PS532 Final Paper: Exploring the Relationship Between Perceived Discrimination and Political Participation: Evidence from Asian Americans

Qiaoyi Nie

2023-12-19

```
#load dataset
setwd("/Users/qiaoyinie/Desktop/PS532Quant III/PS532-Final-Project")
getwd()
```

```
## [1] "/Users/qiaoyinie/Desktop/PS532Quant III/PS532-Final-Project"
```

```
data <- read.csv("Asian Americans' Experiences of Discrimination and Political Participation Survey_Augu
View(data)

###data cleaning
fdata <- data[-c(1,2,3,4,12,27,28,50,59,178,205,250,267,268,269,270,271,276,278,279,281,286,288,289,293
View(fdata)

install.packages('tidyverse')
```

```
##
## The downloaded binary packages are in
##  /var/folders/9m/8_zrd0b55yvbt7g8szv9rr5c0000gn/T//Rtmp0E81D1/downloaded_packages
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
#mutate new variables
fdata$X1<- as.numeric(fdata$X1)
fdata$X2<- as.numeric(fdata$X2)
fdata$X3<- as.numeric(fdata$X3)
fdata$X4<- as.numeric(fdata$X4)
fdata$X5<- as.numeric(fdata$X5)

fdata1 <- fdata %>%
  mutate(MPP = rowMeans(select(fdata, c(25:29)), na.rm = TRUE))
```

```r
fdata1 <- fdata1 %>%
  mutate_at(vars(c(30:129)), as.numeric)

fdata1 <- fdata1 %>%
  mutate(APP = rowMeans(select(fdata1, c(30:129)), na.rm = TRUE))

###subset treatment
fdata1 <- fdata1 %>%
  mutate_at(vars(c(15:24)), as.numeric)

fdata1 <- fdata1 %>%
  mutate(Treatment = case_when(
    PD1 >= 1 ~ "1",
    SD1 >= 1 ~ "2",
    P1 >= 1 ~ "3",
    PD1.1 >= 1 ~ "4",
    TRUE ~ "F"  # Default case if none of the conditions match
  ))

fulldata <- fdata1 %>%
  select(c(1:24,130,132,133,134))


###preps for analysis

#assign meaning
install.packages("Hmisc")
```

```
##
## The downloaded binary packages are in
##  /var/folders/9m/8_zrd0b55yvbt7g8szv9rr5c0000gn/T//Rtmp0E81D1/downloaded_packages
```

```r
library(Hmisc)
```

```
##
## Attaching package: 'Hmisc'
##
## The following objects are masked from 'package:dplyr':
##
##     src, summarize
##
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

```r
label(fulldata$Race.and.Ethnicity)   <- "Ethnicity"
label(fulldata$Q19)   <- "Frequency of Dis"

###descriptive analysis
#ethnicity
unique(fulldata$Asian)
```

```
##  [1] "1"  "6"  "20" "7"  "11" "2"  "8"  "10" "19" "5"  "17" "3"  "12" "15" "9"
## [16] "18" "13" "14"
```

```r
fulldata <- fulldata[order(fulldata$Asian, decreasing = TRUE), ]

fulldata <- fulldata %>%
  mutate(Ethcinity = case_when(
    Asian== 1 ~ "Asian Indian",
    Asian== 2 ~ "Bangladeshi",
    Asian== 3 ~ "Bhutanese",
    Asian== 4 ~ "Burmese",
    Asian== 5 ~ "Cambodian",
    Asian== 6 ~ "Chinese",
    Asian== 7 ~ "Philippino",
    Asian== 8 ~ "Hmong",
    Asian== 9 ~ "Indonesian",
    Asian== 10 ~ "Japanese",
    Asian== 11 ~ "Korean",
    Asian== 12 ~ "Laotian",
    Asian== 13 ~ "Mongolian",
    Asian== 14 ~ "Malaysian",
    Asian== 15 ~ "Nepalese",
    Asian== 16 ~ "Okinawan",
    Asian== 17 ~ "Pakistani",
    Asian== 18 ~ "Sri Lankan",
    Asian== 19 ~ "Thai",
    Asian== 20 ~ "Vietnamese",
    TRUE ~ "F"  # Default case if none of the conditions match
  ))
```

```r
##linear regression without matching
  ##subset
  fulldata <- fulldata %>%
    mutate(Treat = case_when(
      Treatment== 3 ~ "1",
      Treatment== 1 ~ "2",
      Treatment== 2 ~ "3",
      Treatment== 4 ~ "4",
    ))

  PD <- fulldata[fulldata$Treat ==2, ]
  SD <- fulldata[fulldata$Treat ==3, ]
  Placebo <- fulldata[fulldata$Treat ==1, ]

  ## combine the treated and control/placebo group
  PD <- subset(fulldata, Treat < 3)
  SD <- subset(fulldata, Treat ==1 | Treat == 3)

   PD <- PD %>%
    mutate(Treat = case_when(
      Treatment== 1 ~ "1",
      Treatment== 3 ~ "0",
    ))

  SD <- SD %>%
    mutate(Treat = case_when(
      Treatment== 2 ~ "1",
```

```r
    Treatment== 3 ~ "0",
  ))


#remove NAs columns and rows
PD <- subset(PD, select = c(Age,Gender,Race.and.Ethnicity,Birthplace,Education,Residence.area.1,Househ
PD <- PD[-4,]
PD <- PD[-40,]
PD <- PD[-75,]

SD <- subset(SD, select = c(Age,Gender,Race.and.Ethnicity,Birthplace,Education,Residence.area.1,Househ
SD <- SD[-5,]
SD <- SD[-77,]


#Check the observations
table(PD$Treat) #control: 79 Treated:77
```

```
##
##  0  1
## 79 77
```

```r
table(SD$Treat) #control: 79 Treated:81
```

```
##
##  0  1
## 79 81
```

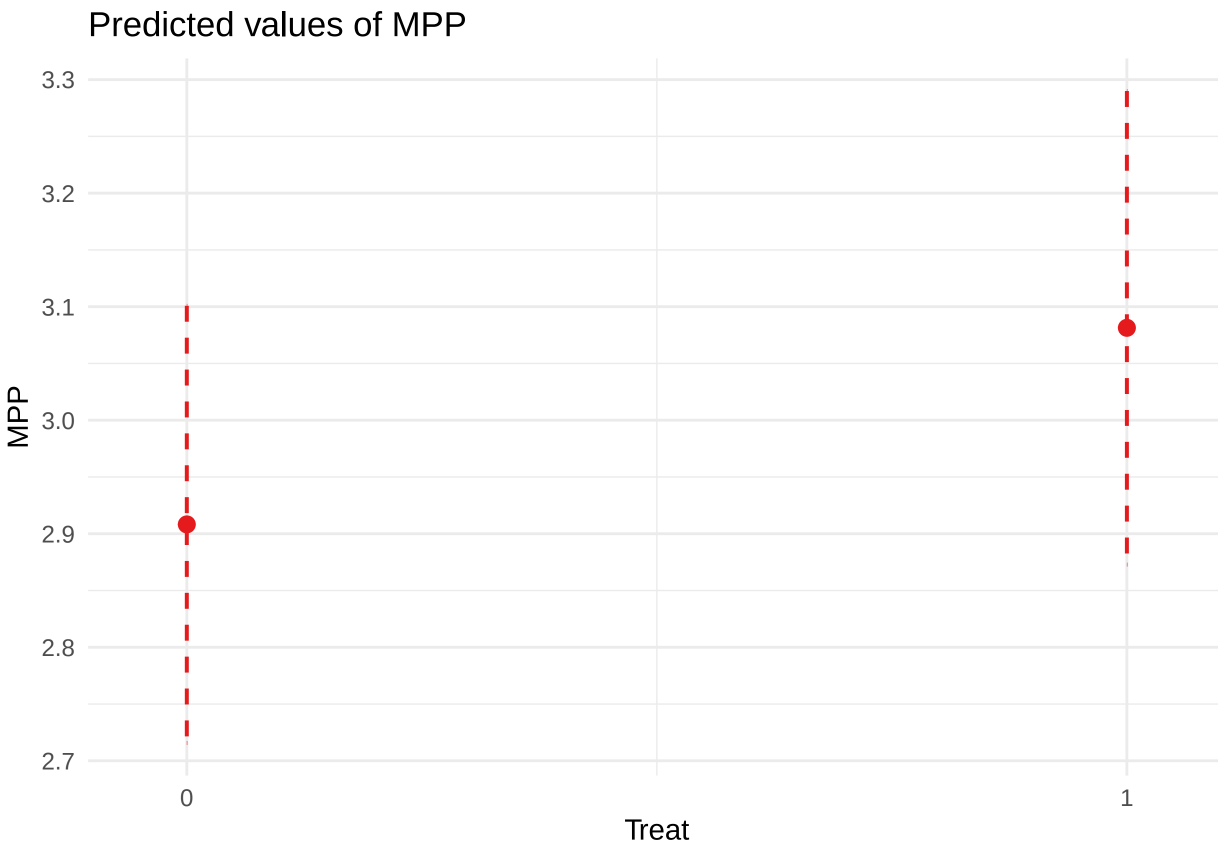
```r
#regression analysis
install.packages("estimatr")
```

```
##
## The downloaded binary packages are in
##  /var/folders/9m/8_zrd0b55yvbt7g8szv9rr5c0000gn/T//Rtmp0E81D1/downloaded_packages
```

```r
library(estimatr)
lm1 <- lm_robust(MPP ~ Treat, data=PD)
summary(lm1)
```

```
##
## Call:
## lm_robust(formula = MPP ~ Treat, data = PD)
##
## Standard error type:  HC2
##
## Coefficients:
##             Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper  DF
## (Intercept)   2.9082    0.09747  29.838 6.768e-66   2.7157    3.101 154
## Treat1        0.1732    0.14369   1.205 2.300e-01  -0.1107    0.457 154
##
## Multiple R-squared:  0.009357 ,  Adjusted R-squared:  0.002925
## F-statistic: 1.452 on 1 and 154 DF,  p-value: 0.23
```

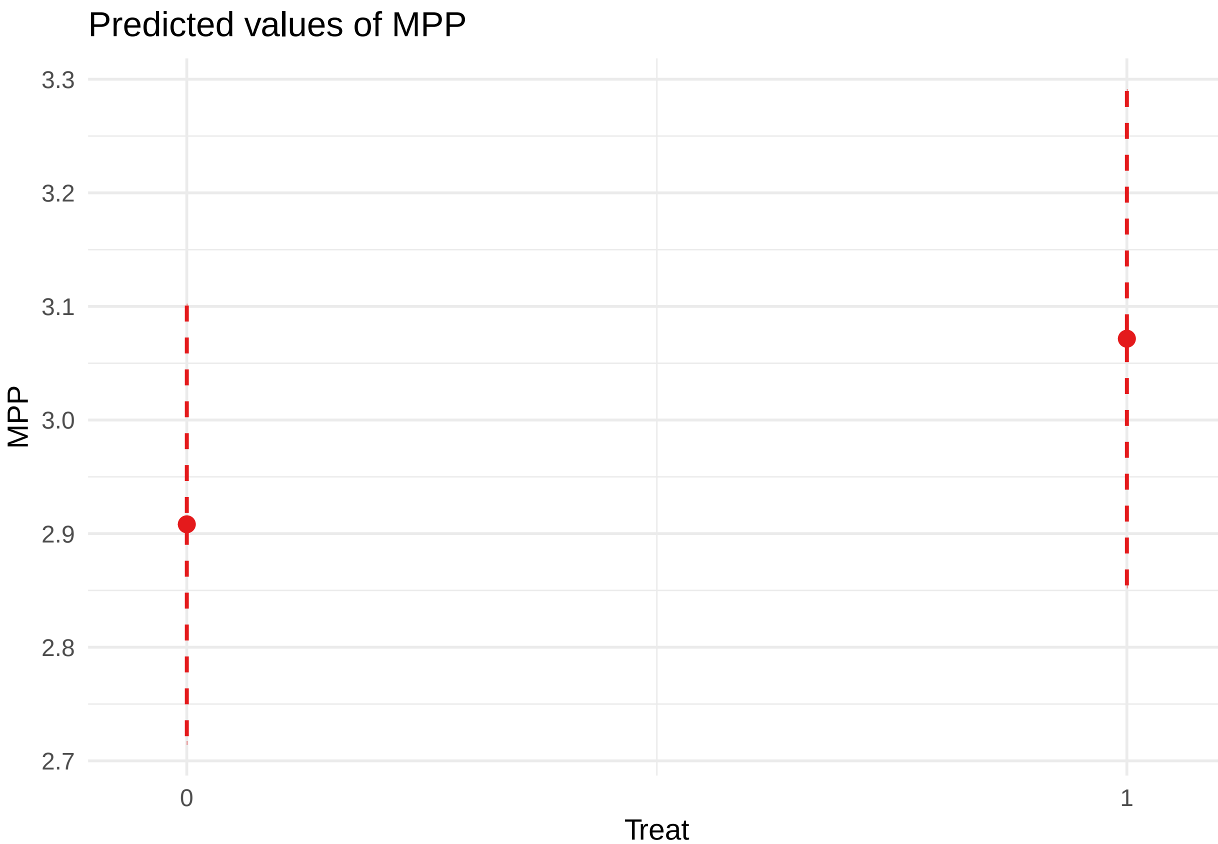
```r
sjPlot::plot_model(lm1, type = "pred")+
  theme(plot.background = element_rect(fill = "white"))+
  theme_minimal()
```




```r
lm2 <- lm_robust(MPP ~ Treat, data=SD)

summary(lm2)
```

```
##
## Call:
## lm_robust(formula = MPP ~ Treat, data = SD)
##
## Standard error type:  HC2
##
## Coefficients:
##             Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper  DF
## (Intercept)   2.9082    0.09747  29.838 8.161e-67   2.7157   3.1007 158
## Treat1        0.1634    0.14726   1.109 2.689e-01  -0.1275   0.4542 158
##
## Multiple R-squared:  0.007704 ,  Adjusted R-squared:  0.001423
## F-statistic: 1.231 on 1 and 158 DF,  p-value: 0.2689
```

```r
sjPlot::plot_model(lm2, type = "pred", )+
  theme_minimal()
```




```
##matching

  # Display the result
  #print(PD)
  #print(SD)

  ### Propensity Score Matching
  install.packages("Matching")
```

```
##
## The downloaded binary packages are in
##  /var/folders/9m/8_zrd0b55yvbt7g8szv9rr5c0000gn/T//Rtmp0E81D1/downloaded_packages
```

```
  library(Matching)
```

```
## Loading required package: MASS
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

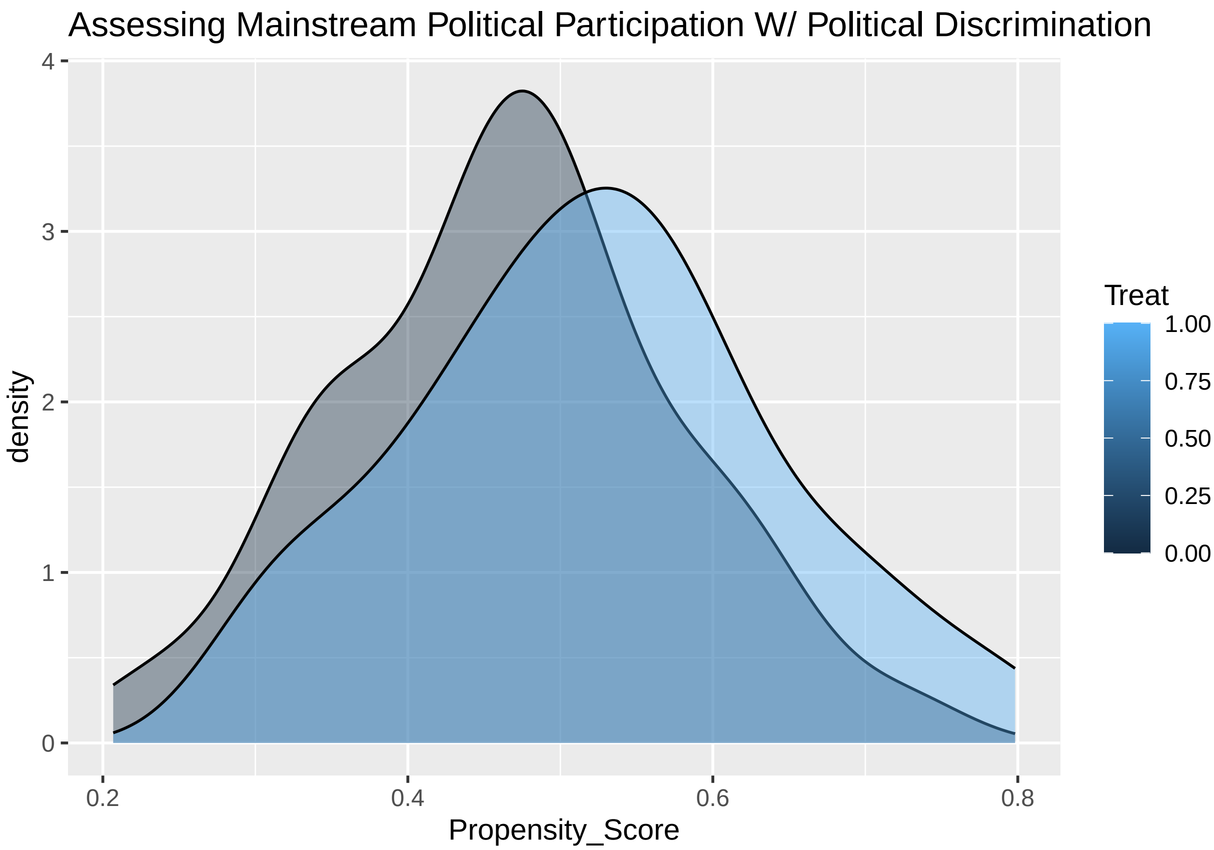
```
## ##
## ##  Matching (Version 4.10-14, Build Date: 2023-09-13)
## ##  See https://www.jsekhon.com for additional documentation.
## ##  Please cite software as:
## ##   Jasjeet S. Sekhon. 2011. ``Multivariate and Propensity Score Matching
## ##   Software with Automated Balance Optimization: The Matching package for R.''
```

```
## ##   Journal of Statistical Software, 42(7): 1-52.
## ##
```

```
library(ggplot2)
library(Matching)


# dependent variable
y1 <- PD$MPP
y2 <- SD$MPP

# treatment
tr1 <- PD$Treat
tr2 <- SD$Treat

# Convert character variable to numeric
PD$Treat <- as.numeric(PD$Treat)
PD$Age <- as.numeric(PD$Age)
PD$Race.and.Ethnicity <- as.numeric(PD$Race.and.Ethnicity)
PD$Gender <- as.numeric(PD$Gender)
PD$Birthplace <- as.numeric(PD$Birthplace)
PD$Education <- as.numeric(PD$Education)
PD$Household.Income.Lev <- as.numeric(PD$Household.Income.Lev)
PD$Discriminatory.Exper <- as.numeric(PD$Discriminatory.Exper)

SD$Treat <- as.numeric(SD$Treat)
SD$Age <- as.numeric(SD$Age)
SD$Race.and.Ethnicity <- as.numeric(SD$Race.and.Ethnicity)
SD$Gender <- as.numeric(SD$Gender)
SD$Birthplace <- as.numeric(SD$Birthplace)
SD$Education <- as.numeric(SD$Education)
SD$Household.Income.Lev <- as.numeric(SD$Household.Income.Lev)
SD$Discriminatory.Exper <- as.numeric(SD$Discriminatory.Exper)


# create a propensity score model for the political discrimination treatment
glm1 <- glm(Treat~Age+Race.and.Ethnicity+Gender+Birthplace+Education+Residence.area.1+Household.Income
            family=binomial, data=PD)

#make sure the length of fitted value is the same with the treated ones
length(glm1$fitted)
```

```
## [1] 156
```

```
joinpscore1 = cbind(PD, glm1$fitted)

colnames(joinpscore1)[12] = "Propensity_Score"

# With transparency
 ggplot(data=joinpscore1, aes(x=Propensity_Score, group=Treat, fill=Treat)) + geom_density(adjust=1, a
```




```r
# default options: estimand="ATT", M=1, exact=NULL,caliper=NULL, replace=TRUE


#one-to-one matching for PD
rr1 <- Match(Y=y1, Tr=tr1, X=glm1$fitted)
summary(rr1)
```

```
##
## Estimate...  -0.1513
## AI SE......  0.1955
## T-stat.....  -0.77392
## p.val......  0.43898
##
## Original number of observations.............  156
## Original number of treated obs...............  77
## Matched number of observations...............  77
## Matched number of observations  (unweighted).  80
```

```r
atc <- Match(Y=y1, Tr=tr1, X=glm1$fitted, estimand = "ATC")
summary(atc)
```

```
##
## Estimate...  0.11709
## AI SE......  0.18916
## T-stat.....  0.61898
## p.val......  0.53593
##
## Original number of observations.............  156
```

```
## Original number of control obs...............  79
## Matched number of observations...............  79
## Matched number of observations  (unweighted).  81
```

```
matched <- PD[c(rr1$index.treated,rr1$index.control),]

Matchdefinition  <- function(Y=NULL,Tr,X,Z=X,V=rep(1,length(Y)), estimand="ATT", M=1, BiasAdjust=FALSE


# Assess balance for one covariate
# Do we have balance on discrimination experience after matching?
qqout1 <- qqstats(PD$Discriminatory.Exper[rr1$index.treated], PD$Discriminatory.Exper[rr1$index.contro
print(qqout1)
```

```
## $meandiff
## [1] 0.04375
##
## $mediandiff
## [1] 0.04375
##
## $maxdiff
## [1] 0.0875
```

```
# check balance
postmatchbal1 = MatchBalance(Treat~Age+Gender+Race.and.Ethnicity+Birthplace+Education+Residence.area.
```

```
##
## ***** (V1) Age *****
##                        Before Matching        After Matching
## mean treatment........     2.3896             2.3896
## mean control..........     2.5443             2.2511
## std mean diff.........    -10.718             9.5982
##
## mean raw eQQ diff.....    0.19481             0.1875
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  2
##
## mean eCDF diff........   0.031317            0.03125
## med  eCDF diff........   0.023673             0.0375
## max  eCDF diff........   0.091731              0.075
##
## var ratio (Tr/Co).....     1.0578             1.2786
## T-test p-value........    0.49847            0.53096
## KS Bootstrap p-value..      0.486              0.681
## KS Naive p-value......    0.47666            0.66761
## KS Statistic..........   0.091731              0.075
##
##
## ***** (V2) Gender *****
##                        Before Matching        After Matching
## mean treatment........     1.7143             1.7143
## mean control..........     1.5823             1.7468
## std mean diff.........     25.919            -6.3748
##
## mean raw eQQ diff.....    0.14286             0.0875
```

```
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
## mean eCDF diff........   0.044002           0.029167
## med  eCDF diff........   0.025974              0.025
## max  eCDF diff........    0.10603             0.0625
##
## var ratio (Tr/Co).....      1.053             1.3538
## T-test p-value........    0.10327            0.59904
## KS Bootstrap p-value..      0.167              0.433
## KS Naive p-value......    0.18646            0.48209
## KS Statistic..........    0.10603             0.0625
##
##
## ***** (V3) Race.and.Ethnicity *****
##                        Before Matching        After Matching
## mean treatment........      7.961              7.961
## mean control..........     8.0759             7.3506
## std mean diff.........    -2.0758             11.027
##
## mean raw eQQ diff.....    0.49351             0.5125
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          4                  3
##
## mean eCDF diff........    0.02833           0.033929
## med  eCDF diff........     0.0217              0.025
## max  eCDF diff........   0.070689              0.075
##
## var ratio (Tr/Co).....    0.96086             1.0882
## T-test p-value........    0.89805            0.47383
## KS Bootstrap p-value..      0.787              0.801
## KS Naive p-value......    0.83278            0.81233
## KS Statistic..........   0.070689              0.075
##
##
## ***** (V4) Birthplace *****
##                        Before Matching        After Matching
## mean treatment........     1.3636             1.3636
## mean control..........     1.4051             1.3961
## std mean diff.........    -8.5557            -6.7054
##
## mean raw eQQ diff.....   0.038961              0.025
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
## mean eCDF diff........   0.020713             0.0125
## med  eCDF diff........   0.020713             0.0125
## max  eCDF diff........   0.041427              0.025
##
## var ratio (Tr/Co).....    0.96056            0.96739
## T-test p-value........    0.59762            0.66663
##
##
## ***** (V5) Education *****
```

```
##                         Before Matching        After Matching
## mean treatment........     3.7792             3.7792
## mean control..........     3.9367             3.7251
## std mean diff.........    -10.255             3.5235
##
## mean raw eQQ diff.....    0.11688               0.15
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
## mean eCDF diff........   0.026248              0.025
## med  eCDF diff........   0.031317            0.01875
## max  eCDF diff........    0.04455              0.075
##
## var ratio (Tr/Co).....     1.1171             1.2056
## T-test p-value........    0.51179            0.81315
## KS Bootstrap p-value..      0.913              0.647
## KS Naive p-value......    0.90096            0.66057
## KS Statistic..........    0.04455              0.075
##
##
## ***** (V6) Residence.area.12 *****
##                         Before Matching        After Matching
## mean treatment........    0.68831            0.68831
## mean control..........    0.72152             0.7013
## std mean diff.........    -7.1227            -2.7856
##
## mean raw eQQ diff.....   0.025974               0.05
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
## mean eCDF diff........   0.016604              0.025
## med  eCDF diff........   0.016604              0.025
## max  eCDF diff........   0.033207               0.05
##
## var ratio (Tr/Co).....     1.0681             1.0242
## T-test p-value........    0.65195            0.84197
##
##
## ***** (V7) Residence.area.13 *****
##                         Before Matching        After Matching
## mean treatment........   0.064935           0.064935
## mean control..........   0.025316           0.051948
## std mean diff.........     15.974             5.2361
##
## mean raw eQQ diff.....   0.051948             0.0125
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
## mean eCDF diff........   0.019809            0.00625
## med  eCDF diff........   0.019809            0.00625
## max  eCDF diff........   0.039619             0.0125
##
## var ratio (Tr/Co).....     2.4615             1.2329
## T-test p-value........    0.23768            0.56457
```

```
##
##
## ***** (V8) Household.Income.Lev *****
##                        Before Matching        After Matching
## mean treatment........     3.5584             3.5584
## mean control..........     3.8228             3.2468
## std mean diff.........     -14.32             16.885
##
## mean raw eQQ diff.....     0.2987               0.35
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
## mean eCDF diff........   0.046312               0.05
## med  eCDF diff........   0.029919             0.0375
## max  eCDF diff........     0.1491             0.1375
##
## var ratio (Tr/Co).....     1.1286             1.1116
## T-test p-value........    0.35879            0.20164
## KS Bootstrap p-value..      0.127              0.188
## KS Naive p-value......    0.14538            0.19908
## KS Statistic..........     0.1491             0.1375
##
##
## ***** (V9) Discriminatory.Exper *****
##                        Before Matching        After Matching
## mean treatment........     1.5065             1.5065
## mean control..........     1.3924             1.4221
## std mean diff.........     22.671             16.775
##
## mean raw eQQ diff.....    0.11688             0.0875
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
## mean eCDF diff........   0.057044            0.04375
## med  eCDF diff........   0.057044            0.04375
## max  eCDF diff........    0.11409             0.0875
##
## var ratio (Tr/Co).....     1.0487             1.0247
## T-test p-value........    0.15412            0.22269
##
##
## Before Matching Minimum p.value: 0.10327
## Variable Name(s): Gender  Number(s): 2
##
## After Matching Minimum p.value: 0.188
## Variable Name(s): Household.Income.Lev  Number(s): 8
```

```r
# Transform to a tabular format

SMDeepExtract <- function(PD, col = c("mean.Tr",  "mean.Co", "var.Tr", "var.Co", "var.ratio", "p.valu
  cbind(
    t(sapply(PD[[ "BeforeMatching" ]], "[", col)),
    t(sapply(PD[[ "AfterMatching" ]], "[", col))
  )
```

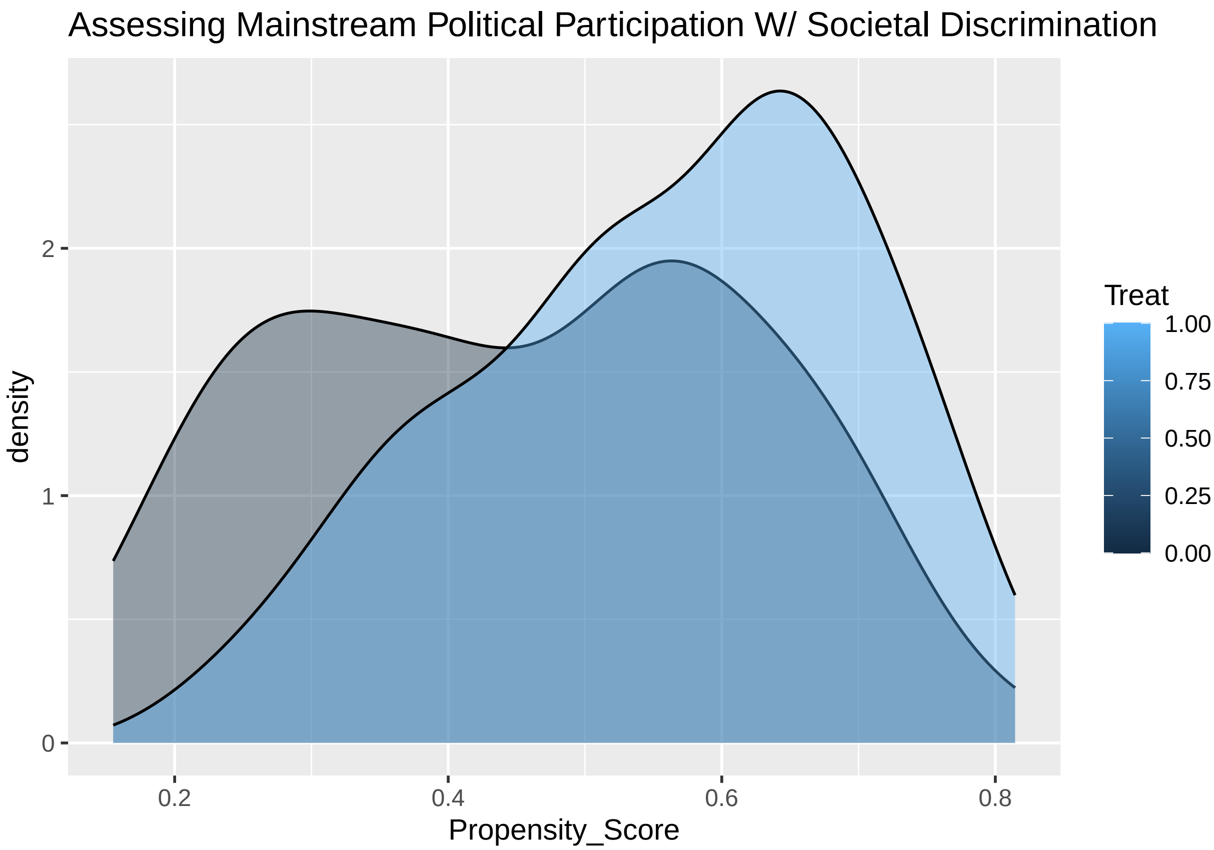
```
}

# extract
res1 <- SMDeepExtract(postmatchbal1)

rownames(res1) = c("Age", "Gender", "Race and Ethnicity", "Birthplace", "Education", "Suburban", "Rura
colnames(res1) = c("mean.Tr",  "mean.Co", "var.Tr", "var.Co", "var.ratio", "p.value", "sdiff", "A.mea

res1[] <- lapply(res1,round,2)
kableExtra::kable(res1[])
```

	mean.Tr	mean.Co	var.Tr	var.Co	var.ratio	p.value	sdiff	A.mean.Tr	A.mean
Age	2.39	2.54	2.08	1.97	1.06	0.5	-10.72	2.39	2.25
Gender	1.71	1.58	0.26	0.25	1.05	0.1	25.92	1.71	1.75
Race and Ethnicity	7.96	8.08	30.64	31.89	0.96	0.9	-2.08	7.96	7.35
Birthplace	1.36	1.41	0.23	0.24	0.96	0.6	-8.56	1.36	1.4
Education	3.78	3.94	2.36	2.11	1.12	0.51	-10.25	3.78	3.73
Suburban	0.69	0.72	0.22	0.2	1.07	0.65	-7.12	0.69	0.7
Rural	0.06	0.03	0.06	0.02	2.46	0.24	15.97	0.06	0.05
Household Income Level	3.56	3.82	3.41	3.02	1.13	0.36	-14.32	3.56	3.25
Discriminatory Experiences	1.51	1.39	0.25	0.24	1.05	0.15	22.67	1.51	1.42

```
# create a propensity score model for the second treatment
glm2 <- glm(Treat~Age+Gender+Race.and.Ethnicity+Birthplace+Education+Residence.area.1+Household.Incom
            family=binomial, data=SD)

#check the length
length(glm2$fitted)
```

```
## [1] 160
```

```
joinpscore2 = cbind(SD, glm2$fitted)

colnames(joinpscore2)[12] = "Propensity_Score"


# With transparency
ggplot(data=joinpscore2, aes(x=Propensity_Score, group=Treat, fill=Treat)) + geom_density(adjust=1, a
```




```
#one-to-one matching for SD
rr2 <- Match(Y=y2, Tr=tr2, X=glm2$fitted)
summary(rr2)
```

```
##
## Estimate...  0.1037
## AI SE......  0.19647
## T-stat.....  0.52784
## p.val......  0.59761
##
## Original number of observations..............  160
## Original number of treated obs...............  81
## Matched number of observations...............  81
## Matched number of observations  (unweighted).  85
```

```
atc <- Match(Y=y2, Tr=tr2, X=glm2$fitted, estimand = "ATC")
summary(atc)
```

```
##
## Estimate...  0.12722
## AI SE......  0.18559
## T-stat.....  0.68545
## p.val......  0.49306
##
## Original number of observations..............  160
## Original number of control obs...............  79
## Matched number of observations...............  79
## Matched number of observations  (unweighted).  79
```

```r
matched <- PD[c(rr2$index.treated,rr2$index.control),]

Matchdefinition  <- function(Y=NULL,Tr,X,Z=X,V=rep(1,length(Y)), estimand="ATT", M=1, BiasAdjust=FALSE


# Assess balance for one covariate
# Do we have balance on discrimination experience after matching?
qqout2 <- qqstats(SD$Race.and.Ethnicity[rr2$index.treated], SD$Race.and.Ethnicity[rr2$index.control])
print(qqout2)
```

```
## $meandiff
## [1] 0.05568627
##
## $mediandiff
## [1] 0.03529412
##
## $maxdiff
## [1] 0.2117647
```

```r
# check balance
postmatchbal2 = MatchBalance(Treat~Age+Gender+Race.and.Ethnicity+Birthplace+Education+Residence.area.
```

```
##
## ***** (V1) Age *****
##                        Before Matching        After Matching
## mean treatment........     2.5062             2.7273
## mean control..........     2.5443             2.1147
## std mean diff.........    -2.5837             38.612
##
## mean raw eQQ diff.....    0.12658             0.6875
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  3
##
## mean eCDF diff........    0.02167            0.11458
## med  eCDF diff........   0.020238              0.125
## max  eCDF diff........   0.060947             0.1875
##
## var ratio (Tr/Co).....     1.1061             2.3569
## T-test p-value........    0.86718          0.0043625
## KS Bootstrap p-value..      0.762              0.038
## KS Naive p-value......    0.78988           0.026277
## KS Statistic..........   0.060947             0.1875
##
##
## ***** (V2) Gender *****
##                        Before Matching        After Matching
## mean treatment........     1.6914             1.6364
## mean control..........     1.5823             1.6883
## std mean diff.........     22.217            -10.173
##
## mean raw eQQ diff.....    0.10127                0.1
## med  raw eQQ diff.....          0                  0
## max  raw eQQ diff.....          1                  1
##
```

```
## mean eCDF diff........    0.03636          0.033333
## med  eCDF diff........   0.012346            0.0125
## max  eCDF diff........   0.096734            0.0875
##
## var ratio (Tr/Co).....    0.97849            1.1997
## T-test p-value........    0.16429           0.48024
## KS Bootstrap p-value..      0.208             0.286
## KS Naive p-value......    0.25168           0.31796
## KS Statistic..........   0.096734            0.0875
##
##
## ***** (V3) Race.and.Ethnicity *****
##                        Before Matching        After Matching
## mean treatment........     7.0123            7.1299
## mean control..........     8.0759            7.5779
## std mean diff.........    -20.745           -9.1391
##
## mean raw eQQ diff.....     1.2152            0.7625
## med  raw eQQ diff.....          1                 0
## max  raw eQQ diff.....          6                 6
##
## mean eCDF diff........   0.056134          0.041346
## med  eCDF diff........   0.065479            0.0375
## max  eCDF diff........    0.11346             0.125
##
## var ratio (Tr/Co).....    0.82427           0.84474
## T-test p-value........     0.2145             0.594
## KS Bootstrap p-value..      0.396             0.313
## KS Naive p-value......     0.3867           0.30394
## KS Statistic..........    0.11346             0.125
##
##
## ***** (V4) Birthplace *****
##                        Before Matching        After Matching
## mean treatment........     1.5062            1.4545
## mean control..........     1.4051            1.4307
## std mean diff.........     20.098            4.7506
##
## mean raw eQQ diff.....    0.10127                 0
## med  raw eQQ diff.....          0                 0
## max  raw eQQ diff.....          1                 0
##
## mean eCDF diff........   0.050555                 0
## med  eCDF diff........   0.050555                 0
## max  eCDF diff........    0.10111                 0
##
## var ratio (Tr/Co).....     1.0369            1.0111
## T-test p-value........     0.2015            0.7833
##
##
## ***** (V5) Education *****
##                        Before Matching        After Matching
## mean treatment........     3.9753             4.026
## mean control..........     3.9367            3.8506
```

```
## std mean diff.........     2.7298            12.652
##
## mean raw eQQ diff.....    0.11392               0.2
## med  raw eQQ diff.....          0                 0
## max  raw eQQ diff.....          1                 1
##
## mean eCDF diff........   0.021279          0.033333
## med  eCDF diff........   0.022582           0.03125
## max  eCDF diff........   0.037975            0.0625
##
## var ratio (Tr/Co).....    0.94698           0.93888
## T-test p-value........    0.86504           0.44251
## KS Bootstrap p-value..      0.937             0.784
## KS Naive p-value......    0.93121           0.79246
## KS Statistic..........   0.037975            0.0625
##
##
## ***** (V6) Residence.area.12 *****
##                        Before Matching        After Matching
## mean treatment........    0.64198            0.7013
## mean control..........    0.72152           0.65584
## std mean diff.........    -16.489            9.8666
##
## mean raw eQQ diff.....   0.088608             0.025
## med  raw eQQ diff.....          0                 0
## max  raw eQQ diff.....          1                 1
##
## mean eCDF diff........   0.039772            0.0125
## med  eCDF diff........   0.039772            0.0125
## max  eCDF diff........   0.079544             0.025
##
## var ratio (Tr/Co).....     1.1435           0.92808
## T-test p-value........    0.28287           0.50529
##
##
## ***** (V7) Residence.area.13 *****
##                        Before Matching        After Matching
## mean treatment........   0.037037          0.051948
## mean control..........   0.025316                 0
## std mean diff.........     6.1678            23.256
##
## mean raw eQQ diff.....          0              0.05
## med  raw eQQ diff.....          0                 0
## max  raw eQQ diff.....          0                 1
##
## mean eCDF diff........  0.0058603             0.025
## med  eCDF diff........  0.0058603             0.025
## max  eCDF diff........   0.011721              0.05
##
## var ratio (Tr/Co).....     1.4449               Inf
## T-test p-value........    0.67176          0.043409
##
##
## ***** (V8) Household.Income.Lev *****
```

```
##                         Before Matching        After Matching
## mean treatment........     3.2099              3.4416
## mean control..........     3.8228              3.1061
## std mean diff.........    -41.627              19.855
##
## mean raw eQQ diff.....    0.68354              0.2375
## med  raw eQQ diff.....          1                   0
## max  raw eQQ diff.....          2                   1
##
## mean eCDF diff........   0.091889            0.033929
## med  eCDF diff........   0.077825              0.0375
## max  eCDF diff........    0.19456                0.05
##
## var ratio (Tr/Co).....    0.71797              1.1585
## T-test p-value........   0.017403             0.24645
## KS Bootstrap p-value..      0.018               0.937
## KS Naive p-value......   0.027495             0.94356
## KS Statistic..........    0.19456                0.05
##
##
## ***** (V9) Discriminatory.Exper *****
##                         Before Matching        After Matching
## mean treatment........     1.5432              1.5195
## mean control..........     1.3924              1.3961
## std mean diff.........     30.087              24.533
##
## mean raw eQQ diff.....    0.13924              0.1375
## med  raw eQQ diff.....          0                   0
## max  raw eQQ diff.....          1                   1
##
## mean eCDF diff........   0.075402             0.06875
## med  eCDF diff........   0.075402             0.06875
## max  eCDF diff........     0.1508              0.1375
##
## var ratio (Tr/Co).....     1.0404              1.0435
## T-test p-value........   0.056445             0.16231
##
##
## Before Matching Minimum p.value: 0.017403
## Variable Name(s): Household.Income.Lev  Number(s): 8
##
## After Matching Minimum p.value: 0.0043625
## Variable Name(s): Age  Number(s): 1
```

```r
# Transform to a tabular format

SMDeepExtract <- function(SD, col = c("mean.Tr",  "mean.Co", "var.Tr", "var.Co", "var.ratio", "p.valu
  cbind(
    t(sapply(SD[[ "BeforeMatching" ]], "[", col)),
    t(sapply(SD[[ "AfterMatching" ]], "[", col))
  )
}

# extract
```

```r
res2 <- SMDeepExtract(postmatchbal2)

rownames(res2) = c("Age", "Gender", "Race and Ethnicity", "Birthplace", "Education", "Suburban", "Rura
colnames(res2) = c("mean.Tr",  "mean.Co", "var.Tr", "var.Co", "var.ratio", "p.value", "sdiff", "A.mea

res2[] <- lapply(res2,round,2)
kableExtra::kable(res2[])
```

	mean.Tr	mean.Co	var.Tr	var.Co	var.ratio	p.value	sdiff	A.mean.Tr	A.mean
Age	2.51	2.54	2.18	1.97	1.11	0.87	-2.58	2.73	2.11
Gender	1.69	1.58	0.24	0.25	0.98	0.16	22.22	1.64	1.69
Race and Ethnicity	7.01	8.08	26.29	31.89	0.82	0.21	-20.74	7.13	7.58
Birthplace	1.51	1.41	0.25	0.24	1.04	0.2	20.1	1.45	1.43
Education	3.98	3.94	2	2.11	0.95	0.87	2.73	4.03	3.85
Suburban	0.64	0.72	0.23	0.2	1.14	0.28	-16.49	0.7	0.66
Rural	0.04	0.03	0.04	0.02	1.44	0.67	6.17	0.05	0
Household Income Level	3.21	3.82	2.17	3.02	0.72	0.02	-41.63	3.44	3.11
Discriminatory Experiences	1.54	1.39	0.25	0.24	1.04	0.06	30.09	1.52	1.4