EPA2023 Program Report

Nick Zambrotta

2023-03-03

Introduction

This is a program report for my EPA2023 research. I will be presenting my findings on March 3rd, 2023 in Boston, MA.

The goal of my research is to identify if LGBT contact mitigates prejudicial attitudes held by conservatives and religious people. I measure the effects of these predictors on two outcome variables: support for marriage equality and LGBT favorability.

Results show that while political ideology and religiosity predict support marriage equality independently, when interacting with LGBT contact, there is no moderation effect. Alternatively, political ideology predicts LGBT favorability independently and when interacting with LGBT contact, suggesting that knowing an LGBT person mitigates the effects of prejudice on LGBT favorability. There were no significant effects for religiosity.

These findings imply that intergroup contact may mitigate prejudicial attitudes towards LGBT persons regarding how they are viewed, but has little effect on policy implementation. Future implications can be made, particularly in developing durable interventions aimed at lessening bias towards LGBT persons.

Pre-processing and Cleaning

From my GitHub repository, import EPA2023_RawEnviron.RData file containing the data frames anes, gss, and anes_gss2020rawdata.

First, select the variables of interest from anes_gss2020rawdata. Then, rename these variables so they are easier to understand (see GitHub repo for details and descriptions). Reclassify all variables to numerics for later analysis.

```
data <- anes_gss2020rawdata

library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
# select vars
data <- select(data, c(AGE_1A, SEX_1A, RACE_1A, SEXORNT_1A, ATTEND_1A, MARHOMO1_1A,
                       POLVIEWS_1A, V202166, V202172, V202471, V202473))
```

```
# rename vars
data <- dplyr::rename(data, age = AGE_1A, sex = SEX_1A, race = RACE_1A, sexori = SEXORNT_1A, relig_atten

# columns to numeric
data[,1:11] <- lapply(data[,1:11], as.numeric)
```

Next, recode equal_mar variable from low to high support (e.g., reverse code) and clean/revalue age, feel_LGB, feel_T, know_LGB, and know_T.

```
# recode equal mar so low to hi support

data$equal_mar <- (6-data$equal_mar)

# wrangle missing data

data$age[data$age < 18] <- NA
data$feel_LGB[data$feel_LGB < 0] <- NA
data$feel_T[data$feel_T < 0] <- NA

data$know_LGB[data$know_LGB < 0] <- NA
data$know_LGB[data$know_LGB == 2] <- 0

data$know_T[data$know_T < 0] <- NA
data$know_T[data$know_T == 2] <- 0
```

Last, create aggregate variables feel_LGBT and know_LGBT derived from the paired measures (e.g., feel_LGB and feel_T, know_LGB and feel_T).

```
# create vars feel_LGBT and know_LBGT as avg. of paired measures

data$feel_LGBT <- ((data$feel_LGB + data$feel_T)/2)
data$know_LGBT <- ((data$know_LGB + data$know_T)/2)
```

Clean R environment to prevent clutter:

```
ls()
```

```
## [1] "anes"                "anes_gss2020rawdata" "data"
## [4] "gss"
```

```
rm(anes, gss, anes2020rawdata)
```

```
## Warning in rm(anes, gss, anes2020rawdata): object 'anes2020rawdata' not found
```

Correlations

Now that pre-processing is complete, create a data frame that consists of the variables of interest. Rename the variables so they are clear and summarize the measure.

```
# select vars, reorder them
cor <- data[,c(5:7, 12:13)]
cor <- select(cor, c(poli, everything()))
```

```
# rename vars
cor <- rename(cor, `Political Ideology` = poli, `Religiosity` = relig_attend,
              `Pro-Marriage Equality` = equal_mar, `LGBT Favorability` = feel_LGBT,
              `Know LGBT` = know_LGBT)
```

Next, load in EPA2023_Environ.RData from GitHub. This loads in the `correlation_matrix()` custom function.

Using the `correlation_matrix()` function, create a correlation matrix with this data frame.

```
## Loading required package: Hmisc

## Loading required package: lattice

## Loading required package: survival

## Loading required package: Formula

## Loading required package: ggplot2

##
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:dplyr':
##
##     src, summarize

## The following objects are masked from 'package:base':
##
##     format.pval, units
```

Export this object as a .pdf file, use the following code:

```
library(gridExtra)
pdf("~/Desktop/cor_mat.plt.pdf", height = 4, width = 12)
grid.table(cor.matrix)
dev.off()
```

Next, determine the *r*-, *t*-, and *p-values* of the correlations of interest, particularly how relig_attend and poli relate to each other, feel_LGBT, and know_LGBT.

```
attach(data) # or attach(anes_gss2020clean) ... either works! #

# relig x poli
cor.test(relig_attend, poli)
```

```
##
##  Pearson's product-moment correlation
##
## data:  relig_attend and poli
## t = 7.5697, df = 523, p-value = 1.702e-13
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.2349759 0.3893387
## sample estimates:
##       cor
## 0.3142327
```

```
# relig + poli x equal_mar
cor.test(relig_attend, equal_mar)
```

```
##
##  Pearson's product-moment correlation
##
## data:  relig_attend and equal_mar
## t = -9.7433, df = 346, p-value < 2.2e-16
```

```
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.5426608 -0.3772741
## sample estimates:
##        cor
## -0.4640013
```

```
cor.test(poli, equal_mar)
```

```
##
##  Pearson's product-moment correlation
##
## data:  poli and equal_mar
## t = -8.8483, df = 343, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.5133126 -0.3410248
## sample estimates:
##        cor
## -0.4310898
```

```
# relig + poli x feel_LGBT
cor.test(relig_attend, feel_LGBT)
```

```
##
##  Pearson's product-moment correlation
##
## data:  relig_attend and feel_LGBT
## t = -3.5792, df = 513, p-value = 0.0003773
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.23926626 -0.07063984
## sample estimates:
##        cor
## -0.1560902
```

```
cor.test(poli, feel_LGBT)
```

```
##
##  Pearson's product-moment correlation
##
## data:  poli and feel_LGBT
## t = -9.0831, df = 510, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.4454080 -0.2960727
## sample estimates:
##        cor
## -0.3731548
```

Regression

After determining how the variables of interest correlate to one another, prepare the data for linear regression by creating a new data frame that considers complete cases only.

```
data.na <- data[complete.cases(data),]
# or don't do this - anes_gss2020regression works too! #
```

Next, create two models that predict equal_mar; one that doesn't contain an interaction with know_LGBT and one that does. Once regressions are processed, conduct an ANOVA analysis to see if there is a difference between the two models.

```
# without interaction
mar_equality <- lm(equal_mar ~ relig_attend + poli + age + sex + race + sexori, data.na)
summary(mar_equality)
```

```
##
## Call:
## lm(formula = equal_mar ~ relig_attend + poli + age + sex + race +
##     sexori, data = data.na)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9490 -0.4276  0.0804  0.6094  1.9218
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   6.21356    0.56724  10.954  < 2e-16 ***
## relig_attend -0.16146    0.03115  -5.183 6.06e-07 ***
## poli         -0.25957    0.05684  -4.566 9.42e-06 ***
## age          -0.01817    0.00526  -3.454 0.000694 ***
## sex           0.32765    0.15937   2.056 0.041302 *
## race         -0.12021    0.15182  -0.792 0.429544
## sexori       -0.06516    0.16356  -0.398 0.690844
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.018 on 172 degrees of freedom
## Multiple R-squared:  0.3694, Adjusted R-squared:  0.3474
## F-statistic: 16.79 on 6 and 172 DF,  p-value: 3.284e-15
```

```
# with interaction
mar_equality2 <- lm(equal_mar ~ relig_attend*know_LGBT + poli*know_LGBT + age + sex + race + sexori, da
summary(mar_equality2)
```

```
##
## Call:
## lm(formula = equal_mar ~ relig_attend * know_LGBT + poli * know_LGBT +
##     age + sex + race + sexori, data = data.na)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.1584 -0.3643  0.0850  0.6024  1.8868
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              6.512386   0.688985   9.452  < 2e-16 ***
## relig_attend            -0.148097   0.051010  -2.903 0.004186 **
## know_LGBT               -0.696446   0.604749  -1.152 0.251101
## poli                    -0.364777   0.107329  -3.399 0.000845 ***
## age                     -0.018568   0.005346  -3.473 0.000654 ***
## sex                      0.310773   0.160208   1.940 0.054068 .
## race                    -0.104300   0.153018  -0.682 0.496412
## sexori                  -0.044120   0.164695  -0.268 0.789109
```

```
## relig_attend:know_LGBT -0.003999   0.092735  -0.043 0.965651
## know_LGBT:poli          0.232444   0.171182   1.358 0.176314
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.02 on 169 degrees of freedom
## Multiple R-squared:  0.3782, Adjusted R-squared:  0.3451
## F-statistic: 11.42 on 9 and 169 DF,  p-value: 6.788e-14
```

```
anova(mar_equality, mar_equality2)
```

```
## Analysis of Variance Table
##
## Model 1: equal_mar ~ relig_attend + poli + age + sex + race + sexori
## Model 2: equal_mar ~ relig_attend * know_LGBT + poli * know_LGBT + age +
##     sex + race + sexori
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    172 178.19
## 2    169 175.68  3     2.507 0.8039 0.4933
```

Results show that the dependent variables relig_attend and poli are significant predictors of equal_mar when controlling for other factors. However, know_LGBT does not moderate the relationship between the dependent variables and the independent variable as shown by the ANOVA output.

Create two models that predict feel_LGBT; one that doesn't contain an interaction with know_LGBT and one that does. Once regressions are processed, conduct an ANOVA analysis to see if there is a difference between the two models.

```
# without interaction
LGBT_favor <- lm(feel_LGBT ~ relig_attend + poli + age + sex + race + sexori, data.na)
summary(LGBT_favor)
```

```
##
## Call:
## lm(formula = feel_LGBT ~ relig_attend + poli + age + sex + race +
##     sexori, data = data.na)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -72.140 -15.408   0.936  16.822  45.896
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  115.17440   12.74717   9.035 3.29e-16 ***
## relig_attend   0.04026    0.69999   0.058  0.95420
## poli          -6.69446    1.27739  -5.241 4.64e-07 ***
## age           -0.27043    0.11820  -2.288  0.02336 *
## sex           10.43333    3.58140   2.913  0.00405 **
## race          -6.75622    3.41164  -1.980  0.04926 *
## sexori        -7.14754    3.67569  -1.945  0.05346 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.87 on 172 degrees of freedom
## Multiple R-squared:  0.2326, Adjusted R-squared:  0.2058
## F-statistic: 8.687 on 6 and 172 DF,  p-value: 2.9e-08
```

```r
# with interaction
LGBT_favor2 <- lm(feel_LGBT ~ relig_attend*know_LGBT + poli*know_LGBT + age + sex + race + sexori, data
summary(LGBT_favor2)
```

```
##
## Call:
## lm(formula = feel_LGBT ~ relig_attend * know_LGBT + poli * know_LGBT +
##     age + sex + race + sexori, data = data.na)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -61.288 -14.937   2.931  16.380  41.554
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)            116.0266    14.9552   7.758 7.77e-13 ***
## relig_attend             1.3081     1.1072   1.181  0.23909
## know_LGBT              -14.7649    13.1268  -1.125  0.26227
## poli                   -10.3165     2.3297  -4.428 1.70e-05 ***
## age                     -0.2544     0.1161  -2.192  0.02976 *
## sex                      9.3704     3.4775   2.695  0.00776 **
## race                    -5.4984     3.3214  -1.655  0.09969 .
## sexori                  -6.1069     3.5749  -1.708  0.08942 .
## relig_attend:know_LGBT  -1.7023     2.0129  -0.846  0.39892
## know_LGBT:poli           9.6416     3.7157   2.595  0.01030 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.13 on 169 degrees of freedom
## Multiple R-squared:  0.2941, Adjusted R-squared:  0.2565
## F-statistic: 7.823 on 9 and 169 DF,  p-value: 1.322e-09
```

```r
anova(LGBT_favor, LGBT_favor2)
```

```
## Analysis of Variance Table
##
## Model 1: feel_LGBT ~ relig_attend + poli + age + sex + race + sexori
## Model 2: feel_LGBT ~ relig_attend * know_LGBT + poli * know_LGBT + age +
##     sex + race + sexori
##   Res.Df   RSS Df Sum of Sq      F   Pr(>F)
## 1    172 89985
## 2    169 82773  3    7212.1 4.9084 0.002695 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results show that poli is a significant predictor of feel_LGBT when controlling for other factors. Additionally, the interaction between poli and know_LGBT is statistically significant, suggesting that LGBT contact moderates this relationship. There is no significant effect for relig_attend in either model. The ANOVA output indicates that the models are statistically different.

Data Visualization

I was having issues with the `predict()` function in my analyses, so I opted to create four custom functions that consist of the regression equations of each model. Using these equations, I included each term and created four new columns containing their predicted values.

```r
attach(data.na) # or attach(anes_gss2020regression) ... either works!
```

```
## The following objects are masked from data:
##
##     age, equal_mar, feel_LGB, feel_LGBT, feel_T, know_LGB, know_LGBT,
##     know_T, poli, race, relig_attend, sex, sexori
```

```r
equal_mar1 <- function(a,b,c,d,f,g){
  6.124 - .161*a -.260*b -.019*c + .328*d - .120*f - .064*g}
equal_mar2 <- function(a,b,c,d,f,g,h,i,j){
  6.512 - .148*a - .696*b - .365*c - .019*d + .311*f - .104*g - .044*h - .004*i + .232*j}
feel_LGBT1 <- function(a,b,c,d,f,g){
  115.174 + .040*a - 6.694*b - .270*c + 10.433*d - 6.756*f - 7.148*g}
feel_LGBT2 <- function(a,b,c,d,f,g,h,i,j){
  116.027 + 1.308*a - 14.765*b - 10.317*c - .254*d + 9.370*f - 5.498*g - 6.107*h -  1.702*i + 9.642*j}
```

```r
# equal_mar functions
data.na$predicted <- equal_mar1(relig_attend,poli,age,sex,race,sexori)
data.na$predicted2 <- equal_mar2(relig_attend,know_LGBT,poli,age,sex,race,sexori,relig_attend*know_LGBT

# feel_LGBT functions
data.na$predicted3 <- feel_LGBT1(relig_attend,poli,age,sex,race,sexori)
data.na$predicted4 <- feel_LGBT2(relig_attend,know_LGBT,poli,age,sex,race,sexori,relig_attend*know_LGBT
```

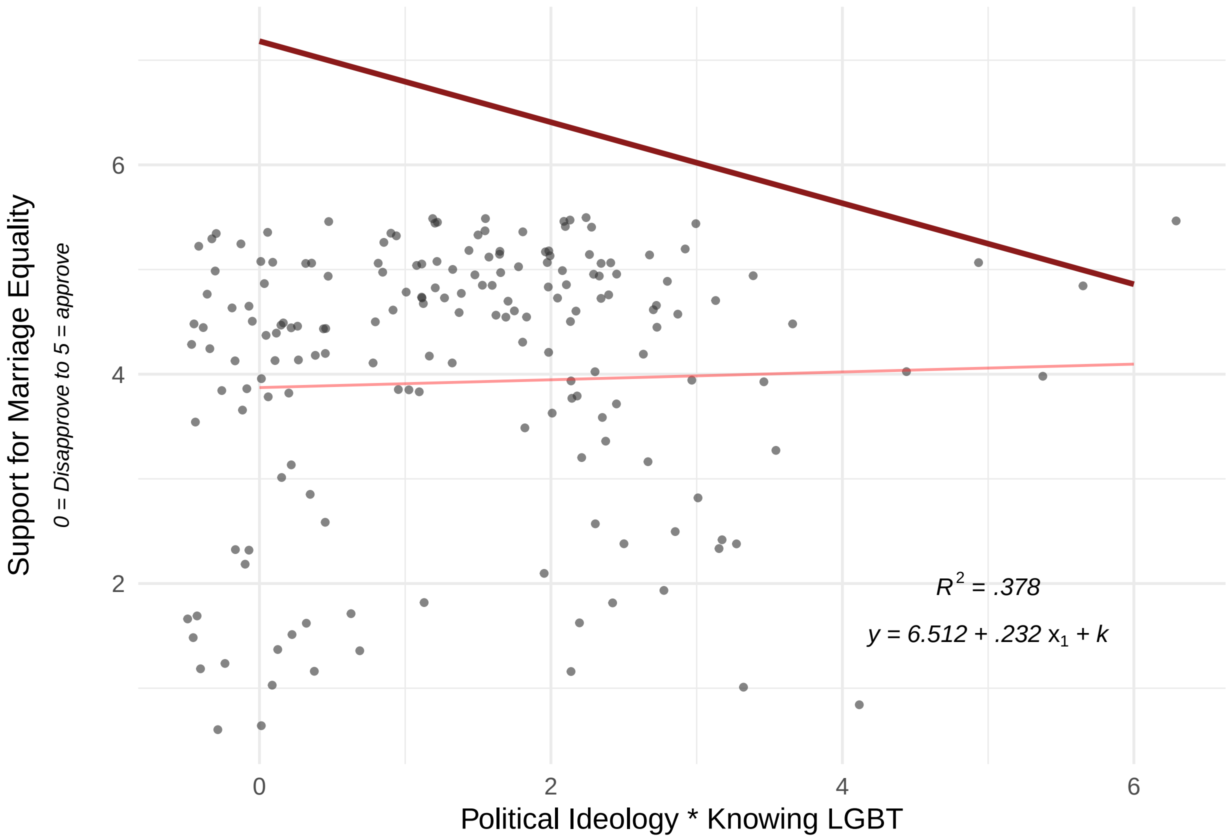
With these predicted values the data can be visiualized using the package `ggplot2`. In the first plot, I create a jitter-plot containing the regression lines for equal_mar. The line including the interaction is in bold.

```r
library(ggplot2)
```

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'

## Warning in is.na(x): is.na() applied to non-(list or vector) of type
## 'expression'

## Warning in is.na(x): is.na() applied to non-(list or vector) of type
## 'expression'
```

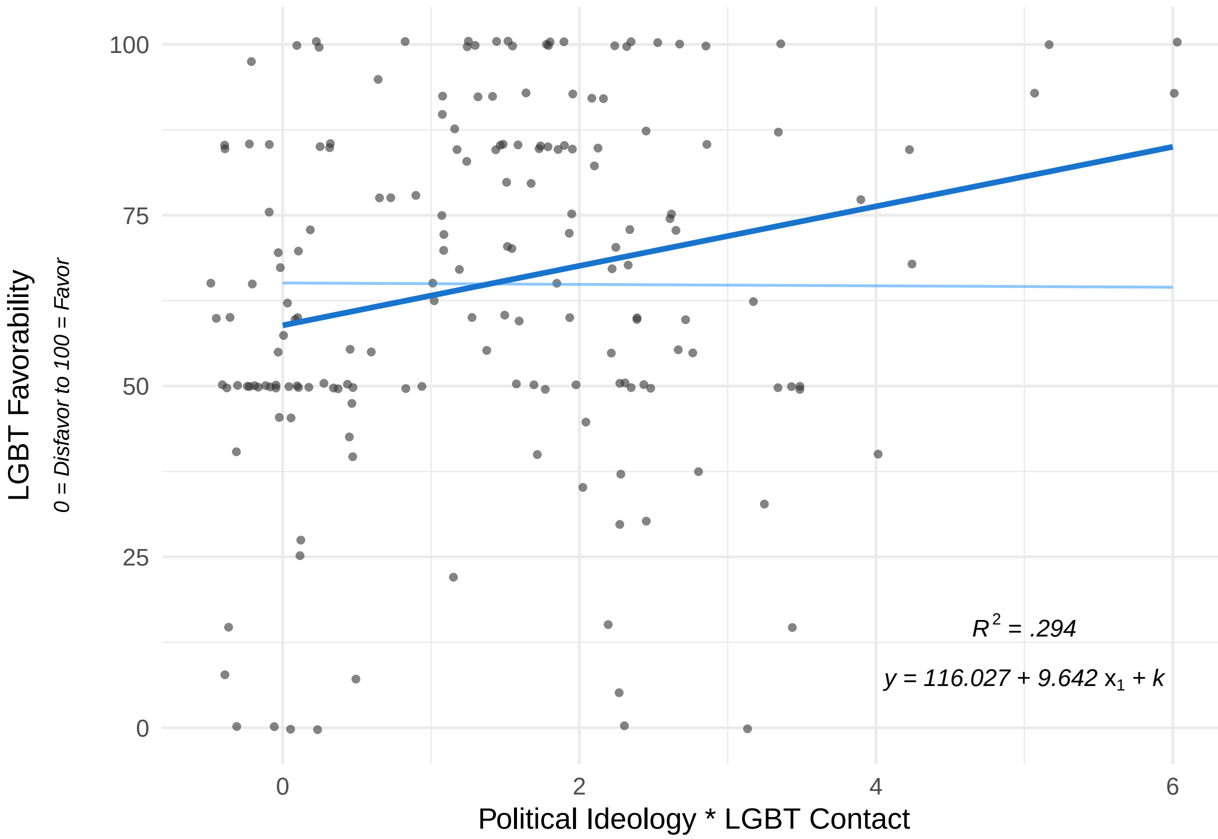


Similarly, plot the regression lines for feel_LGBT. This plot displays a significant moderator effect! The bold line is the one with the interaction.

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning in is.na(x): is.na() applied to non-(list or vector) of type
## 'expression'
```

```
## Warning in is.na(x): is.na() applied to non-(list or vector) of type
## 'expression'
```


LGBT Favorability
0 = Disfavor to 100 = Favor
100
75
50
25
0
0
2
4
6
Political Ideology * LGBT Contact
R^2 = .294
$y = 116.027 + 9.642\ x_1 + k$