Framing an Unconditional Cash Transfer: Pre-Analysis Plan

Author[*]

July 10, 2017

Abstract

This document describes the pre-analysis plan for a randomized experiment examining the effects of framing of welfare payments on self-concept and economic behavior. In this study, we will provide small, unconditional cash transfers to residents of an informal settlement in Nairobi and vary the way in which the transfers are framed to participants. Participants will be randomly assigned to one of three treatment groups: the transfer framed as a means toward poverty alleviation, individual empowerment, or collective support. We will then collect self-reported measures of self-efficacy, judgement, and affect and observed measures of temporal discounting and investment. 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 source files to be used for data analysis.

[*]The Busara Center for Behavioral Economics

Contents

1. Introduction

2. Research Design

2.1 Sampling

This study will be conducted in conjunction with the Busara Center for Behavioral Economics in Nairobi with 450 participants residing in Kibera, one of Kenya's largest urban slums (Haushofer et al. 2014). We will draw a random sample of participants using SMS and phone calls from the Busara Center's active pool of over 12,000 Nairobi residents. During the recruitment call, participants will be informed that they would be paid KES 300 for their participation and have the opportunity to earn more during the study. The sample will include males and females at least eighteen years old. [[Verify this and edit]]

2.2 Manipulation

At the outset of the survey, eligible and consenting participants will be told they are receiving an unconditional cash transfer from an organization unaffiliated with the Busara Center. Participants will be randomly assigned to receive one of three messages introducing the purpose of the cash transfer. All frames are identical in content and structure save for the described purpose of the cash transfer. In the poverty alleviation framing, the payment is described as a means to meet basic needs. The individual empowerment framing describes the payment as a means toward individual goals and the collective support framing as a means toward goals regarding family and the community. For the individual empowerment and collective support treatments, enumerators will also elicit participants to list either individual or collective goals and beliefs about the purpose of the payment.

2.2.1 Poverty alleviation framing

"As part of this project, we're working with a poverty alleviation organization. This organization is not Busara. It is a different organization. If you could give the aid a name to represent your needs and the organization's goal of poverty alleviation, what would it be? For example, it could be the food fund or emergency fund. Great. This is your [] fund. If other people you interact with regularly knew that you received aid from the [], how would they view you? How would receiving this aid affect your relationships? Can you describe to me in your own words the reasons why you think this organization is giving out aid to people like you?"

2.2.2 Individual empowerment framing

"As part of this project, we're working with an individual empowerment organization. This organization is not Busara. It is a different organization. As a reminder, these resources are intended to help you help yourself and to choose how to live your own life. Can you tell me: what are your most important goals for yourself? If you could give the resources a name to represent your individual goals, and the organization's goal of [], what would it be? For example, it could be your self-investment fund or your business fund. Great. This is your [] fund. How would receiving this aid affect your relationships? Can you describe to me in your own words the reasons why you think this organization is giving out aid to people like you?"

2.2.3 Collective support framing

"As part of this project, we're working with a group support organization. This organization is not Busara. It is a different organization. As a reminder, these resources are intended to help you find ways to support and empower people in your family and in your community that you care about most. Can you tell me: what are the most important goals you have for helping your family and the people you care most about? Who would you be helping? If you could give these resources a name to represent the goals you have for helping your community and the organization's goal of [], what would it be? For example, it could be your education fund or your growing together fund. Great. This is your [] fund. If other people you interact with regularly knew that you received these resources from the [], how would they view you? How would receiving these resources affect your relationships? Can you describe to me in your own words the reasons why you think this organization is giving out resources to people like you?"

2.3 Data collection

2.3.1 Field protocol

Respondent payments, including participation fees and cash transfers will be delivered via the mobile money system M-Pesa.[1]

2.3.2 Experimental procedure

The following summarizes the

1. **Consent agreement**

2. **Cash transfer framing**:

3. **Self-efficacy questionnaire**:

4. **Judgement questionnaire**:

5. **Affect questionnaire**:

6. **Video selection task**:

7. **Savings task**:

8. **Subjective social status ladder scale**:

9. **Sociodemographic questionnaire**:

10. **Frame evaluation**:

11. **Message of support**:

12. **Enumerator response**:

13. **Video task playback**:

---

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

3. Empirical Analysis

3.1 Treatment effect of cash transfer frames

We will use the following reduced-form specification to estimate the treatment effect of different frames.[2]

$$Y_i = \beta_0 + \beta_1 \text{IND}_i + \beta_2 \text{COL}_i + \varepsilon_i \tag{1}$$

Y_i refers to the outcome variables for individual i measured after the manipulation. The outcome variables described in Table 1 will be the focus of this analysis. IND_i indicates assignment to the individual empowerment frame while COL_i indicates assignment to the collective support frame. The reference category in this model is the poverty alleviation frame. We will estimate cluster-robust standard errors at the individual level. Table 2 lists the hypotheses we will test using Equation 1.

Table 1: Primary outcome variables

Variable	Description
Self-efficacy	
Affect	
Video selection	
Savings	

Table 2: Primary hypothesis tests

Null hypothesis	Description
$H_0 : \beta_1 = 0$	Effect of individual empowerment frame relative to poverty alleviation frame
$H_0 : \beta_2 = 0$	Effect of collective support frame relative to poverty alleviation frame
$H_0 : \beta_1 = \beta_2$	Effect of collective support frame relative to individual empowerment frame

To improve precision, we will also apply covariate adjustment with a vector of baseline indicators $\mathbf{X}_i$. We obtain the covariate-adjusted treatment effect estimate by estimating Equation 1 including the demeaned covariate vector $\dot{\mathbf{X}}_i = \mathbf{X}_i - \bar{\mathbf{X}}_i$ as an additive term and as an interaction with the treatment indicator.

$$Y_i = \beta_0 + \beta_1 \text{IND}_i + \beta_2 \text{COL}_i + \gamma_0 \dot{\mathbf{X}}_i' + \gamma_1 \text{IND}_i \dot{\mathbf{X}}_i' + \gamma_2 \text{COL}_i \dot{\mathbf{X}}_i' + \varepsilon_i \tag{2}$$

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 2 including the control variables listed in Table 3.

3.2 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).[3] The

---

[2]We will conduct the data analysis outlined in this section using the R programming language with the scripts included in Appendix C.

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

Table 3: Control variables for covariate adjustment

Variable	Description
Age	Dummy variable indicating participant is over 25
Gender	Dummy variable indicating participant is female
Employment	Dummy variable indicating participant is employed
Education	Dummy variable indicating participant completed std. 8
Children	Dummy variable indicating participant has children

basis for this inferential framework is that the distribution of test statistics comes from random treatment assignment rather than from drawing a finite sample from a super-population. This method produces exact p-values which do not rely on asymptotic theorems for valid inference. We perform Monte Carlo approximations of the exact p-values using $M = 10,000$ permutations of the treatment assignment. We will then estimate our primary specification within each mth 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:

$$\hat{p}_\beta = \frac{1}{10,000} \sum_{m=1}^{10,000} \mathbf{1} \left[\hat{\beta}'_m V(\hat{\beta}_m)^{-1} \hat{\beta}_m \geq \hat{\beta}'_{obs.} V(\hat{\beta}_{obs.})^{-1} \hat{\beta}_{obs.} \right] \tag{3}$$

Following Young (2015), we will permute the data and calculate the regressions for all outcomes within each draw.

3.3 Heterogeneous treatment effects

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

$$Y_i = \beta_0 + \beta_1 \text{IND}_i + \beta_2 \text{COL}_i + \delta_0 x_i + \delta_1 \text{IND}_i x_i + \delta_2 \text{COL}_i x_i + \varepsilon_i \tag{4}$$

x_i is the binary dimension of heterogeneity measured before treatment assignment. δ_1 and δ_2 identify the heterogeneous treatment effects of the individual empowerment and collective support frames relative to the poverty alleviation frame. Testing $\delta_1 = \delta_2$ identifies heterogeneous effects between the former two frames. Standard errors are clustered at the individual level. We estimate this model with the baseline variables summarized in Table 4.

Table 4: Dimensions of heterogeneity

Variable	Description
Age	Dummy variable indicating participant is over 25
Gender	Dummy variable indicating participant is female
Marital status	Dummy variable indicating participant is married or co-habitating 2
Education	Dummy variable indicating participant completed std. 8
Children	Dummy variable indicating participant has children

3.4 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 1 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 separately for each hypothesis test and will report both standard p-values and minimum q-values in our analysis.

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. 24, 2014.

Jack, William and Tavneet Suri. *Mobile money: The economics of M-PESA*. National Bureau of Economic Research, 2011.

Lin, Winston. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique". 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*. National Bureau of Economic Research, 2011.

Young, Alwyn. *Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results*. Technical Report, Working paper, 2015.

A. Consent Form

B. Survey Instrument

C. Data Analysis Scripts

C.1 Generating data

```
#Create locals for simulation
  OBS <- 1000
  TE  <- .8
  HET <- .4

#Generate treatment
  Treat <- sample(0:1,OBS,rep = TRUE,prob = c(.5,.5)) %>%
  factor(levels = c(0,1), labels = c("Control","Treatment"))

#Generate gender
 Gen <- sample(0:1,OBS,rep = TRUE,prob = c(.5,.5))  %>%
  factor(levels = c(0,1), labels = c("Male","Female"))

#Generate factor variable measuring highest level of education
 Edu <- sample(1:3,OBS,rep = TRUE,prob = c(.5,.3,.2)) %>%
  factor(levels = c(1,2,3), labels = c("Primary school","High school","University & above"

#Generate income
 LnInc <- rnorm(OBS, mean = 5, sd = 1)
 Inc <- exp(LnInc)

#Generate y with notreatment effect
  y_nottreat <- rnorm(OBS, 0, 1)

#Generate outcome with noisy treatment effect of ___
  y_Teffect <- rnorm(OBS, TE, 1)
  y_Teffect[Treat == "Control"] <- 0
  y_treated = y_nottreat + y_Teffect

#Generate outcome with noisy treatment effect of ___ and noisy het of ___ gender
  y_GenTeffect <- rnorm(OBS, HET, 1)
  y_GenTeffect[Treat == "Control"] <- 0
  y_GenTeffect[Gen == "Male"] <- 0
  y_HetTreated <- y_treated + y_GenTeffect

#Create, save dataframe
  SimTreat <- data.frame(Treat, Gen, Edu, Inc, y_nottreat, y_treated, y_HetTreated)
  save(SimTreat,file = "SimTreat.Rda")
  attach(SimTreat)

## The following objects are masked _by_ .GlobalEnv:
##
##    Edu, Gen, Inc, Treat, y_HetTreated, y_nottreat, y_treated
```

Table 5: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
Inc	1,000	249.790	302.516	5.358	3,069.520
y_nottreat	1,000	−0.021	0.985	−2.950	3.653
y_treated	1,000	0.340	1.271	−3.672	4.591
y_HetTreated	1,000	0.423	1.380	−3.806	5.455

Table 6: Treatment assignment

	TreatCount	TreatProp
Control	499	49.90
Treatment	501	50.10

Table 7: Gender

	GenCount	GenProp
Male	499	49.90
Female	501	50.10

Table 8: Education

	EduCount	EduProp
Primary school	515	51.50
High school	299	29.90
University & above	186	18.60

C.1.1 Regression

```
#Reg with no treatment effect
NoTreReg <- lm(y_nottreat ~ Treat)
summary(NoTreReg)

##
## Call:
## lm(formula = y_nottreat ~ Treat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9210 -0.6664 -0.0259  0.6351  3.6650
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.02936    0.04410  -0.666    0.506
## TreatTreatment  0.01705    0.06230   0.274    0.784
##
## Residual standard error: 0.985 on 998 degrees of freedom
## Multiple R-squared:  7.508e-05,Adjusted R-squared:  -0.0009268
## F-statistic: 0.07494 on 1 and 998 DF,  p-value: 0.7843

#Reg with treatment effect
TreReg <- lm(y_treated ~ Treat)
summary(TreReg)

##
## Call:
## lm(formula = y_treated ~ Treat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3797 -0.7766 -0.0405  0.8476  3.8834
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.02936    0.05447  -0.539     0.59
## TreatTreatment  0.73720    0.07696   9.579   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.217 on 998 degrees of freedom
## Multiple R-squared:  0.0842,Adjusted R-squared:  0.08328
## F-statistic: 91.76 on 1 and 998 DF,  p-value: < 2.2e-16

#Controlling for Education
TreRegEdu <- lm(y_treated ~ Treat + Gen + Edu)
summary(TreRegEdu)

##
## Call:
```

```
## lm(formula = y_treated ~ Treat + Gen + Edu)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3513 -0.7757 -0.0460  0.8384  3.8557
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          -0.068711   0.078158  -0.879    0.380
## TreatTreatment        0.738469   0.077203   9.565   <2e-16 ***
## GenFemale             0.009617   0.077253   0.124    0.901
## EduHigh school        0.056244   0.088575   0.635    0.526
## EduUniversity & above 0.091814   0.104331   0.880    0.379
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.218 on 995 degrees of freedom
## Multiple R-squared:  0.08505,Adjusted R-squared:  0.08137
## F-statistic: 23.12 on 4 and 995 DF,  p-value: < 2.2e-16
```

```r
  #Het effects with Gen
  TreRegHet <- lm(y_HetTreated ~ Treat*Gen + Edu)
  summary(TreRegHet)
```

```
##
## Call:
## lm(formula = y_HetTreated ~ Treat * Gen + Edu)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.8059 -0.8222 -0.0342  0.8717  4.3369
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -0.09572    0.09387  -1.020   0.3081
## TreatTreatment         0.71083    0.11658   6.097 1.54e-09 ***
## GenFemale             -0.01715    0.11664  -0.147   0.8831
## EduHigh school         0.18130    0.09443   1.920   0.0552 .
## EduUniversity & above  0.11771    0.11120   1.059   0.2901
## TreatTreatment:GenFemale 0.40228  0.16461   2.444   0.0147 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.298 on 994 degrees of freedom
## Multiple R-squared:  0.1202,Adjusted R-squared:  0.1157
## F-statistic: 27.15 on 5 and 994 DF,  p-value: < 2.2e-16
```

```r
  #Het effects with Gen and robust SE
  vce <- vcovHC(TreRegHet, type = "HC1")
  TreRegHetRobSE <- sqrt(diag(vce))
  TreRegHetRobF <- waldtest(TreRegHet, vcov = vce)
  TreRegHetRobF
```

```
## Wald test
##
## Model 1: y_HetTreated ~ Treat * Gen + Edu
## Model 2: y_HetTreated ~ 1
##   Res.Df Df     F     Pr(>F)
## 1    994
## 2    999 -5 25.655 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 9: Regression results

| | *Dependent variable:* | | | | |
| | Untreated Y | Treated Y | | Het. Treated Y | |
	(1)	(2)	(3)	(4)	(5)
Treatment	0.017	0.737***	0.738***	0.711***	0.711***
	(0.062)	(0.077)	(0.077)	(0.117)	(0.107)
Female			0.010	−0.017	−0.017
			(0.077)	(0.117)	(0.088)
High school			0.056	0.181*	0.181*
			(0.089)	(0.094)	(0.093)
University			0.092	0.118	0.118
			(0.104)	(0.111)	(0.115)
Female × Treatment				0.402**	0.402**
				(0.165)	(0.165)
Constant	−0.029	−0.029	−0.069	−0.096	−0.096
	(0.044)	(0.054)	(0.078)	(0.094)	(0.071)
Observations	1,000	1,000	1,000	1,000	1,000
Adjusted R^2	−0.001	0.083	0.081	0.116	0.116
F Statistic	0.075	91.758***	23.124***	27.153***	27.153***

$^*p < 0.1;^{**}p < 0.05;^{***}p < 0.01$. Standard errors in parentheses.