Biden-might-win-the-2020-election

Haoming Hu, Ziyu Hao, QianWen Shen, SongYuan Luo

11/2/2020

#Abstract The 2020 American presidential election holds great uncertainty due to the pandemic. For the purpose of forecasting the results of the 2020 american presidential election, we applied Multilevel regression with post-stratification to the 2020 american presidential election survey and census data and found that Trump has a higher probability of winning among the two potential candidates, Trump and Biden.

keyword:Forecasting; Trump; Biden; Multilevel regression with post-stratification

#Intro For the purpose of forecasting the winner of the 2020 American presidential election, we obtained the survey data and post stratification data from Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape and the work of American Community Surveys ACS respectively. There are 6479 observations in the survey data containing 265 variables and 613777 observations in the post stratification data.

The variables of interest are 'gender', 'race', 'education', 'state', 'age' and 'employment'. Among all the observations, we only intend to include eligible and valid data. Therefore , we filtered all the non-responses select groups of interest, with 5300 and 11307 observations left in the survey and post stratification data respectively.

We used the brm (Bayesian generalized non-linear multilevel model) to help predict the winner of the 2020 american presidential election. A variable 'support_Trump' was coined to show if someone has voted for trump, with 1 representing people having voted for Trump and 0 representing voted for Biden, another potential candidate in the election. Several plots have been generated and labelled appropriately.

The main purpose of this report is to predict the winner of the 2020 American presidential election. By applying the multilevel regression with post stratification model to the survey and census data, we found out that Trump has a higher probability of winning. In order to see how the changes of state affect the intercept, we applied one layer to the variable 'state' and created MRP estimate and data comparison plots.

#data

```
## # A tibble: 6,479 x 265
##    response_id start_date          right_track economy_better interest
##    <chr>       <dttm>              <fct>       <fct>          <fct>
##  1 05000005    2020-06-25 07:21:56 Generally ~ About the same Some of~
##  2 05000006    2020-06-25 07:21:25 Off on the~ Worse          Some of~
##  3 05000007    2020-06-25 07:21:32 Off on the~ Worse          Some of~
##  4 05000009    2020-06-25 07:21:14 Off on the~ Worse          Most of~
##  5 05000010    2020-06-25 07:23:26 Generally ~ Better         Most of~
##  6 05000011    2020-06-25 07:22:28 Off on the~ Worse          Only no~
##  7 05000012    2020-06-25 07:22:21 Off on the~ About the same Some of~
##  8 05000013    2020-06-25 07:21:47 Off on the~ Worse          Some of~
##  9 05000014    2020-06-25 07:22:57 Generally ~ About the same Some of~
## 10 05000015    2020-06-25 07:23:04 Off on the~ Worse          Most of~
## # ... with 6,469 more rows, and 260 more variables: registration <fct>,
```

```
## #   news_sources_facebook <fct>, news_sources_cnn <fct>,
## #   news_sources_msnbc <fct>, news_sources_fox <fct>,
## #   news_sources_network <fct>, news_sources_localtv <fct>,
## #   news_sources_telemundo <fct>, news_sources_npr <fct>,
## #   news_sources_amtalk <fct>, news_sources_new_york_times <fct>,
## #   news_sources_local_newspaper <fct>, news_sources_other <fct>,
## #   news_sources_other_TEXT <chr>, pres_approval <fct>, vote_2016 <fct>,
## #   vote_2016_other_text <chr>, consider_trump <fct>, not_trump <fct>,
## #   vote_intention <fct>, vote_2020 <fct>, vote_2020_other_text <chr>,
## #   vote_2020_lean <fct>, primary_party <fct>, group_favorability_whites <fct>,
## #   group_favorability_blacks <fct>, group_favorability_latinos <fct>,
## #   group_favorability_asians <fct>, group_favorability_evangelicals <fct>,
## #   group_favorability_socialists <fct>, group_favorability_muslims <fct>,
## #   group_favorability_labor_unions <fct>, group_favorability_the_police <fct>,
## #   group_favorability_undocumented <fct>, group_favorability_lgbt <fct>,
## #   group_favorability_republicans <fct>, group_favorability_democrats <fct>,
## #   group_favorability_white_men <fct>, group_favorability_jews <fct>,
## #   cand_favorability_trump <fct>, cand_favorability_obama <fct>,
## #   cand_favorability_biden <fct>, rep_prim_vote <fct>,
## #   rep_prim_vote_TEXT <chr>, dem_prim_vote <fct>, house_intent <fct>,
## #   senate_intent <fct>, governor_intent <fct>, trump_biden <fct>,
## #   trump_sanders <fct>, pence_biden <fct>, pence_sanders <fct>,
## #   primary_sen_barrasso <fct>, primary_sen_blackburn <fct>,
## #   primary_sen_blunt <fct>, primary_sen_cassidy <fct>,
## #   primary_sen_collins <fct>, primary_sen_cornyn <fct>,
## #   primary_sen_cotton <fct>, primary_sen_daines <fct>,
## #   primary_sen_ernst <fct>, primary_sen_gardner <fct>,
## #   primary_sen_graham <fct>, primary_sen_hoeven <fct>,
## #   primary_sen_hydesmith <fct>, primary_sen_inhofe <fct>,
## #   primary_sen_lee <fct>, primary_sen_mcconnell <fct>,
## #   primary_sen_mcsally <fct>, primary_sen_moorecapito <fct>,
## #   primary_sen_moran <fct>, primary_sen_perdue <fct>,
## #   primary_sen_portman <fct>, primary_sen_risch <fct>,
## #   primary_sen_rounds <fct>, primary_sen_rubio <fct>, primary_sen_sasse <fct>,
## #   primary_sen_shelby <fct>, primary_sen_sullivan <fct>,
## #   primary_sen_tillis <fct>, primary_sen_toomey <fct>,
## #   primary_sen_young <fct>, primary_sen_boozman <fct>,
## #   primary_sen_braun <fct>, primary_sen_cramer <fct>, primary_sen_crapo <fct>,
## #   primary_sen_cruz <fct>, primary_sen_fischer <fct>,
## #   primary_sen_grassley <fct>, primary_sen_hawley <fct>,
## #   primary_sen_lankford <fct>, primary_sen_murkowski <fct>,
## #   primary_sen_neelykennedy <fct>, primary_sen_paul <fct>,
## #   primary_sen_romney <fct>, primary_sen_scott_rick <fct>,
## #   primary_sen_scott_tim <fct>, primary_sen_thune <fct>,
## #   primary_sen_wicker <fct>, cand_truth_donald_trump <fct>, ...
```

```
##     gender          race              education           state
##  Female:2586   Length:5200        Length:5200        Length:5200
##  Male  :2614   Class :character   Class :character   Class :character
##                Mode  :character   Mode  :character   Mode  :character
##
##
##
##      age             employment            trump
```

```
##  Length:5200        Length:5200        Min.   :0.0000
##  Class :character   Class :character   1st Qu.:0.0000
##  Mode  :character   Mode  :character   Median :0.0000
##                                        Mean   :0.4771
##                                        3rd Qu.:1.0000
##                                        Max.   :1.0000
```

The survey data was provided generously by Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape, October 10-17, 2019 (version 20200814). Retrieved from [https://www.voterstudygroup.org/downloads?key=1c2ff38b-5ade-4f75-856b-616f859becff].

Data collection and methodology:

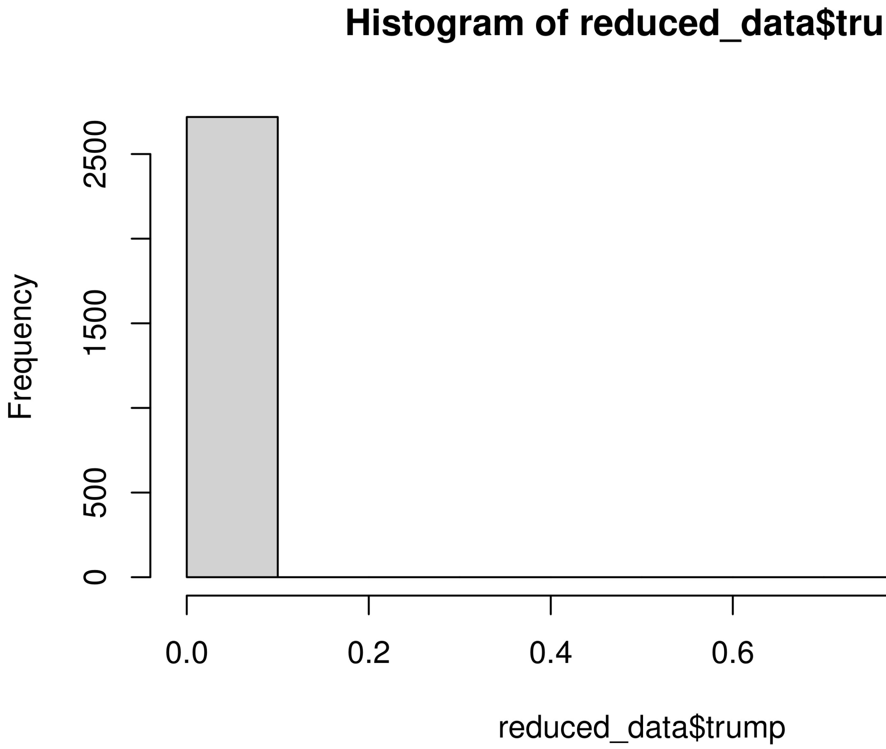
Nowadays, true random-digit-dial surveys now typically have such low response rates that theorems based on random sampling do little to ensure the representativeness of the set of people who actually respond to the survey (Kennedy and Deane 2019). Hence, Nationscape has a convenience sample selected on a set of demographic criteria. Purposive sampling method was used, selection respondents based upon their characteristics to obtain a sample that is constructed to be representative of the population in terms of a specified set of characteristics. The survey conducted 500,000 interviews of Americans from July 2019 through December 2020, covering the 2020 campaign and election. The survey has been in the field since July 10, 2019, and it includes interviews with roughly 6,250 people per week. All respondents take the survey online and must complete an attention check before taking the survey. The survey is conducted in English.

There are 6479 observations in the original dataset containing 265 variables, and for the purpose of this study, we kept 5200 observations and 7 variables. Our response variable is 'Supports Trump':

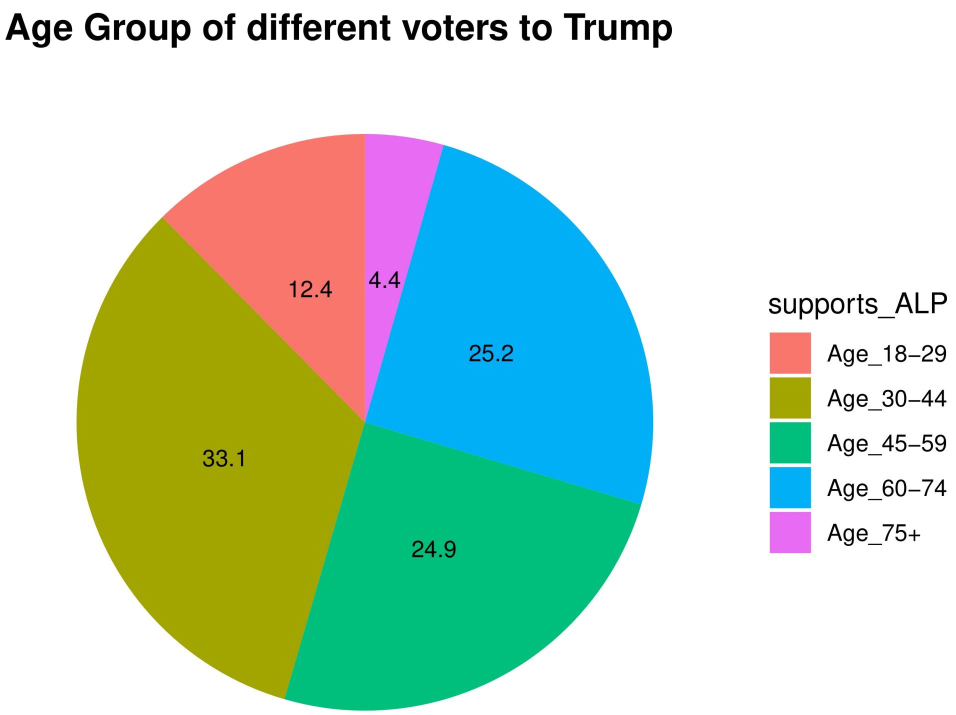
1 represents the respondent supports Donald Trump 0 represents the respondent supports Joe Biden

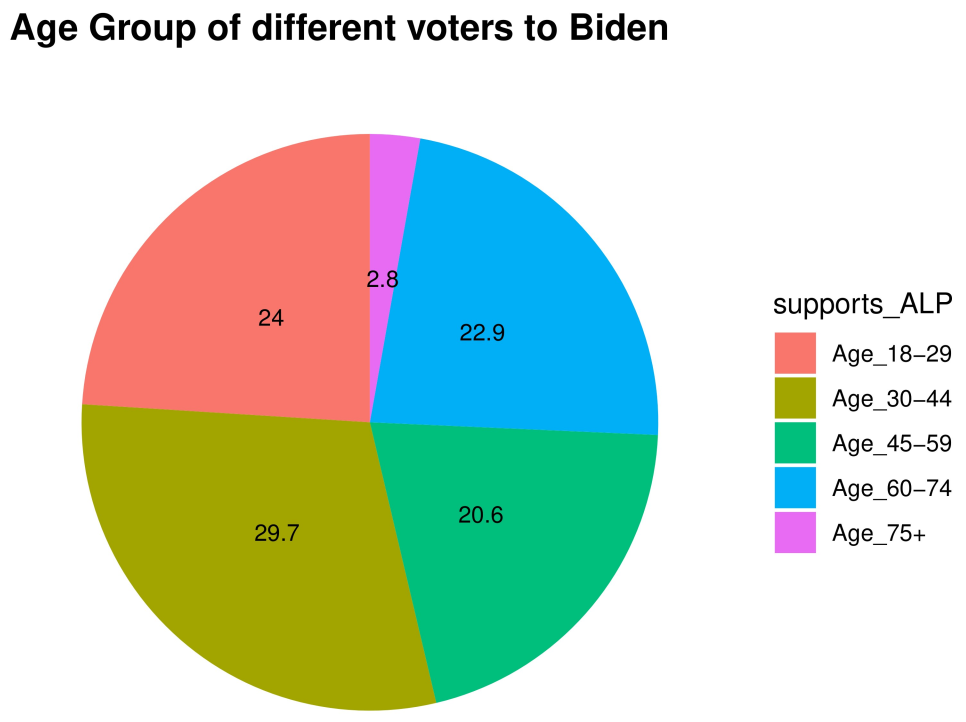
The predictors are our respondent's employment status, gender, racial ethnicity, education level, state of residence, and age group.

Age were categorized into 5 age groups: 18-29, 30-44, 45-59, 60-74 and 75+ Race was categorized into 4 groups, Asian: White, Black and Others Employment status was categorized into 2 groups: employed and unemployed Education level was categorized into 4 groups:Less than highschool, Highschool, Undergraduates or



similar degree and Graduates or higher degree


Age Group of different voters to Trump
12.4
4.4
25.2
33.1
24.9
supports_ALP
Age_18–29
Age_30–44
Age_45–59
Age_60–74
Age_75+



From Figure 1. the histogram, we can see that Joe Biden has slightly higher supporters in our respondents compared to Donald Trump.

From Figure 2 and 3, the pie chart, we can see that respondents aged 30 to 44 are the primary supporters of Donald Trump, with 33.1%, compared to 29.7% of the age group supporting Joe Biden. Respondents in the age group 18 - 29 made up 12.4% of Donald Trump's supporters compared to 24% supporting Joe Biden.

The post stratification data was provided generously from the work of American Community Surveys ACS and Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. The data we used was AMERICAN COMMUNITY SURVEY 2018 SAMPLE. You can find it at https://usa.ipums.org/usa/sampdesc.shtml#us2018a.

The Integrated Public Use Microdata Series (IPUMS USA) collects the data from more than fifty high-precision samples of the American population drawn from fifteen federal censuses and from the American Community Surveys of 2000-Present to the present. However, because different researchers created these samples at different times and employed different record layouts, coding schemes, and documentation. The IPUMS assigns unified all the samples and brings relevant documentation into a coherent form to enable others a better analysis of social and economic change. And the data we used is only a part of it – ACS 2018 version.

The population is all citizens of The United State that are eligible to vote for their president candidates. And the Frame population is all the citizens that took part of the Censors. To enable a better accuracy of our model, we filtered out those non-response, such as those under 18.

The strength of these data is that it is very comprehensive, which makes our filtered data set more accurate. However, there are some drawbacks. For instance, in the original data, the labforce status only includes two values: 'yes' and 'no'. However, in the survey dataset, there are many other values, such as 'retired' and 'full-time' and 'part-time' and so on. To match these two setdata, we curtailed the survey data, and this might cause the loss of accuracy. If in the post-strat dataset, it has more genres, we can have more cells and more accurate data.

The dataset contains 613777 observations of 28 variables. After categorizing and removing undesired data, 11307 observations remained. Variables kept includes sex, race, education level, state of residence, age and labour force. These predictors are categorized and renamed in response to the survey data.

Age were categorized into 5 age groups, 18-29, 30-44, 45-59, 60-74 and 75+ Race was categorized into 4 groups, Asian, White, Black and Others Employment status was categorized into 2 groups, employed and unemployed Education level was categorized into 4 groups: Less than highschool, Highschool, Undergraduates or similar degree and Graduates or higher degree

To be more specific, in our total sampling population, we included 236586 males and 252954 females. Among them, 20408 have less than highschool degree, and 168297 have highschool education. Besides, 242190 and 58645 have undergraduate and graduate educational backgrounds respectively. In terms of the age group, most of the people sit in the age group between 45-59 (124169), while people between 18 and 29 take the least proportion in the overall population.

In this research, we used multilevel regression with the post-stratification method, and it is useful in terms of survey modelling. Our survey dataset cannot cover all the population, and it is a non-representative sample. Therefore, we used post-stratification to adjust the sampling weights to sum to the population sizes within each subset. This will result in a decreased bias due to the nonresponse and underrepresented groups in the population. Although it will automatically adjust for the underrepresented groups, it is still troublesome when we have small-sized subsets.

#model

```
##   X gender   race                        education state       age employment
## 1 1 Female Others                       Highschool    AK Age_45-59 unemployed
## 2 2 Female  White                       Highschool    AK Age_30-44   employed
## 3 3 Female  White Undergraduates or similar degree    AK Age_18-29   employed
## 4 4 Female  White Undergraduates or similar degree    AK Age_18-29 unemployed
## 5 5 Female  White Undergraduates or similar degree    AK Age_30-44 unemployed
## 6 6   Male  White Undergraduates or similar degree    AK   Age_75+ unemployed
##   trump
## 1     1
## 2     0
## 3     0
## 4     1
## 5     1
## 6     1
```

```
## Warning: Parts of the model have not converged (some Rhats are > 1.05). Be
## careful when analysing the results! We recommend running more iterations and/or
## setting stronger priors.
```

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: supports_ALP ~ gender + age + state + employment + race + education
##    Data: example_poll (Number of observations: 1045)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                                         Estimate Est.Error   l-95% CI
## Intercept                                  78.00     78.52       0.73
```

```
## genderMale                                 0.88      0.17      0.51
## ageAge_30M44                               0.40      0.22     -0.01
## ageAge_45M59                               0.60      0.22      0.18
## ageAge_60M74                               0.65      0.30      0.09
## ageAge_75P                                 0.66      0.76     -0.96
## stateAL                                  -81.27     78.52   -205.99
## stateAR                                  -80.83     78.57   -205.82
## stateAZ                                  -78.15     78.51   -202.44
## stateCA                                  -79.84     78.51   -203.93
## stateCO                                  -79.48     78.52   -203.81
## stateCT                                  -80.77     78.50   -205.63
## stateDC                               -21277.99  18650.45  -66171.84
## stateDE                              -106812.44  83003.50 -309674.58
## stateFL                                  -79.57     78.51   -203.68
## stateGA                                  -79.63     78.52   -203.83
## stateHI                                  -80.01     78.48   -204.37
## stateIA                              -124424.84 106229.92 -345882.59
## stateID                               -93488.77  83493.00 -289710.60
## stateIL                                  -79.87     78.49   -204.09
## stateIN                                  -78.88     78.52   -203.27
## stateKS                                  -78.13     78.51   -202.43
## stateKY                                  -79.18     78.52   -203.75
## stateLA                                  -80.65     78.47   -205.07
## stateMA                                  -81.33     78.47   -205.88
## stateMD                                  -80.74     78.53   -205.22
## stateMI                                  -79.62     78.48   -203.91
## stateMN                                  -77.28     78.62   -201.97
## stateMO                               -10373.33   4782.30  -21643.44
## stateMS                                  -79.87     78.54   -204.42
## stateMT                               505493.93 503600.42    582.03
## stateNC                                  -79.20     78.52   -203.34
## stateND                              -124089.87  88393.28 -325914.39
## stateNE                              1105629.81 728065.56  14389.17
## stateNH                              -143219.36 101822.42 -342382.57
## stateNJ                                  -79.70     78.52   -203.95
## stateNM                               -47632.23  54373.81 -197702.68
## stateNV                                  -79.41     78.53   -203.75
## stateNY                                  -80.05     78.53   -204.25
## stateOH                                  -81.23     78.49   -205.48
## stateOK                                  -78.08     78.54   -202.44
## stateOR                                  -79.21     78.50   -203.70
## statePA                                  -78.77     78.48   -202.98
## stateRI                              -133942.22 178573.05 -647407.79
## stateSC                                  -79.66     78.53   -204.01
## stateSD                              -259714.06 268715.43 -867455.54
## stateTN                                  -79.15     78.48   -203.55
## stateTX                                  -79.07     78.51   -203.17
## stateUT                                  -76.78     78.56   -201.82
## stateVA                                  -79.48     78.45   -203.58
## stateVT                                  -79.51     78.40   -204.30
## stateWA                                  -79.09     78.55   -203.41
## stateWI                                  -79.40     78.50   -203.91
## employmentunemployed                      -0.08      0.21     -0.46
## raceBlack                                 -1.30      0.20     -1.69
```

```
## raceOthers                                      0.41      0.16      0.05
## educationHighschool                            -0.05      0.29     -0.65
## educationLessthanhighschool                    -0.46      1.10     -3.05
## educationUndergraduatesorsimilardegree          0.04      0.25     -0.44
##                                             u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept                                     202.14 3.35        4       17
## genderMale                                      1.18 1.18       31       21
## ageAge_30M44                                    0.79 1.22       14      100
## ageAge_45M59                                    1.01 1.23       14       36
## ageAge_60M74                                    1.32 1.18       17       25
## ageAge_75P                                      1.99 1.15       25       29
## stateAL                                        -3.81 3.30        4       13
## stateAR                                        -3.29 3.26        4       17
## stateAZ                                        -0.97 3.36        4       16
## stateCA                                        -2.85 3.37        4       17
## stateCO                                        -2.31 3.37        4       16
## stateCT                                        -3.01 3.34        4       15
## stateDC                                      -637.85 1.86        6       25
## stateDE                                     -9601.06 1.78        6       12
## stateFL                                        -2.49 3.38        4       17
## stateGA                                        -2.49 3.37        4       17
## stateHI                                        -2.88 3.37        4       16
## stateIA                                      -821.75 1.83        6       15
## stateID                                     -2829.17 3.18        5       15
## stateIL                                        -2.79 3.38        4       16
## stateIN                                        -1.74 3.32        4       16
## stateKS                                        -0.80 3.32        4       15
## stateKY                                        -1.99 3.31        4       14
## stateLA                                        -3.26 3.25        4       15
## stateMA                                        -4.16 3.28        4       15
## stateMD                                        -3.45 3.38        4       16
## stateMI                                        -2.52 3.35        4       16
## stateMN                                        -0.06 3.35        4       18
## stateMO                                     -2836.71 1.56        7       17
## stateMS                                        -2.56 3.35        4       15
## stateMT                                   1841785.42 3.10        5       19
## stateNC                                        -2.10 3.39        4       16
## stateND                                    -17780.59 1.55       11       34
## stateNE                                   2265917.73 2.39        5       28
## stateNH                                    -23722.23 1.86        6       59
## stateNJ                                        -2.47 3.36        4       16
## stateNM                                     -1110.40 1.60        7       15
## stateNV                                        -2.26 3.34        4       16
## stateNY                                        -3.01 3.37        4       17
## stateOH                                        -3.96 3.42        4       15
## stateOK                                        -1.00 3.39        4       16
## stateOR                                        -1.95 3.30        4       16
## statePA                                        -1.67 3.39        4       15
## stateRI                                     -4762.74 2.79        5       11
## stateSC                                        -2.45 3.34        4       14
## stateSD                                    -13260.30 2.86        5       12
## stateTN                                        -1.99 3.38        4       16
## stateTX                                        -2.01 3.39        4       18
## stateUT                                         1.06 3.23        4       16
```

```
## stateVA                                       -2.50 3.38         4        15
## stateVT                                       -2.22 3.13         5        15
## stateWA                                       -2.04 3.39         4        16
## stateWI                                       -2.33 3.33         4        15
## employmentunemployed                           0.34 1.48         8        46
## raceBlack                                     -0.89 1.18        19        57
## raceOthers                                     0.69 1.12        31        53
## educationHighschool                            0.44 1.37         9        55
## educationLessthanhighschool                    1.45 1.19        17        66
## educationUndergraduatesorsimilardegree         0.51 1.44         8        17
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

The model we used in this research is Bayesian Multilevel regression with post-stratification.In this case, we have a Bernoulli distributed response variable, and that means we set the dependent variable as being binary in a Bayesian setting. We are using the MRP model to predict Trump or Biden who will win the 2020 election of the USA. As in the dataset, we have categorized our response outcome into 2 categories, support Trump or support Biden, where 1 represents the respondent supports Donald Trump 0 represents the respondent supports Joe Biden Gender,age,state,employment,race,and education are predictors. The equation is:

$$n_{[i]}^{supportALP=1} BernoulliLogit(N_{[i]}, \alpha[i])$$

$$alpha_[i] = \alpha_{g[i]}^{gender} + \alpha_{a[i]}^{age} + \alpha_{s[i]}^{state} + \alpha_{e[i]}^{employment} + \alpha_{r[i]}^{race} + \alpha_{edu[i]}^{education}$$

Discussion and Results

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## # A tibble: 6 x 10
##      X1 gender race  education state age   employment people_in_each_~
##   <dbl> <chr>  <chr> <chr>     <chr> <chr> <chr>                 <dbl>
## 1     1 Female Asian Graduate~ AK    Age_~ employed                980
## 2     2 Female Asian Highscho~ AK    Age_~ employed                980
## 3     3 Female Asian Highscho~ AK    Age_~ unemployed              980
## 4     4 Female Asian Highscho~ AK    Age_~ unemployed              980
## 5     5 Female Asian Highscho~ AK    Age_~ employed                980
## 6     6 Female Asian Highscho~ AK    Age_~ unemployed              980
## # ... with 2 more variables: number_in_cell <dbl>,
## #   cell_prop_of_division_total <dbl>
```

```
## # A tibble: 49 x 4
##    state  mean    lower upper
##    <chr> <dbl>    <dbl> <dbl>
##  1 AK    0.992 0.881    1.
##  2 AL    0.152 0        0.447
##  3 AR    0.231 0        0.643
##  4 AZ    0.710 0.474    0.898
```

```
##   5 CA    0.341 0.204    0.490
##   6 CO    0.447 0.150    0.761
##   7 CT    0.239 0.000714 0.631
##   8 DC    0     0        0
##   9 DE    0     0        0
##  10 FL    0.400 0.230    0.583
## # ... with 39 more rows
```

```
## [1] 36
```

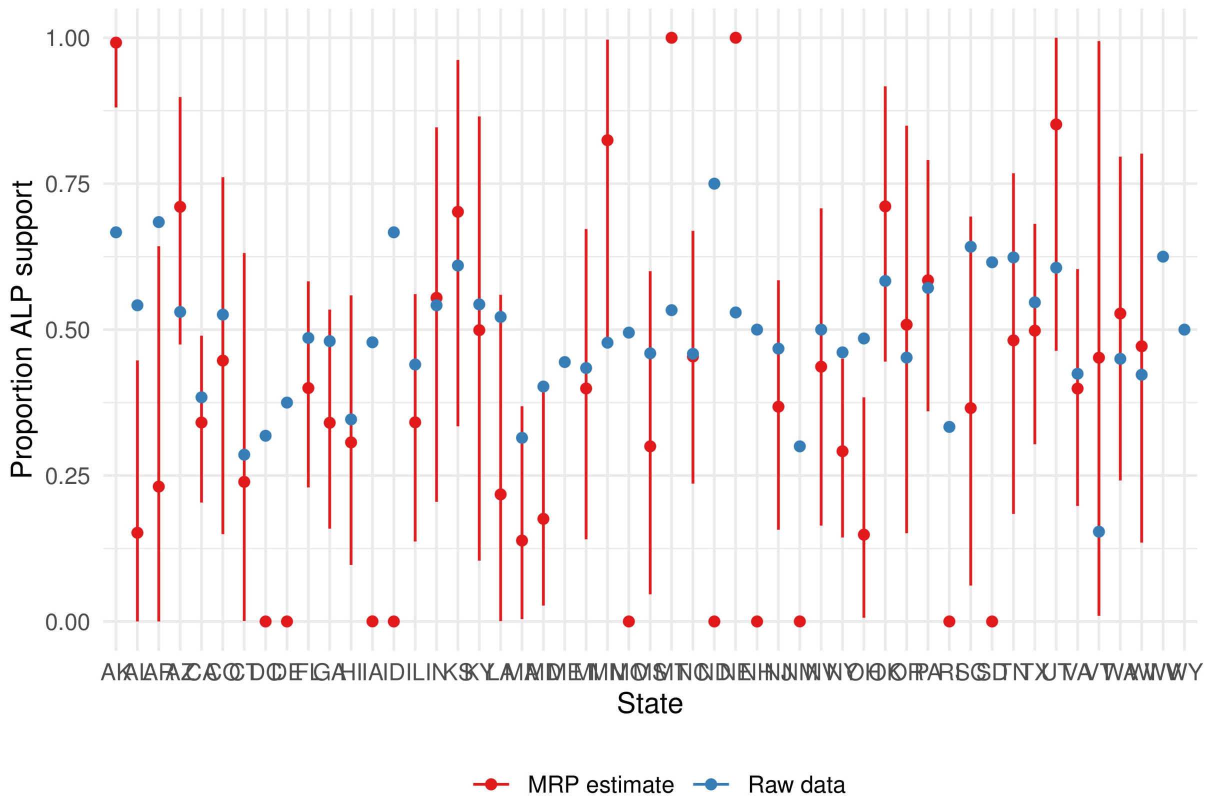
```
##  Family: bernoulli
##   Links: mu = logit
## Formula: supports_ALP ~ gender + age + (1 | state)
##    Data: example_poll (Number of observations: 1045)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Group-Level Effects:
## ~state (Number of levels: 48)
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.66      0.17     0.35     1.03 1.00     1327     2213
##
## Population-Level Effects:
##              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept       -1.87      0.21    -2.29    -1.47 1.00     2761     3051
## genderMale       0.81      0.16     0.51     1.13 1.00     7696     2582
## ageAge_30M44     0.45      0.20     0.05     0.84 1.00     5026     3341
## ageAge_45M59     0.54      0.22     0.11     0.97 1.00     5005     2782
## ageAge_60M74     0.42      0.25    -0.06     0.91 1.00     4943     3301
## ageAge_75P       0.28      0.73    -1.29     1.61 1.00     6160     2771
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
##    state   n
## 1     AK   6
## 2     AL  72
## 3     AR  38
## 4     AZ 132
## 5     CA 565
## 6     CO  78
## 7     CT  56
## 8     DC  22
## 9     DE  24
## 10    FL 420
## 11    GA 152
## 12    HI  26
## 13    IA  46
## 14    ID  24
## 15    IL 234
## 16    IN  96
## 17    KS  41
```

```
## 18    KY  81
## 19    LA  69
## 20    MA  89
## 21    MD  82
## 22    ME  18
## 23    MI 152
## 24    MN  67
## 25    MO  97
## 26    MS  37
## 27    MT  15
## 28    NC 181
## 29    ND   4
## 30    NE  17
## 31    NH  16
## 32    NJ 169
## 33    NM  20
## 34    NV  60
## 35    NY 423
## 36    OH 231
## 37    OK  48
## 38    OR  73
## 39    PA 210
## 40    RI   9
## 41    SC  81
## 42    SD  13
## 43    TN  93
## 44    TX 364
## 45    UT  33
## 46    VA 172
## 47    VT  13
## 48    WA 100
## 49    WI  97
## 50    WV  32
## 51    WY   2
```

Then we use post-stratified estimates for each division. Our new Bayesian approach shows trump's support rate for each state. In our post-stratified estimates, there will be 16 states (out of 49 states) which are strongly opposing Trump to be the next president. They will not vote for Trump definitely. Also, there are 14 states which are strongly in favor of voting for Trump. So there are 19 swing states. And those state's estimated mean is lower than 0.5. It is hard for Trump to get the vote from those states.

```
## Warning: Removed 1 rows containing missing values (geom_point).
```



According to the post-stratified estimates, we plot a graph.Asthe graph shows, Trump is in a disadvantage. For state model we used the bayesian approach and a layer is added. By changing the second level group-parameter 'state', we can get different intercept. We want to show how the model affects the results and a graph is created to compare the raw estimate to the model estimate. By observing the proportion ALP support vs State graph we can see that state KY, NC, PA and WI, the raw estimates are the closest to the model estimate whereas in state ND a big difference is observed.

Appendices

1. You can find our codes in:https://github.com/songyuan-l/Biden-might-win-the-2020-election

References

1.R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

@Manual{, title = {here: A Simpler Way to Find Your Files}, author = {Kirill Müller}, year = {2017}, note = {R package version 0.1}, url = {https://CRAN.R-project.org/package=here}, }

2.To cite package 'broom' in publications use:

David Robinson, Alex Hayes and Simon Couch (2020). broom: Convert Statistical Objects into Tidy Tibbles. R package version 0.7.2. https://CRAN.R-project.org/package=broom 3.To cite brms in publications use:

Paul-Christian Bürkner (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80(1), 1-28. doi:10.18637/jss.v080.i01

Paul-Christian Bürkner (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal, 10(1), 395-411. doi:10.32614/RJ-2018-017

To see these entries in BibTeX format, use 'print(, bibtex=TRUE)', 'toBibtex(.)', or set 'options(citation.bibtex.max=999)'.

4.To cite package 'here' in publications use:

Kirill Müller (2017). here: A Simpler Way to Find Your Files. R package version 0.1. https://CRAN.R-project.org/package=here

5.Kay M (2020). *tidybayes: Tidy Data and Geoms for Bayesian Models.* doi: 10.5281/zenodo.1308151 (URL: https://doi.org/10.5281/zenodo.1308151), R package version 2.1.1, <URL: http://mjskay.github.io/tidybayes/>.

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

7.To cite package 'haven' in publications use:

Hadley Wickham and Evan Miller (2020). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.3.1. https://CRAN.R-project.org/package=haven

8.data resource https://www.voterstudygroup.org/publication/nationscape-data-set https://www.voterstudygroup.org/downloads?key=9bcb7391-219e-44b8-85f1-d411c753d0b0

9.week 6 lec note