

More Than Money: Effects of Cash Transfer Narratives on Agency and Self-Investment

Justin Abraham ^{*†‡}, Nicholas Otis^{†‡}, Catherine Thomas^{‡‡}, Hazel Markus [§], Greg Walton [¶]

August 16, 2017

Abstract

This document describes the pre-analysis plan for a randomized experiment examining the effects of narratives accompanying unconditional cash transfers on self-concept and economic behavior. We provided one-time, unconditional cash transfers to residents of two informal settlements in Nairobi and randomly assign participants to receive one of three messages. Respondents will receive a non-binding message stating that the cash is intended for 1) poverty alleviation, 2) individual empowerment, or 3) community empowerment. We then collected self-reported measures of self-efficacy, stigma, and affect and behavioral measures of future-orientation, self-investment, and program support. This pre-analysis plan outlines our hypotheses, the schedule of experimental tasks, and our empirical strategy. In order to guarantee transparency and bind ourselves from fishing for results, we will pre-register the scripts to be used for data analysis.

*University of California, San Diego. ^{††}Contributed equally.

[†]University of California, Berkeley. ^{††}Contributed equally.

[‡]Stanford University. ^{††}Contributed equally.

[§]Stanford University.

[¶]Stanford University.

Contents

1	Research Design	3
1.1	Sampling	3
1.2	Statistical power	3
1.3	Experimental procedure	3
1.4	Treatment	4
2	Data	5
2.1	Self-efficacy module	5
2.2	Stigma module	5
2.3	Affect module	6
2.4	Video selection task	6
2.5	Savings decision task	7
2.6	Support for organization and message	7
2.7	Messages evaluation	7
2.8	Sociodemographic questionnaire (9 items)	8
3	Empirical Analysis	8
3.1	Randomization balance checks	8
3.2	Treatment effect of cash transfer messages	8
3.3	Covariate adjustment	9
3.4	Heterogeneous treatment effects	10
3.5	Randomization inference	10
3.6	Multiple testing adjustment	10
A	Consent Form	13
B	Source Code	14
B.1	Packages	14
B.2	User-defined functions	14
B.3	Data cleaning	16
B.4	Randomization checks	19
B.5	Treatment effect	19
B.6	Covariate-adjustment	20
B.7	Heterogeneous treatment effects	21

1. Research Design

1.1 Sampling

This study was conducted in conjunction with the Busara Center for Behavioral Economics (Busara) in Nairobi with 565 participants residing in Kibera and Kawangware, two of Kenya’s largest informal settlements (Haushofer et al. 2014). Treatment and data collection were conducted by Busara Center enumerators with participants from Kibera and Kawangware in lab and field settings, using tablets to display audio and video media and record participant responses. This section outlines the sampling procedure used in the experiment.

Participants were recruited from the Busara participant pool and were asked to participate in the survey in one of the lab settings. There were seven survey locations used throughout the study period. Table 1 summarizes these areas.

Table 1: Survey locations

Area	Survey location
Kibera	AIC Church
Kibera	Kibera Immanuel Technical Institute
Kibera	Kibera Labour Hall
Kibera	Kibera Chonesus Hall
Kibera	Busara Center (Lab)
Kawangware	Kawangware Pastor Ken’s Hall
Kawangware	Kawangware CDF Hall

Participants were recruited to participate in the study if they met the following eligibility criteria:

1. Member of the Busara Center’s participant pool
2. Resident of Kibera or Kawangware
3. Owns a working phone and an M-Pesa account registered under the participant’s name

1.2 Statistical power

To achieve power of 80% for an estimated effect size of 0.30 SD, the required sample size is 525 participants, with 175 in each of the treatment arms.

1.3 Experimental procedure

The survey questionnaire was delivered by enumerators to participants in Swahili or English, as preferred by the participant. The following summarizes the schedule of tasks in the questionnaire.¹

1. Consent agreement
2. Cash transfer and message (randomized)
3. Self-efficacy module
4. Stigma module
5. Affect module

¹We will use a single survey instrument, programmed with Qualtrics, for treatment delivery and subsequent data collection.

6. Video selection task
7. Savings task
8. Message evaluation
9. Support for organization and message
10. MacArthur Subjective Social Status Ladders (normalized)
11. Sociodemographic module

1.4 Treatment

At the outset of the survey, eligible and consenting participants were told they would be receiving an unconditional cash transfer of KES 400 (USD PPP 10.5) from an organization unaffiliated with the Busara Center.²

Participants were randomly assigned by the survey software within enumerator³ to receive one of three messages introducing the purpose of the cash transfer. The three messages had a similar structure, but we experimentally varied the described purpose of the cash transfer. Specifically, we changed the stated goals of the organization, rationale for providing money, assumptions about recipients, and expectations and goals for the use of the transfer. In the poverty alleviation message, the payment was described as a means to meet basic needs. The individual empowerment message described the payment as a means toward individual goals and advancement. The community empowerment message described the payment as a means toward goals advancing one’s family and the community for community advancement. Participants listened to the message twice in their preferred language (English or Swahili) with pre-recorded audio clips or as read by the enumerator.

After hearing the message once, senior enumerators were alerted to use a project MPESA account to send USD PPP 10.5 to the participant via the mobile money system M-Pesa.⁴ Enumerators were instructed to confirm receipt of the payment on the respondent’s phone, after which enumerators played the message a second time.⁵ Then, enumerators led the respondent’s through a series of questions on how they view the transfer. In particular they are asked questions on their current needs (in the “poverty alleviation” arm) or goals (in the “individual empowerment” and “community empowerment” arms), the name they would assign to these funds (for example “education fund”), how receipt of these funds would affect their relationship with others, and their perceived goal of the organization.

Below, we list the three treatment messages that respondents received:

Poverty alleviation message: *The goal of this Poverty Alleviation Organization is to alleviate poverty and reduce financial hardship among the poor. This organization believes that people living in poverty should be given income support to help them meet their basic needs. This organization aims to help promote a decent standard of living among the poor and help them deal with emergencies. Thus, the Poverty Alleviation Organization gives financial assistance to people like you, to help them make ends meet. For example, with the financial assistance, people*

²This study was conducted with Kenyan shillings (KES). We report USD values calculated at purchasing power parity using a conversion factor for private consumption of 38.15 in 2013. The price level ratio of PPP conversion factor (GDP) to KES market exchange rate for 2011 was 0.444.

³We evenly assigned treatment groups to achieve balance in group size.

⁴For more information on M-Pesa, we refer the reader to Jack and Suri (2011) and Mbiti and Weil (2011).

⁵For the first day and a half of the survey period (for approximately approximately 100 respondents), we used a system in which the respondent texted a code which enabled the direct transfer of the money to their account. Due to technical difficulties, we were required to change to the above system.

might be able to struggle less to afford basic needs, like paying off debts, paying rent, and buying clothes and food. Now we are going to send you 400 KSh. Please note that this is a one-time transfer of financial assistance.

Individual empowerment message: *The goal of this Individual Empowerment Organization is to promote individuals' potential to create a better future for themselves. The organization believes that individuals are wise and know best how to help themselves become self-reliant/independent if they have the financial resources to do so. This organization aims to empower individuals to pursue their personal interests and create their own path to independence. Thus, the Individual Empowerment Organization gives financial resources to individuals, like you, to enable them to invest in their personal goals. For example, people might use their unique talents to start a self-run business, invest in job training courses, or create art. Now we are going to send you 400 KSh. Please note that this is a one-time transfer of financial resources.*

Community empowerment message: *The goal of this Community Empowerment Organization is to enable people to help promote better futures for those they care about and want to support most. The organization believes that people know best how to support each other and grow together if they have financial resources to do so. This organization aims to empower people to improve their own lives and those of the people and communities they care about most. Thus, the Community Empowerment Organization gives financial resources to community members, like you, to enable them to contribute positively to the lives of people important to them. For example, when people can invest in themselves, they are better able to expand employment opportunities for others, provide valuable services to their community, or teach others, including children, useful skills and knowledge. Now Community Empowerment Organization is going to send you 400 KSh. Please note that this is a one-time transfer of financial resources.*

2. Data

This section describes the data collected following the cash transfer and messaging.

2.1 Self-efficacy module

This module assesses the extent to which respondents feel capable of improving their lives and the lives of others important to them in the current moment.

- In this moment, how much do you feel in control of your financial situation, such as your success in your business or employment, or other income generating activities.
- In this moment, how much do you feel capable of making progress towards your goals.
- In this moment, how much do you feel capable of making progress towards goals for your community, such as helping and empowering others you care about.
- In this moment, how much do you feel that life will get better?

2.2 Stigma module

This module assesses the ways in which respondents feel that they or other recipients feel judged by others.

- People may negatively judge others for various reasons. How much do you feel that other people in Kenya make judgments about you based on your economic status? By economic status, I mean things like the place where you live, your job, or the amount of money you have.

- How much would other people feel embarrassed if they received money from the [ORGANIZATION NAME]?
- If your neighbors found out that you received money from the [ORGANIZATION NAME], how upset or jealous would they be with you?
- In this moment, how much do you feel like a good family member, whatever that means to you?
- In this moment, how much do you feel like a good community member, whatever you means to you?

2.3 *Affect module*

The affect module assesses the extent of experienced positive and negative emotional states.

- Recall that you just received some cash from [ORGANIZATION], which has a goal of [ORGANIZATION].
- Different experiences and conversations can make people feel different emotions, sometimes sad, sometimes excited, sometimes nervous. We want to know how you're feeling right now, from the time we started this conversation. Please answer as quickly and honestly as you can because we want to get your immediate thoughts.
- In this moment, how bad or good do you feel?
- In this moment, how embarrassed do you feel?
- In this moment, how empowered do you feel?
- In this moment, how much do you feel worried/concerned about your finances?

2.4 *Video selection task*

This task asked participants to make a choice about watching 3-4 minute video clips. Enumerators described the following six videos and the participant chose to watch two at the end of the survey. Participant could not select the same clip more than once. Video clips were played after the completion of the sociodemographic questionnaire.

- A video from the Mark Angel comedy group, featuring Emanuela (leisure)
- A trailer for the Nigerian movie, featuring Ramsey Noah (leisure)
- A Noa Ubongo video on math skills for business or CBO management (self-investment)
- A video of football highlights from around the world (leisure)
- A Noa Ubongo video on using equity and debt for financing business development (self-investment)
- A Naswa prank skit (leisure)

This task provided information on participants' willingness to engage in self-investment (i.e. skills building) activities over leisurely activities. We collected data on the participant's ordered first and second choices. We classified each clip as either for leisure or for self-investment and observe the number of self-investment videos (0, 1, or 2) the participant chooses to watch.

2.5 Savings decision task

This task allowed participants to invest a portion (either one-quarter or one-half of their initial endowment) in savings with an interest rate of 50%, to be paid out in two weeks. Enumerators reminded the participant about receiving KES 400 and present the participant with the following two choices.

1. “If you send us 100 right now, after two weeks you will get back 150 KSh.”
2. “If you send us 200 right now, after two weeks you will get back 300 KSh.”

If the participant chose to save, enumerators instruct them to send the appropriate amount of money to a project phone number a project phone number using M-Pesa. We also use M-Pesa to complete transfers scheduled in two weeks. To further reduce uncertainty regarding the delayed payment, we provided a phone number for participants to call to follow up on the transaction.

In addition to observing the participants’ intertemporal allocation, we employed the framework of Johnson, Häubl, and Keinan (2007) and elicited thoughts the participant may have regarding the choice to delay payment. This was done prior to the participant making the choice to delay payment. Enumerators asked participants to list up to five ‘queries’ regarding the decision. They were then asked to classify each query as either in favor of or against choosing to save the money. We collected data on both the content of the queries and their classification. We calculated for each participant a standardized median rank difference of aspect types to summarize the tendency to produce delay-favored queries before opposed queries.

$$\frac{2(MR_p - MR_i)}{n} \quad (1)$$

MR_p is the median rank of queries supporting delayed payment, MR_i is the median rank against delayed payment, and n is the total number of queries listed.

2.6 Support for organization and message

Participants were reminded of the organization’s goal by listening to the audio message treatment once more. They were then asked to evaluate the message and were asked whether they would want to show their support for the organization by recording the organization’s message themselves.

1. How empowering is this recorded message?
2. Overall do you like or dislike this audio message?
3. This [ORGANIZATION] is asking recipients whether they want to help promote their goal of [ORGANIZATION] by recording the voices of recipients saying their message. They want to share these recordings with possible future recipients as a show of support from current recipients. If you support their goal, you could stay after the survey ends to record the message you heard earlier. Would you like to end after watching the videos, or to continue and do a recording to show support for this organization?

2.7 Messages evaluation

We ask participants to forecast how many business videos other participants in different treatment arms would watch. Respondents make forecasts for each treatment arm, starting with the treatment message they received, and tell us their level of confidence, allowing us to roughly calculate how well participants are able to forecast treatment effects, and how these forecasts vary by participant confidence.

2.8 Sociodemographic questionnaire (9 items)

The final portion of the survey asked participants to report various sociodemographic characteristics including:

1. Participant is female
2. Participant completed standard 8
3. Participant is Christian
4. Age
5. Participant employment status
6. Average monthly income (in KSh, log transformed, and Winsorized at the top 1%)
7. Consumption in the last seven days (in KSH, log transformed, and Winsorized at the top 1%)
8. Participant has KSh 1000 stored away
9. Difficulty in raising KSh 3000 within 2 days (normalized)

3. Empirical Analysis

3.1 Randomization balance checks

Although the randomization of the treatment ensures balance across groups in expectation, we test for differences in sociodemographic characteristics using the following specification.⁶

$$Y_i = \beta_0 + \beta_1 \text{IND}_i + \beta_2 \text{COM}_i + \varepsilon_i \quad (2)$$

Y_i refers to the sociodemographic variables listed in Table 4 for individual i measured at the end of the survey. IND_i indicates assignment to the individual empowerment message while COM_i indicates assignment to the community empowerment message. The reference category in this model is the poverty alleviation message. We will estimate cluster-robust standard errors at the individual level.

We include any sociodemographic variable for which we reject balance as a control variable when estimating treatment effects.

3.2 Treatment effect of cash transfer messages

We will use the following reduced-form specification to estimate the treatment effect of different messages.

$$Y_i = \beta_0 + \beta_1 \text{IND}_i + \beta_2 \text{COM}_i + \varepsilon_i \quad (3)$$

Y_i refers to the outcome variables for individual i measured after the manipulation. The outcome variables described in Table 2 will be the focus of this analysis.

IND_i indicates assignment to the individual empowerment message while COM_i indicates assignment to the community support message. The reference category in this model is the poverty alleviation message. We will estimate cluster-robust standard errors at the individual level. Table 3 lists the hypotheses we will test using Equation 3.

⁶We will conduct the data analysis outlined in this section using the R programming language with the scripts included in Appendix B. Preliminary data cleaning, including data download, appending survey versions, inclusion of location data, and removal of personally identifiable information, is not included in the source code.

Table 2: Primary outcome variables

Variable	Description
Video selection	Number of self-investment videos chosen (0, 1, 2 out of 6)
Savings choice	Amount saved (0 KSh, 100 KSh, 200 KSh)
Message recording	Dummy variable for decision to record message of support

Table 3: Hypothesis tests

Null hypothesis	Description
$H_0 : \beta_1 = 0$	Effect of individual empowerment message relative to poverty alleviation message
$H_0 : \beta_2 = 0$	Effect of community empowerment message relative to poverty alleviation message
$H_0 : \beta_1 = \beta_2$	Effect of community empowerment message relative to individual empowerment message

In addition to our primary outcomes, we estimate the effect of the treatment on self-efficacy, stigma, affect, support of message, MacArthur subjective social status ladders, and standardized mean rank difference of thoughts in favor of and against saving. We will analyze these variables by both looking at individual items and constructing summary indices. Indices of self-efficacy, stigma, and affect will be constructed by taking the sum of the normalized constituent items and standardizing the index by its mean and SD.

3.3 Covariate adjustment

To improve precision, we will apply covariate adjustment with a vector of baseline indicators \mathbf{X}_i . We obtain the covariate-adjusted treatment effect estimate by estimating Equation 3 including the demeaned covariate vector $\dot{\mathbf{X}}_i = \mathbf{X}_i - \bar{\mathbf{X}}$ as an additive term and as an interaction with the treatment indicator.

$$Y_i = \beta_0 + \beta_1 \text{IND}_i + \beta_2 \text{COM}_i + \gamma_0 \dot{\mathbf{X}}'_i + \gamma_1 \text{IND}_i \dot{\mathbf{X}}'_i + \gamma_2 \text{COM}_i \dot{\mathbf{X}}'_i + \varepsilon_i \quad (4)$$

The set of indicators partitions our sample so that our estimate for β_j remains unbiased for the average treatment effect (Lin 2013). We will estimate cluster-robust standard errors at the individual level. We use this model to test the hypotheses detailed in Table 3 including the control variables listed in Table 4. These variables were chosen based on prior expectations about their predictive power in relation to the outcomes. Because these covariates were measured after treatment, we exclude any variables found to be significantly correlated with the treatment. Equation 3 without covariate adjustment remains our preferred specification and report both estimates for robustness.

Table 4: Control variables for covariate adjustment

Variable	Description
Gender	Participant is female
Education	Participant completed standard 8
Age	Participant age
Unemployed	Participant is unemployed
Income	Average monthly income (in KSh, log transformed, and Winsorized at the top 1%)
Consumption	Consumption in the last seven days (in KSh, log transformed, and Winsorized at the top 1%)
Savings	Participant has KSh 1000 stored away

3.4 Heterogeneous treatment effects

We will analyze the extent to which the policy messages produced heterogeneous treatment effects with the following specification.

$$Y_i = \beta_0 + \beta_1 \text{IND}_i + \beta_2 \text{COM}_i + \delta_0 x_i + \delta_1 \text{IND}_i x_i + \delta_2 \text{COM}_i x_i + \varepsilon_i \quad (5)$$

x_i is the binary dimension of heterogeneity. δ_1 and δ_2 identify the heterogeneous treatment effects of the individual empowerment and community empowerment messages relative to the poverty alleviation message. Testing $\delta_1 = \delta_2$ identifies heterogeneous effects between the former two messages. Standard errors are clustered at the individual level. We estimate this model with the variables summarized in Table 5. Because these variables were measured after treatment, we exclude any found to be significantly correlated with treatment.

Table 5: Dimensions of heterogeneity

Variable	Description
Gender	Participant is female
Savings	Participant has KSh 1000 stored away
Education	Participant completed standard 8

3.5 Randomization inference

One potential concern is that inference might be invalidated by finite sample bias in estimates of the standard errors. To address this issue, we will conduct randomization inference to test the Fisherian sharp null hypothesis of no treatment effect for every participant (Fisher 1935).⁷ We perform Monte Carlo approximations of the exact p -values using 10,000 permutations of the treatment assignment. We will then estimate the treatment effect within each m^{th} permutation and calculate the standard Wald statistics for each of our hypothesis tests. We will compare the Wald statistics from the original sample with the distribution of permuted statistics to produce approximations of the exact p -values:

$$\hat{p}_\beta = \frac{1}{10,000} \sum_{m=1}^{10,000} \mathbf{1} \left[\hat{\beta}'_m V(\hat{\beta}_m)^{-1} \hat{\beta}_m \geq \hat{\beta}'_{\text{obs.}} V(\hat{\beta}_{\text{obs.}})^{-1} \hat{\beta}_{\text{obs.}} \right] \quad (6)$$

Following Young (2015), we will permute the data and calculate the regressions for all outcomes within each draw. We will conduct the permutation test for Equations 3, 4, and 5. While we will highlight analytic p -values as primary, we report these bootstrapped p -values for robustness.

3.6 Multiple testing adjustment

Given that our survey instrument included several items related to a single behavior or dimension, we will calculate sharpened q -values over outcomes in Table 2 and 4 to control the false discovery rate (Benjamini, Krieger, and Yekutieli 2006). Rather than specifying a single q , we will report the minimum q -value at which each hypothesis is rejected (Anderson 2008). We will apply this correction over each set of outcomes but separately for each hypothesis test and equation. When estimating Equation 5, we correct over different dimensions of heterogeneity separately. We will report standard p -values and use minimum q -values as primary in our analysis. Table 6 summarizes the specified models and methods of statistical inference.

⁷Note that this is more restrictive than the null hypothesis of zero average treatment effect we will test in the previous section.

Table 6: Summary of models

	Treatment effect	Heterogeneous effects
Equation 3	Yes*	Yes*
Equation 3 with imbalanced covariates	Yes*	Yes*
Equation 4	Yes*	No

*Inference using robust standard errors, approximations of the exact p -value, and p -values controlling for the FDR.

References

- Anderson, Michael L. “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects”. In: *Journal of the American Statistical Association* 103.484 (2008), pp. 1481–1495. ISSN: 0162-1459. DOI: 10.1198/016214508000000841.
- Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. “Adaptive Linear Step-up Procedures That Control the False Discovery Rate”. In: *Biometrika* (2006), pp. 491–507.
- Fisher, Ronald Aylmer. *The Design of Experiments*. Includes index. Bibliography: p. 245. Edinburgh: Oliver & Boyd, 1935.
- Haushofer, Johannes et al. *A Methodology for Laboratory Experiments in Developing Countries: Examples from the Busara Center*. SSRN Scholarly Paper ID 2155217. Rochester, NY: Social Science Research Network, Feb. 2014.
- Jack, William and Tavneet Suri. *Mobile money: The economics of M-PESA*. Tech. rep. National Bureau of Economic Research, 2011.
- Johnson, Eric J., Gerald Häubl, and Anat Keinan. “Aspects of endowment: a query theory of value construction.” In: *Journal of experimental psychology: Learning, memory, and cognition* 33.3 (2007), p. 461.
- Lin, Winston. “Agnostic notes on regression adjustments to experimental data: Reexamining Freedman’s critique”. EN. In: *The Annals of Applied Statistics* 7.1 (Mar. 2013), pp. 295–318. ISSN: 1932-6157, 1941-7330. DOI: 10.1214/12-AOAS583.
- Mbiti, Isaac and David N. Weil. *Mobile banking: The impact of M-Pesa in Kenya*. Tech. rep. National Bureau of Economic Research, 2011.
- Young, Alwyn. *Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results*. Tech. rep. Technical Report, Working paper, 2015.

A. Consent Form

DESCRIPTION: You are invited to participate in a research study on how people view social programs conducted by the Busara Center for Behavioral Economics. You have been asked to participate in this study because you live in Nairobi, Kenya. If you volunteer to participate in this study, you will be asked to complete a survey about your beliefs and opinions. You will also be asked to provide basic information about yourself, your family. Before you decide to participate in this study, it is important that you understand why the research is being done and what it will involve. Please ask questions about anything you do not understand before deciding whether or not to participate. Please ask me if there is anything that is not clear or if you need more information.

TIME INVOLVEMENT: Your participation will take approximately 40-60 minutes.

RISKS AND BENEFITS: There are no direct benefits to you from participating in this research and minimal foreseeable risks to you or your family if you choose to participate in this study. We cannot and do not guarantee or promise that you will receive any benefits from this study.

CONFIDENTIALITY AND DATA STORAGE: Your participation in this study will remain confidential, and your identity will not be stored with your data. Your responses will be assigned a code number, and the list connecting your name with this number will be kept in a locked room and will be destroyed once all the data have been collected and analyzed. Presenting results will never reveal your identity. Confidentiality of subjects' data will be maintained to the extent allowed by law. No information about you, or provided by you during the research will be disclosed to others without your written permission, except: if necessary to protect your rights or welfare, or if required by law.

PAYMENTS: If you do decide to participate, you will receive a token of appreciation for agreeing to take part in this survey. You may also receive more money from our partner organization. If you choose to participate but decide to withdraw part way through you will be compensated accordingly.

PARTICIPANT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your participation is VOLUNTARY and you have the right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. The alternative is not to participate. You have the right to refuse to answer particular questions. Since these questionnaires might include personal questions about your feelings and finances, it is important that you know you may skip any questions that make you feel uncomfortable or choose to withdraw from the study at any time. The results of this research study may be presented at scientific or professional meetings or published in scientific journals. To be eligible to participate, you must be 18 years old and above. You may choose to participate but then revoke your permission for us to use the data we collected while still being compensated for your participation.

CONTACT INFORMATION:

Questions: If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, contact the Protocol Director, Jane Atieno at 072896461 or Catherine Thomas, at +1 (601) 750-1960.

Independent Contact: If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the research or your rights as a participant, please contact the Stanford Institutional Review Board (IRB) to speak to someone independent of the research team at (650)-723-2480 or toll free at 1-866-680-2906, or email at IRB2-Manager@lists.stanford.edu. You can also write to the Stanford IRB, Stanford University,

3000 El Camino Real, Five Palo Alto Square, 4th Floor, Palo Alto, CA 94306. If you have any questions or concerns about your rights and treatment as a research subject, you may contact The Secretary, KEMRI SERU, PO Box 54840-00200, Nairobi; Telephone numbers 020-2722541, 0717719477. Email address seru@kemri.org.

The extra copy of this signed and dated consent form is for you to keep.

I understand the information that was presented and that:

A.

My participation is voluntary, and I may withdraw my consent and discontinue

B. Source Code

B.1 Packages

```
setwd("/Users/Justin/Google Drive/UBIF/UBIF_Deliverables/UBIF_PAP/K1_PAP")
set.seed(47269801)

required.packages <- c("dplyr", "multiwayvcov", "multcomp", "reshape2", "knitr")
packages.missing <- required.packages[!required.packages %in% installed.packages()[,"Package"]]

if(length(packages.missing) > 0) {install.packages(required.packages, repo="https://cran.cnr.berkeley.edu/"
lapply(required.packages, library, character.only = TRUE)
```

B.2 User-defined functions

```
## RegTest conducts asymptotic tests from linear model ##

RegTest <- function(equation, clustvars, hypotheses, data) {

  model <- lm(equation, data = data, na.action = na.omit)

  if (missing(clustvars)) model$vcov <- vcov(model)
  else model$vcov <- cluster.vcov(model, cluster = clustvars)

  model$test <- summary(glht(model, linfct = hypotheses, vcov = model$vcov))$test

  numhyp <- length(hypotheses)

  EST <- matrix(nrow = numhyp, ncol = 4)

  for (i in 1:numhyp) {

    EST[i, 1] <- model$test$coefficients[i]
    EST[i, 2] <- model$test$tstat[i]
    EST[i, 3] <- model$test$sigma[i]
    EST[i, 4] <- model$test$pvalues[i]

  }

  colnames(EST) <- c("Estimate", "Tstat", "SE", "P")

  return(EST)

}

## PermTest returns MC approximations of the exact p-value ##

PermTest <- function(equation, treatvars, clustvars, hypotheses, iterations, data) {

  stopifnot(length(hypotheses) <= 1)

  obsEST <- RegTest(equation, clustvars, hypotheses, data)
  obsStat <- obsEST[1, 2]

  simEST <- matrix(ncol = 4)

  for (i in 1:iterations) {

    simTreat <- data[, treatvars, drop = FALSE]
    simTreat <- simTreat[sample(nrow(simTreat)),]
```

```

simData <- cbind(simTreat, data[, !(names(data) %in% treatvars), drop = FALSE])
colnames(simData)[1:length(treatvars)] <- treatvars

simEST <- rbind(simEST, RegTest(equation, clustvars, hypotheses, data = simData))
}

simSTAT <- simEST[2:nrow(simEST), 2]
countSTAT <- matrix(abs(simSTAT) >= abs(obsStat), ncol = 1)

ExactP <- colSums(countSTAT) / iterations

EST <- cbind(obsEST, ExactP)

colnames(EST) <- c("Estimate", "Tstat", "SE", "P", "ExactP")

return(EST)
}

## FDR returns minimum q-values ##

FDR <- function(pvals, step) {

  if (sum(is.na(pvals)) == FALSE) <= 1) {return(pvals)}
  if (missing(step)) {step <- 0.001}

  allpvals <- cbind(as.matrix(pvals), matrix(1:nrow(as.matrix(pvals)), ncol = 1))

  pvals <- na.omit(allpvals)
  nump <- nrow(pvals)

  pvals <- pvals[order(pvals[, 1]), ]
  rank <- matrix(1:nump, ncol = 1)
  pvals <- cbind(pvals, rank, matrix(0, nrow = nump, ncol = 1))

  qval <- 1

  while (qval > 0) {

    qfirst <- qval / (1 + qval)
    fdrtemp <- (qfirst * rank) / nump

    subrank <- which(fdrtemp >= as.matrix(pvals[, 1]))

    if (length(subrank) < 1) {
      numreject <- 0
    } else numreject <- max(subrank)

    qsec <- qfirst * (nump / (nump - numreject))
    fdrtemp <- (qsec * rank) / nump

    subrank <- which(fdrtemp >= as.matrix(pvals[, 1]))

    if (length(subrank) < 1) {
      numreject <- 0
    } else numreject <- max(subrank)

    pvals[which(pvals[, 3] <= numreject), 4] <- qval

    qval <- qval - step
  }
}

```

```

}

pvals <- pvals[order(pvals[, 2]), ]

qvals <- matrix(nrow = nrow(allpvals), ncol = 1)
qvals[match(pvals[, 2], allpvals[, 2]), 1] <- pvals[, 4]

return(as.matrix(qvals))

}

## Interact returns a string of interacted variables ##

Interact <- function(d, x) {

  catstring <- ""

  for (var in x) {

    catstring <- paste(catstring, " + ", d, "*", var, sep = "")

  }

  return(substr(catstring, 3, nchar(catstring)))

}

```

B.3 Data cleaning

```

varnames <- as.vector(read.delim(file = "K1__Field_Survey_v34+35_Appended.csv", sep = ",", header = FALSE,
k1_df <- read.delim(file = "K1__Field_Survey_v34+35_Appended.csv", sep = ",", header = FALSE, stringsAsFactors = FALSE))

## Survey meta data ##

k1_df$start.time.mst <- as.POSIXct(as.character(k1_df$V3), format = "%m/%d/%y %H:%M")
k1_df$start.time.eat <- k1_df$start.time.mst + (3600 * 9)

k1_df$end.time.mst <- as.POSIXct(as.character(k1_df$V4), format = "%m/%d/%y %H:%M")
k1_df$end.time.eat <- k1_df$end.time.mst + (3600 * 9)

## Participant ID ##

k1_df$survey.id <- k1_df$V1
k1_df <- k1_df[complete.cases(k1_df$survey.id), ]

nonentry <- c("R_5ZkChbiXDIj6KsK", "R_7N8rNtLsxF5a9on", "R_97Tx2cAjy30fqMR", "R_1YnRv0AJhuax6LS", "R_8dky4i")
k1_df <- k1_df[! k1_df$survey.id %in% nonentry, ]

## Treatment assignment ##

k1_df$treat[k1_df$condition == "poor"] <- 0
k1_df$treat[k1_df$condition == "individual"] <- 1
k1_df$treat[k1_df$condition == "community"] <- 2
k1_df$treat <- factor(k1_df$treat, labels = c("Pov", "Ind", "Com"))

k1_df$pov <- ifelse(k1_df$condition == "poor", 1, 0)
k1_df$ind <- ifelse(k1_df$condition == "individual", 1, 0)
k1_df$com <- ifelse(k1_df$condition == "community", 1, 0)

k1_df$msg1 <- recode(as.numeric(as.factor(k1_df$ORG_MESSAGE)), `2` = 0, `3` = 2)
k1_df$msg2 <- recode(as.numeric(as.factor(k1_df$ORG_MESSAGE_2)), `2` = 0)

```



```

k1_df$msg3 <- recode(as.numeric(as.factor(k1_df$ORG_MESSAGE_3)), `3` = 2)

## Self-efficacy ##

for (var in c(k1_df$sel.con, k1_df$sel.pers, k1_df$sel.com, k1_df$sel.prob, k1_df$sel.bett)) {

  var[var < 0] <- NA

}

k1_df$sel.score <- scale(k1_df$sel.con) + scale(k1_df$sel.pers) + scale(k1_df$sel.com) + scale(k1_df$sel.pr
k1_df$sel.score.z <- scale(k1_df$sel.score)

## Stigma ##

for (var in c(k1_df$jud.fam, k1_df$jud.com, k1_df$jud.judg, k1_df$jud.emb, k1_df$jud.ups)) {

  var[var < 0] <- NA

}

k1_df$sti.score <- scale(6 - k1_df$jud.fam) + scale(6 - k1_df$jud.com) + scale(k1_df$jud.judg) + scale(k1_d
k1_df$sti.score.z <- scale(k1_df$sti.score)

## Affect ##

for (var in c(k1_df$affect.pos, k1_df$affect.ash, k1_df$affect.pow, k1_df$affect.fina)) {

  var[var < 0] <- NA

}

k1_df$affect.score <- scale(k1_df$affect.pos) + scale(k1_df$affect.pow) + scale(7 - k1_df$affect.ash) + scale(7 - k1_df
k1_df$affect.score.z <- scale(k1_df$affect.score)

## Video selection ##

k1_df$vid.imp1 <- k1_df$vid.dec1 %in% c(3, 5)
k1_df$vid.imp2 <- k1_df$vid.dec2 %in% c(3, 5)
k1_df$vid.num <- k1_df$vid.imp1 + k1_df$vid.imp2

## Intertemporal choice ##

k1_df$sav.save <- k1_df$sav.dec > 1
k1_df$sav.save[k1_df$sav.dec < 0] <- NA

k1_df$sav.amt[k1_df$sav.dec == 1] = 0
k1_df$sav.amt[k1_df$sav.dec == 2] = 100
k1_df$sav.amt[k1_df$sav.dec == 3] = 200
k1_df$sav.amt[k1_df$sav.dec < 0] <- NA

## Query theory (savings) ##

que_df <- k1_df[names(k1_df) %in% c("survey.id", "que.rat1", "que.rat2", "que.rat3", "que.rat4", "que.rat5")]

k1_df$que.nonm <- apply(que_df[, 1:5], 1, function(x) length(x[is.na(x) == FALSE]))

que_df <- melt(que_df, id = c("survey.id"))
que_df$variable <- as.numeric(que_df$variable)
que_df <- dcast(que_df[is.na(que_df$value) == FALSE & que_df$value > 0, ], survey.id ~ value, median, value
names(que_df) <- c("survey.id", "que.mri", "que.mrp")
k1_df <- merge(k1_df, que_df, all.x = TRUE)

```

```

k1_df$que.smrdr <- (2 * (k1_df$que.mrp - k1_df$que.mri)) / k1_df$que.nonm
k1_df$que.smrdr[is.na(k1_df$que.mrp)] <- 1
k1_df$que.smrdr[is.na(k1_df$que.mri)] <- -1

# dealing with missing by filling in upper/lower bounds for now

## Message of support ##

k1_df$msg.dec[k1_df$msg.dec < 0] = NA
k1_df$msg.dec <- k1_df$msg.dec - 1

k1_df$msg.emp[k1_df$msg.emp < 0] = NA

k1_df$msg.lik[k1_df$msg.lik < 0] = NA
k1_df$msg.lik <- 7 - k1_df$msg.lik

k1_df$msg.avg <- (k1_df$msg.emp + k1_df$msg.lik) / 2

## Frame evaluation ##

k1_df$eva.poor[k1_df$msg1 == 0] <- k1_df$eva.msg1[k1_df$msg1 == 0]
k1_df$eva.poor[k1_df$msg2 == 0] <- k1_df$eva.msg2[k1_df$msg2 == 0]
k1_df$eva.poor[k1_df$msg3 == 0] <- k1_df$eva.msg3[k1_df$msg3 == 0]

k1_df$eva.ind[k1_df$msg1 == 1] <- k1_df$eva.msg1[k1_df$msg1 == 1]
k1_df$eva.ind[k1_df$msg2 == 1] <- k1_df$eva.msg2[k1_df$msg2 == 1]
k1_df$eva.ind[k1_df$msg3 == 1] <- k1_df$eva.msg3[k1_df$msg3 == 1]

k1_df$eva.com[k1_df$msg1 == 2] <- k1_df$eva.msg1[k1_df$msg1 == 2]
k1_df$eva.com[k1_df$msg2 == 2] <- k1_df$eva.msg2[k1_df$msg2 == 2]
k1_df$eva.com[k1_df$msg3 == 2] <- k1_df$eva.msg3[k1_df$msg3 == 2]

k1_df$eva.vid.poor <- k1_df$eva.rank.vid_8
k1_df$eva.vid.ind <- k1_df$eva.rank.vid_9
k1_df$eva.vid.com <- k1_df$eva.rank.vid_10

k1_df$eva.conf[k1_df$eva.conf < 0] <- NA

k1_df$eva.emp.poor <- k1_df$eva.rank.emp_5
k1_df$eva.emp.ind <- k1_df$eva.rank.emp_6
k1_df$eva.emp.com <- k1_df$eva.rank.emp_7

## Ladder scales ##

k1_df$ses.lad.now[k1_df$ses.lad.now < 0] <- NA
k1_df$ses.lad.now.z <- scale(k1_df$ses.lad.now)

k1_df$ses.lad.y2[k1_df$ses.lad.y2 < 0] <- NA
k1_df$ses.lad.y2.z <- scale(k1_df$ses.lad.y2)

k1_df$ses.lad.diff <- k1_df$ses.lad.y2 - k1_df$ses.lad.now
k1_df$ses.lad.avg <- (k1_df$ses.lad.y2 + k1_df$ses.lad.now) / 2

## Sociodemographics ##

k1_df$soc.age[k1_df$soc.age < 0] <- NA
k1_df$soc.pri <- as.numeric(k1_df$soc.edu > 3)
k1_df$soc.fem <- k1_df$soc.gen - 1
k1_df$soc.chr <- k1_df$soc.rel %in% c(1, 2)
k1_df$ses.unemp <- k1_df$ses.emp %in% c(1, 2)

k1_df$soc.inc[k1_df$soc.inc < 0] <- NA
k1_df$soc.inc.wins[k1_df$soc.inc <= quantile(k1_df$soc.inc, .99)] <- k1_df$soc.inc[k1_df$soc.inc <= quantile(k1_df$soc.inc, .99)]

```

```

k1_df$soc.inc.wins.ln <- log(k1_df$soc.inc.wins + sqrt(k1_df$soc.inc.wins^2 + 1))

k1_df$soc.con[k1_df$soc.con < 0] <- NA
k1_df$soc.con.wins[k1_df$soc.con <= quantile(k1_df$soc.con, .99)] <- k1_df$soc.con[k1_df$soc.con <= quantile(k1_df$soc.con, .99)]
k1_df$soc.con.wins.ln <- log(k1_df$soc.con.wins + sqrt(k1_df$soc.con.wins^2 + 1))

k1_df$soc.sav <- k1_df$soc.sav - 1

k1_df$soc.eme.z <- scale(k1_df$soc.eme)

## Survey validity ##

k1_df$end.hear <- k1_df$end.hear - 1
k1_df$end.hear[k1_df$end.hear < 0] <- NA

attach(k1_df)

## Center covariates ##

soc.fem.c <- scale(soc.fem, scale = FALSE)
soc.pri.c <- scale(soc.pri, scale = FALSE)
soc.age.c <- scale(soc.age, scale = FALSE)
ses.unemp.c <- scale(ses.unemp, scale = FALSE)
soc.inc.wins.ln.c <- scale(soc.inc.wins.ln, scale = FALSE)
soc.con.wins.ln.c <- scale(soc.con.wins.ln, scale = FALSE)
soc.sav.c <- scale(soc.sav, scale = FALSE)

```

B.4 Randomization checks

```

## Randomization balance checks ##

hypotheses <- c("treatInd = 0", "treatCom = 1", "treatInd - treatCom = 0")
depvars <- c("soc.fem", "soc.pri", "soc.age", "ses.unemp", "soc.inc.wins.ln", "soc.con.wins.ln", "soc.sav")

for (h in hypotheses) {

  RES <- matrix(nrow = 1, ncol = 5)

  for (devar in depvars) {

    eqn <- paste(devar, "~ treat", sep = " ")
    RES <- rbind(RES, PermTest(eqn, treatvars = c("treat", "pov", "ind", "com"), clustvars = k1_df$surv))

  }

  RES <- RES[2:nrow(RES), 1:ncol(RES)]
  RES <- cbind(RES, FDR(RES[, 4]))

  rownames(RES) <- depvars
  colnames(RES)[6] <- "Min. Q"

  print("-----", quote = FALSE)
  print(paste("H_0:", h), quote = FALSE)
  print(RES, quote = FALSE)

}

```

B.5 Treatment effect

```

## Plain OLS for primary outcomes ##

hypotheses <- c("treatInd = 0", "treatCom = 1", "treatInd - treatCom = 0")
depvars <- c("vid.num", "sav.amt", "msg.dec")

for (h in hypotheses) {

  RES <- matrix(nrow = 1, ncol = 5)

  for (depvar in depvars) {

    eqn <- paste(depvar, "~ treat", sep = " ")
    RES <- rbind(RES, PermTest(eqn, treatvars = c("treat", "pov", "ind", "com"), clustvars = k1_df$surv

  }

  RES <- RES[2:nrow(RES), 1:ncol(RES)]
  RES <- cbind(RES, FDR(RES[, 4]))

  rownames(RES) <- depvars
  colnames(RES)[6] <- "Min. Q"

  print("-----", quote = FALSE)
  print(paste("H_0:", h), quote = FALSE)
  print(RES, quote = FALSE)

}

## Plain OLS for secondary outcomes ##

hypotheses <- c("treatInd = 0", "treatCom = 1", "treatInd - treatCom = 0")
depvars <- c("sel.score.z", "sti.score.z", "aff.score.z", "msg.avg", "que.smrdr", "ses.lad.now", "ses.lad.y2

for (h in hypotheses) {

  RES <- matrix(nrow = 1, ncol = 5)

  for (depvar in depvars) {

    eqn <- paste(depvar, "~ treat", sep = " ")
    RES <- rbind(RES, PermTest(eqn, treatvars = c("treat", "pov", "ind", "com"), clustvars = k1_df$surv

  }

  RES <- RES[2:nrow(RES), 1:ncol(RES)]
  RES <- cbind(RES, FDR(RES[, 4]))

  rownames(RES) <- depvars
  colnames(RES)[6] <- "Min. Q"

  print("-----", quote = FALSE)
  print(paste("H_0:", h), quote = FALSE)
  print(RES, quote = FALSE)

}

```

B.6 Covariate-adjustment

```

## Covariate adjustment for primary outcomes ##

hypotheses <- c("treatInd = 0", "treatCom = 1", "treatInd - treatCom = 0")

```

```

depvars <- c("vid.num", "sav.amt", "msg.dec")
covariates <- c("soc.fem.c", "soc.pri.c", "soc.age.c", "ses.unemp.c", "soc.inc.wins.ln.c", "soc.con.wins.ln.c")

for (h in hypotheses) {

  RES <- matrix(nrow = 1, ncol = 5)

  for (depvar in depvars) {

    eqn <- paste(depvar, "~", Interact("treat", covariates), sep = " ")
    RES <- rbind(RES, PermTest(eqn, treatvars = c("treat", "pov", "ind", "com"), clustvars = k1_df$surv))

  }

  RES <- RES[2:nrow(RES), 1:ncol(RES)]
  RES <- cbind(RES, FDR(RES[, 4]))

  rownames(RES) <- depvars
  colnames(RES)[6] <- "Min. Q"

  print("-----", quote = FALSE)
  print(paste("H_0:", h), quote = FALSE)
  print(RES, quote = FALSE)

}

## Covariate adjustment for secondary outcomes ##

hypotheses <- c("treatInd = 0", "treatCom = 1", "treatInd - treatCom = 0")
depvars <- c("sel.score.z", "sti.score.z", "aff.score.z", "msg.avg", "que.smrdr", "ses.lad.now", "ses.lad.y2")
covariates <- c("soc.fem", "soc.pri", "soc.age", "ses.unemp", "soc.inc.wins.ln", "soc.con.wins.ln", "soc.sad")

for (h in hypotheses) {

  RES <- matrix(nrow = 1, ncol = 5)

  for (depvar in depvars) {

    eqn <- paste(depvar, "~", Interact("treat", covariates), sep = " ")
    RES <- rbind(RES, PermTest(eqn, treatvars = c("treat", "pov", "ind", "com"), clustvars = k1_df$surv))

  }

  RES <- RES[2:nrow(RES), 1:ncol(RES)]
  RES <- cbind(RES, FDR(RES[, 4]))

  rownames(RES) <- depvars
  colnames(RES)[6] <- "Min. Q"

  print("-----", quote = FALSE)
  print(paste("H_0:", h), quote = FALSE)
  print(RES, quote = FALSE)

}

```

B.7 Heterogeneous treatment effects

```

## Heterogeneous effects for primary outcomes ##

depvars <- c("vid.num", "sav.amt", "msg.dec")
hetvars <- c("soc.fem", "soc.sav", "soc.pri")

```

```

for (hetvar in hetvars) {

  hypotheses <- c(paste("treatInd:", hetvar, " = 0", sep = ""), paste("treatCom:", hetvar, " = 0", sep = ""))

  for (h in hypotheses) {

    RES <- matrix(nrow = 1, ncol = 5)

    for (depvar in depvars) {

      eqn <- paste(depvar, " ~ treat*", hetvar, sep = "")
      RES <- rbind(RES, PermTest(eqn, treatvars = c("treat", "pov", "ind", "com"), clustvars = k1_df$))

    }

    RES <- RES[2:nrow(RES), 1:ncol(RES)]
    RES <- cbind(RES, FDR(RES[, 4]))

    rownames(RES) <- depvars
    colnames(RES)[6] <- "Min. Q"

    print("-----", quote = FALSE)
    print(paste("H_0:", h), quote = FALSE)
    print(RES, quote = FALSE)

  }

}

## Heterogeneous effects for secondary outcomes ##

depvars <- c("sel.score.z", "sti.score.z", "aff.score.z", "msg.avg", "que.smr", "ses.lad.now", "ses.lad.y2")
hetvars <- c("soc.fem", "soc.sav", "soc.pri")

for (hetvar in hetvars) {

  hypotheses <- c(paste("treatInd:", hetvar, " = 0", sep = ""), paste("treatCom:", hetvar, " = 0", sep = ""))

  for (h in hypotheses) {

    RES <- matrix(nrow = 1, ncol = 5)

    for (depvar in depvars) {

      eqn <- paste(depvar, " ~ treat*", hetvar, sep = "")
      RES <- rbind(RES, PermTest(eqn, treatvars = c("treat", "pov", "ind", "com"), clustvars = k1_df$))

    }

    RES <- RES[2:nrow(RES), 1:ncol(RES)]
    RES <- cbind(RES, FDR(RES[, 4]))

    rownames(RES) <- depvars
    colnames(RES)[6] <- "Min. Q"

    print("-----", quote = FALSE)
    print(paste("H_0:", h), quote = FALSE)
    print(RES, quote = FALSE)

  }

}

```