Hope is on Joe Biden, Pressure is on Donald Trump

Ying Xiong(1004795885),Yixin Liang(1005549998),Huiyi Lu(1004754615),Deyu Meng(1004739991)

Code and data supporting this analysis is available at: https://github.com/huiyilu99/304-a3

11/02/2020

Contents

1 Model

We aim to predict the popular vote outcome and electoral vote outcome of the 2020 American federal election[1]. We apply a post-stratification technique with a logistic model generated in R. In the subsections, we will describe the data cleaning process, model specifics, and the post-stratification method.

1.1 Data Cleaning

To correspond our survey data variables to the census data variables for modeling and predictions, we recategorize the data within several variables, race, age, state, and sex in both data sets.

- race

Initially, in the census data, the variable has the following 9 categories: <white, chinese, black/african american/negro, two major races, other race, nec, american indian or alaska native, other asian or pacific islander, three or more major races, japanese>. In the survey data, the variable contains the following 15 categories: <White, Black, or African American, Asian (Asian Indian), Asian (Vietnamese), Asian (Chinese), Asian (Korean), Asian (Japanese), Some other race, Asian (Filipino), Asian (Other), Pacific Islander (Native Hawaiian), American Indian or Alaska Native, Pacific Islander (Other), Pacific Islander (Samoan), Pacific Islander (Guamanian)>. To match the categories, we reclassify the observations into 6 categories. Similar categories are retained while the others are recorded as *other races*.

<white, black, other asian or pacific islander, chinese, japanese, other races, american indian or alaska native>.

- age

Initially, in both data, the age variable is expressed numerically. We classify them into 6 groups that would contain all the eligible ages. Specifically, we identify the young (18[2] - 25[3]) and the retired group (66[4] to 97^[The maximum age in both survey data and census data. Referring to **Table 3(Maximum Age in Survey data & Census data)** in Appendix 4.1), the remaining ages are grouped with a 10-year difference.

<age_46_55, age_36_45, age_66_97, age_18_25, age_56_65, age_26_35>.

- state

The categories (51 states) in the survey data exactly match those of the census data. Notably, the survey data uses the abbreviations for each state while the census data uses the full name. We rename each category with the abbreviations in the census data for simplicity.

<WI, VA, TX, WA, OH, MA, CA, NC, MD, FL, WV, NY, KY, IN, MI, IA, SC, MN, GA, PA, NJ, AZ, IL, AR, OK, NV, OR, CT, DE, MO, CO, DC, NM, TN, HI, MT, VT, UT, NE, KS, NH, LA, ME, AL, ID, MS, SD, WY, ND, AK, RI>.

- sex

The categories (2 genders) in the survey data exactly match those of the census data. There are some spelling differences, such as the capitalization of the first letter in the survey data. We apply no capitalized letters for categories.

<female, male>.

[1]https://www.cnn.com/politics/live-news/us-election-news-11-02-2020/index.html More details about the election are included here.

[2]https://www.usa.gov/voter-registration-age-requirements Notably, different states have different voter registration age standards. We use 18 as the unified minimum age to be eligible to vote here for simplicity.

[3]It is common knowledge that most people will finish their education at age around 25 when they finish graduate studies.

[4]https://www.nasi.org/learn/socialsecurity/retirement-age 66 is the retirement age in the U.S.

1.2 Model Specifics

We are interested in predicting the popular vote and electoral vote results of the 2020 American federal election. To predict the outcome, we employ the post-stratification technique on the census data. Before predictions, we estimate the probability of each post-stratification cell voting for Trump with our model. The model is created through the survey data collected from the Voter Study Group.

We employ a logistic regression model to calculate the probabilities, which is specific for binary response modeling. We apply a frequentist method because people's minds would change from year to year, and there is no such precise prior information to use for the Bayesian inference.

The following mathematical notation is the logistic regression model we create in R with four predictors, *race_factor*, *age_factor*, *state*, and *sex*.

$$\log \frac{p_i}{1-p_i} = \beta_0 + \beta_{state} X_{state} + \beta_{race} X_{race} + \beta_{age} X_{age} + \beta_{sex} X_{sex}$$

p_i: the probability of the respondent voting for Trump in the 2020 American federal election

β_0: the intercept value; represents the log odds of an 18-25 ages old female being in Alaska and is identified as American Indian or Alaska Native. These information in each variable would be the reference groups in the following discussion.

β_{state}: the slope values for states; represent the difference of preference between people in Alaska and the other 50 states

X_{state}: = 1 if the respondent is in states except for Alaska

β_{race}: the slope values for race groups; represent the difference of preference between people of American Indian or Alaska Native and the other 6 race groups

$X_{racefactor}$: = 1 if the respondent is identified as races other than American Indian or Alaska Native

β_{age}: the slope values for age groups; represent the difference of preference between people aged 18-25 and the other 5 age groups

$X_{agefactor}$: = 1 if the respondent's age is in range other than 18 - 25

β_{sex}: the slope values for states; represent the difference of preference between females and males

X_{sex}: = 1 if the respondent is identified as a male

Given the above variable information, we can effectively predict a person's probability of a citizen voting for Trump. The variables separate the information into different categories that allow each post-stratification cell to be classified precisely and give back a relatively correct probability.

We employ 4 predictors in our final prediction model, *race_factor*, *age_factor*, *state*, and *sex*. To avoid potential errors that may mislead the probability, we choose to use the existing variables that are matchable. This is the model with the smallest AIC[5] we create. Since most of the betas for the variable state have large p-values, we try another model that excludes the state variable. Referring to **Table 5 (AIC comparison for Trump's Models)** in Appendix 4.4, it turns out that the AIC of the model would be reduced if we include the variable. The other three variables all have statistically significant p-values, and we retain them as predictors.

1.3 Post-Stratification

The post-stratification method partitions the census data into demographic cells according to the chosen variables. Typically, the number of cells should equal all possible combinations of the variables unless no observations satisfy the cells' characteristics.

[5] Akaike's information criterion. A method used to select models. The smaller, the better.

To estimate the candidates' voting rate, state, age, race, and sex are used to create cells to implement the post-stratification analysis. The census data set is separated into 4241 different demographic cells. The response variable, proportion of voters, is estimated within each demographic cell. The population-level estimate is then calculated by weighting each cell based on its relative proportion in the population. Post-stratification is useful because it helps to reduce the bias resulting from non-probability based sampling.

In the United States, fifty-one individual states may have unique cultures, characteristics, or even local laws. These differences lead to the variation in voter approval, which is worthy of being analyzed. Different age groups may have various attitudes toward the candidates, and the policies candidates commit to implementing after becoming the president. The gender gap is always a significant issue because men tend to vote for republicans, while women tend to vote for democrats (Thompson, 2020). Also, the candidates' behavior, personality, and communication style may trigger gender bias. The candidate's policy proposals may dramatically raise support from a specific race. Also, candidates' attitudes toward immigration and ethnicity may affect the support rate (Cohn, 2020).

1.4 Eligibility for Votes

When cleaning the data, juveniles and children are removed from the data set because 18-years-old is the required age to gain the legal voting right (USAGov, 2020). Though a few states may have a different age restriction, 18-years-old is assumed to be the required age to simplify our analysis. Also, observations above 90 years old in 1980 and 1990 are removed because they are less likely to attend the in-person voting due to personal or health conditions. They would be around 120 years old now if not passed away, which is less likely.

1.5 Electoral Vote

This study focuses on analyzing both the electoral vote and popular vote. According to the special presidential election procedure, most states comply with a winner-takes-all system (National Archives, 2019). The candidate who wins a plurality of votes wins the state's total allocated electoral votes. The electoral rate measures the possibility that the candidate would win the election after considering the winner-takes-all system. However, the popular rate only measures the probability that the presidential candidate has higher total votes in the U.S. Therefore, this study calculates both electoral rate and popular rate to make comparisons. This is essential because some candidates win the popular vote but lose in the electoral vote, which leads to the final failure in the presidential election.

For the electoral vote analysis, it is notable that states Maine and Nebraska do not use a winner-takes-all method. They give two votes to the statewide winner and one vote to each Congressional district winner. Due to the absence of Congressional districts' data in two states, this study assumes these two states also follow the winner-takes-all method.

2 Results

Table 1: Predicted Popular Vote for Trump v.s. Biden

Trump	Biden
0.395	0.4149

- Referring to **Table 1(Predicted Popular Vote for Trump vs. Biden)** in Appendix 4.6, we estimate that Trump will likely receive 39.5% of the total number of votes, and Biden will likely receive 41.5% of the total number of votes.

Table 2: Predicted Electoral Vote for Trump v.s. Biden

Trump	Biden
0.4535	0.5465

- Referring to **Table 2(Predicted Electoral Vote for Trump vs. Biden)** in Appendix 4.5, we estimate that the proportion of electoral votes that Trump will be likely to receive is 45.4%. In comparison, the proportion of electoral votes that Biden will be likely to receive is 54.6%.

This is based on our post-stratification analysis modeled by logistic regression, which accounts for the state, age, race, and sex. The same process is done in both the Trump and Biden models to predict the election winner. Referring to Appendix 4.3(**Table 4(Model for Trump)**) and 4.4(**Table 5(Model for Biden)**), most categories under state have large p-values in both the Trump and Biden models, which means the state is statistically insignificant. However, we still believe that states can influence the outcome as the U.S has blue (Democratic) states and red (Republican) states. Age is significant in the Trump model but insignificant in the Biden model. We include age in both models since Millennials are more Democratic-leaning. 54% of Millennials identify with the Democratic Party or lean Democratic, while 38% identify with or lean to the GOP (Pew Research Center, 2020). Overall, according to our prediction of the popular vote and electoral vote, we believe that Biden has a higher chance of winning.

3 Discussion

3.1 Summary

We collect the individual-level survey data from the Democracy Fund and 2018 5-year American Community Surveys from IPUMS. The individual-level survey data consists of some personal information and the interviewees' intention to vote, while the 2018 5-year American Community Surveys only consist of some personal information.

We clean and create four variables(age, state, race, sex) in both data sets to assure the categories under each categorical variable are consistent. Logistic regression is used on individual-level survey data to assess the impact of each explanatory variable we chose.

We apply a post-stratification method to partition the census data into demographic cells according to the selected variables. Thus the census data is split into 4241 different demographic cells. We estimate the proportion of voters in each state and follow the "winner-takes-all" principle to calculate the electoral vote. Then we weigh each proportion estimate by the respective population size of that state and sum those values and divide that by the entire population size into the census data to calculate the popular vote.

The results are that Biden may gain 54.6% of the electoral vote and 41.5% of the popular vote, and Trump may acquire 45.4% of the electoral vote and 39.5% of the popular vote. Our model might be biased since we only include four variables and split the data into 4241 cells — generally, the more cells, the better. Computation errors might occur during the record of the survey. Interviewees might answer the survey or the census without thinking thoroughly.

3.2 Conclusion

Based on the statistical analysis conducted in this report, it is estimated that 41.5% of the overall popular vote and 54.6% of the Electoral College vote are in favor of Biden. Simultaneously, for Trump, the statistics are 39.5% and 45.4%, respectively, referring to **Table 1 (Predicted Popular Vote for Trump v.s. Biden)** and **Table 2 (Predicted Electoral Vote for Trump v.s. Biden)**. The results of the two vote-counting approaches are consistent — Biden is taking the lead. We can conclude that Joe Biden is predicted to have a higher probability of winning the 2020 U.S. presidential election at the current moment using our data.

The line of reasoning behind including the results from two vote-counting approaches is that most states comply with the "winner-takes-all" procedure for the Electoral College vote. (National Archives, 2019) There are occasions when the two results may not be in line with one another. For instance, on historical precedent, the candidate would still lose the election even though the overall popular vote was in favor of him/her, which was the case for Hillary Clinton, who outcompeted Trump for about 3 million popular votes but lost the previous U.S. presidential election in 2016. (Robertson, Kirk, & Hulley-Jones, 2020) Estimating the results of both the overall popular vote and the Electoral College vote would enhance the completeness of our analysis.

Besides predicting the winner, the results mentioned above certainly are influential to the two candidates in various aspects. Further inferences and predictions could be made from the results. To elaborate, amid the COVID-19 pandemic, the candidates should quickly adapt to the current situation. It is reported that Biden exceeded Trump in terms of fundraising because Biden and his team had fully utilized the online platform early on, which proved to be cost-effective. (Wilkie, 2020) In our case, for the party being at a disadvantage — Trump and his team would probably reflect upon the results and make revisions to his campaign to gain more votes based on the results.

3.3 Weaknesses

Throughout our report, we have made the assumptions of the eligible voting age being 18, and all states following the "winner-takes-all" procedure for the Electoral College vote for simplicity. However, in reality, the age restrictions for voting in different states vary. (USAGov, 2020) Moreover, two states — Maine and Nebraska, have their own rules for processing the Electoral College vote. (National Archives, 2019) Unfortunately, we did not consider all the variations when conducting the analysis, which may constrain our prediction results.

3.4 Next Steps

To enhance our logistic model's prediction accuracy, we can gather more information and find more common variables between the two data sets to include more predictors for further analysis. Alternatively, the multilevel regression model could be another suitable method to analyze our data due to group-level data such as states. Applying other models may generate smaller AIC. We can also evaluate using other measurements such as Bayesian Information Criterion (BIC) and Residual Sum of Squares (RSS) for model selection.

Furthermore, we can compare our model results with the actual outcomes to see what causes the difference. Then, we can conduct another statistical analysis with more comprehensive information to validate our methodology. This would lay the basis for a more accurate future election prediction.

4 Appendix

4.1 Distribution of Age

Table 3: Maximum Age in Survey data & Census data

survey_max	census_max
93	97

4.2 Model for Trump

Table 4: Summary of Logit Regression for Trump

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.754	0.751	-1.003	0.316
stateAL	-0.145	0.756	-0.192	0.848
stateAR	0.140	0.781	0.179	0.858
stateAZ	-0.434	0.739	-0.587	0.557
stateCA	-0.844	0.725	-1.164	0.244
stateCO	-0.521	0.751	-0.694	0.488
stateCT	-1.476	0.777	-1.898	0.058
stateDC	-0.663	0.875	-0.757	0.449
stateDE	-0.910	0.834	-1.091	0.275
stateFL	-0.443	0.727	-0.609	0.542
stateGA	-0.078	0.740	-0.106	0.916
stateHI	-0.545	0.842	-0.647	0.518
stateIA	-0.717	0.773	-0.927	0.354
stateID	-0.093	0.814	-0.115	0.909
stateIL	-0.678	0.732	-0.926	0.354
stateIN	-0.503	0.746	-0.674	0.500
stateKS	-0.265	0.778	-0.341	0.733
stateKY	-0.283	0.753	-0.376	0.707
stateLA	-0.160	0.761	-0.210	0.834
stateMA	-1.313	0.754	-1.741	0.082
stateMD	-0.626	0.754	-0.830	0.406
stateME	-0.892	0.855	-1.043	0.297
stateMI	-0.723	0.738	-0.980	0.327
stateMN	-0.392	0.762	-0.515	0.606
stateMO	-0.537	0.746	-0.720	0.472
stateMS	-0.104	0.793	-0.131	0.896
stateMT	-0.444	0.872	-0.509	0.611
stateNC	-0.398	0.736	-0.542	0.588
stateND	-0.511	1.058	-0.483	0.629
stateNE	-0.586	0.847	-0.692	0.489
stateNH	-0.755	0.862	-0.876	0.381
stateNJ	-0.596	0.736	-0.810	0.418
stateNM	-1.442	0.861	-1.675	0.094
stateNV	-0.312	0.764	-0.408	0.683
stateNY	-0.651	0.727	-0.896	0.370
stateOH	-0.622	0.731	-0.851	0.395
stateOK	-0.319	0.765	-0.416	0.677
stateOR	-0.818	0.754	-1.084	0.278

	Estimate	Std. Error	z value	Pr(>|z|)
statePA	-0.445	0.732	-0.609	0.543
stateRI	-1.089	1.000	-1.089	0.276
stateSC	-0.044	0.751	-0.058	0.954
stateSD	-0.399	0.882	-0.453	0.651
stateTN	-0.115	0.746	-0.154	0.877
stateTX	-0.244	0.727	-0.335	0.737
stateUT	-0.651	0.777	-0.837	0.403
stateVA	-0.655	0.737	-0.888	0.374
stateVT	-2.090	1.051	-1.989	0.047
stateWA	-0.774	0.746	-1.038	0.299
stateWI	-0.885	0.748	-1.183	0.237
stateWV	-0.231	0.791	-0.292	0.770
stateWY	-1.817	1.336	-1.359	0.174
age_factorage_26_35	0.491	0.107	4.581	0.000
age_factorage_36_45	0.758	0.104	7.261	0.000
age_factorage_46_55	0.865	0.111	7.777	0.000
age_factorage_56_65	0.820	0.109	7.543	0.000
age_factorage_66_97	0.849	0.111	7.616	0.000
race_factorblack	-1.869	0.263	-7.110	0.000
race_factorchinese	-0.988	0.375	-2.633	0.008
race_factorjapanese	-0.936	0.633	-1.479	0.139
race_factorother asian or pacific islander	-0.404	0.276	-1.461	0.144
race_factorother races	-0.539	0.256	-2.110	0.035
race_factorwhite	0.193	0.232	0.832	0.406
sexmale	0.455	0.056	8.189	0.000

4.3 Model for Biden

Table 5: Summary of Logit Regression for Biden

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.344	0.848	-1.585	0.113
stateAL	0.372	0.849	0.438	0.662
stateAR	-0.203	0.886	-0.229	0.819
stateAZ	0.600	0.835	0.718	0.472
stateCA	0.954	0.822	1.160	0.246
stateCO	0.604	0.845	0.714	0.475
stateCT	1.320	0.854	1.546	0.122
stateDC	1.673	0.947	1.767	0.077
stateDE	1.187	0.905	1.312	0.190
stateFL	0.685	0.824	0.832	0.406
stateGA	0.431	0.834	0.517	0.605
stateHI	0.969	0.910	1.065	0.287
stateIA	0.939	0.863	1.087	0.277
stateID	-0.116	0.919	-0.126	0.899
stateIL	0.803	0.827	0.971	0.332
stateIN	0.547	0.841	0.650	0.516
stateKS	0.325	0.875	0.371	0.710
stateKY	0.810	0.847	0.957	0.339
stateLA	0.562	0.852	0.660	0.509
stateMA	1.189	0.840	1.415	0.157

	Estimate	Std. Error	z value	Pr(>|z|)
stateMD	0.839	0.844	0.995	0.320
stateME	1.282	0.934	1.373	0.170
stateMI	0.961	0.832	1.155	0.248
stateMN	1.138	0.853	1.334	0.182
stateMO	0.594	0.840	0.708	0.479
stateMS	0.300	0.875	0.343	0.732
stateMT	0.796	0.956	0.832	0.405
stateNC	0.747	0.830	0.899	0.369
stateND	-0.682	1.359	-0.501	0.616
stateNE	0.444	0.932	0.476	0.634
stateNH	0.899	0.943	0.953	0.341
stateNJ	0.763	0.831	0.918	0.359
stateNM	1.057	0.904	1.169	0.242
stateNV	0.517	0.854	0.605	0.545
stateNY	0.820	0.823	0.996	0.319
stateOH	0.653	0.827	0.789	0.430
stateOK	0.173	0.863	0.201	0.841
stateOR	0.881	0.846	1.042	0.297
statePA	0.339	0.829	0.409	0.682
stateRI	1.090	1.005	1.085	0.278
stateSC	-0.056	0.849	-0.065	0.948
stateSD	0.462	0.981	0.471	0.638
stateTN	0.083	0.844	0.099	0.921
stateTX	0.273	0.824	0.331	0.741
stateUT	0.012	0.880	0.013	0.989
stateVA	0.907	0.831	1.092	0.275
stateVT	2.093	1.008	2.077	0.038
stateWA	0.862	0.838	1.028	0.304
stateWI	0.962	0.839	1.146	0.252
stateWV	0.375	0.890	0.421	0.674
stateWY	-0.022	1.386	-0.016	0.987
age_factorage_26_35	-0.004	0.092	-0.039	0.969
age_factorage_36_45	-0.023	0.092	-0.253	0.800
age_factorage_46_55	-0.239	0.100	-2.388	0.017
age_factorage_56_65	0.049	0.097	0.509	0.611
age_factorage_66_97	0.154	0.100	1.536	0.125
race_factorblack	1.563	0.253	6.190	0.000
race_factorchinese	0.988	0.329	3.006	0.003
race_factorjapanese	1.489	0.561	2.654	0.008
race_factorother asian or pacific islander	0.697	0.275	2.534	0.011
race_factorother races	0.637	0.258	2.465	0.014
race_factorwhite	0.296	0.242	1.222	0.222
sexmale	-0.321	0.053	-6.001	0.000

4.4 Model for Trump (exclude *state*)

Table 6: Summary of Logit Regression for Trump without variable state

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.253	0.236	-5.306	0.000

	Estimate	Std. Error	z value	Pr(>|z|)
age_factorage_26_35	0.490	0.106	4.630	0.000
age_factorage_36_45	0.718	0.103	6.970	0.000
age_factorage_46_55	0.850	0.110	7.754	0.000
age_factorage_56_65	0.786	0.107	7.329	0.000
age_factorage_66_97	0.835	0.110	7.614	0.000
race_factorblack	-1.788	0.257	-6.969	0.000
race_factorchinese	-1.174	0.370	-3.176	0.001
race_factorjapanese	-1.027	0.604	-1.700	0.089
race_factorother asian or pacific islander	-0.457	0.269	-1.696	0.090
race_factorother races	-0.583	0.249	-2.338	0.019
race_factorwhite	0.180	0.225	0.800	0.424
sexmale	0.422	0.054	7.747	0.000

Table 7: AIC comparison for Trump's Models

With_state	Without_state
7940.1	7950.6

5 References

1. Britannica. (2016, November 7). United States Electoral College Votes by State. Encyclopædia Britannica. https://www.britannica.com/topic/United-States-Electoral-College-Votes-by-State-1787124.
2. Cohn, N. (2020, October 28). The Election's Big Twist: The Racial Gap Is Shrinking. The New York Times. https://www.nytimes.com/2020/10/28/upshot/election-polling-racial-gap.html.
3. 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
4. IPUMS USA, University of Minnesota, www.ipums.org.
5. Macaya, M., & Hayes, M. (2020, November 2). US election 2020: Latest news on Biden, Trump and voting. https://www.cnn.com/politics/live-news/us-election-news-11-02-2020/index.html.
6. NASI. What is the Social Security Retirement Age? National Academy of Social Insurance. https://www.nasi.org/learn/socialsecurity/retirement-age.
7. National Archives. (2019, December 23). What is the Electoral College? National Archives and Records Administration. https://www.archives.gov/electoral-college/about.
8. Pew Research Center.(2020, August). Changing U.S. Electorate, Race and Education Remain Stark Dividing Lines. 28).https://www.pewresearch.org/politics/2020/06/02/in-changing-u-s-electorate-race-and-education-remain-stark-dividing-lines/.
9. Robertson, H., Kirk, A., & Hulley-Jones, F. (2020, October 30). Electoral college explained: How Biden faces an uphill battle in the US election. https://www.theguardian.com/us-news/ng-interactive/2020/oct/30/electoral-college-explained-how-biden-faces-an-uphill-battle-in-the-us-election.
10. Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0
11. Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape, October 10-17, 2019 (version 20200814). https://www.voterstudygroup.org/downloads?key=8b7d3ae0-ae6b-4bbe-ab55-738726c9b607.
12. Thompson, D. (2020, February 9). Why Men Vote for Republicans, and Women Vote for Democrats. https://www.theatlantic.com/ideas/archive/2020/02/how-women-became-democratic-partisans/606274/.
13. USAGov. (2020, August 12). Voter Registration Age Requirements: USAGov. https://www.usa.gov/voter-registration-age-requirements.
14. Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
15. Wilkie, C. (2020, September 29). Coronavirus changed the 2020 money race - and helped Biden erase Trump's massive cash advantage. https://www.cnbc.com/2020/09/29/election-money-coronavirus-impact-biden-trump.html.
16. Yihui Xie (2020). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.30.
17. Yihui Xie (2015) Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963
18. Yihui Xie (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch and Roger D. Peng, editors, Implementing Reproducible Computational Research. Chapman and Hall/CRC. ISBN 978-1466561595