	(0.88)	(0.47)
Control Group Mean	3.99	2.69
Observations	737	737

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Envelopes is the pooled treatment of MA and MAD, where MA and MAD differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Remittances Received and Given are self-reported remittances received and given to/from neighbors, friends, and family in the last 30 days. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

E.9 Decomposing Timing and Type of Investment: MA vs. MAD

Table 40: Decomposing Timing and Investment Type

	(1) Agriculture	(2) Livestock	(3) Enterprise
Panel A. Midline			
MA	5.99**	12.63	-47.11**
	(2.39)	(18.56)	(22.96)
MAD	1.21	0.88	-43.91**
	(2.16)	(15.67)	(21.80)
Control Group Mean	7.50	134.65	96.79
Panel B. Endline			
MA	-1.49	-28.46*	65.62**
	(3.67)	(16.26)	(29.37)
MAD	-5.03	25.00	53.19*
	(3.32)	(24.96)	(27.43)
Control Group Mean	18.59	126.42	51.85
N	737	737	737
Panel C. Combined			
MA	1.52	-15.75	71.29*
	(4.62)	(31.86)	(40.46)
MAD	-2.00	34.42	45.38
	(4.56)	(35.37)	(36.34)
t-test MA vs. MAD	0.93	0.14	0.51
Control Group Mean	26.26	249.97	119.28
N	707	707	707

Notes: Intention to Treat estimates. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. MA and MAD refer to the two treatments, which differ because households in MAD were first shown a default recommended allocation of the cash transfer across the four envelope categories. Agriculture, Livestock, and Enterprise refer to pre-specified investments in each of the three categories. Control mean and standard deviation refer to the mean value and standard deviation of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

F Cost-Effectiveness Calculations

Costs:

- Procurement of Envelopes
  - 500 Envelopes cost $13.05 ⇒ one envelope cost $0.026
  - Each household received (4*6) = 24 envelopes. ⇒ total cost of envelopes = $0.623
- Procurement of Stickers
  - 24 stickers cost $0.392 ⇒ one sticker cost $0.016
  - Each household received (6*6) = 36 stickers. ⇒ total cost of stickers = $0.587
- Total Procurement Costs: $0.623 + $0.587 = $1.21
- Labour Costs: Putting Stickers on Envelopes
  - On average, 6 stickers per minute ⇒ 360 stickers her hour
  - Every household receives (6*6) = 36 stickers ⇒ 6 minutes
  - Daily wage: 50,000 UGX = $13.05, for 8 hours ⇒ 6 minutes = $0.163
- Labour Costs: Cash Distribution
  - 2 employees, 1 minute per household
  - Daily wage: 50,000 UGX = $13.05, for 8 hours
  - 1 minute, for 2 employees, for 6 cash transfers ⇒ $0.326

Combining the costs per household:

- Procurement Envelopes: $0.623
- Procurement Stickers: $0.587
- Labour costs: stickers on envelopes: $0.163
- Labour costs: handing out envelopes: $0.326
- **Total Cost: $1.70**

Total costs could have been reduced by only having 3 stickers, instead of 6. This would have reduced the 'Procurement of Stickers costs', and 'Labour Costs: Putting Stickers on Envelopes' by half. Then the costs would have equalled: $1.33

Cost-Effectiveness Calculation for Aggarwal et al. (2023)

Each lockbox cost $3.40, with an additional lock and key costing an additional $1.00.

In the single-box treatment, households were only offered one lockbox. In the multiple-box treatment, households were offered up to three lockboxes. 24% only took 1 lockbox, 33% took 2 lockboxes, and 42% took all three lockboxes. 1% did not take any lockboxes. Hence, in the 'multiple lockboxes' treatment, households on average took $(0.01 * 0 + 0.24 * 1 + 0.33 * 2 + 0.42 * 3 =)$ 2.16 lockboxes. With each lockbox costing $4.40, the average cost of the intervention is $2.16 * 4.40 = 9.50$.

Column 9 in Tables 4 and 5 of (Aggarwal et al., 2023) report treatment effects of the multiple lockboxes on total deposits, with treatment effects of 0.84 and 0.49 (with a Control group standard deviation of 3.93 and 1.80, respectively). Hence treatment effects, expressed in terms of standard deviations, are $0.84/3.93 = 0.21$ and $0.49/1.80 = 0.27$,

G Heterogeneous Treatment Effects

G.1 Persistent Users

Table 41: 2SLS Regression: Persistent Envelope Users, Midline

	(1) Savings	(2) Loans	(3) Durable Good	(4) Productive Investments	(5) Monthly Income
Persistent	95.01***	51.43***	-35.65	-80.08	-1.87
	(16.94)	(16.00)	(166.42)	(96.16)	(10.44)
Control Group Mean	41.44	22.97	392.95	244.60	35.15
Control Group S.D.	56.88	47.10	694.58	428.83	44.99
N	810	810	810	810	810

Notes: 2SLS IV estimates for 'Persistent'. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. All regressions include strata variables, imbalanced baseline variables, and the baseline value of the outcome, where available. Productive investments refer to investments in a pre-specified list of productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Lumpy investments are self-reported large investments made in the previous year in physical and human capital. Savings are total savings held by the household. Monthly income is self-reported monthly income, while Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Control mean refers to the meal value of the outcome in the control group at endline. Robust standard errors are in parentheses. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 42: Endline: Persistent users (Propensity Score Matching)

	(1) Productive Investments	(2) Lumpy Investments	(3) Monthly Income	(4) Savings	(5) Loans	(6) Durable Goods
Persistent	67.38	50.27***	40.74**	49.08***	-10.50	19.37*
	(41.83)	(14.09)	(19.13)	(14.38)	(8.51)	(10.73)
N	158	158	158	158	158	158

Notes: Propensity Score Matching based on LASSO-selected control variables. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. Persistent is coded as a dummy variable equal to one if the household is still using the envelopes at endline, and zero otherwise.Productive investments refer to investments in a pre-specified list of productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Lumpy investments are self-reported large investments made in the previous year in physical and human capital. Savings are total savings held by the household. Monthly income is self-reported monthly income, while Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements.. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 43: Midline: Persistent users (Propensity Score Matching)

	(1) Savings	(2) Loans	(3) Durable Good	(4) Productive Investments	(5) Monthly Income
Persistent	24.30***	11.43*	-27.25	5.39	-6.25
	(8.05)	(6.00)	(103.38)	(34.96)	(6.14)
N	291	291	291	291	291

Notes: Propensity Score Matching based on LASSO-selected control variables. Monetary outcomes are winsorized at the 99th percent level, separately per experimental group, and converted into 2021 USD PPP. Persistent is coded as a dummy variable equal to one if the household is still using the envelopes at endline, and zero otherwise.Productive investments refer to investments in a pre-specified list of productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements. Lumpy investments are self-reported large investments made in the previous year in physical and human capital. Savings are total savings held by the household. Monthly income is self-reported monthly income, while Durable Goods refer to the acquisition of a pre-specified list of durable, non-productive assets, multiplied by the median market price obtained from three vendors in the refugee settlements.. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.