Analysing the effects of higher education on voters' preferences

Xinyu Zhong

December 21, 2020

github: https://github.com/Amyzhongkinu/STA304-Final-Project

Abstract

In this study, the dataset retrieved from Democracy Fund Voter Study Group was used to investigate how propensity score matching is used to make causal inferences between higher education and voting preferences of voters for 2020 U.S. presidential election. Although there are some bias in propensity score matching, the results of the propensity score model indicated that highly educated voters are more likely to choose Biden as the next president of the United States.

Keywords

Key words: Propensity Score, Causal Inference, Higher education, President election

Introduction

Today, people who receive higher education generally have more opportunities for jobs and higher overall quality of life. They might have realized the benefits of higher education and thus may be more concerned about this field. Policies formulated by the president, such as cutting college costs, play a very important role in the development of higher education. People who care about higher education might vote for the one that will advance a comprehensive higher education agenda.

The purpose of this project is to analyze the effects of higher education on voters' preferences for candidates of U.S. Presidential election 2020. Firstly, data retrieved from Democracy Fund Voter Study Group will be organized by Propensity Score Matching. To be more specific, treated and controlled observations will be matched on the estimated probability of being treated while the treatment is the higher education. Then, a logistic regression model with age, gender, household income, employment, state, race_ethnicity, and higher education assumed as correlated variables will be conducted to see whether this treatment is statistically significant. If the answer is yes, then we will be able to make a causal inference that voters with higher education levels are more likely or less likely to choose Biden as the next U.S. president.

The following sections explain how to use propensity score matching to make causal inferences between higher education and voting preferences. The data, propensity score matching method and propensity score analysis model will be described in the "Methodology" section. The analysis results will be provided in the "Results" section, and the causal inferences and conclusions of the data will be provided in the "Conclusions" section.

Methodology

Data:

The dataset is obtained from Democracy Foundation + UCLA Nationscape survey data (2020) conducted by Democracy Foundation + UCLA Nationscape. The Democracy Foundation + UCLA Nationscape survey data (2020) was collected by interviewing people in almost every county, congressional district, and medium-sized city in the United States before the 2020 U.S.presidential election. There are 6,479 observations in the dataset and 265 variables. Not all variables are needed, so only a few of them are selected for this project. The selected variables and their categories are shown in the following table.

variable	category
Age	discrete
Registration	categorical
Vote_intention	categorical
Vote_2020	categorical
Gender	categorical
Race_ethnicity	categorical
Education	categorical
Employment	categorical
State	categorical
household_income	categorical

Table.1

```
##                                                 Stratified by edu_high
##                                                  0
##   n                                                 3132
##   registration (%)
##      Registered                                     2514 (80.3)
##      Not registered                                  522 (16.7)
##      Don't know                                       96 ( 3.1)
##   vote_intention (%)
##      Yes, I will vote                               2430 (77.6)
##      No, I will not vote but I am eligible           333 (10.6)
##      No, I am not eligible to vote                     0 ( 0.0)
##      Not sure                                        369 (11.8)
##   vote_2020 (%)
##      Donald Trump                                   1199 (38.3)
##      Joe Biden                                      1247 (39.8)
##      Someone else                                    140 ( 4.5)
##      I would not vote                                197 ( 6.3)
##      I am not sure/don't know                        349 (11.1)
##   age (mean (SD))                                  43.85 (17.33)
##   gender = Male (%)                                 1411 (45.1)
##   race_ethnicity (%)
##      White                                          2256 (72.0)
##      Black, or African American                      451 (14.4)
##      American Indian or Alaska Native                 47 ( 1.5)
##      Asian (Asian Indian)                             33 ( 1.1)
##      Asian (Chinese)                                  20 ( 0.6)
##      Asian (Filipino)                                 11 ( 0.4)
##      Asian (Japanese)                                  4 ( 0.1)
##      Asian (Korean)                                    5 ( 0.2)
##      Asian (Vietnamese)                                3 ( 0.1)
##      Asian (Other)                                    15 ( 0.5)
##      Pacific Islander (Native Hawaiian)                7 ( 0.2)
##      Pacific Islander (Guamanian)                      0 ( 0.0)
##      Pacific Islander (Samoan)                         2 ( 0.1)
##      Pacific Islander (Other)                          6 ( 0.2)
##      Some other race                                 272 ( 8.7)
##   education (%)
##      3rd Grade or less                                11 ( 0.4)
##      Middle School - Grades 4 - 8                     19 ( 0.6)
```

```
##      Completed some high school                      552 (17.6)
##      High school graduate                            904 (28.9)
##      Other post high school vocational training      297 ( 9.5)
##      Completed some college, but no degree          1214 (38.8)
##      Associate Degree                                  0 ( 0.0)
##      College Degree (such as B.A., B.S.)               0 ( 0.0)
##      Completed some graduate, but no degree            0 ( 0.0)
##      Masters degree                                    0 ( 0.0)
##      Doctorate degree                                135 ( 4.3)
##   employment (%)
##      Full-time employed                              925 (29.5)
##      Homemaker                                       243 ( 7.8)
##      Retired                                         549 (17.5)
##      Unemployed or temporarily on layoff             383 (12.2)
##      Part-time employed                              339 (10.8)
##      Permanently disabled                            224 ( 7.2)
##      Student                                         179 ( 5.7)
##      Self-employed                                   232 ( 7.4)
##      Other:                                           58 ( 1.9)
##   state (%)
##      AK                                                7 ( 0.2)
##      AL                                               51 ( 1.6)
##      AR                                               27 ( 0.9)
##      AZ                                               81 ( 2.6)
##      CA                                              341 (10.9)
##      CO                                               50 ( 1.6)
##      CT                                               25 ( 0.8)
##      DC                                                9 ( 0.3)
##      DE                                               20 ( 0.6)
##      FL                                              228 ( 7.3)
##      GA                                               92 ( 2.9)
##      HI                                               11 ( 0.4)
##      IA                                               32 ( 1.0)
##      ID                                               17 ( 0.5)
##      IL                                              135 ( 4.3)
##      IN                                               71 ( 2.3)
##      KS                                               22 ( 0.7)
##      KY                                               52 ( 1.7)
##      LA                                               44 ( 1.4)
##      MA                                               47 ( 1.5)
##      MD                                               42 ( 1.3)
##      ME                                               10 ( 0.3)
##      MI                                               84 ( 2.7)
##      MN                                               23 ( 0.7)
##      MO                                               76 ( 2.4)
##      MS                                               24 ( 0.8)
##      MT                                                7 ( 0.2)
##      NC                                              117 ( 3.7)
##      ND                                                3 ( 0.1)
##      NE                                                6 ( 0.2)
##      NH                                               10 ( 0.3)
##      NJ                                               81 ( 2.6)
##      NM                                               16 ( 0.5)
##      NV                                               33 ( 1.1)
```

```
##      NY                                                 196 ( 6.3)
##      OH                                                 153 ( 4.9)
##      OK                                                  39 ( 1.2)
##      OR                                                  53 ( 1.7)
##      PA                                                 126 ( 4.0)
##      RI                                                   9 ( 0.3)
##      SC                                                  58 ( 1.9)
##      SD                                                   6 ( 0.2)
##      TN                                                  70 ( 2.2)
##      TX                                                 257 ( 8.2)
##      UT                                                  28 ( 0.9)
##      VA                                                  79 ( 2.5)
##      VT                                                   9 ( 0.3)
##      WA                                                  65 ( 2.1)
##      WI                                                  57 ( 1.8)
##      WV                                                  29 ( 0.9)
##      WY                                                   4 ( 0.1)
##   household_income (%)
##      Less than $14,999                                  738 (23.6)
##      $15,000 to $19,999                                 230 ( 7.3)
##      $20,000 to $24,999                                 232 ( 7.4)
##      $25,000 to $29,999                                 215 ( 6.9)
##      $30,000 to $34,999                                 190 ( 6.1)
##      $35,000 to $39,999                                 185 ( 5.9)
##      $40,000 to $44,999                                 140 ( 4.5)
##      $45,000 to $49,999                                 149 ( 4.8)
##      $50,000 to $54,999                                 157 ( 5.0)
##      $55,000 to $59,999                                  80 ( 2.6)
##      $60,000 to $64,999                                  80 ( 2.6)
##      $65,000 to $69,999                                  61 ( 1.9)
##      $70,000 to $74,999                                  87 ( 2.8)
##      $75,000 to $79,999                                  82 ( 2.6)
##      $80,000 to $84,999                                  39 ( 1.2)
##      $85,000 to $89,999                                  25 ( 0.8)
##      $90,000 to $94,999                                  32 ( 1.0)
##      $95,000 to $99,999                                  47 ( 1.5)
##      $100,000 to $124,999                               129 ( 4.1)
##      $125,000 to $149,999                                77 ( 2.5)
##      $150,000 to $174,999                                37 ( 1.2)
##      $175,000 to $199,999                                26 ( 0.8)
##      $200,000 to $249,999                                54 ( 1.7)
##      $250,000 and above                                  40 ( 1.3)
##   edu_high = 1 (%)                                        0 ( 0.0)
##   .fitted (mean (SD))                                  0.36 (0.20)
##   treated (mean (SD))                                  0.00 (0.00)
##   match.ind (mean (SD))                             3228.01 (1887.97)
##   cnts (mean (SD))                                     0.86 (0.35)
##   pairs (mean (SD))                                 1347.50 (777.84)
##                                                    Stratified by edu_high
##                                                     1                  p      test
##   n                                                    2694
##   registration (%)                                                  <0.001
##      Registered                                        2529 ( 93.9)
##      Not registered                                     142 (  5.3)
```

```
##      Don't know                                     23 (  0.9)
##    vote_intention (%)                                              NaN
##      Yes, I will vote                             2477 ( 91.9)
##      No, I will not vote but I am eligible         110 (  4.1)
##      No, I am not eligible to vote                   0 (  0.0)
##      Not sure                                      107 (  4.0)
##    vote_2020 (%)                                                 <0.001
##      Donald Trump                                 1100 ( 40.8)
##      Joe Biden                                    1250 ( 46.4)
##      Someone else                                   89 (  3.3)
##      I would not vote                               57 (  2.1)
##      I am not sure/don't know                      198 (  7.3)
##    age (mean (SD))                               47.73 (15.40)   <0.001
##    gender = Male (%)                              1492 ( 55.4)   <0.001
##    race_ethnicity (%)                                            <0.001
##      White                                        2140 ( 79.4)
##      Black, or African American                    238 (  8.8)
##      American Indian or Alaska Native               28 (  1.0)
##      Asian (Asian Indian)                           44 (  1.6)
##      Asian (Chinese)                                45 (  1.7)
##      Asian (Filipino)                               25 (  0.9)
##      Asian (Japanese)                               12 (  0.4)
##      Asian (Korean)                                  8 (  0.3)
##      Asian (Vietnamese)                              9 (  0.3)
##      Asian (Other)                                  15 (  0.6)
##      Pacific Islander (Native Hawaiian)              3 (  0.1)
##      Pacific Islander (Guamanian)                    1 (  0.0)
##      Pacific Islander (Samoan)                       0 (  0.0)
##      Pacific Islander (Other)                        1 (  0.0)
##      Some other race                               125 (  4.6)
##    education (%)                                                 <0.001
##      3rd Grade or less                               0 (  0.0)
##      Middle School - Grades 4 - 8                    0 (  0.0)
##      Completed some high school                      0 (  0.0)
##      High school graduate                            0 (  0.0)
##      Other post high school vocational training      0 (  0.0)
##      Completed some college, but no degree           0 (  0.0)
##      Associate Degree                              517 ( 19.2)
##      College Degree (such as B.A., B.S.)          1363 ( 50.6)
##      Completed some graduate, but no degree        207 (  7.7)
##      Masters degree                                607 ( 22.5)
##      Doctorate degree                                0 (  0.0)
##    employment (%)                                                <0.001
##      Full-time employed                           1478 ( 54.9)
##      Homemaker                                      79 (  2.9)
##      Retired                                       475 ( 17.6)
##      Unemployed or temporarily on layoff           181 (  6.7)
##      Part-time employed                            189 (  7.0)
##      Permanently disabled                           70 (  2.6)
##      Student                                        55 (  2.0)
##      Self-employed                                 149 (  5.5)
##      Other:                                         18 (  0.7)
##    state (%)                                                     <0.001
##      AK                                              1 (  0.0)
```

```
##       AL                                                   33 (  1.2)
##       AR                                                   15 (  0.6)
##       AZ                                                   68 (  2.5)
##       CA                                                  309 ( 11.5)
##       CO                                                   34 (  1.3)
##       CT                                                   42 (  1.6)
##       DC                                                   11 (  0.4)
##       DE                                                    4 (  0.1)
##       FL                                                  211 (  7.8)
##       GA                                                   75 (  2.8)
##       HI                                                   15 (  0.6)
##       IA                                                   19 (  0.7)
##       ID                                                   10 (  0.4)
##       IL                                                  128 (  4.8)
##       IN                                                   42 (  1.6)
##       KS                                                   25 (  0.9)
##       KY                                                   27 (  1.0)
##       LA                                                   31 (  1.2)
##       MA                                                   57 (  2.1)
##       MD                                                   48 (  1.8)
##       ME                                                   10 (  0.4)
##       MI                                                   80 (  3.0)
##       MN                                                   39 (  1.4)
##       MO                                                   37 (  1.4)
##       MS                                                   20 (  0.7)
##       MT                                                    9 (  0.3)
##       NC                                                   72 (  2.7)
##       ND                                                    3 (  0.1)
##       NE                                                   13 (  0.5)
##       NH                                                    8 (  0.3)
##       NJ                                                  107 (  4.0)
##       NM                                                    9 (  0.3)
##       NV                                                   33 (  1.2)
##       NY                                                  277 ( 10.3)
##       OH                                                  110 (  4.1)
##       OK                                                   21 (  0.8)
##       OR                                                   32 (  1.2)
##       PA                                                  118 (  4.4)
##       RI                                                    2 (  0.1)
##       SC                                                   39 (  1.4)
##       SD                                                    9 (  0.3)
##       TN                                                   34 (  1.3)
##       TX                                                  175 (  6.5)
##       UT                                                   16 (  0.6)
##       VA                                                  103 (  3.8)
##       VT                                                    5 (  0.2)
##       WA                                                   51 (  1.9)
##       WI                                                   48 (  1.8)
##       WV                                                    8 (  0.3)
##       WY                                                    1 (  0.0)
##    household_income (%)                                                  <0.001
##       Less than $14,999                                   141 (  5.2)
##       $15,000 to $19,999                                   78 (  2.9)
##       $20,000 to $24,999                                   79 (  2.9)
```

```
##      $25,000 to $29,999                              105 (  3.9)
##      $30,000 to $34,999                              130 (  4.8)
##      $35,000 to $39,999                               98 (  3.6)
##      $40,000 to $44,999                               95 (  3.5)
##      $45,000 to $49,999                              110 (  4.1)
##      $50,000 to $54,999                              167 (  6.2)
##      $55,000 to $59,999                               84 (  3.1)
##      $60,000 to $64,999                               87 (  3.2)
##      $65,000 to $69,999                               65 (  2.4)
##      $70,000 to $74,999                               97 (  3.6)
##      $75,000 to $79,999                               98 (  3.6)
##      $80,000 to $84,999                               71 (  2.6)
##      $85,000 to $89,999                               66 (  2.4)
##      $90,000 to $94,999                               40 (  1.5)
##      $95,000 to $99,999                              102 (  3.8)
##      $100,000 to $124,999                            323 ( 12.0)
##      $125,000 to $149,999                            237 (  8.8)
##      $150,000 to $174,999                            140 (  5.2)
##      $175,000 to $199,999                             89 (  3.3)
##      $200,000 to $249,999                             99 (  3.7)
##      $250,000 and above                               93 (  3.5)
##   edu_high = 1 (%)                                  2694 (100.0)   <0.001
##   .fitted (mean (SD))                               0.58 (0.20)    <0.001
##   treated (mean (SD))                               1.00 (0.00)    <0.001
##   match.ind (mean (SD))                          2503.91 (1390.08) <0.001
##   cnts (mean (SD))                                  1.00 (0.00)    <0.001
##   pairs (mean (SD))                              1347.50 (777.84)   1.000
```

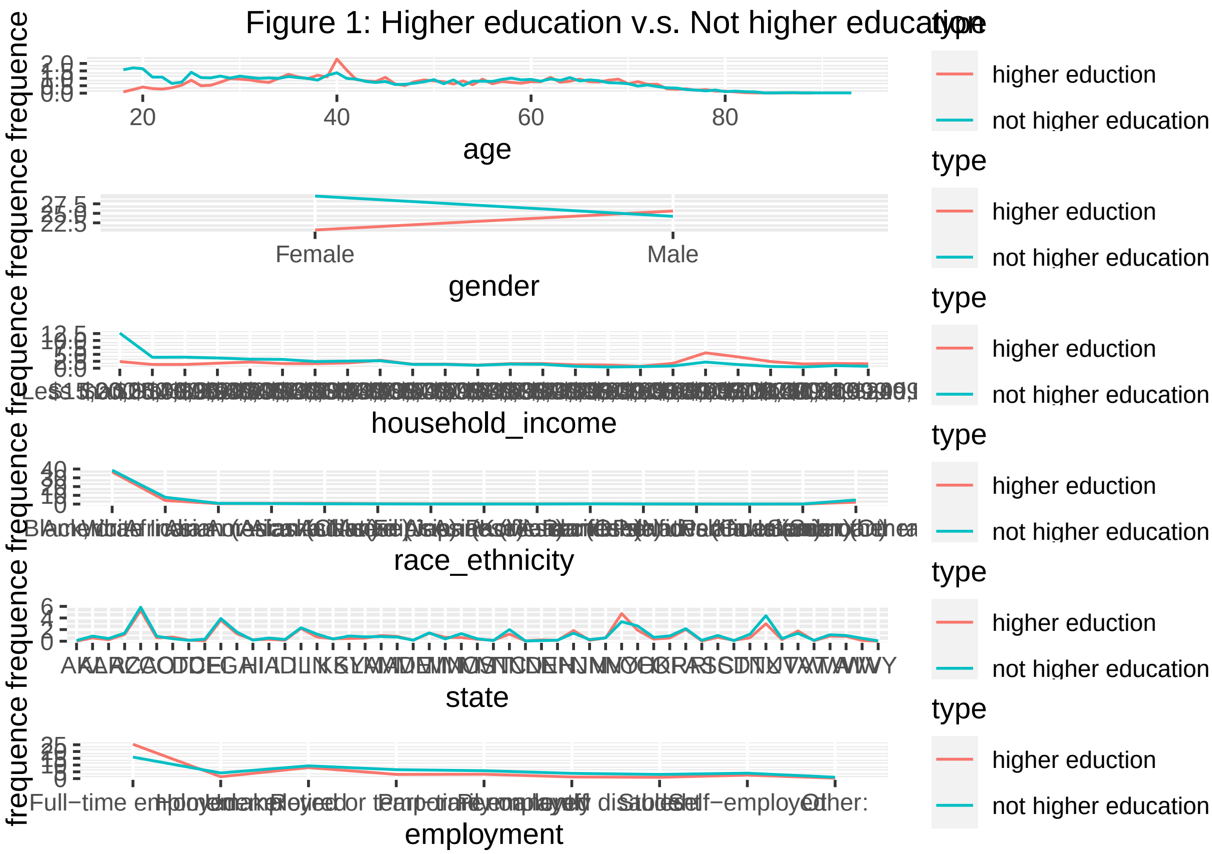

Table.1 reports the baseline characteristics of the data separated by treatment groups(higher education). In the table, the discrete variable, age, is summarized by means and standard deviations while categorical variables, such as gender and state, are summarized by frequency, The p-value printed along with the table is generated using hypothesis test: chisq.test() for categorical variables and oneway.test() for continuous variables, to compare the characteristics between two groups. From Table.1, it can be seen that the p-value of all variables in the table, except for Vote_intention, are smaller than 0.001, which is a threshold for statistical significance. Thus, these variables are good fits for explanatory variables of the regression model. From Figure.1, the differences in characteristics between the treatment group and the control group(higher education and not higher education) can also be seen more intuitively. It is shown that there are not so much differences between the two groups.

Propensity Score matching

Propensity score matching is a statistical technique that matches treated and controlled group on the propensity score, which is the estimated probability of being treated. In this case, higher education is the treatment and the propensity score was calculated in the following logistic regression model:

$$log(\frac{\widehat{p}}{1-\widehat{p}}) = \hat{\beta}_0 + \hat{\beta}_1 x_{age} + \hat{\beta}_2 x_{gender} + \hat{\beta}_3 x_{householdincome} + \hat{\beta}_4 x_{raceethnicity} + \hat{\beta}_5 x_{state}$$

where p represents the probability to have higher education. x_i(i=1~5) correspond to the values of age and the levels of gender, household income, state, and Race ethnicity. Age, gender, household income, state,and race_ethnicity are the predictor variables. β_i(i=2~5) are estimate coefficients, β_0 is an intercept. Then, for every person who was treated with higher education, we find the untreated person who was considered as the closest match based on the propensity score. After that, the dataset was reduced to just those that are matched. At last, the effect of being treated on Vote preference in the 2020 presidential election could be examined in the 'usual' way.

Model:

The logistic regression model that we are interested in estimating is:

$$log(\frac{\widehat{p}}{1-\widehat{p}}) = \hat{\beta}_0 + \hat{\beta}_1 x_{age} + \hat{\beta}_2 x_{gender} + \hat{\beta}_3 x_{householdincome} + \hat{\beta}_4 x_{raceethnicity} + \hat{\beta}_5 x_{state} + \hat{\beta}_6 x_{highereducation}$$

where p represents the probability to vote for Biden during 2020 presidential election. x_i(i=1~6) correspond to the values of age and the levels of gender, household income, Race ethnicity, state, and higher education. Age, gender, household income, state, race_ethnicity, and higher education are the predictor variables. β_i(i=2~6) are estimate coefficients, β_0 is an intercept.

Results

```
##                                                          beta
## (Intercept)                                       0.750753399 6.00e-01
## age                                              -0.013834024 2.96e-08
## genderMale                                       -0.439542068 2.81e-12
## household_income$15,000 to $19,999                0.067646076 6.93e-01
## household_income$20,000 to $24,999               -0.284649582 8.90e-02
## household_income$25,000 to $29,999               -0.156786394 3.39e-01
## household_income$30,000 to $34,999               -0.135065099 4.07e-01
## household_income$35,000 to $39,999               -0.143335374 4.04e-01
## household_income$40,000 to $44,999               -0.154051355 3.87e-01
## household_income$45,000 to $49,999               -0.255180409 1.38e-01
## household_income$50,000 to $54,999               -0.308245475 5.89e-02
## household_income$55,000 to $59,999               -0.337912990 9.40e-02
## household_income$60,000 to $64,999               -0.146131404 4.69e-01
## household_income$65,000 to $69,999               -0.222766901 3.11e-01
## household_income$70,000 to $74,999               -0.245152000 2.07e-01
## household_income$75,000 to $79,999               -0.420344224 3.09e-02
## household_income$80,000 to $84,999               -0.028225253 9.06e-01
## household_income$85,000 to $89,999               -0.201400962 4.26e-01
## household_income$90,000 to $94,999               -0.437018244 1.07e-01
## household_income$95,000 to $99,999               -0.122212250 5.63e-01
## household_income$100,000 to $124,999             -0.537573359 3.78e-04
## household_income$125,000 to $149,999             -0.496317602 2.73e-03
## household_income$150,000 to $174,999             -0.383102296 5.26e-02
## household_income$175,000 to $199,999             -0.912887111 9.00e-05
## household_income$200,000 to $249,999             -1.184968294 2.48e-08
## household_income$250,000 and above               -0.792712117 3.35e-04
## employmentHomemaker                              -0.170357662 2.69e-01
## employmentRetired                                 0.277658125 7.87e-03
## employmentUnemployed or temporarily on layoff     0.257697890 2.70e-02
## employmentPart-time employed                      0.131213453 2.61e-01
## employmentPermanently disabled                    0.361440656 2.30e-02
## employmentStudent                                 1.051764071 5.59e-06
## employmentSelf-employed                           0.300091845 2.00e-02
## employmentOther:                                  0.350044846 2.29e-01
## stateAL                                          -0.126045211 9.31e-01
## stateAR                                          -0.218017838 8.82e-01
## stateAZ                                           0.154182777 9.15e-01
## stateCA                                           0.561416839 6.95e-01
## stateCO                                           0.248781495 8.64e-01
## stateCT                                           1.162094929 4.26e-01
```

```
## stateDC                                             0.348687150 8.19e-01
## stateDE                                             1.048140378 5.12e-01
## stateFL                                             0.130459848 9.27e-01
## stateGA                                            -0.248706561 8.63e-01
## stateHI                                             0.393106780 7.95e-01
## stateIA                                             0.514714307 7.25e-01
## stateID                                             0.001504026 9.99e-01
## stateIL                                             0.441285191 7.58e-01
## stateIN                                             0.265856972 8.54e-01
## stateKS                                            -0.278518359 8.49e-01
## stateKY                                             0.275204584 8.50e-01
## stateLA                                            -0.113969179 9.37e-01
## stateMA                                             1.150079653 4.27e-01
## stateMD                                             0.511423889 7.24e-01
## stateME                                             0.588425614 6.95e-01
## stateMI                                             0.454054996 7.52e-01
## stateMN                                             0.148842957 9.18e-01
## stateMO                                             0.251506288 8.62e-01
## stateMS                                            -0.265418747 8.57e-01
## stateMT                                             0.256774199 8.66e-01
## stateNC                                             0.088751349 9.51e-01
## stateND                                            -0.002088997 9.99e-01
## stateNE                                             0.546767370 7.17e-01
## stateNH                                             0.566117381 7.10e-01
## stateNJ                                             0.377708206 7.93e-01
## stateNM                                             0.858689293 5.71e-01
## stateNV                                             0.023643227 9.87e-01
## stateNY                                             0.438998273 7.59e-01
## stateOH                                             0.365128289 7.99e-01
## stateOK                                             0.127876624 9.30e-01
## stateOR                                             0.604110206 6.77e-01
## statePA                                             0.227508774 8.74e-01
## stateRI                                             1.894197129 2.92e-01
## stateSC                                            -0.368269580 7.99e-01
## stateSD                                            -0.027470259 9.86e-01
## stateTN                                            -0.116046770 9.36e-01
## stateTX                                             0.030849722 9.83e-01
## stateUT                                             0.642235134 6.62e-01
## stateVA                                             0.509463517 7.23e-01
## stateVT                                             2.439458380 1.68e-01
## stateWA                                             0.576352642 6.89e-01
## stateWI                                             0.677531238 6.39e-01
## stateWV                                            -0.344066848 8.17e-01
## stateWY                                             1.525386833 4.08e-01
## race_ethnicityBlack, or African American            2.071562782 3.58e-43
## race_ethnicityAmerican Indian or Alaska Native     -0.079677953 7.63e-01
## race_ethnicityAsian (Asian Indian)                  0.419739966 1.13e-01
## race_ethnicityAsian (Chinese)                       1.382496596 1.06e-04
## race_ethnicityAsian (Filipino)                     -0.161147859 6.47e-01
## race_ethnicityAsian (Japanese)                      0.798496030 1.98e-01
## race_ethnicityAsian (Korean)                        0.429634124 4.92e-01
## race_ethnicityAsian (Vietnamese)                    0.770852740 2.25e-01
## race_ethnicityAsian (Other)                         0.756980817 1.15e-01
## race_ethnicityPacific Islander (Native Hawaiian)    0.601702855 5.12e-01
```

```
## race_ethnicityPacific Islander (Guamanian)       -12.556274519 9.69e-01
## race_ethnicityPacific Islander (Other)            12.147317660 9.40e-01
## race_ethnicitySome other race                      0.763548833 5.98e-08
## edu_high1                                          0.267783655 4.24e-05
```

The Propensity score regression is based on age, gender, household income, employment, state, race_ethinicity, and higher education and it looks for the proportion of voters who are more willing to vote for Biden.

The results of the model show that age, gender, household income, employment, race_ethnicity, and higher education are significant predictors of voters' preferences because their p-value is less than 0.05. Moreover, age, gender, and higher education are especially significant in influencing voters' preferences because their p-value is less than 0.001.

The propensity score adjusted data can be used to assess the causality of treatment (higher education). It can be seen from the table that the coefficient of edu_high is 0.268. This means that for the person with higher education, the log odds of the percentage of voters who are more willing to vote for Biden will increase by 0.268.

Discussion

Summary

The Democracy Foundation + UCLA Nationscape survey data (2020) was used to analyze the effects of higher education on voters' preferences for candidates of the U.S. Presidential election 2020. Firstly, a few variables were selected from the original dataset. Then, the dataset was adjusted by propensity score matching while the treatment is higher education. Finally, a propensity score regression was conducted to draw causal inferences about the treatment.

Conclusions

The results of the propensity score regression indicated that the treatment of higher education is statistically significant. It also assessed the causality of treatment, that is, highly educated voters are more likely to choose Biden as the next president of the United States. The results show that for those with higher education, the log odds of the proportion of voters who are more willing to vote for Biden will increase by 0.268.

People with higher education may have realized the benefits brought by higher education. They may vote for the one that will formulate the policies that benefit education. Therefore, inferences about the causal relationship between higher education and voting preferences can provide references to government agencies in advancing a comprehensive higher education agenda. Also, today, higher education is an important contribution to the country's competitiveness in the global market and is vital to economic strength, social well-being, and world leadership. In other words, the development of high education is closely related to the future of society. Hence, it is necessary for the government to pay more attention to higher education.

Weakness & Next Steps:

There are some flaws in propensity score matching. One of them is modeling. The basic assumption for using propensity scores is that all confounders of treatment options and outcomes have been measured and included in the propensity model. The results tend to be specific to the model that is used. If the model does not include variables strongly related to both outcome and assignment, it may increase bias. From the results of the final model, it is seen that the variable of state is not statistically significant. Thus, this possibly increases bias. Also, there may be bias due to confounding factors that cannot be measured. However, if propensity scores are constructed based only on treatment covariates that are statistically significantly different between the pre-treatment and comparison groups, they will not be able to consider the relationships between the covariates and the results.

Therefore, for the next step, our goal should be creating as rich a propensity score model as possible while including a strong covariate.

References

1. Lanza, S., Moore, J., & Butera, N. (2013, December). Drawing causal inferences using propensity scores: A practical guide for community psychologists. Retrieved December 20, 2020, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4098642/

2. The Importance of Higher Education in the 21st Century. (2020, October 15). Retrieved December 06, 2020, from https://www.vistacollege.edu/blog/resources/higher-education-in-the-21st-century/

3. Vote Biden: Inside Higher Ed. (n.d.). Retrieved December 06, 2020, from https://www.insidehighered.com/blogs/leadership-higher-education/vote-biden

4. Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape, October 10-17, 2019 (version 20200814). Retrieved from https://www.voterstudygroup.org/downloads?key=9e6f71ed-8c3b-4238-be7b-9d332bf90590

5. Republic, T., Reason, & Insider, B. (2020, December 15). New: Second Nationscape Data Set Release. Retrieved December 20, 2020, from https://www.voterstudygroup.org/publication/nationscape-data-set

6. Yoshida, K. (2020, July 25). Retrieved December 20, 2020, from https://cran.r-project.org/web/packages/tableone/vignettes/introduction.html

7. Difference in differences. (2020, November 05). Retrieved December 20, 2020, from https://www.tellingstorieswithdata.com/06-03-matching_and_differences.html

8. Teague, L.W. (2015). Higher Education Plays Critical Role in Society: More Women Leaders Can Make a Difference.

9. Hadley Wickham and Evan Miller (2020). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. http://haven.tidyverse.org, https://github.com/tidyverse/haven, https://github.com/WizardMac/ReadStat.

10. Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

11. David Robinson, Alex Hayes and Simon Couch (2020). broom: Convert Statistical Objects into Tidy Tibbles. https://broom.tidymodels.org/, https://github.com/tidymodels/broom.

12. Andrew Gelman and Yu-Sung Su (2020). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.11-2. https://CRAN.R-project.org/package=arm

13. Joseph Larmarange (2020). labelled: Manipulating Labelled Data. R package version 2.7.0. http://larmarange.github.io/labelled/

14. Alboukadel Kassambara (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. https://rpkgs.datanovia.com/ggpubr/

15. Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

16. Kazuki Yoshida and Alexander Bartel (2020). tableone: Create 'Table 1' to Describe Baseline Characteristics with or without Propensity Score Weights. R package version 0.12.0. https://github.com/kaz-yos/tableone