

FRAMING A UNIVERSAL BASIC INCOME POLICY: PRE-ANALYSIS PLAN

Author*

June 14, 2017

Abstract

This document describes the pre-analysis plan for a randomized experiment examining the relationship between policy framing of a universal basic income (UBI) policy, welfare stigma, and political affiliation. Using the Amazon Mechanical Turk (AMT), participants will be asked to read a randomly assigned prompt concerning a UBI and complete questionnaires on demographics, welfare stigma, political affiliation, and support for the policy. The pre-analysis plan outlines our hypotheses, schedule of tasks, and empirical strategy. In order to guarantee transparency and bind ourselves from “fishing for results” we will pre-register the program producing our main results along with this pre-analysis plan.

*The Busara Center for Behavioral Economics

1. Introduction

Large scale social welfare interventions like universal basic income provide both social services and a political, social, and psychological narrative about the recipients and the existing social system. This study will recruit nine-hundred participants over age 18 from the US using the online platform Amazon Mechanical Turk. Participants will be presented with different frames of Universal Basic Income and asked to rate their levels of support of the program. We present participants with a number of frames highlighting different aspects of universal basic income, exploring framings related to justice, human rights, universality, and freedom. These messages will be coded on the social cognitive orientations that they tap and that are associated with political affiliation. In particular, messages that target liberal foundations will address the values of equality and welfare outcomes and the regulatory orientation of promotion goals while messages built on conservative foundations will be oriented towards the values of harm and security, vigilant to threat, and focused on prevention goals (maintenance of a satisfactory current state) (Sanoff-Bulman 2009). First, it will assess which messages have the most support overall, across all political affiliations. Second, it will assess the messages that generate the largest gaps in support between liberals and conservatives.

It will assess two mediation pathways: moral fit and the activation of stereotype threat. First, we hypothesize that liberal participants will be more supportive of policy messages that match liberal moral foundations while conservatives will be more supportive of messages that match conservative moral foundations. Secondly, we hypothesize that messages that increase the saliency of negative stereotypes about welfare and welfare applicants will lead to lower levels of policy support. We are also interested to examine whether some messages that meet the moral foundations of liberals (e.g. fairness, harm, equality) will be more likely to lead to activation of welfare stereotypes than messages based in conservative foundations (e.g. purity). For example, by discussing equality and welfare outcomes, such as poverty reduction, liberal messages may weaken support for UBI from more conservative populations, who may be particularly sensitive to welfare stereotypes. One implication may be that, because of these differing moral foundations and orientations, the messages that liberal advocates support and use may in fact undermine public support by failing to meet the moral foundations of conservative populations as well as exacerbate perceived stigma.

Three out of the eight messages that generate relatively high levels of support and that generate low levels of stereotype activation, compared to other framings, will then be used a separate, follow-up experiment.

2. Research Design

2.1 Sampling

The study was conducted at the Busara Center for Behavioral Economics (Busara) in Nairobi, Kenya, a facility specially designed for experimental economics and psychology studies. Busara maintains an active participant pool of more than 12,000 Nairobi residents. For the present study, 449 participants who had previously signed up to be part of the Busara participant pool were recruited from Kibera, two informal settlements in Nairobi. Participants were recruited using phone calls. In the recruitment phone call, participants were told that they were invited to participate in a study being conducted by behavioral economics researchers. They were informed that they would be paid KES 300 for their participation and have the opportunity to earn more during the study. To minimize demand effects, they were also told clearly that no government or outside organization was sponsoring the study. The sample included both males and females over eighteen years of age.

2.2 Manipulation

2.2.1 Frame 1

2.2.2 Frame 2

2.2.3 Control group

2.3 Data collection

2.3.1 Field protocol

2.3.2 Survey

3. Empirical Analysis

3.1 Treatment effect of policy frames

We will use the following reduced-form specification to estimate the treatment effect of policy frames.¹

$$Y_i = \beta_0 + \sum_{j=1}^2 \beta_j \text{FRAME}_{ij} + \varepsilon_i \quad (1)$$

Y_i refers to the outcome variables for individual i measured after the manipulation. FRAME_{ij} indicates assignment to one of the 2 policy frames. The reference category in this model is the control group. We will estimate cluster-robust standard errors at the individual level. Table 1 lists the hypotheses we will test using Equation 1.

Table 1: Primary hypothesis tests

Null hypothesis	Description
$H_0 : \beta_1 = 0$	Effect of frame 1 relative to control group
$H_0 : \beta_2 = 0$	Effect of frame 2 relative to control group
$H_0 : \beta_1 = \beta_2$	Effect of frame 1 relative to frame 2

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

$$Y_i = \beta_0 + \sum_{j=1}^2 \beta_j \text{FRAME}_{ij} + \dot{\mathbf{X}}_i' \gamma_0 + \sum_{j=1}^2 \text{FRAME}_{ij} \dot{\mathbf{X}}_i' \gamma_j + \varepsilon_i \quad (2)$$

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 1 including the control variables listed in Table 2.

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 conduct randomization inference to test the

¹We will conduct the data analysis outlined in this section using the R programming language with the scripts included in Appendix C.

Table 2: Control variables for covariate adjustment

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-habiting 2
Education	Dummy variable indicating participant completed std. 8
Children	Dummy variable indicating participant has children

Fisherian sharp null hypothesis of no treatment effect for every participant (Fisher 1935).² The 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 then estimate our primary specification within each m^{th} permutation and calculate the standard Wald statistics for each of our hypothesis tests. We compare the Wald statistics from the original sample with the distribution of permuted statistics to produce approximations of the exact p -values:

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

Following Young (2015), we 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.

$$Y_i = \beta_0 + \sum_{j=1}^2 \beta_j \text{FRAME}_{ij} + \delta_0 x_i + \sum_{j=1}^2 \text{FRAME}_{ij} x_i \delta_j + \varepsilon_i \quad (4)$$

x_i is the binary dimension of heterogeneity measured before treatment assignment. δ_j identify the heterogeneous treatment effects of treatment j . Standard errors are clustered at the individual level. We estimate this model with the baseline variables summarized in Table.

Table 3: 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-habiting 2
Education	Dummy variable indicating participant completed std. 8
Children	Dummy variable indicating participant has children

²Note that this is more restrictive than the null hypothesis of zero average treatment effect we test in the previous section.

3.4 Multiple testing adjustment

Given that our survey instrument included several items related to a single behavior or dimension, we calculate sharpened q -values over all outcomes following Benjamini, Krieger, and Yekutieli (2006) to control the false discovery rate (FDR). Rather than specifying a single q , we report the minimum q -value at which each hypothesis is rejected (Anderson 2008). We apply this correction separately for each hypothesis test. We report both standard p -values and minimum q -values in our analysis.

3.5 Outcomes of interest

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.
- 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.
- 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 Instruments

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 4: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
Inc	1,000	263.633	352.852	4.961	4,645.345
y_nottreat	1,000	-0.019	0.976	-3.015	3.147
y_treated	1,000	0.345	1.195	-3.244	4.801
y_HetTreated	1,000	0.493	1.356	-3.244	5.485

Table 5: Treatment assignment

	TreatCount	TreatProp
Control	506	50.60
Treatment	494	49.40

Table 6: Gender

	GenCount	GenProp
Male	482	48.20
Female	518	51.80

Table 7: Education

	EduCount	EduProp
Primary school	473	47.30
High school	305	30.50
University & above	222	22.20

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
## -3.0099 -0.6352  0.0094  0.6348  3.1819
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.004591   0.043420  -0.106   0.916
## TreatTreatment -0.029861   0.061777  -0.483   0.629
##
## Residual standard error: 0.9767 on 998 degrees of freedom
## Multiple R-squared:  0.0002341, Adjusted R-squared:  -0.0007677
## F-statistic: 0.2336 on 1 and 998 DF,  p-value: 0.6289

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

##
## Call:
## lm(formula = y_treated ~ Treat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.9465 -0.7431  0.0342  0.7423  4.0982
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.004591   0.050770  -0.090   0.928
## TreatTreatment  0.707244   0.072234   9.791 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.142 on 998 degrees of freedom
## Multiple R-squared:  0.08764, Adjusted R-squared:  0.08672
## F-statistic: 95.86 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
## -3.8814 -0.7483  0.0352  0.7415  4.0856
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.00536    0.07271   0.074   0.941
## TreatTreatment      0.70993    0.07238   9.809 <2e-16 ***
## GenFemale          0.05664    0.07234   0.783   0.434
## EduHigh school    -0.07775    0.08402  -0.925   0.355
## EduUniversity & above -0.07613    0.09301  -0.819   0.413
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.143 on 995 degrees of freedom
## Multiple R-squared:  0.08924, Adjusted R-squared:  0.08558
## F-statistic: 24.37 on 4 and 995 DF,  p-value: < 2.2e-16

#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
## -3.9102 -0.7847 -0.0151  0.7925  4.2205
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.009142    0.086546   0.106 0.915899
## TreatTreatment      0.697805    0.113844   6.130 1.27e-09 ***
## GenFemale          0.044991    0.110785   0.406 0.684749
## EduHigh school    -0.080227    0.091422  -0.878 0.380402
## EduUniversity & above -0.062941    0.101737  -0.619 0.536279
## TreatTreatment:GenFemale 0.592317    0.158299   3.742 0.000193 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.241 on 994 degrees of freedom
## Multiple R-squared:  0.1667, Adjusted R-squared:  0.1625
## F-statistic: 39.77 on 5 and 994 DF,  p-value: < 2.2e-16

#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 34.779 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Error: <text>:12:13: unexpected symbol
## 11:          add.lines = (c("Adjusted F", TreRegHetRobF))
## 12:          df
##          ^
```