Pre-Analysis Plan for June 2015 Miami Transgender Experiment

David Broockman and Joshua Kalla

July 7, 2015

Contents

Background **2**

Scope of This PAP **3**

Outcomes **3**

Main Outcome: General Trans Acceptance Attitudes . . . 3

Secondary Outcome: Gender Non-Conformity . . . 4

Secondary Outcome: Laws Only . . . 4

Secondary Outcome: Thermometer . . . 4

Other . . . 4

Estimation Procedures and Assumptions **4**

Covariates to use in Regression Adjustment . . . 5

Factor Analysis on Outcome to Increase Power . . . 6

Note: Fake treatment indicator in this PAP . . . 8

OLS with Clustered Robust Standard Errors . . . 8

One-tailed p-values . . . 8

Code for ATE Estimation **8**

Direct Effect . . . 8

Indirect Effect . . . 9

Other **10**

Exclusions and Non-Compliance With Protocol . . . 10

Missing Values . . . 10

Ceiling Effects . . . 10

Unadjusted Point Estimates . . . 10

Test for Proper Placebo Delivery . . . 11

Tests for Differential Attrition . . . 11

Not in This PAP **12**

Canvasser Heterogeneity 12

Treatment Effect Heterogeneity by Subject Attributes 12

Follow-Up Surveys 13

Note: the data analyzed below is the real data but does not contain the real treatment vector, only a simulated treatment vector for the purpose of preparing the analysis. One cannot draw any conclusions about the effect of the experiment from this document.

Background

The experiment has six steps:

1. We recruited voters to join an online survey panel, which we called the 2015 Miami-Dade Opinion Study. The voters were recruited via mail sent to the address at which they were registered to vote. This mail directed them to an online survey. The online survey was available in English and Spanish. We robo-called voters the day the mail hit and, for voters who did not respond, again the last day the survey was open.
2. Among voters who completed the baseline survey, we randomly assigned voters to receive a transgender equality conversation or a placebo recycling conversation. This random assignment took place at the household level, within blocks defined by household size and the first factor from factor analysis on a number of pre-treatment covariates.
3. The partner group implements the canvass, delivering either a transgender equality conversation or a placebo conversation depending on the voter's random assignment. Voters do not realize the canvassing is connected to the survey in any way.
4. We invite all voters who live in a household in which anyone was successfully reached to take a follow-up survey. If voters do not respond to the follow-up survey within five days, we mail them a letter asking them to complete it. For those who do not complete it within four days of receiving this letter, we call them.
5. We analyze the results of this survey to assess the effects of the conversations. In particular, we compare the views of voters who received the transgender equality conversations to those who received the placebo conversations. How we plan to do so is described below.
6. We will assess how long-lasting the persuasive effects of the canvass are by continuing to send follow-up survey waves for up to a year.

The experiment was also divided into three stages, for each of which we conducted each of those steps. We first conducted a pilot to test our procedure and identify the best ways to recruit Miami voters to the survey. This data is not considered part of the experiment and we have not analyzed it for any treatment/control differences. The data below for the full study comes from the remaining two stages. We conducted the first half of the full study for a pool of voters to be canvassed in early June, focused around a June 6 canvass, and finally have recruited a pool of voters to be canvassed in late June, focused around a June 20 canvass. There were also three canvasses during each of these stages, one main canvass on the original date (June 6 and June 20) and another two canvasses on the following Wednesday and Saturday.

The table below gives each of the 6 canvass dates and the dates when voters were solicited to take each survey. Note that there were two 'starting universes' as described above. Voters who were not contacted on June 6 from the June 6 universe were again attempted on June 10, and then again on June 17. Then, the June 20th canvass started with an entirely new universe.

Canvass Contact Date	t0 Survey Closed	"Universe"	t1 Survey Opened
June 6	June 4	June 6	June 9

Canvass Contact Date	t0 Survey Closed	"Universe"	t1 Survey Opened
June 10	June 4	June 6	June 13
June 17	June 4	June 6	June 20
June 20	June 17	June 20	June 24
June 24	June 17	June 20	June 27
June 27	June 17	June 20	June 30

We will pool the data from all six canvasses.

Scope of This PAP

1825 respondents completed the pre-survey and, as of this writing, 508 voters have responded to the t1 survey. We have also just launched the t2 survey. To write this PAP we have been analyzing a version of the merged outcome data without access to the real treatment indicator. We have not looked at the data with the real treatment indicator. This has allowed us to develop strategies for extracting the most power out of the pre-treatment survey.

We plan to analyze the t1 data beginning on Tuesday, July 7 after the submission of this PAP to EGAP but will allow for t1 responses to continue to come in until Monday, July 12.

Outcomes

The following variables will be combined into one scale measuring the overall effect of the conversations on prejudice towards transgender people.

Main Outcome: General Trans Acceptance Attitudes

```
main.dv.names <- c('miami_trans_law_t1', 'miami_trans_law2_t1', 'therm_trans_t1',
'gender_norm_sexchange_t1', 'gender_norm_moral_t1', 'gender_norm_abnormal_t1',
'gender_norm_trans_moral_wrong_t1')
```

These items are as follows:

- Miami-Dade county recently passed a law that prohibits discrimination in housing, employment and public accommodations based on gender identity and expression, a category that includes transgender men and women. Do you favor or oppose this new law?
- Some people say it's important to protect transgender people from discrimination in housing and employment. Other people have concerns about society becoming too accepting of transgender people, and do not want transgender people included in our non-discrimination law. What do you think? Do you agree or disagree that Miami law should protect transgender people from discrimination?
- Feeling thermometer towards trans people (0-100)
- I would support a friend choosing to have a sex change.
- It is morally wrong for a man to present himself as a woman in public.
- A man who identifies as a woman is psychologically abnormal.
- Saying you are a gender that is different than the one you were born with is morally wrong.

Secondary Outcome: Gender Non-Conformity

We will also analyze the below secondary DVs as a different scale capturing effects on gender norms.

```
secondary.dv.names <- c('gender_norm_rights_t1', 'gender_norm_looks_t1')
```

These items are as follows:

- To keep children from being confused, it's better when men look and act like men, and women look and act like women.
- Men and women should have equal rights, but men and women are not the same; it's normal for men to act like men, and women to act like women.

Secondary Outcome: Laws Only

We will also analyze the below DVs, from the main DVs, as capturing the effect on support for an anti-discriminational law only.

```
legal.dv.names <- c('miami_trans_law_t1', 'miami_trans_law2_t1')
```

As a reminder, these items were:

- Miami-Dade county recently passed a law that prohibits discrimination in housing, employment and public accommodations based on gender identity and expression, a category that includes transgender men and women. Do you favor or oppose this new law?
- Some people say it's important to protect transgender people from discrimination in housing and employment. Other people have concerns about society becoming too accepting of transgender people, and do not want transgender people included in our non-discrimination law. What do you think? Do you agree or disagree that Miami law should protect transgender people from discrimination?

Secondary Outcome: Thermometer

We will also analyze the feeling thermometer on its own.

```
therm.name <- 'therm_trans_t1'
```

Other

This question was included in the 'gender norm battery' but we do not consider it an outcome.

```
not.dv.names <- c('gender_norm_daugher_t1')
```

This item was:

- Parents usually maintain stricter control over their daughters than their sons, and they should.

Estimation Procedures and Assumptions

We will use the procedures and assumptions below to calculate ATE estimate, standard errors / confidence intervals, and p-values.

Covariates to use in Regression Adjustment

We will use the following variables from the baseline survey and administrative data as covariates to predict the outcome and increase power.

```
# Recode age for small number of observations where it is missing.
# Variables that begin with vf_ come from the voter file.
# Variables that end with _t0 were collected on the baseline survey.
data$vf_age[which(is.na(data$vf_age))] <- mean(data$vf_age, na.rm=TRUE)

# Language of interview
data$survey_language_es[is.na(data$survey_language_es)] <-
    data$survey_language_t0[is.na(data$survey_language_es)] == "ES"

# Insert the average of the baseline scale for blocking at the cluster level among respondents
all.blocks.and.clusters <- read.dta(paste0(wd, 'miami_blocks_hhs.dta'), convert.factors=FALSE)
data$hh_id <- all.blocks.and.clusters$hh_id[
  match(data$vf_vanid, all.blocks.and.clusters$vf_vanid)]
data$cluster_level_t0_scale_mean <- as.vector(
  by(data$scale_for_blocking_t0[data$respondent_t1==1],
     data$hh_id[data$respondent_t1==1], mean)[data$hh_id]
)


t0.covariate.names <- c('miami_trans_law_t0', 'miami_trans_law2_t0', 'therm_trans_t0',
'gender_norms_sexchange_t0', 'gender_norms_moral_t0', 'gender_norms_abnormal_t0',
'ssm_t0', 'therm_obama_t0', 'therm_gay_t0',
'vf_catvprop', 'vf_democrat', 'ideology_t0', 'religious_t0', 'exposure_gay_t0',
'exposure_trans_t0', 'vf_clalgbtsup1', 'pid_t0', 'sdo_scale', 'gender_norm_daugher_t0',
'gender_norm_looks_t0', 'gender_norm_rights_t0', 'therm_afams_t0', 'vf_female', 'vf_hispanic',
'vf_black', 'vf_age', 'survey_language_es', 'cluster_level_t0_scale_mean')
```

```
#Covariate matrix for the regression model
x <- data[,c(t0.covariate.names)]
x <- as.matrix(x, dimnames = list(NULL, names(x)))
```

These are all the variables in the dataset that we do not plan to use as dependent variables or to perform regression adjustment.

```
used.variable.names <- c(main.dv.names, secondary.dv.names, not.dv.names, t0.covariate.names)
names(data)[which(!names(data) %in% used.variable.names)]
```

```
##   [1] "gender_norms_moral2_t0"          "therm_rubio_t0"
##   [3] "therm_marijuana_t0"              "therm_immigrant_t0"
##   [5] "therm_muslims_t0"                "respondent_t0"
##   [7] "scale_for_blocking_t0"           "survey_language_t0"
##   [9] "survey_language_t1"              "therm_obama_t1"
##  [11] "therm_rubio_t1"                  "therm_gay_t1"
##  [13] "therm_marijuana_t1"              "therm_afams_t1"
##  [15] "therm_immigrant_t1"              "comments_t1"
##  [17] "payment_cash_t1"                 "payment_amazon_t1"
##  [19] "payment_acs_charity_t1"          "payment_bro_charity_t1"
##  [21] "payment_redcross_charity_t1"     "payment_unitedway_charity_t1"
```

```
##  [23] "respondent_t1"                  "vf_vanid"
##  [25] "contacted"                      "contacted_date"
##  [27] "enteredemail"                   "yob"
##  [29] "respondent_entered_maddress"    "respondent_entered_city"
##  [31] "respondent_entered_state"       "respondent_entered_zip5"
##  [33] "phone"                          "howwouldyouliketobepaidforyourpa"
##  [35] "vf_dwid"                        "vf_maddress"
##  [37] "vf_mcity"                       "vf_mstate"
##  [39] "vf_mzip5"                       "vf_mzip4"
##  [41] "vf_sex"                         "vf_vaddress"
##  [43] "vf_city"                        "vf_state"
##  [45] "vf_zip5"                        "vf_zip4"
##  [47] "vf_lastname"                    "vf_firstname"
##  [49] "vf_middlename"                  "vf_suffix"
##  [51] "vf_cd"                          "vf_sd"
##  [53] "vf_hd"                          "vf_phone"
##  [55] "vf_cellphone"                   "vf_dob"
##  [57] "vf_homephone"                   "vf_datereg"
##  [59] "vf_party"                       "vf_precinctname"
##  [61] "vf_racename"                    "vf_deadwood"
##  [63] "vf_catalistgenact"              "vf_ideolgy12"
##  [65] "_partisanshipscr"               "vf_relattend12"
##  [67] "vf_jobcrea"                     "vf_obama"
##  [69] "vf_lathost"                     "vf_socprog"
##  [71] "vf_aca_subsidy"                 "vf_fiscalpolicy"
##  [73] "vf_catideology"                 "vf_catpartisan"
##  [75] "vf_uninsured"                   "vf_catvotepropv2"
##  [77] "vf_vvnmunicount"                "vf_aflcristsup"
##  [79] "vf_catid"                       "vf_catpar"
##  [81] "vf_cattcksplt"                  "vf_clalgbtact1"
##  [83] "vf_clalgbtmov1"                 "vf_clalgbtopp1"
##  [85] "_4catcnhh"                      "_4catincome"
##  [87] "vf_white"                       "vf_vg_14"
##  [89] "vf_vg_12"                       "vf_vg_10"
##  [91] "vf_republican"                  "hh_size"
##  [93] "pre_incentive"                  "post_incentive"
##  [95] "mailing_version"                "login"
##  [97] "speaks_spanish_only"            "tranche"
##  [99] "useragent"                      "june20"
## [101] "contacted_hh"                   "contacted_indirect"
## [103] "hh_id"
```

Factor Analysis on Outcome to Increase Power

We will create the dependent variables using factor analysis as follows. The factors will be rescaled to mean 0 and standard deviation 1 to allow for a natural interpretation of the size of the effects in standard deviations.

```
compute.factor.dv <- function(dv.names, print.loadings = TRUE){
  responders <- subset(data, respondent_t1 == 1)

  factor.obj <- princomp(responders[, dv.names], cor=TRUE)
  if(print.loadings) print(loadings(factor.obj))
  dv <- factor.obj$scores[,1]
```

```
  if(cor(dv, responders$miami_trans_law_t1) < 0) dv <- -1 * dv # Make sure it points the right way.
  dv <- scale(dv) #rescale to mean 0 sd 1
  return(dv[match(data$vf_vanid, responders$vf_vanid)])
}

main.dv <- compute.factor.dv(main.dv.names)
```

```
##
## Loadings:
##                                 Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## miami_trans_law_t1              -0.365  0.526 -0.215  0.387
## miami_trans_law2_t1             -0.383  0.504 -0.171
## therm_trans_t1                  -0.397  0.106        -0.880 -0.128
## gender_norm_sexchange_t1        -0.331         0.926  0.112  0.124
## gender_norm_moral_t1             0.386  0.435        -0.169  0.328 -0.720
## gender_norm_abnormal_t1          0.390  0.313  0.232        -0.831
## gender_norm_trans_moral_wrong_t1 0.390  0.408        -0.171  0.405  0.690
##                                 Comp.7
## miami_trans_law_t1               0.623
## miami_trans_law2_t1             -0.751
## therm_trans_t1                   0.200
## gender_norm_sexchange_t1
## gender_norm_moral_t1
## gender_norm_abnormal_t1
## gender_norm_trans_moral_wrong_t1
##
##                Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7
## SS loadings     1.000  1.000  1.000  1.000  1.000  1.000  1.000
## Proportion Var  0.143  0.143  0.143  0.143  0.143  0.143  0.143
## Cumulative Var  0.143  0.286  0.429  0.571  0.714  0.857  1.000
```

```
secondary.dv <- compute.factor.dv(secondary.dv.names)
```

```
##
## Loadings:
##                       Comp.1 Comp.2
## gender_norm_rights_t1  0.707 -0.707
## gender_norm_looks_t1   0.707  0.707
##
##                Comp.1 Comp.2
## SS loadings       1.0    1.0
## Proportion Var    0.5    0.5
## Cumulative Var    0.5    1.0
```

```
legal.dv <- compute.factor.dv(legal.dv.names)
```

```
##
## Loadings:
##                     Comp.1 Comp.2
## miami_trans_law_t1  -0.707  0.707
## miami_trans_law2_t1 -0.707 -0.707
```

```
##
##                 Comp.1 Comp.2
## SS loadings        1.0    1.0
## Proportion Var     0.5    0.5
## Cumulative Var     0.5    1.0
```

Note: Fake treatment indicator in this PAP

For the analysis below in the pre-analysis plan we are using a fake treatment indicator, not the real treatment indicator for the study.

OLS with Clustered Robust Standard Errors

For estimating treatment effects we will use OLS with standard errors clustered at the household level and with the covariates mentioned above, as shown below.

```
# Function for calculating clustered standard errors
# Requires library sandwich and lmtest
cl   <- function(fm, cluster){
           M <- length(unique(cluster))
           N <- length(cluster)
           K <- fm$rank
           dfc <- (M/(M-1))*((N-1)/(N-K))
           uj  <- apply(estfun(fm), 2, function(x) tapply(x, cluster, sum))
           vcovCL <- dfc*sandwich(fm, meat=crossprod(uj)/N)
           coeftest(fm, vcovCL) }
# Function for estimating ATE with regression adjustment and clustered standard errors
est.ate <- function(dv, include.obs){
  include.obs <- include.obs & !is.na(dv) # remove missing values so cl() doesn't break
  lm.result <- lm(dv[include.obs] ~ data$treat_ind_FAKE_FOR_PAP[include.obs] + x[include.obs,])
  return(cl(lm.result, data$hh_id[include.obs])[2,]) # Return just treatment coefficient.
}
```

Note that we do not include treatment-by-covariate interactions in the regression model because the design is balanced (see Lin, "Agnostic Notes").

One-tailed p-values

If the ATE estimate is positive we plan to use one-tailed p-values with a rejection threshold of 0.04 (which would be 0.08 two-tailed). If the ATE estimate is negative we will use a two-tailed p-value with a rejection threshold of 0.02 two-tailed, or 0.01 in one-tailed terms. This preserves the property that only 5% of the sampling distribution allows us to reject the null, but in this case the bottom 1% and the top 4% instead of the typical bottom 2.5% and top 2.5% (see Olken, "Pre-Analysis Plans in Economics").

Code for ATE Estimation

Direct Effect

These are the main analyses. The most important among them is the test on the factor analysis we called "main.dv" above.

To examine the direct effect we will examine only compliers. Recall that due to the placebo design we are able to observe compliance in both the treatment and control groups. 369 individuals who were contacted have taken the t1 survey as of this writing.

```
est.ate(main.dv, data$contacted == 1) # MAIN TEST FOR PAPER
```

```
##    Estimate  Std. Error     t value    Pr(>|t|)
## 0.005606365 0.047585623 0.117816367 0.906282935
```

```
est.ate(secondary.dv, data$contacted == 1)
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
## 0.02819146 0.07260002 0.38831197 0.69802892
```

```
est.ate(legal.dv, data$contacted == 1)
```

```
##    Estimate  Std. Error     t value    Pr(>|t|)
## -0.09526483  0.07084521 -1.34468974  0.17962445
```

```
est.ate(data$therm_trans_t1, data$contacted == 1)
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
##  0.6736545  2.0173311  0.3339336  0.7386362
```

Indirect Effect

To examine the indirect effect we will examine those who live with compliers but were not themselves contacted. We do not expect this analysis to be very informative because only 139 individuals in the t1 survey lived in a contacted household but were not contacted themselves.

```
est.ate(main.dv, data$contacted_indirect == 1)
```

```
##    Estimate  Std. Error     t value    Pr(>|t|)
## -0.04360444  0.12857388 -0.33913915  0.73515741
```

```
est.ate(secondary.dv, data$contacted_indirect == 1)
```

```
##    Estimate  Std. Error     t value    Pr(>|t|)
## -0.07480248  0.14952999 -0.50025068  0.61790708
```

```
est.ate(legal.dv, data$contacted_indirect == 1)
```

```
##    Estimate  Std. Error     t value    Pr(>|t|)
## -0.01474384  0.14369023 -0.10260847  0.91846225
```

```
est.ate(data$therm_trans_t1, data$contacted_indirect == 1)
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
## -4.9964053  4.3301474 -1.1538649  0.2510793
```

Other

Exclusions and Non-Compliance With Protocol

Canvassers marked what conversations they actually had, although we are analyzing the data on an intent-to-treat basis. Sometimes canvassers deviated from their assigned treatmetns and delivered the wrong script.

For most canvassers, we will analyze the data as follows. We will make the no-defiers assumption. We will adjust the point estimates above to account for this non-compliance using the usual formula, dividing the point estimates by the share of treated compliers in the contacted treatment group minus the share of treated compliers in the contacted placebo group. We will use global correct delivery rates and not canvasser-specific rates. (We do not yet have individual-level data on actual treatment delivery)

A very small number of canvassers were especially likely to fail to follow the protocol, delivering the wrong script at least half the time in at least one of the conditions. We will drop the observations of the subjects assigned to these canvassers for the main analysis, but report the results in a supplementary appendix.

A small number of canvassers who do not speak Spanish also reported that they marked voters as contacted when they encountered monolingual Spanish speakers with whom they would not have been able to have a treatment conversation. These individuals were thus marked as compliers in control but not in treatment. We anticipate receiving data from the partner group on which control observations fall in this category and excluding these observations from the analysis – that is, excluding control observations where a monolingual English speaker reached a monolingual Spanish speaker. In the treatment group this would have been marked as "language barrier" but sometimes in the control group this was not done.

Missing Values

We recode missing values to their means. These are rare.

For the feeling thermometer, Qualtrics leaves the values as missing if the subjects do not move the slider. We plan to recode these to 50 as well, as they appear at 50 at start.

Ceiling Effects

We register in advance our expectation that we may encounter ceiling effects on the items related to the law and the indicies that include these items, as baseline support for the law is high.

Unadjusted Point Estimates

We expect to draw a great deal of power from the baseline survey and combinations of the outcomes so will not put much stock in the unadjusted estimates. However, the design-based p-values can be computed by using randomization inference and the blocking and clustering in the design.

```
pick.treatment.hh.one.block <- function(block.id) sample(hh.ids[block.ids==block.id], 1)
cluster.assign <- function(){
  assigned.households <- unlist(lapply(unique.blocks, pick.treatment.hh.one.block))
    return(data$hh_id %in% assigned.households)
}
perms <- replicate(50, cluster.assign()) # We will use 50,000 for the real analysis.

# General function to compare an observed test statistic to a null distribution.
diff.means <- function(treat, dv){
```

```
  return(mean(dv[treat], na.rm=TRUE) - mean(dv[!treat], na.rm=TRUE))
}
ri.est <- function(treat, dv){
  observed <- diff.means(treat, dv)
  dist.under.sharp.null <- apply(perms, 2, diff.means, dv)
  return(mean(observed <= dist.under.sharp.null)) # One-tailed p-value.
}
ri.est(data$treat_ind_FAKE_FOR_PAP, main.dv)
```

```
## [1] 0.46
```

Test for Proper Placebo Delivery

We will first compare contact rates in treatment and placebo.

```
cl(lm(contacted ~ treat_ind_FAKE_FOR_PAP, data), data$hh_id)
```

```
##
## t test of coefficients:
##
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.2702104  0.0152390 17.7315   <2e-16 ***
## treat_ind_FAKE_FOR_PAPTRUE 0.0085315  0.0217556  0.3922    0.695
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We will also test that compliers in the placebo group have similar baseline values to compliers in the treatment group.

```
compliers <- subset(data, contacted==1)
cl(lm(contacted ~ treat_ind_FAKE_FOR_PAP, compliers), compliers$hh_id)
```

```
##
## t test of coefficients:
##
##                               Estimate Std. Error    t value Pr(>|t|)
## (Intercept)                 1.0000e+00 5.2858e-31 1.8918e+30   <2e-16 ***
## treat_ind_FAKE_FOR_PAPTRUE 2.3206e-16 2.3218e-16 9.9950e-01    0.318
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tests for Differential Attrition

The below tests for average differential attrition.

```
table(data$treat_ind_FAKE_FOR_PAP, data$respondent_t1)
```

```
##
##           0   1
##   FALSE 661 242
##   TRUE  656 266
```

```
cl(lm(data$respondent_t1 ~ data$treat_ind_FAKE_FOR_PAP), data$hh_id) #two-tailed p-value
```

```
##
## t test of coefficients:
##
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     0.267996   0.019117 14.0185   <2e-16 ***
## data$treat_ind_FAKE_FOR_PAPTRUE 0.020508   0.027175  0.7546   0.4506
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The below tests for differential attrition by covariates.

```
get.F.stat <- function(treat){
  reduced.model <- lm(data$respondent_t1 ~ x + treat)
  xXt <- matrix(nrow = nrow(x), ncol = ncol(x))
  for(col in 1:ncol(x)) xXt[,col] <- as.numeric(treat) * x[,col]
  full.model <- lm(data$respondent_t1 ~ x + treat + xXt)
  return(anova(reduced.model, full.model)$F[2])
}
f.distribution.under.null <- apply(perms, 2, get.F.stat)
mean(f.distribution.under.null <= get.F.stat(data$treat_ind_FAKE_FOR_PAP))
```

```
## [1] 0.54
```

Not in This PAP

Canvasser Heterogeneity

We have not yet received data from the partner group on which canvassers were responsible for canvassing each individual. When we receive this data we intend to conduct tests of whether the ATE of canvassing is larger for trans canvassers than for all other canvasers and whether the ATE of canvassing is larger for trans and gender non-conforming canvassers than for all other canvassers.

Canvassers were randomly assigned to turf using a complicated scheme that we will detail in a follow-up PAP. We plan to account for the uncertainty associated with these estimaties using randomization inference that replicates the process of randomly assigning canvassers to turf.

We will also account for the probability of assignment to each by using inverse probability weighting, with the probabilities determined by simulations of the assignment process.

When a transgender or gender-non-conforming canvasser canvassed together with a cisgender/gender conforming canvasser as a pair, we will count the pair as a transgender / gender-non-conforming observation because such a person was present at the door. The treatment for this analysis will thus be defined as the presence of such a person at the door versus the absence of such a person.

Treatment Effect Heterogeneity by Subject Attributes

We have three theoretical predictions about treatment effect heterogeneity: 1. Democrats will be more treatment responsive than Republicans as a result of partisan cues being more aligned, 2. subjects higher on the baseline support scale will be more treatment responsive as a result of being more open to outgroup

rights in general, and 3. subjects higher in need for cognition will show larger and longer-lasting effects (we did not ask this item on the baseline, t1, or t2 surveys; please see the "Wave 3 Survey" section).

We also plan two other steps that we will likely not include in the academic paper: 1. We will use BART to detect any patterns of effects that may assist in targeting, 2. We will validate the ability of the partner group's "mover model" to predict heterogeneity.

The code for evaluating the mover model is below.

```
include.indicator <- data$contacted == 1 & data$respondent_t1 == 1 & !is.na(data$vf_clalgbtmov1)
movmodellm <- lm(main.dv[include.indicator] ~
                   data$treat_ind[include.indicator]*data$vf_clalgbtmov1[include.indicator] +
                   x[include.indicator,])
cl(movmodellm, data$hh_id[include.indicator])[c(2:3,32),]
```

```
##                                                                                 Estimate
## data$treat_ind[include.indicator]TRUE                                        -0.40214972
## data$vf_clalgbtmov1[include.indicator]                                        0.01496077
## data$treat_ind[include.indicator]TRUE:data$vf_clalgbtmov1[include.indicator]  0.01144489
##                                                                              Std. Error
## data$treat_ind[include.indicator]TRUE                                       0.370530092
## data$vf_clalgbtmov1[include.indicator]                                      0.009431906
## data$treat_ind[include.indicator]TRUE:data$vf_clalgbtmov1[include.indicator] 0.010397427
##                                                                               t value
## data$treat_ind[include.indicator]TRUE                                       -1.085336
## data$vf_clalgbtmov1[include.indicator]                                       1.586187
## data$treat_ind[include.indicator]TRUE:data$vf_clalgbtmov1[include.indicator]  1.100742
##                                                                              Pr(>|t|)
## data$treat_ind[include.indicator]TRUE                                       0.2786001
## data$vf_clalgbtmov1[include.indicator]                                      0.1136970
## data$treat_ind[include.indicator]TRUE:data$vf_clalgbtmov1[include.indicator] 0.2718466
```

Follow-Up Surveys

We anticipate filing mini-PAPs for the follow-up surveys but for now register the following.

Wave 2 Survey

The wave 2 survey contains the following two items that we plan to add to the list of main dependent variables in that survey to be factor analyzed as part of the rest.

- Transgender women (people who identify as women but were born men) should not be allowed to serve as public school teachers.
- It would be wrong to allow a transgender woman (a person who identifies as a woman but was born as a man) to use the woman's restroom.

We also plan to pool these two items and analyze them as a measure of 'successful innoculation against the opposition message.'

The following item will also be added to the list of secondary dependent variables capturing gender non-conformity.

- Men should dress like men and women should dress like women.

Wave 3 Survey

In addition to assessing persistence, we currently plan to test the below two additional hypotheses on the Wave 3 survey:

1. Does a need for cognition battery predict the size of the treatment effect?
2. Will canvassed voters be less persuaded by the opposition message? We plan to operationalize this with some kind of exposure to the opposition message in the survey, such as by asking voters to watch a short video or showing them pictures of a piece of mail.