Using Lotteries to Encourage Saving: A Pre-Analysis Plan

Merve Akbas[*], Dan Ariely[†], and Chaning Jang[‡]

September 30, 2015

Abstract

This paper describes the analysis plan for a randomized controlled trial evaluating the impact of a lottery-linked savings program on informal workers in Nairobi, Kenya. In 2014, we designed and administered a mobile savings program to 311 randomly sampled residents from the Kibera and Viwandani settlements. This study aims to test the effect of using lotteries to increase demand for commitment savings by comparing the lottery-linked savings program with a standard interest-bearing program. This plan outlines our experimental design, identification strategy, and outcomes of interest.

[*]Department of Economics, Duke University. merve.akbas@duke.edu

[†]Fuqua School of Business, Duke University. dan@danariely.com

[‡]Department of Psychology, Princeton University and Busara Center for Behavioral Economics. cjang@princeton.edu

1 Introduction

Although access to savings accounts for the poor has improved in recent years, demand for savings and usage of savings accounts is a stumbling block for full financial inclusion (Dupas et al. 2014). Researchers have developed and tested new and innovative ways to spur demand – including text messages, puzzle pieces, lockboxes and deposit collectors, with varying degrees of success (Ashraf, Karlan, and Yin 2006a) (Ashraf, Karlan, and Yin 2006b) (Dupas and Robinson 2013) (Karlan et al. 2010). Our study proposes to test a scheme that has been around for thousands of years and seen as anathema to savers – lotteries. Although lotteries are derisively referred to as a "tax on the poor", lottery-based savings products have attracted savers for hundreds of years (Murphy 2005). Versions of these accounts exist in various forms around the globe (Kearney et al. 2010), but none have been rigorously evaluated as a method of attracting deposits from the poor. Recent experimental evidence has shown that lottery incentives holds promise in increasing savings among the poor, unbanked, and gambling prone population (Atalay et al. 2014) (Filiz-Ozbay et al. 2014).

This study aims to test the effect of using lotteries to increase demand for commitment savings. We created a savings mechanism whereby the act of saving enters customers into a lottery. Lottery-linked deposit accounts work by pooling interest payments and distributing them in bulk to fewer recipients rather than piecemeal to all. In our mechanism, winners of the lottery receive additional funds into their savings account, while those who do not win receive no interest.

We are primarily interested in learning if we can use lottery devices as a way to encourage the poor to save. Specifically, we are interested in understanding how behavioral nudges and the attractiveness of gambling changes a person's perception towards act of saving. Our first hypothesis is that lottery linked deposit accounts will have significantly higher take up and usage rates compared to similarly structured interest bearing accounts with an equivalent expected return. Our second hypothesis is that framing matters. We believe that framing the product that maximizes regret will significantly increase take-up and usage rates compared to an equivalent product that does not mention regret.

2 Savings Program

We implemented our mobile-phone based savings program over Safaricom's *Sambaza* airtime sharing service. Using *Sambaza*, Safaricom users can send airtime to each other free of charge. Subjects saved into our program by sending airtime to a designated project phone that held the airtime in an account for each user. We chose to run our savings program via airtime rather than M-Pesa to avoid the fees involved with conversion to and from M-Pesa.

Subjects received two SMS messages every morning after the first morning of the project period. The first message was an end-of-day message that reported how much the subject saved the previous day, how much the subject earned through interest or winnings, and their total balance. An hour later, subjects received a beginning-of-day message encouraging them to save that day. Subjects were allowed to send in savings at any time but any savings sent in after the end-of-day message would be counted towards the next day's total. We used a custom-developed administrative system to manage the savings program. This system logged airtime sent to our project phone, maintained an internal ledger of balances, sent automated SMS confirmations after every transaction, and conducted the daily lottery game.

Subjects were enrolled in the savings program for a total of 60 days, split into consecutive 30-day periods. After the first 30 days, subjects were allowed to withdraw any amount of their savings up to the total balance. Outside of this opportunity, regular withdrawals were not allowed. If a subject wished to withdraw their balance, they were required to withdraw from the entire experiment.

At the end of our experiment, we returned subjects' savings and accumulated interest or winnings via an M-Pesa transfer. This M-Pesa transfer included the extra withdrawal fees needed to cash out an amount equal to the subject's full account balance. Therefore, subjects paid no explicit fees to participate in our program.

2.1 Treatment groups

Subjects were randomized into one of three treatment groups and had the chance to earn either daily interest (in the form of matching) or play a daily lottery depending on assignment.

2.1.1 Interest (control group)

Subjects in the control group participated in a standard, interest-bearing savings program. Subjects earned a 5% matching contribution on any amount that they saved in a particular day.

2.1.2 Lottery

After saving a non-zero amount, subjects earned a lottery ticket - transmitted via text message, which could win a cash prize in proportion to the amount they saved. A lottery ticket was a random sequence of four numbers between 1 and 9, inclusive. Each day, our administrative system randomly generated a winning sequence of four numbers. Prizes were awarded according to how well a subject's lottery numbers matched the winning numbers. If the first *or* second numbers matched, a 10% match of savings was awarded. If *both* the first *and* second numbers matched, a 100% match of savings was awarded. Finally if all numbers matched, a prize of 200 times daily savings was awarded. The earnings on this lottery ticket were equal in expectation to the 5% match earned in the control group. Our system processed the matching of lottery numbers and entered winnings into the internal ledger. Subjects could only earn one lottery ticket per day.

2.1.3 Lottery+Regret

This scheme is similar to the Lottery treatment but subjects in this third group were sent lottery tickets in their beginning-of-day text message. These tickets only became redeemable, however, after subjects had saved a non-zero amount that day. Subjects with winning lottery tickets who did not save that day did not win money from their lottery ticket. However, they were informed of the amount they *would have* won in their end-of-day message the next morning.

3 Study Design

This study was conducted in conjunction with the Busara Center for Behavioral Economics in Nairobi, Kenya. We recruited 311 subjects through Busara's subject pool. Participants were first invited to the lab at Busara where they completed a computerized questionnaire and played various behavioral games. The following outlines the schedule of tasks during the lab portion of the study:

1. Risk preference elicitation
2. Time preference elicitation
3. Lottery task
4. Locus of control
5. Baseline questionnaire

Following the lab session, subjects were randomly assigned to one of the three treatment groups. Subjects were then taught about their assigned savings program. Each subject was given KSH 20 airtime credit and asked to practice saving using the Sambaza mechanism. Subjects were then sent home with business-card sized handouts which described their savings program. We provided subjects simple instructions for saving and listed the number to our project phone. This was the number through which the savings program operated that also functioned as a help line for subjects.

Lab sessions took place over five weeks in May and June of 2014. Subjects were enrolled in the our savings program for two consecutive periods of 30 days starting from the day of a subject's lab session. On a subject's 30th day, a field officer called them and asked if they wished to withdraw any amount of their balance. Subjects who requested withdrawals were sent M-Pesa tranfers equal to their request plus the M-Pesa withdrawal fee. These withdrawals were recorded in our system's ledger.

Following this, subjects moved on to their second 30-day savings period. Subjects were called and notified a few days before the end of their second 30-day period that the program would be ending soon. After receiving the end-of-day message on their 60th day, subject were unenrolled from the program and were no longer allowed to save. Field officers called subjects to confirm final balances and sent M-Pesa transfers equal to total balance plus withdrawal fee shortly after. All subjects had completed the program by August 2014. In September 2014, we called subjects and conducted an endline survey. We obtained endline surveys for all but 27 of the 311 subjects.

4 Identification Strategy

Our study has two main goals. The first goal is to measure treatment effects on a number of key outcome variables outlined in section 5. The second goal is to explore the link between actual saving behavior and lab measures such as risk and time preferences. Below we explain our analysis plan for each of these goals separately.

4.1 Treatment effects

We use the following econometric specification for basic identification of the treatment effect.

$$y_{i,t} = \beta_0 + \beta_1 LOTTERY_i + \beta_2 REGRET_i + \delta y_{i,t=0} + \varepsilon_i \tag{1}$$

Where $Y_{i,t}$ refers to a host of outcome variables for individual i, $LOTTERY$ indicates assignment to the lotter group, and $REGRET$ indicates assignment to the lottery with regret framing group. The omitted group is the interest group. β_1 and β_2 identify the treatment effect. Following McKenzie (2012), where possible we will estimate equation 1 conditional on the baseline level of the individual outcome $y_{i,t=0}$ to improve statistical power. We will use an F-test to compare the joint effect of both lottery treatments to the comparison group. We also extend this analysis by re-analyzing the treatment effects on a restricted timeframe of the first 30 days (the first part of the trial), 15 days and 10 days periods.

4.1.1 Heterogeneous treatment effects

We explore the extent to which the savings program produces heterogeneous treatment effects. We use the following specification for this analysis:

$$
\begin{aligned}
y_{i,t} =& \beta_0 + \beta_1 LOTTERY_i + \beta_2 REGRET_i \\
& + \beta_3 LOTTERY_i \times X_{i,t=0} + \beta_4 REGRET_i \times X_{i,t=0} \\
& + \beta_5 X_{i,t=0} + \delta y_{i,t=0} + \varepsilon_i
\end{aligned}
$$

Where $X_{i,t=0}$ refers to the interaction term for individual i at baseline $t = 0$. To do this, we first look at the interaction between treatment effects and gender, income level, and previous savings patterns, education level, marital status, as well as risk, time and gambling preferences. As a second step, we investigate whether there is any subgroup for which we observe particularly different treatment effects than the rest of the sample. To this end, we identify subgroups based on key independent variables such as income level and risk preferences and run restricted regressions. Restricted regressions will take the form:

$$
(y_{i,t} \mid G_{i,t} = 1) = \beta_0 + \beta_1 LOTTERY_i + \beta_2 REGRET_i + \delta y_{i,t=0} + \varepsilon_i \tag{2}
$$

Where $G_{i,t}$ refers to the restricted subgroup. The subgroups will be defined after we analyze the general features of our sample.

4.1.2 Seemingly unrelated regression

We might expect that the errors for each of these regression are correlated. Instead of estimating these equations separately, we can estimate the system of seemingly related regressions (SUR) to improve the precision of the coefficient estimates (Zellner 1962). SUR estimation is equivalent to OLS when the error terms are in fact uncorrelated between regressions or when each equation contains the same set of regressors. Simultaneous estimation allows us to perform Wald tests of joint significance on the treatment coefficients.

4.1.3 Multiple inference

We will account for multiple inference by using a family-wise p-value adjustment. We will report both unadjusted p-values as well as p-values corrected for multiple comparisons using the Family-Wise Error Rate. We adjust the p-values of our coefficients of interest for multiple statistical inference following the procedure developed in Anderson (2008).

4.1.4 Sample attrition

To assess whether attrition potentially confounds our results, we proceed as follows. First, we define $attrit_i = 1$ if individual i was surveyed at baseline but not at endline, and zero otherwise. We then assess the severity of attrition using three approaches. First, equation

3 estimates whether the magnitude of attrition is different for treatment groups and the control group:

$$attrit_i = \beta_0 + \beta_1 LOTTERY_i + \beta_2 REGRET_i + \varepsilon_i \tag{3}$$

Where $LOTTERY$ refers to the group assigned to the lottery treatment, and $REGRET$ refers to the group assigned to the lottery+regret treatment. The omitted group is the interest (control) group. Second, equation 4 assesses whether attrition individuals are different in terms of a comprehensive range of baseline characteristics:

$$y_{i,t=0} = \beta_0 + \beta_1 attrit_i + \varepsilon_{i,t=0} \tag{4}$$

And third, equation 5 measures whether the baseline characteristics of attrition individuals in the treatment group are significantly different from those in the control group. The sample for regression will be restricted to attrition individuals:

$$(y_{i,t=0} \mid attrit_i = 1) = \beta_0 + \beta_1 LOTTERY_i + \beta_2 REGRET_i + \varepsilon_{i,t=0} \tag{5}$$

If worrying levels of attrition are found, we will adjust for the potential effect of such attrition using bounding techniques, matching, or modeling the selection process.

4.2 Transaction analysis

With daily data on savings behavior, we will be able to explore the temporal dynamics of the treatment effect. That is, we can observe how the intervention impacts the dynamics of savings behavior over the sixty-day period. This analysis will be used to explore the following questions, for example. In the panel data, we have outcome measures y_{it} for individual i for $t = 1, \ldots, 60$, where $t = 1$ indexes the first day of the intervention and $t = 60$ indexes the end of the savings period. We will estimate the following autoregressive model:

$$y_{it} = \beta_0 + \beta_1 LOTTERY_i + \beta_2 REGRET_i + \sum_{k=1}^{K} [\beta_{k+2} y_{i,t-k} + \beta_{k+3} LOTTERY_i y_{i,t-k} + \beta_{k+4} REGRET_i y_{i,t-k}] + \varepsilon_{it} \tag{6}$$

Where y_{it} is decision to deposit, amount of deposit, or withdrawals at day t. We will use the Akaike information criterion to determine number of lags K to include. The standard errors will be corrected for auto-correlation by clustering at the individual level. We might also be interested in whether or not there is evidence for the hot-hand fallacy or gambler's fallacy amongst our participants in the lottery conditions. This can be estimated using the following distributed lag model for the subsample of respondents in the lottery and regret groups:

$$y_{it} = \beta_0 + \sum_{k=1}^{K} X_{i,t-k} + \varepsilon_{it} \tag{7}$$

Where y_{it} is decision to deposit, amount of deposit, or withdrawals at day t and $X_{i,t-k}$ measures lottery winnings or whether subject i won the lottery at day $t-k$.

4.3 Baseline correlates of savings behavior

A second goal of our study is to explore the correlation between preferences that can be measured in the lab and savings behavior. To do this, we will run a regression using baseline outcomes as predictors of total savings:

$$y_{it} = \beta_0 + \delta y_{i,t=0} + \varepsilon_{it} \tag{8}$$

Baseline indicators that we are interested include risk, time, and gambling preferences. Where applicable, we will use data reduction methods such as principal component analysis and factor analysis. These analyses are correlational and will be used to inform potential mechanisms or future avenues of research.

5 Outcomes of Interest

We estimate treatment effects on measured savings behavior. The main outcome variables we are interested in are:

1. Average savings over the entire study period.

2. Average savings over the first and second 30-day period.

3. Average number of active days and average number of transactions.

4. Average length of the streaks, i.e. the highest number of consecutive days with a positive balance for each person.

Aside from the overall savings behavior, we additionally estimate the effect of the program on:

1. Monthly self-reported savings balance

2. Savings engagement (whether you currently save, engagment with ROSCA)

3. Displacement of other saving mechanisms

4. How often subject discussed savings program with family and friends

5. Trust in the savings program
6. Satisfaction with saving behavior in the program
7. Continuation with the savings program
8. Self-perception as a saver
9. Self-reported gambling behavior
10. Trust in the savings program

References

Anderson, Michael L. 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 (484): 1481–1495.

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006a. "Deposit collectors." *Advances in Economic Analysis & Policy* 5, no. 2.

———. 2006b. "Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines." *The Quarterly Journal of Economics*, pp. 635–672.

Atalay, Kadir, Fayzan Bakhtiar, Stephen Cheung, and Robert Slonim. 2014. "Savings and prize-linked savings accounts." *Journal of Economic Behavior & Organization* 107:86–106.

Dupas, Pascaline, Sarah Green, Anthony Keats, and Jonathan Robinson. 2014. "Challenges in banking the rural poor: Evidence from Kenya's western province." In *African Successes: Modernization and Development.* University of Chicago Press.

Dupas, Pascaline, and Jonathan Robinson. 2013. "Why don't the poor save more? Evidence from health savings experiments." *The American Economic Review* 103 (4): 1138–71.

Filiz-Ozbay, Emel, Jonathan Guryan, Kyle Hyndman, Melissa Schettini Kearney, and Erkut Y. Ozbay. 2014. "Do Lottery Payments Induce Savings Behavior: Evidence from the Lab (Digest Summary)."

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2010. "Getting to the top of mind: How reminders increase saving." Technical Report, National Bureau of Economic Research.

Kearney, Melissa Schettini, Peter Tufano, Jonathan Guryan, and Erik Hurst. 2010. "Making savers winners: An overview of prize-linked savings products." Technical Report, National Bureau of Economic Research.

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

Murphy, Anne L. 2005. “Lotteries in the 1690s: Investment or Gamble?” *Financial History Review* 12 (02): 227–246.

Zellner, Arnold. 1962. “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias.” *Journal of the American Statistical Association* 57 (298): 348–368 (June).