Loose Knots:
Strong versus Weak Commitments to Save for Education in Uganda*

January 2018

Dean Karlan
Leigh L. Linden

Abstract:
Commitment devices offer an opportunity to restrict future choices. However, strict commitments may deter participation. Using a school-based commitment savings program for educational expenses in Uganda, we compare an account fully-committed to educational expenses to an account with a weaker commitment (funds withdrawn in cash, rather than a voucher). The weaker commitment generates higher account savings, and when combined with a parental outreach generates higher educational supplies expenditures and 0.11 standard deviation higher math and language test scores. We find no effect from the fully-committed account, and no effect for either account on attendance, enrollment, or non-cognitive skills.

JEL Codes: D12, D91, I21, O12
Key Words: Commitment Savings, Micro-Savings, Educational Resources, School Participation

* Contact information: Dean Karlan, karlan@northwestern.edu, Northwestern University, IPA, M.I.T. J-PAL, NBER, and Center for Global Development; Leigh Linden, leigh.linden@austin.utexas.edu, The University of Texas at Austin, BREAD, IPA, IZA, M.I.T. J-PAL, and NBER. We are grateful for the assistance of a number of individuals and organizations without whom this project could not have been successfully completed. Caton Brewster, Carl Brinton, Sarah Kabay, Sana Khan, Selvan Kumar, Kelsey Larson, Doug Parkerson, Pia Raffler, Elana Safran and Glynis Startz from Innovations for Poverty Action all provided valuable assistance with the research or project management. We gratefully acknowledge the financial support of the Yale Savings and Payments Research Fund at Innovations for Poverty Action, sponsored by a grant from the Bill & Melinda Gates Foundation. The United States Agency for International Development (Development Innovations Venture Award AID-OAA-G-12-00008) generously provided additional funding for the project. All opinions are those of the authors, and not necessarily that of any of the funders or collaborating institutions.

"Make it easy" – Richard Thaler, co-author of *Nudge: Improving Decisions about Health, Wealth, and Happiness* (Clement 2013)

Section I. Introduction

Commitment devices offer individuals an opportunity to restrict future choices. Self-aware individuals may prefer such restrictions in order to resist future temptations, or to deflect future claims on assets from family or extended social networks. Indeed, prior research has found demand for commitment savings accounts that restrict access to one's money in order to help with self-control issues (Ashraf, Karlan, and Yin 2006; Brune et al. 2016; Dupas and Robinson 2013; Giné et al. 2017), and other research has found demand for commitment devices in other domains.

This project began after qualitative research on household finance in Uganda identified saving for school fees and supply costs as a key barrier for families.[1] We tackle three primary questions within the context of an educational savings intervention: First, a program evaluation question, can a commitment savings program be used to improve student performance through increased educational expenditures? We will discuss below, but note that the commitment savings program engages in more than merely offering a commitment savings account, and also includes effectively reminders via weekly opportunities to deposit and class discussions about the importance of savings. Second, will the commitment savings account work better with strict rule on how the accumulated funds are spent, or a flexible rule? And third, do commitment devices to save merely get unwound through offsetting behavior, and thus not change actual expenditures?

The specifics of what one means by "commitment" on a commitment savings account can vary in critical ways, and accordingly have large differential impacts on whether an account is opened, how much one deposits, when one withdraws, and perhaps most importantly, how ultimate consumption and investment choices differ. The commitment versus flexibility tradeoff exists across many points in the savings and spending process (see Amador, Werning, and Angeletos (2006) for a theoretical analysis of the overall tradeoff between commitment and flexibility). We focus here on one key dimension: whether the funds deposited are locked in for a specific "good" expenditure, or if individuals have freedom to spend withdrawals as they wish, but the "good" item is made easily available.[2] In developing countries, prior work has tested soft savings commitments versus control (Dupas and Robinson 2013), hard savings commitments versus control (Duflo, Kremer, and Robinson 2011); and, in the United States, prior work has tested soft commitment (non-binding pledge and a planning exercise) compared to a hard commitment that restricted withdrawals (Burke, Luoto, and Perez-Arce 2017), as well as harshness of early withdrawal penalties from commitment savings accounts (Beshears et al. 2017). Alas given the challenge in

[1] The initial role for the qualitative work was to collaborate with an insurance company, and identify missing insurance products. Although the insurance work did not proceed, the dominance of low savings for school in the focus groups on insurance led the researchers to engage in this project.

[2] Clearly in a perfect market, specifically one with zero transaction costs, this would make no difference: any items purchased with the locked-in commitment account could simply be sold in exchange for the most desirable item for the same value. In our market, supplies and services associated with primary education in Uganda, there are significant enough transaction costs to make such an exchange quite costly, and thus the original expenditure sticky.

categorizing product features into simplistic “strong” and “weak”, and the changes in context (health to agriculture), it is challenging to compare across studies in order to assess the tradeoffs between soft versus hard commitment. Here we consider “hard” one which locks spending in for specific purposes (although one could argue that in the Dupas and Robinson (2013) study, participants believed their lockboxes had to be spent on health, even though it did not; in this sense however we believe it is accurate to equate their health lockbox to the soft commitment account in our context, since similarly there was no requirement to spend it as such but alas that is how people did use the funds).

In theory the tradeoffs are clear: a strong commitment device may be more effective in enforcing the behavior of the future self, but the current self may be less likely to participate in the contract at all. An individual may want to commit in some, but not all, future states of the world, since emergencies do happen. The challenge is designing a contract in which a third party has the right level of enforcement discretion. If an individual cannot trust any third parties with that discretion, a self-enforcing commitment contract may work instead. In such a contract, the increased price of vice is derived from psychic costs, i.e., disappointment with oneself and one’s lack of adherence to a plan. This is akin to a model put forward by Benabou and Tirole (2004) on how personal rules can shift later behavior, and also could be construed as a test of whether “mental accounting” can be a policy instrument that induces behavior change (Shefrin and Thaler 1992).

Our third question examines whether commitment devices get unwound through offsetting behavior(Karlan, Ratan, and Zinman 2014). More money spent from a commitment account for a particular purpose may simply crowdout spending for that same purpose with funds from other sources. By examining how actual expenditures change for the particular purpose, rather than merely observing whether savings increases, we are able to make stronger statements about welfare outcomes, similar to Ashraf et al (2010) with respect to household durable goods purchases and Dupas and Robinson (2013) with respect to health investments.

We examine these questions in the context of a school-based commitment savings account in Uganda. Specifically, we test whether a stronger versus a weaker savings commitment device helps children and their families save more, spend more on educational expenses, and achieve higher test scores. Relative to the economics of education literature, we thus gain a better understanding of the education production process (Kremer and Holla 2009), building on a growing body of evidence demonstrating the possibly significant effects of basic school supplies – notebooks, uniforms, workbooks, etc. – on student performance (Das et al. 2013 and Hidalgo et al. 2013, in contrast to Kremer and Holla 2009) and parental involvement (Avvisati et al. 2013). Second, the results build on existing evidence of the importance of savings constraints for educational expenses (Barrera-Osorio et al. 2011) as well as mechanisms for tying resources to educational expenses (De Arcangelis et al. 2014).[3]

[3] It is interesting to note that, while we find that relaxing savings constraints improves educational outcomes, we find improvements in academic performance rather than participation. This contradicts the results of Barrera-Osorio et al.

We evaluate the intervention ("Supersavers") as follows: working with a local nonprofit organization Private Education Development Network (PEDN) in the Busoga sub-region of the Eastern region of Uganda, and Innovations for Poverty Action (IPA), we randomly assigned 136 primary schools to one of three groups: a strong commitment savings account (funds could be withdrawn no earlier than the end of the term, and had to be spent on educational items through a voucher that we provided), a weak commitment savings account (funds could be withdrawn no earlier than the end of the term, but were available in cash, to be spent as individuals wished)[4], or control. For both treatments, students could deposit cash into an account. At the end of each trimester they were able to use their cash or vouchers to purchase school supplies at a fair.[5] This thus becomes a test between a stricter commitment device and a weaker "make it easy" nudge of individuals towards a specific behavior (Thaler and Sunstein 2009). Although the accounts were framed as their accounts, we cannot rule out that some of the funds were parent's funds.[6] We developed a brief teacher training component, and also coordinated the transfer of money from a savings box held at the school to a local bank for safekeeping. One year into the implementation, we implemented one sub-treatment in half of the treatment schools, a parental involvement workshop.

The first stage is critical and revealing: students deposit significantly more money into the soft commitment savings account than the hard commitment savings account. And, for those with the parental outreach sub-treatment, the additional money deposited into the account leads to higher investment in school supplies, which then in turn leads to higher test scores. We find a 0.11 standard deviation (se=0.04) improvement in overall scores; this includes effects on each of the covered subjects: grammar (0.15 standard deviations, se=0.05), reading (0.11 standard deviations, se=0.05), and math (0.00 standard deviation, se=0.04). The implication for the school production function is simple: for a student to learn basic skills, having a pen, paper and workbook matters. Furthermore, the treatment effect on educational outcomes is sizable, as large as many direct educational interventions, and consistent with other estimates of the effects of such supplies (Das et al. 2013) We find no effect on student participation (either attendance or enrollment) or on a set of non-cognitive outcomes.

One critical gap we leave in our understanding of the underlying mechanics of why this program found the effects it did: whose money is used to contribute to the accounts, the child's or the

(2011) which finds that distributing funds at the time that families have to pay enrollment expenses improves enrollment rates. The difference may, in part, be due to the fact that unlike Uganda, Colombian schools still charged official fees for enrollment.

[4] The weak commitment treatment arm is thus most similar to the SEED account in Ashraf et al (2006), i.e., a commitment merely to not withdraw funds until a certain future point in time.

[5] Note that control schools did not receive a fair. Thus analysis of "any treatment" is also bundling the presence of these fairs when examining school supplies and thus test score outcomes. However, we observe no positive treatment effect on school supplies for the voucher treatment arm, thus we conclude that the fairs themselves did not have an observable direct treatment effect.

[6] As we show below, both the children and other family members contribute to the accounts, raising the possibility that multiple household mechanisms are involved.

parent's? About half of our sample reports engaging in some work, and saving some from the money earned from work. This question has important implications for asserting whether the children were time inconsistent, or the parents, or whether the account shifted power across individuals with different preferences within the household. This is true, of course, in most studies on savings of individuals who live within a household. For example, in a typical "commitment savings" account test (e.g., see Ashraf, Karlan, and Yin 2006; Dupas and Robinson 2013; Brune et al. 2016), accounts are offered to individuals, and outcomes tracked at some combination of individual and household. Yet given fungibility of money within the household, it is difficult if not impossible to assert where the funds came from that got deposited into the account. This gap is made poignant in our setting because of the power dynamics between parent and child, yet the gap exists for any study of a savings intervention which targets individuals which live within a household of multiple adults.

Section II. Background

A. Ugandan Primary Education System

Uganda abolished most primary school fees in 1997.[7] In the same year, the gross primary enrollment rate[8] ballooned from 87 percent in the early 1990s, to 123 percent in 1997. Between 1996 and 1997, 2.3 million children enrolled in primary school, increasing total enrollment to 5.7 million (Murphy, Bertoncino, and Wang 2002).

Unfortunately, while most children now enroll in primary school, the majority of them fail to graduate. In 2008, for example, the gross enrollment rate[9] in lower secondary was 33 percent–11 percentage points below the average for Sub-Saharan Africa (UNESCO 2013). The transition from primary to secondary is a challenge, as in many countries. However, the majority of Ugandan students simply fail to complete primary school. As of 2010, only 32 percent of students entering primary school completed the seventh grade ("Opportunities Lost: The Impact of Grade Repetition and Early School Leaving" 2012).

While the poor quality of primary education is a likely factor (Piper 2010),[10] students still face financial barriers. Students no longer pay enrollment fees, but they face other expenses. Many schools require uniforms, and families are responsible for providing food and school supplies, such as paper, writing instruments, and workbooks. See Appendix Table 1 for a summary of educational

[7] Initially, up to four children per family could attend school without paying tuition fees (Murphy, Bertoncino, and Wang 2002).

[8] The gross primary school enrollment ratio is the ratio of the number of enrolled primary school children, regardless of age, to the total number of primary school-aged children in the population.

[9] The gross enrollment rate for lower secondary school is the ratio of the number of children enrolled in lower secondary school regardless of age relative to the total number of children in the population who are of age to attend secondary school.

[10] The dramatic increases in enrollment have strained existing resources. In the average school in 2005, three children had to share the same textbook and 94 children crammed into a single classroom (Independent Evaluation Group (IEG) 2007).

fees and expenses expected of households. With the approval of the parent-teacher association and school management committee, schools can also charge fees for ancillary services such as supplementary lessons, practice exams and feeding programs. Official policy prohibits preventing a child from enrolling due to an inability to pay, but the majority of dropouts cite financial concerns. In the baseline survey described below, families paid an average of 5,790 UGX (2.30 USD) to send a child to school for a year, 0.5 percent of Uganda's per capita income in 2010 (UN data 2013).

Confusion and suspicion create additional complications. As we discovered through qualitative interviews and feedback from parents, politicians try to drum up support by claiming school fees are illegal. The terms "universal" and "free" education are sometimes used interchangeably. Many parents do not understand the official financing rules. Some believe that the government should provide for all school related expenses. Finally, rumors of corruption can make even knowledgeable parents reluctant to pay.

B. Description of the Intervention

To facilitate families' and children's saving for school, we evaluated four variations of a school-based savings program. The intervention had two primary objectives. First, it sought to facilitate and encourage the practice of children saving for education, and through saving, improve overall academic performance and support students' continued enrollment. The program targeted students in grades five, six and seven, i.e. the last three years of primary school, in order to target students at high risk for dropping out of school.[11] At baseline, the mean student age was 12 (sd dev = 1.52).

We developed and implemented the programs in partnership with the Private Education Development Network (PEDN). PEDN is a Ugandan non-profit organization focusing on youth financial and entrepreneurial education. PEDN comprises five full and part time employees, often supplemented by project specific staff hired as needed. For the savings programs, IPA worked with PEDN to hire a local implementation team of about 10 people.[12]

Each treatment variation included the same core component: a savings account administered through the school, and a program to support and encourage children to use the accounts. During an introductory meeting, the implementation team described the program to a joint meeting of the Parent Teacher Associate, the School Management Committee, and other interested parents. If they all voted to participate, we provided each school with metal lock boxes. A designated teacher assisted by student-elected[13] representatives from each class then managed the program. The implementation team conducted weekly visits to each school to encourage saving and to assist

[11] Uganda follows a 7+2+2 grade structure. Students attend primary school for seven years followed by two years each of lower and then upper secondary school.

[12] This includes only those individuals hired to implement the described programs. It does not include the research staff who conducted the surveys and monitoring visits described below.

[13] The Ugandan educational system classifies children enrolled in primary school as "pupils" and those in secondary school as "students". In this article, we refer to all enrolled children as students.

with accounting procedures. Interested students received a passbook in which their individual savings were recorded, and the designated teacher and the implementation team maintained an official register. Depending on a school's preference, students then deposited money into the lockboxes on a daily or weekly basis.

To provide security and transparency, two padlocks secured each box. Parents elected a representative to keep the key to one lock, while the bank held the other. At the end of each trimester,[14] the two key holders opened the box. The bank representative provided a deposit slip and deposited the funds into the school's account.[15] The accounts did not earn interest. Inflation varied but averaged around 10% in this time period, thus the accounts had a negative real interest rate. After the break between trimesters, the implementation team and bank representatives returned to the school for the payout of the funds. Two representatives signed a withdrawal slip to confirm the withdrawal. The designated teacher, student representatives and our team then distributed the money according to the savings register. At the same time, the implementation team organized a small market at each school where students could purchase school supplies or school services such as practice exams or tutoring sessions (most of the funds went to school supplies, although detailed data are unfortunately not available).[16]

Thus in net, treatment effects from the program (irrespective of variations discussed in a moment) compared to control schools could be a result of several factors. The commitment device is motivated by theories about time inconsistent preferences, but the weekly meetings also serve as a mere reminder to save. Reminders have bene shown to generate higher savings, albeit on a sample of adults (Karlan et al. 2016). The meetings also constituted effectively curriculum, i.e. lessons from the school about the importance of savings. In an evaluation of an in-school program to promote savings in Ghana, based on a program by Aflatoun, school children were found to save more in school but no downstream effects were found on attitudes, aggregate savings, or education outcomes (Berry, Karlan, and Pradhan 2018). One study is of course not dispositive, particularly given it is a different curriculum and setting, and thus it is important to note that the program as a whole could be shifting behavior due to the commitment aspect, attention, or information/signaling mechanisms.

On top of the core treatment above, there were four treatment variations, a 2x2 design: "cash" or "voucher" for the withdrawals, and "Parent Outreach" or "No Parent Outreach".

For the cash treatment arm, students received, in cash, their savings from one trimester at the beginning of the next trimester. They could then spend the funds at their discretion—at the markets

[14] The academic year starts in February and follows a trimester system. Schools run for 12 weeks at a time. Students receive a three week break after the first and second terms, and schools are closed in December, January and February.
[15] Working with the bank, FINCA Uganda, we designed an account for the intervention modeled on a traditional group savings account. We also provided the minimum 5,000 UGX deposit and worked with the school's elected signatories to obtain the documentation required to open the accounts.
[16] Students were allowed to rollover vouchers to future terms, and upon completion of the final year (P7), were allowed to withdraw any remaining balance in cash.

provided on the disbursement day (thus "making it easy" to spend on school supplies) or elsewhere. The voucher treatment arm, on the other hand, employed a stronger commitment — students had to buy educational products or services at the market, on the disbursement day.[17] In both variants, children could also re-deposit their savings for the next trimester.

The Parent Outreach component (implemented halfway through the study) was motivated from qualitative feedback from schools that parents wanted more information about the program, and that there were some incorrect beliefs by some parents as well about the program. The implementation team hosted a meeting for sixth and seventh grade parents. The meetings began by identifying the various stakeholders in primary education, their roles and responsibilities. PEDN then discussed the various ways in which parents could support their children's education. In particular, PEDN explained that in addition to providing a student learning experience, the savings program provided an opportunity for the household. It could be a tool to help families finance their children's education. A snack and soda were provided to encourage attendance.

Section III. Design of the Evaluation

A. Research Design

Figure 1 depicts the timeline for the randomized trial and data collection. We selected 136 primary schools from the Jinja, Iganga, Mayuge, and Luuka districts of the Busoga Region because they predominantly comprised poor rural and peri-urban schools. We then administered a baseline survey and test during the final trimester of 2009. Finally, we randomly assigned schools to receive either the cash treatment (39 schools), voucher treatment (39 schools), or no treatment (58 schools), stratifying by the total normalized score on the baseline exam and by geographic regions called sub-counties.[18]

Following the first randomization, school outreach began. It took two trimesters to recruit the majority of schools, but by the beginning of the third trimester of 2010, 95 percent of the treatment schools had agreed to participate (77 joined, 1 refused).[19] The school that refused to participate did, however, permit data collection, and is thus included for all analysis in the treatment group.

[17] Early in the intervention there was concern that the teachers and community members mobilized to manage the supplies fair were marking up prices to take advantage of the situation. To avoid this, the supplies markets were taken over as part of the intervention. In collaboration with a wholesale distributer, prices were set to match typical alternative prices available to students, and the fairs were organized by the implementing NGOs directly. Managing the fairs as part of the intervention also ensured the essentials supplies were there. This does have implications for scale-up attempts, i.e., whether through explicit management or alternative approach, one likely needs to have a competitive market for supplies available for the students.

[18] In 2010, Uganda included four major jurisdictions called "regions." Spread across the four regions, were 111 "districts." Each district was divided into urban areas known as "municipalities" or rural areas called "counties." Counties were further sub-divided into sub-counties. Depending on the population, a district could have as few as three or as many as thirty or more sub-counties.

[19] When they were not canceled, meetings had to be held with school administrators, the school management committee, and the parent-teacher association for each school. Many were initially reluctant to hold additional meetings.

In 2011, we conducted a second randomization for the parent sensitization program. To isolate the effect of the program while still treating all of the schools, we assigned schools either to the Parent Outreach group who received the intervention in the first trimester of 2011 or to the No Parent Outreach group who received the intervention too late to affect student behavior – immediately before the endline survey in second trimester. Half of the schools in each treatment were assigned to each group. We stratified assignment by the schools' initial treatment group and sub-county, and checked for balance using the demeaned savings rates from 2010.

Finally, we conducted the endline survey and exam during the beginning of the third trimester of 2011.[20]

B. Description of the Datasets

Appendix Table 2 presents the specifics (year of survey, year of student, sample size) of the three datasets created.

We utilize two samples of students, as well as data at the classroom level. The "Attendance Survey" includes all students present in class at baseline and then tracks their attendance in subsequent rounds of data.

Second, we created a representative, longitudinal sample of students identified prior to treatment assignment (the "Student Survey"). These students were tracked regardless of whether or not they continued to be enrolled in the original schools.

The classroom-level data included all classes in grades five, six, and seven. Enumerators counted the number of children present, enrolled and possessing notebooks, math set, uniform, or shoes.[21,22] We conducted these monitoring visits prior to the randomization as part of the baseline and at least once a trimester after the randomization.

The Student Survey includes 4,716 students who completed a baseline survey and aptitude test prior to the randomization. To identify the students for the second (longitudinal) student sample, we compiled a list of all students of the correct ages and grades in September of 2009 (P4 and P5, so that this constituted the students who would be in P5 and P6 for the start of the study).[23] Teachers then classified each student using a five-point scale to rate frequency of attendance. In particular, this allowed us to identify students on the rosters who did not attend school. From the

[20] In 2012, we conducted a second, smaller experiment in which we randomly assigned a fraction of the original control group to receive the cash with sensitization program. We also collected the classroom-level data described below. However, the remaining control group proved too small. The point estimates are consistent with those presented here, but the standard errors are too large to provide meaningful information. These results are available upon request.
[21] The enumerator only counted a student as having each item if the enumerator could see it.
[22] Notebooks cost approximately 200 UGX (0.08 $USD) each. In Uganda, they are usually called "exercise books." A math set costs approximately 1,000 UGX (0.40 $USD) and includes such tools as a ruler, protractor and compass. Uniform and shoes each cost about 6,000 UGX. (2.39 $USD) They are a traditional school requirement.
[23] For a small number of classes, rosters were unavailable. We had to create a list of students based on the students present in class and information provided by the teacher.

set of attending children, we randomly selected 35 students from each school, except for two schools in which we included all students because fewer than 35 students had enrolled.

The baseline survey completed by the students in the longitudinal sample was a 40 minute survey that included questions about their education history, experiences with saving, time preferences, and demographic information. Students also completed an hour-long, 35-question exam covering math, grammar, and reading comprehension. Students in each grade took separate exams based on the national curriculum for their grade.[24]

Students completed an endline survey about two years after the baseline survey. The 40 minute survey included questions about saving behavior, possession of resources like those in the class-level survey, such as uniforms, books, math sets, and shoes. It also included a 60 minute exam in the same three subjects as the baseline exam. The grade level of the endline exam was based on the students' grade at baseline, and all baseline and endline scores are standardized within grade and subject relative to the contemporaneous control distribution. We tracked students regardless of their enrollment status, finding 3,838 of the original respondents.

Finally, we verified the presence of each student in the longitudinal sample during each class-level monitoring visit. This provided an objective measure of students' attendance rates as well as whether students were still enrolled in school in the appropriate grade.

Unfortunately, we lack two datasets which would have been fruitful, but were not feasible to collect: individual level savings data, and specific purchase decisions from the fairs in which school supplies were sold.

C. Econometric Model

Since the random assignment should ensure the orthogonality of treatment assignment and other student characteristics, we estimate the effects of the treatments via ordinary least squares using the following specification:

$$Y_{ijk} = \alpha + \boldsymbol{\tau}'\boldsymbol{treat_j} + \boldsymbol{\delta}'\boldsymbol{X_{ik}} + \varepsilon_{ij}. \quad (1)$$

The variable Y_{ijk} is the dependent variable of interest. We perform estimates at the student and class level. The index i then represents either the student or class in school j and sub-county k. The vector $\boldsymbol{treat_j}$ is a vector of indicator variables for each treatment, and $\boldsymbol{X_{ik}}$ is a vector of control variables. For each estimate, we control for baseline test scores, sub-county fixed effects and the baseline value of the outcome if available. We cluster standard errors by the unit of randomization, the school.

[24] For both the baseline and endline exams, all scores are normalized within grade and subject relative to the contemporaneous control distribution.

Section IV. Internal Validity

In Table 1, we verify the effectiveness of the randomization in creating observably similar treatment and control groups on average. Each row presents estimates for the indicated baseline characteristic. Columns 1-3 provide the sample size for each variable,[25] the pooled treatment mean and standard deviation, and the control mean and standard deviation. Column 4 provides the regression estimates of the difference between the combined treatment group and control group, while Columns 5-8 provide regression estimates of the difference between each treatment group and the control group. All differences are estimated using equation (1), controlling for the sub-counties in which the schools were stratified.

Overall, the differences are minimal, i.e., the assignment to each treatment is orthogonal to a series of baseline variables. Of the 77 estimates presented, only eight are statistically significant: one at the one-percent level, four at the five percent level, and three at the ten percent level. And the overall joint test of significance presented in the bottom row is not significant for any treatment group. Most importantly, the magnitudes of the estimated differences are also all relatively small. Regardless, the main outcome tables are repeated in the appendix with a full set of controls for baseline value of the outcome variable, and results are qualitatively similar.

Table 2 analyzes attrition. First, Row 1 presents the basic test for whether treatment led to differential attrition rates overall. Columns 2 and 3 show that we have identical survey completion rates in treatment and control (81 percent), and Row 1 Columns 5-8 show that we also find no differences in attrition rates across the four treatment groups. However, even though the overall attrition rate is not affected by assignment to treatment, differential attrition could result in differences in the analysis sample frame (i.e., those who complete the endline survey, or take the endline exams). To test for this, we replicate Table 1 analysis on various baseline measures (rows 2 onward). The table is organized similarly to Table 1 (except that the classroom variables are omitted, since there is no attrition at the classroom level). Overall, we find no evidence of compositional effects from differential attrition. Only five tests are statistically significant (out of 63), and the only differences from Table 1 are the estimates for days missed per school term and the time preference measures.

Section V. Results

First, we assess students' savings behavior. In Table 3, we provide two measures of total program savings over the two years: the total per school and per student (using two measures of the latter). Columns 1-4 provide the average for each research group. With a less restrictive measure of the student body (attendance at any point during the two year study period), the two cash payout treatment groups produce average per student savings of 2,196 UGX and 2,342 UGX in the parent outreach and no parent outreach groups, respectively. Using average attendance, these results approximately double to 4,212 and 4,560, respectively. In comparison, the two voucher treatments,

[25] Sample sizes vary because subjects refused to respond to some questions.

with and without parent outreach, show average savings of 1,181 UGX and 1,118 UGX with a less restrictive measure of attendance, and 2,158 UGX and 2,167 UGX using average attendance. The differences between cash and voucher are statistically significant at the 1% level (Column 5), whereas the differences between parent and no parent outreach are not statistically significantly different from zero (Column 6).

We draw three conclusions from the savings data. First, the more restrictive savings vehicle, the voucher treatment, generated significantly less savings than the less restrictive cash treatment. Second, for those in either of the savings treatment groups, we find no additional effect of the parent outreach on savings (and the parental outreach treatment was only implemented within the treatment groups, not within the control group, thus we can estimate its treatment effect in an environment with the savings treatments). Thus this supports the upcoming evidence that while the cash treatment arm led to higher savings, the parent outreach component shifted *how* the funds were spent.[26]

Table 4 examines other key process and intermediate outcomes. First, in Panel A, we examine process outcomes from the program itself, as reported by students in the endline survey. We find that 79 percent of treatment students and only 11 percent of control students were familiar with the Supersavers program. Similarly, 44 percent of treatment group students and only 3 percent of control group students reported saving with Supersavers. There was little difference in program awareness or self-reported participation on the extensive margin across treatment groups. This thus supports the argument that the difference in outcomes is not due to differential marketing or promotion of the program, or differential compliance to experimental protocols, but rather to the attractiveness of the cash versus voucher condition and the parent outreach. We also observe an increase in self-reported in-school savings, and reduction in self-reported out-of-school savings.

In Panel B and C, we then examine intermediate outcomes, i.e., the possession of school supplies (measured both during classroom visits as well as in the endline survey[27]), parental involvement, savings attitudes, and payment of school fees. Analysis of these questions helps to understand the mechanism through which the program worked. We present the results for each, but only find an impact on the possession of school supplies and whether students spent savings on school supplies, suggesting that the other mechanisms are not responsible for the observed impacts, or are poorly measured.

As an indicator of general spending on school related expenses, we collect data on school supplies observable to the survey both in the classroom and endline survey. Panel B presents the results on school supplies that could be observed during classroom visits. The classroom visit school supplies

[26] Both parents and children contributed to the accounts. According to the endline survey, 57 percent of children reported having earned the money that they deposited.

[27] If control group households were buying school supplies earlier than treatment schools, because of the savings accounts and fairs, we would on average observe this because the classroom surveys were conducted during the term, not merely at the end of the term.

index is normalized with respect to the control group mean and standard deviation, and takes the average of four proportions: proportion of students in the classroom possessing uniforms, notebooks, math sets, and shoes.[28] In 2010, none of the treatment groups yields statistically significant increases relative to the control group, and in particular the Cash with Parent Outreach treatment estimate is 0.12 standard deviations (se=0.13). However, oddly, relative to each other the cash parent group is statistically different than the other treatment groups (Column 8). This is partly the result of a negative point estimate for two of the treatment groups. Given that the parental outreach did not occur until the end of 2010, the treatment estimate for Cash with Parent Outreach is not statistically significantly different than zero, and the negative point estimates in the other two treatment groups are not statistically significant, we consider the differential across treatment groups likely spurious.

For 2011, with an additional year of experience implementing the program and after the parent outreach had been fully launched, the Cash with Parent Outreach treatment arm performs considerably better than control, as well as the other three treatments (both when compared individually, as well as when the other treatments are pooled with control). Column 5 shows a 0.32 standard deviation improvement (se=0.14) in the school supplies index, compared just to control. This result is then reinforced by the endline survey, reported in Panel C: The school supplies index from the self-reported survey also shows in Column 5 a 0.11 standard deviation improvement (se=0.06) compared to control, and a statistically significant improvement relative to the other treatment groups (Column 9).[29] We do not however observe any statistically significant shifts in school fees expenditures (albeit with large standard errors), self-reported absence because of failure to pay school fees, or amount paid for most recent tests.[30]

Panel C reports on data from the endline survey on parental involvement and savings attitudes. Although the school supplies and test score impacts are strongest on the Cash with Parent Outreach treatment cell, we do not observe a direct impact on an index of three questions[31] regarding parental involvement in the child's education (or the individual components, as reported in Appendix Table 4b). Furthermore, we do not observe changes in a savings attitudes index of seven

---

[28] Appendix Table 4a presents the results for each of the components of this index.

[29] Appendix Tables 4a, 4b and 4c provide the details for each component of the supplies indexes in Panels B and C. The differences seem to be driven primarily by exercise books, although the individual components analysis is less robust statistically.

[30] This pattern of results is consistent with students' reports on the endline survey regarding the disposition of the saved funds. We observe that students in the cash treatment with the parent sensitization report spending 3.6 percent more of the saved funds on school related expenses than students in the cash treatment without the sensitization. This is not a large enough difference to explain our results, but we are skeptical of the accuracy of these self-reported answers. We observe no differences in the amount of the savings used to purchase food or clothing or given to other family members. The increase in school related expenditures primarily comes from "other" expenses. This difference, however, is likely too small to explain all of the observed increase in school supplies, suggesting that the parent sensitization functioned both to divert students' savings and other unsaved household resources towards school supplies.

[31] The three questions in the Parental Involvement Index are (1) Student thinks parents are responsible for children's education (2) Has your parent come to your school in the past year? (3) Has your parent seen a report of yours from school in the past year? Appendix Table 4b presents the results for each of the components in this index.

questions.[32] This may have implications for long-term change in saving behavior, if one posits that these attitudinal shifts are a necessary component for long-term behavior change, after the active involvement from the NGO and savings program. Alternatively, the measures may be flawed, or the attitudinal changes may be unnecessary; the learned pattern of savings may be possible to change without changing underlying savings attitudes.

Next we turn to test score results in Table 5. We put forward two basic mechanisms here: first, the savings account enables the purchasing of school supplies that are necessary for learning; second, the parental outreach leads the households and children to use the savings accounts to actually spend the saved money on school supplies. This is consistent with the results in Table 4 on the impact on school supplies. And likewise, this mechanism predicts that the Cash with Parent Outreach treatment group should generate the largest (or only) positive impacts. Column 5 indicates that Cash with Parent Outreach improves overall test scores by 0.11 standard deviations (se=0.04). Looking at the components of the test, we find improvements in grammar (0.15 standard deviations, se=0.05) and reading (0.11 standard deviation, se=0.05), but no effect on math. None of the other three treatment groups generates statistically significant improvements compared to the control group, either overall or for any subject. However, we do observe a 0.11 reduction in math test scores for the Voucher with Parent Outreach treatment group. We have no hypothesis, aside from this being a spurious result. Interestingly, these results are consistent with Das et al. (2013) who find similar effects resulting from a $3 per student increase in student supplies. Both sets of results contrast with the traditional view that resources have limited effects on learning (Kremer and Holla 2009). Appendix Table 6 repeats Table 5, with the addition of a control for the baseline value of the outcome variable, and all results hold.

We also examine whether the improved test scores arises through increased attendance or enrollment, but find no evidence for either. Table 6 Panel A reports results on observed attendance as well as an index of three self-reported questions on attendance, and Panel B reports results on enrollment. None of the treatments generates statistically significant improvements relative to the control group.[33]

Last we examine several attitudinal indices, and child labor, in Table 7. Starting with the five attitudinal indexes, we note caution in interpretation: in theory, these may be either intermediate outcomes influenced directly by the treatment(s), or consequences of the shift in resources and test scores. In practice, we observe only two statistically significant shifts out of 20 estimates.[34]

[32] Savings Attitude Index includes 7 statements each of which the student evaluated on a Likert scale, 1-5. All scales were converted after the fact so that higher on the scale meant more positive attitude toward saving. (1) Saving money is not necessary if you live at home with your family. (2) Saving is a good thing to do. (3) Saving is for adults only. (4) My parents or relatives would be proud of me for saving. (5) Managing to save makes me feel happy with myself. (6) It's better to spend money today than to save it for use in the future. (7) Every time I get money I put away some money for saving. Appendix Table 4b presents the results for each of the components in this index.

[33] Appendix Table 4d presents the results for the components of the index.

[34] Appendix Tables 4e – 4f present the results for the components of these indices.

In terms of child labor, critics of financial education for youth posit that introducing children to savings and financial decision-making may have the unintended consequence of focusing their attention on income, and then discourage school attendance in order to work (Varcoe et al. 2005). Berry, Karlan and Pradhan (2018) tests this in Ghana with students of similar age as this study, and finds that a financial education curriculum along with a savings box (but no directive or facilitation of using the savings for education expenses) did lead to higher child labor, whereas if a social values component was added to the financial education curriculum, there was no impact on child labor. In our setting, we find no impact from the program on child labor, either hours worked or total wages. Overall, the estimates from Tables 6 and 7, combined with the other outcomes, indicate that the observed effects on learning occur solely through changes in available supplies rather than changes in attitude or participation.

To explore econometric robustness and robustness to risks of baseline imbalance, Appendix Tables 5 and 6 present the core results (Tables 4 to 5), except with covariates for all baseline values of outcome variables. We find no changes in the core results.

Section VI. Conclusion

Weaker rather than stronger commitments can, in some instances, prove more effective. In the context of an educational savings program, we find that families and children save more under a weaker savings strategy in which funds are not dedicated to educational expenses upon deposit than they do under a strict commitment savings program in which all deposits are dedicated to educational expenses. The purpose of commitment savings devices is to intentionally limit the use of deposited funds. In some contexts, however, such services may need to strike a balance between providing sufficient limitations to make the savings mechanisms useful while being careful not to make the limitations so severe that they deter savings. The stricter limitations may work worse for behavioral reasons (e.g., wanting option value or judgment to change own's mind) or for institutional reasons (e.g., not trusting the institution that is offering the commitment device). In our setup, for example, the voucher treatment may work worse because individuals do not trust that proper and fairly priced school supplies will be available. However, although this seems plausible in the first year, we believe by the second year, after seeing the program work for a year, households should have learned that the right school supplies would be available at a reasonable price. Understanding the nature of this trade-off between strict and loose commitment is an important direction for future research.

When combined with a parent sensitization program, we find that families and children in the cash arm spend their savings on educational expenses (school supplies).[35] This does not, however, alter

[35] Although we find that the voucher treatment led to about half the deposits as that of the cash arm, we do not find that school supplies increased by half. We posit two possible explanations. First, although the point estimate is close to zero, we cannot reject, statistically, a point estimate of half of that of the cash treatment effect. Second, the voucher treatment arm may have led to a reduction in school supplies through an anchoring effect (if the amount saved in vouchers was smaller for some than they would have spent otherwise).

school participation – we find no effects on enrollment or attendance – but does improve students' scores on a basic math and language test by 0.11 standard deviations. This suggests that financial constraints may play an important role in students' academic performance and that understanding the role of families' financial decision process may be an important element in understanding the overall production process of education.

On a practical level, several implementation issues are important to explore. As a program designed to improve student learning, treatment effects of this magnitude are large compared to other evaluations of interventions designed to provide resources to schools or directly to children (Jameel Poverty Action Lab 2014), but they are small relative to many other types of programs (most notably, for example, programs that provide additional resources while also changing pedagogical strategies). Taking the programs relatively low cost (2.24 USD per student per year) into account using the methodology proposed by Dhaliwal et. al. (2014), however, the program delivers learning gains at a cost of 1.49 USD per tenth of a standard deviation or 6.71 standard deviations per 100 USD.[36] This is very competitive relative to other programs. Relative to the 27 studies compared by J-PAL (2014), only four produce improvements in test scores more cost-effectively.

In terms of encouraging family savings, the program costs were high relative to the savings generated. However, if the program generated long term savings behavior change, then between the continued savings and the improvement in educational outcomes, it would surpass typical cost benefit calculations. Because we do not observe changes in attitudes, however, we cannot confidently predict that the long term impacts will sustain themselves (although lack of attitude changes does not mean the results will not sustain themselves: attitudes are difficult to measure and may merely reflect noise, and furthermore the habit and pattern of saving could change and sustain without changing attitudes). On the cost side, it may be possible to reduce costs, particularly with implementation via mobile banking. This would obviate the need for physical transfer of cash to a bank, and would lower the risk of theft from keeping cash in a (albeit locked) box at the school. However, if the group nature of the intervention (i.e., the public and communal training) was an important element for take-up (through mimicking of peers, or learning from peers) and adherence (through monitoring and potential for social recognition), then a mobile banking implementation may lose that visual classroom element. Although these peer mechanisms were not emphasized in the training and implementation of the program, the fact that the savings were done publicly may have had such an effect.

On a more theoretical level, these results open up many related questions. How does the optimality of looser versus stricter commitments depend on whether savings is long term or short-term? If one is saving for potentially short-run needs, such as a buffer stock, looser knots may be optimal; whereas long-term savings, such as for retirement, may require tighter commitments as the benefits from savings are too remote. Also with respect to timing, are external interventions of this sort

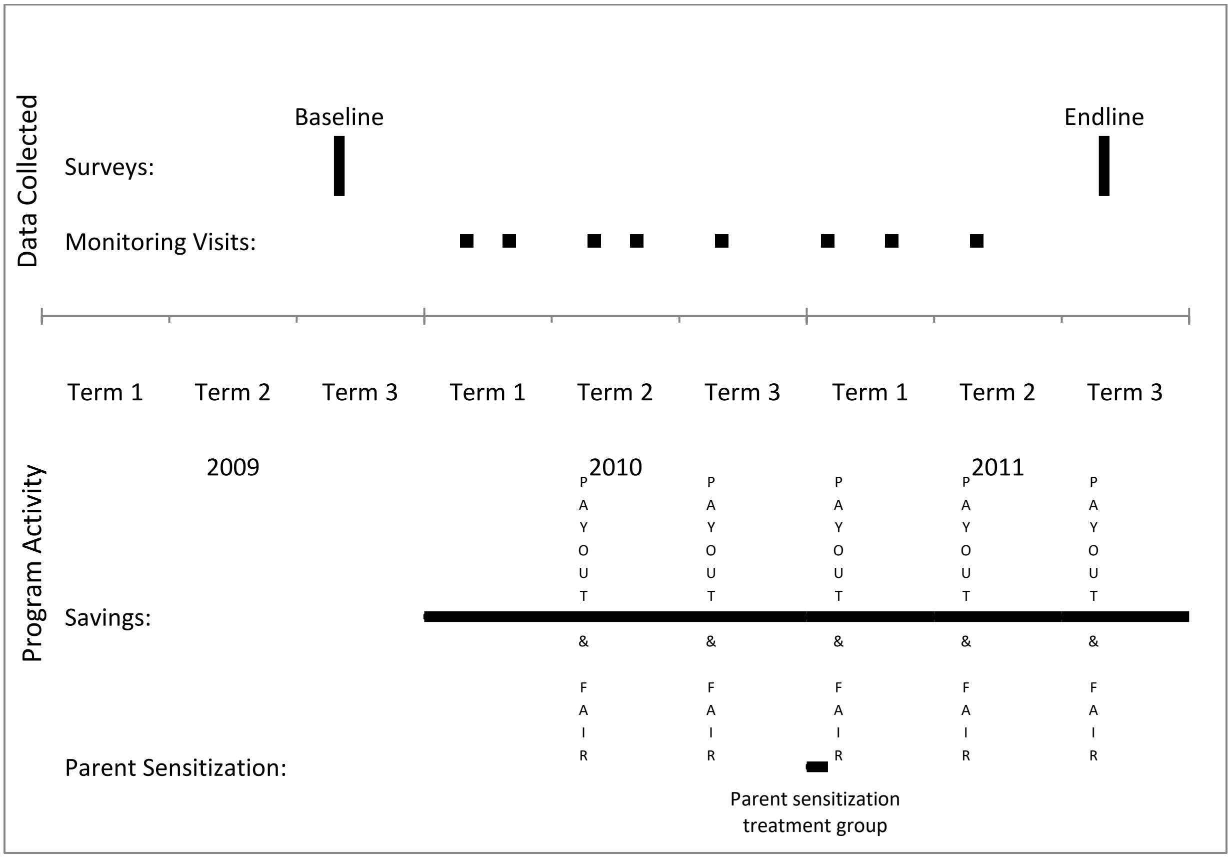
[36] Estimates are provided in 2011 USD.

effective in changing long term behavior, i.e., does the psychic cost of deviation persist, even without an outsider-led intervention?

Questions also persist regarding how such interventions influence intra-household dynamics. Did the intervention shift the preferences of the child, or the parents, or both, and what does this imply for intra-household cross-generational bargaining issues?

Lastly, design issues may be critical for such a program to work. For example, how critical was the timing element of the “soft” commitment device, i.e., the fact that the school supplies were immediately available for purchase at the time of withdrawal? If that was critical, it is a ringing endorsement for the “make it easy” mantra, and also implies that the soft commitment device may have worked for reasons elaborated on in Mullainathan and Shafir (2013), because it increased the attention of individuals to educational expenses at exactly the right moment, when they had cash in their hands.

Figure 1: Research Timeline



References

Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. "Commitment vs. Flexibility." *Econometrica* 74 (2):365–96. https://doi.org/10.1111/j.1468-0262.2006.00666.x.

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

———. 2010. "Female Empowerment: Further Evidence from a Commitment Savings Product in the Philippines." *World Development* 38 (3):333–44.

Avvisati, Francesco, Marc Gurgand, Nina Guyon, and Eric Maurin. 2013. "Getting Parents Involved: A Field Experiment in Deprived Schools." *The Review of Economic Studies*, August, rdt027. https://doi.org/10.1093/restud/rdt027.

Barrera-Osorio, Felipe, Marianne Bertrand, Leigh Linden, and Francisco Perez-Calle. 2011. "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia." *American Economic Journal: Applied Economics* 3 (2):167–95.

Benabou, Roland, and Jean Tirole. 2004. "Willpower and Personal Rules." *Journal of Political Economy* 112 (4):848–87.

Berry, James, Dean Karlan, and Menno Pradhan. 2018. "The Impact of Financial Education for Youth in Ghana." *World Development* 102 (February):71–89. https://doi.org/10.1016/j.worlddev.2017.09.011.

Beshears, John, James J. Choi, Christopher Harris, David Laibson, Brigitte C. Madrian, and Jung Sakong. 2017. "Which Early Withdrawal Penalty Attracts the Most Deposits to a Commitment Savings Account?" *Working Paper*.

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang. 2016. "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi." *Economic Development and Cultural Change* 64 (2):187–220. https://doi.org/10.1086/684014.

Burke, Jeremy, Jill Luoto, and Francisco Perez-Arce. 2017. "Soft versus Hard Commitments: A Test on Savings Behaviors." *Journal of Consumer Affairs*, November, n/a-n/a. https://doi.org/10.1111/joca.12170.

Clement, Douglas. 2013. "Interview with Richard Thaler." *The Region*, September 2013.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Teset Scores." *American Economic Journal: Applied Economics* 5 (2):29–57.

De Arcangelis, Giuseppe, Majlinda Joxhe, David McKenzie, Erwin Tiongson, and Dean Yang. 2014. "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." *The World Bank: Policy Research Working Paper*, May.

Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch. 2014. "Comparative Cost-Effectiveness to Inform Policy in Developing Countries." In *Education Policy in Developing Countries*, edited by Paul Glewwe. Chicago, IL: University of Chicago Press.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101 (6):2350–90. https://doi.org/10.1257/aer.101.6.2350.

Dupas, Pascaline, and Jonathan Robinson. 2013. "Why Don't the Poor Save More? Evidence from Health Savings Experiments." *American Economic Review* 103 (4):1138–71. https://doi.org/10.1257/aer.103.4.1138.

Giné, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang. 2017. "Revising Commitments: Field Evidence on the Adjustment of Prior Choices." *The Economic Journal*, March. https://doi.org/10.1111/ecoj.12378.

Hidalgo, Diana, Mercedes Onofa, Hessel Oosterbeek, and Juan Ponce. 2013. "Can Provision of Free School Uniforms Harm Attendance? Evidence from Ecuador." *Journal of Development Economics* 103:43–51.

Independent Evaluation Group (IEG). 2007. "Fall out from the 'Big Bang' Approach to Universal Primary Education: The Case of Uganda." World Bank. http://www.worldbank.org/oed/education/uganda.html.

Jameel Poverty Action Lab. 2014. "Student Learning." M.I.T. http://www.povertyactionlab.org/policy-lessons/education/student-learning.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science* 62 (12):3393–3411.

Karlan, Dean, Aishwarya Lakshmi Ratan, and Jonathan Zinman. 2014. "Savings by and for the Poor: A Research Review and Agenda." *Review of Income and Wealth* 60 (1):36–78. https://doi.org/10.1111/roiw.12101.

Kremer, Michael, and Alaka Holla. 2009. "Pricing and Access: Lessons from Randomized Evaluations in Education and Health." In *What Works in Development? Thinking Big and Thinking Small*, edited by Jessica Cohen and William Easterly. Brookings Institution Press.

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means so Much*. New York: Times Books, Henry Holt and Company.

Murphy, Paul, Carla Bertoncino, and Lianqin Wang. 2002. "Achieving Universal Primary Education in Uganda: The 'Big Bang' Approach." 24107. Education Notes. Washington, D.C.: The World Bank. http://documents.worldbank.org/curated/en/2002/04/12947917/achieving-universal-primary-education-uganda-big-bang-approach.

"Opportunities Lost: The Impact of Grade Repetition and Early School Leaving." 2012. Global Education Digest 2012. UNESCO Institute for Statistics.

Piper, Benjamin. 2010. "Uganda Early Grade Reading Assessment Findings Report: Literacy Aquisition and Mother Tongue." Makerere Institute for Social Research.

Shefrin, H., and R. Thaler. 1992. "Mental Accounting, Saving, and Self-Control." In *Choice Over Time*. New York: Russell Sage Foundation.

Thaler, Richard H, and Cass R Sunstein. 2009. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New York: Penguin Books.

UN data. 2013. "Uganda." United National Statistical Division. http://data.un.org/CountryProfile.aspx?crName=Uganda.

UNESCO. 2013. "Data Centre." Institute for Statistics. 2013. http://www.uis.unesco.org/Pages/DataCentre.aspx.

Varcoe, Karen P., Allen Martin, Zana Devitto, and Charles Go. 2005. "Using a Financial Education Curriculum for Teens." *Journal of Financial Counseling and Planning* 16 (1):63–71.