RIETI Discussion Paper Series 19-E-025

Hearing the Voice of the Future: Trump vs Clinton

NARITA, Yusuke
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 19-E-025
March 2019

Hearing the Voice of the Future:

Trump vs Clinton*

Yusuke NARITA†
Yale University
Research Institute of Economy, Trade and Industry

Abstract

Many countries face growing concerns that population aging may make voting and policy-making myopic. This concern begs for electoral reform to better reflect voices of the youth, such as weighting votes by voters' life expectancy. This paper predicts the effect of the counterfactual electoral reform on the 2016 U.S. presidential election. Using the American National Election Studies (ANES) data, I find that Hillary Clinton would have won the election if votes were weighted by life expectancy. I also discuss limitations due to data issues.

Keywords: Intergenerational conflicts, aging, U.S. presidential election

*The author is grateful for helpful comments and suggestions by Discussion Paper seminar participants at RIETI. I received outstanding research assistance from Katherine Kwok.
†Yale University, Department of Economics and Cowles Foundation. Email: yusuke:narita@yale:edu. URL: www:yusuke-narita:com. Address: 37 Hillhouse Avenue, New Haven, CT 06511

1 Introduction

Intergenerational conflicts are a long-standing regularity in politics. For example, turnout behavior, party identification, and policy preferences (e.g. liberal vs conservative) vary across generations.[1] At the same time, population aging is a pressing issue in the developed world. The two facts raise the concern that in the aging developed countries, voting and policy-making may become biased toward the elderly. This concern induces policy and media discussions about electoral reforms to better reflect the youth's voice. Such electoral reforms can take many possible forms:[2]

- Giving proxy votes to parents of minor children (Demeny, 1986)[3]
- Creating generational electoral districts that accommodate only particular generations (Ihori and Doi, 1998)
- Weighting votes by voters' life expectancy[4]

This paper studies the effects of these intergenerational electoral reforms on electoral outcomes. Specifically, I focus on weighting-votes-by-life-expectancy and study its counterfactual effect on the 2016 U.S. presidential election.[5]

My analysis proceeds as follows. Imagine the 2016 presidential election weighted votes by voters' life expectancy. For each state, I simulate the life-expectancy-weighted popular vote shares of real candidates, especially Hillary Clinton and Donald Trump, as follows:

$$\begin{aligned}&\text{counterfactual weighted \# popular votes for each candidate}\\ =&\textstyle\sum_a(\text{real \# popular votes among voters of age } a) \times (\text{life expectancy of age } a),\end{aligned}$$

where the real number of popular votes among voters of age a comes from the American National Election Studies (ANES) data. The life expectancy of age a is based on the "United States Life

[1] For such evidence, see Wolfinger and Rosenstone (1980); Leighley and Nagler (2013); Pew Research Center (2018). There is no shortage of stories about intergenerational conflicts in the media. To name a few for different continents, see "In 20 years, British politics went from being about class to being about age" *Washington Post* (June 14, 2017), "India's New Voters: We are connected" *Economist* (April 8, 2014), "Brazil's angry millennials are forming their own Tea Party and Occupy movements" *Washington Post* (March 4, 2018), and "Better off than their parents: Why Russia's youth are backing Putin" *Wall Street Journal* (March 17, 2018). The consequences of intergenerational conflicts are also a subject of many studies (Alesina and Tabellini, 1988; Tabellini, 1991; Poterba, 1998; Bassetto and Sargent, 2006; Song et al., 2012; Halac and Yared, 2014; Bisin et al., 2015).

[2] A common policy response is to try to increase young voter turnout. However, in countries where a large proportion of the electorate is the elderly, increasing young voter turnout may not sufficiently increase the youth's influence. More radical electoral reforms may be a palatable response in such a case.

[3] Phillips, Leigh, "Hungarian mothers may get extra votes for their children in election." *Guardian*, April 2011. `https://www.theguardian.com/world/2011/apr/17/hungary-mothers-get-extra-votes`

[4] Založnik, Maja, "Here's what would have happened if Brexit was weighted by age." *Independent*, July 2016. `https://www.independent.co.uk/news/uk/here-s-what-would-have-happened-if-brexit-vote-was-weighted-by-age-a7120536.html`

[5] Kamijo et al. (2015) provide a laboratory experiment on the effect of giving proxy votes to parents of minor children.

Tables, 2014," published by the U.S. Department of Health and Human Service's *National Vital Statistics Reports*. Aggregating these state-level weighted votes predicts a president under hypothetical generational vote weighting.[6]

This counterfactual simulation suggests that Hillary Clinton would have won the 2016 presidential election if votes were weighted by life expectancy. Clinton's national electoral college vote share would have been over 63% (336 votes) under generational vote weighting, as opposed to the real 42.7% (227 votes). Given that the ANES data are collected from survey samples of the entire voter population, I also quantify statistical confidence in my result: Clinton's counterfactual victory is statistically significant at the 95% level.

A few caveats are in order. First, my analysis assumes that voters' locations, turnout, and voting behavior do not change in response to generational vote weighting. Perhaps more importantly, the ANES data are of limited quantity and quality, as pointed out by recent political science studies (Enamorado et al., 2018; Enamorado and Imai, 2018). To investigate this data issue, again for the 2016 presidential election, I compare the ANES data's prediction for each state-level electoral outcome with its real outcome. The ANES data turn out to correctly predict the winning party in most states (40 out of 50 states + Washington D.C.), but it does not very closely capture the exact vote shares. This data-quality concern motivates me to sketch a plan to improve the analysis using a larger and higher quality proprietary data set.

2 Data

My analysis requires three types of data. First, I need data on each individual voter's turnout, vote choice, age, and state in which the voter is registered to vote. Second, I need data on the life expectancy of U.S. citizens at different ages, in order to construct the weights for the generational vote weighting scheme. I use these pieces of information to construct counterfactual voting outcomes. Finally, I need data on actual election outcomes for evaluating the quality of the above voter data. Below, I describe the data sets I use.

Vote Choice and Age Data: I use self-reported voting data from the American National Election Studies (ANES) Time Series Study. ANES survey individuals who are U.S. citizens aged 18 or above. Their surveys have been conducted before and after presidential elections since 1948, and after most non-presidential elections since 1956. The interviews include questions on partisanship, election candidates and incumbents, government performance, political participation, media use, ideologies, values, and support for specific issues. The survey also collects personal and demographic data. ANES select a sample of 1200 to 2500 individuals for each survey. The sampling process includes

[6]This aggregation of state-level weighted votes into a counterfactual president assumes the electoral college keeps using the same voting system. That is, in Maine and Nebraska, one electoral college vote is allocated to the plurality candidate in each congressional district, and two electoral college votes are allocated to the state-wide plurality candidate. The remaining 48 states use the winner-take-all system, where all electoral college votes go to the state-wide plurality candidate. As the ANES data do not provide information on the congressional districts in which individuals are registered to vote, my analysis excludes the district-level electoral votes for Maine (2 votes) and Nebraska (3 votes).

oversampling racial/ethnic minorities, geographically-stratified cluster sampling, and randomly selecting a member of each household. I use self-reported information on voter registration, turnout, and vote choice from the 2016 ANES data. I limit the data to individuals 18 years old or above, who report that they registered to vote and voted in the 2016 election. Since age is a key variable in my analysis, I also focus on individuals who correctly report their age. The resulting sample size is 4271, with 1181 individuals surveyed through face-to-face interviews and 3090 through online interviews.

Life Expectancy Data: To construct generational vote weights, I use the life expectancy data from the "United States Life Tables, 2014" published by the U.S. Department of Health and Human Service's *National Vital Statistics Reports* in August 2017. The report includes life expectancy estimates based on 2014 census and Medicare data for U.S. citizens at different ages (Aris et al., 2017).

Actual Election Outcome Data: For evaluating the quality of the ANES data, I use the Congressional Quarterly Voting and Elections Collection (CQ) as a benchmark for actual election outcomes. The CQ Voting and Elections Collection is a database that collects data on U.S. elections, parties, and campaigns.

3 Method

To construct the counterfactual 2016 presidential election under generational vote weighting, I weight votes by votes' life expectancy as follows. I first calculate each party's vote share by voter age within each state or district. Let

$$y_{(a,p)j} = \begin{cases} 1 & \text{if } j\text{-th individual is of age } a \text{ and votes for party } p \\ 0 & \text{otherwise,} \end{cases}$$

where j indexes individuals (survey respondents) in the data.[7] The vote count for a given party p within a given age group a for state or district s is

$$\hat{N}_{asp} = \sum_j I_s(j) w_j y_{(s,p)j},$$

where $I_s(j)$ is 1 if the j-th individual in state or district s, and 0 if not. w_j is the j-th individual's *sampling weight*, which ANES and the STATA package "svy" provide for making the data more representative of the national population (DeBell, 2010). (Recall that the ANES study sample was not a simple, random sample from the U.S. population.) As a result, the within-state or within-

[7]ANES collects self-reported information on vote choices in two phases, once prior to Election Day and once after Election Day. The pre-election vote choice captures individuals who submitted their ballot through early voting or absentee voting. If individuals indicated their vote choices in the pre-election survey, they would not be asked about their vote choices in the post-election survey. I construct a variable that indicates each respondent's vote choice, either from the pre-election or post-election survey. If an individual reported that they did not vote in both pre- and post-election surveys, I exclude the individual from the analysis.

district vote share for each party and age group is given by $\hat{X}_{asp} = \dfrac{\hat{N}_{asp}}{\hat{N}_s}$, where $\hat{N}_s = \sum_p \sum_a \hat{N}_{asp}$. To find the counterfactual vote share of each party in each state or district, I multiply $\hat{X}_{asp}$ by the age-specific weights w_a, the expected life years for an average American citizen at age a. I sum the weighted vote shares across ages, and then normalize it to obtain

$$\hat{X}_{sp}^{CF} = \frac{\sum\limits_a w_a \hat{X}_{asp}}{\sum\limits_p \sum\limits_a w_a \hat{X}_{asp}}.$$

I use this formula to find the counterfactual for each party's vote share in each state or district. For major party vote shares, I only include Democrats and Republicans. For the all party vote shares, I include independent candidates and other parties under one category of "Other" since each other party gets only a small number of votes. I also calculate the standard error of each counterfactual vote share estimate, as detailed in Appendix A.1.

I determine the counterfactual president under generational vote weighting as follows. The electoral college determines the final outcome for U.S. presidential elections. There are a total of 538 electoral college votes. For each state, the number of Senate and House of Representative delegates corresponds to the number of electors. In addition, Washington D.C. or the District of Columbia has 3 electoral college votes. Electoral college votes are awarded based on the results of the general elections, in which citizens cast votes for the presidential candidate and vice presidential candidate of their choice, in the state where they are registered as voters. In 48 states and Washington D.C., the electoral votes are awarded to candidates who receive the most popular votes, i.e., plurality. In Maine and Nebraska, one electoral college vote is allocated to the plurality candidate in each district, and then two electoral college votes are allocated to the state-wide plurality candidate. The presidential and vice presidential candidates must have 270 votes out of 538 votes to win the election.

My counterfactual simulation follows the same procedure in allocating electoral college votes, except the following two assumptions. First, I exclude the district-level electoral votes for Maine (2 votes) and Nebraska (3 votes), as the ANES does not provide information on individuals' district of voter registration. Second, I assume that there are no faithless electors. A *faithless elector* is a member of the electoral college who does not vote for the plurality winner in their given state. Typically, electors follow the "winner-takes-all" rule and all electoral college votes go to the candidate who wins the plurality in a state. In the 2016 election, however, 7 electors did not follow the rules.[8] Importantly, the 5 excluded votes and 7 potential faithless electoral votes fall far short of changing the final counterfactual president below.

[8] Schmidt, Kiersten and Wilson, Andrews, "A historic number of electors defected, and most were supposed to vote for Clinton." *New York Times*, December 2016. `https://nyti.ms/2jWW5CY`

4 Results: Counterfactual President

I find that Hilary Clinton would win the 2016 presidential election using generationally weighted votes. As summarized in Figure 1, Hilary Clinton would receive 63% of the electoral college votes (336 votes), while Donald Trump would receive 36% (194 votes). Given that the ANES data are collected from survey samples of the entire voter population, I also quantify statistical confidence in my result: Clinton's counterfactual victory is statistically significant at the 95% level, when I estimate the standard error of the final distribution of electoral college votes, following the method in Appendix A.2.

Heterogeneity across States. I provide further details of my prediction by visualizing the electoral outcomes by state and congressional district in Figures 2 and 3. Figure 3 describes the states in which the winning parties would be changed by generational vote weighting. Several of the key Rust-Belt states, such as Michigan, Ohio, Pennsylvania, and Wisconsin, are predicted to be flipped to Hilary Clinton by generational vote weighting. At the same time, there are also states that are predicted to be flipped to Donald Trump by generational vote weighting. Minnesota and Virginia are such examples.

I further investigate this inter-state heterogeneity in Figure 4. Here, I plot the differences between the counterfactual and actual vote shares of Democrats and Republicans. As shown by its horizontal bars, on average across states, generational vote weighting results in an increase in the vote shares of Democrats. However, there are large variations in the magnitude of change across states, a pattern consistent with Figure 3. More detailed state-level statistics are available in Tables 3 and 4 in the appendix.

Heterogeneity across Generations. Generational vote weighting has such big impacts because of generational differences in voter preferences. I highlight the generational differences in Table 1, which summarizes the vote shares of Democrats and Republicans for each generation. The Republican vote share increases as age-level increases. The trend identified here is similar to the results of existing studies (Pew Research Center, 2018).

5 Limitations

I should acknowledge a few limitations of the above analysis. First, I assume that voters' locations, turnout, and voting behavior do not change in response to generational vote weighting. Perhaps more importantly, the ANES data are of limited quantity and quality due to the following factors.

Misreporting: A problem with the ANES data is that individuals self-report whether they registered to vote, whether they voted, and their vote choice. Enamorado et al. (2018) and Enamorado and Imai (2018) compare the ANES voting data against national voter registration data supplied by L2, a non-partisan voter data collection firm. They found that 20 percent of ANES survey respondents who indicated that they voted in the elections did not actually vote. Many previous studies have explored why individuals would be likely to misreport when answering survey questions on voting. One prominent theory is social desirability bias. Another explanation is non-response bias, the idea that those who respond to the survey differ in a meaningful way from those who do not (Bernstein et al., 2001).

Sample Size: Another problem with the ANES data is sample size. While some states have hundreds of observations, other states have only a handful of observations. I report the sample size for each state of party registration in Table 2. The largest sample size is 302 (California) and the smallest sample size is 2 (Alaska). When I restrict the sample to individuals who reportedly voted in the 2016 elections, the sample size becomes even smaller for some states. As a result of the small sample size, some of the ANES vote shares are prone to bias.

Validation of the ANES Data

To quantify how serious these data issues are, I gauge the accuracy of the ANES data by comparing the ANES data's predicted election outcomes against the CQ election outcome data. For each state, I use the ANES data to calculate the vote shares of parties as well as their confidence intervals, without any generational vote weighting. I find that the ANES vote shares correctly capture the winning parties in 40 out of 50 states and the District of Columbia (about 78% accuracy). However, ANES does not capture the exact vote shares very accurately. I show this point in Figure 5, which reports the plots of the ANES and actual vote shares for Democrat and Republican candidates in the 2016 presidential election. The correlation between the ANES proportions and actual proportions is modest. The root-mean-square error for Democrat and Republican vote shares are 14.15% and 14.25%, respectively.

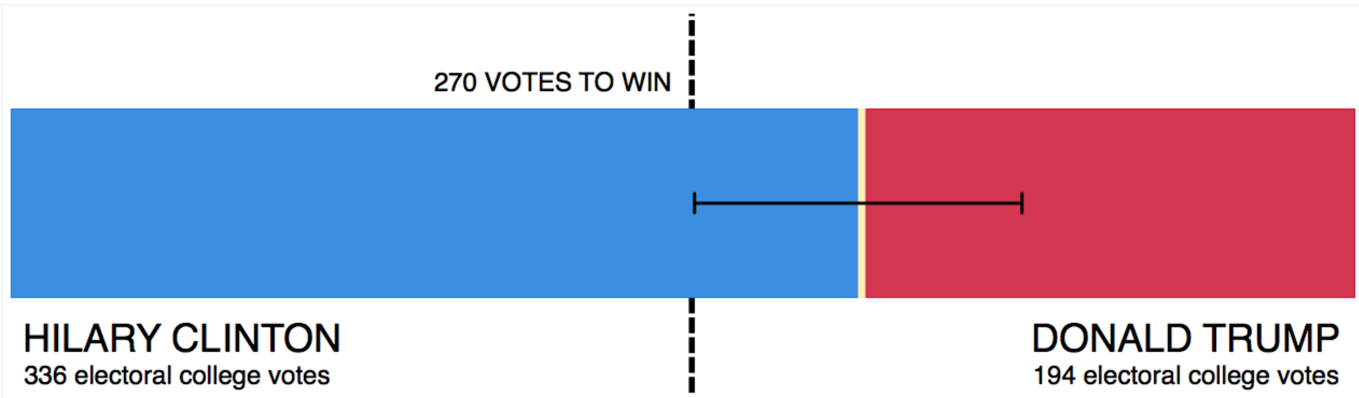
6 Path Forward

I find that generational vote weighting could change the result of a critical election like the 2016 U.S. presidential election. This analysis leads to a variety of avenues for future work. As already noted, it is important to obtain a more credible prediction about the counterfactual presidential election by using a larger and higher quality data set. Such potential data sets include Catalist data, L2 data, and Cooperative Congressional Election Study (CCES) data. Catalist and L2 both provide national voter files of hundreds of millions of individuals and contain data on voter registration, turnout, age, and predicted partisanship. CCES surveys the opinions of over 50,000 individuals across the U.S. during election years, through YouGov (a public opinion data company).

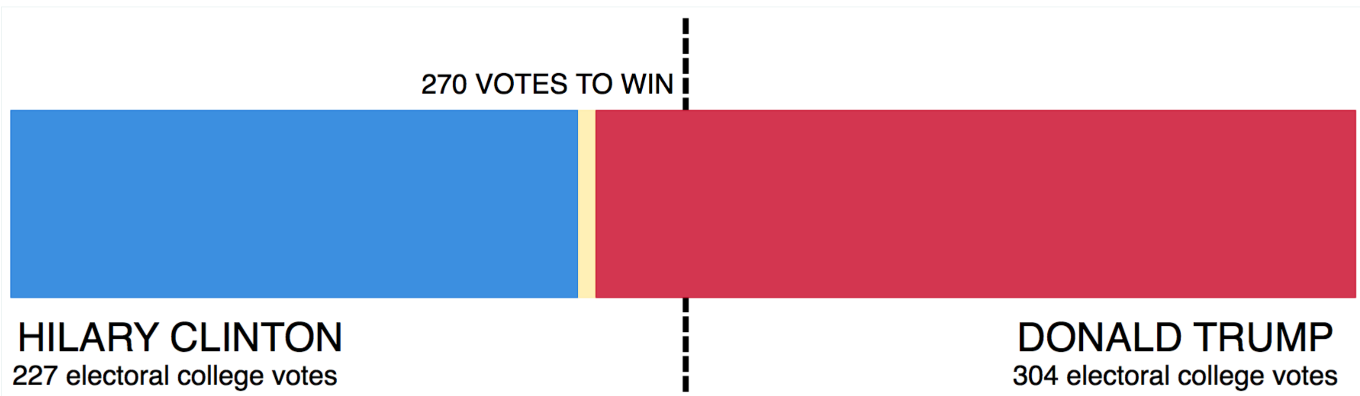
More conceptually, I plan to measure the effects of giving more votes to the young on a wider range of outcomes. Especially intriguing are policy outcomes and the welfare of different generations. I also plan to explore other weighting methods, such as generational electoral districts and giving proxy votes to parents of minor children. I leave these challenging directions to future work.

Figure 1: Counterfactual and Actual Electoral College Voting Outcomes

(a) Counterfactual Outcome

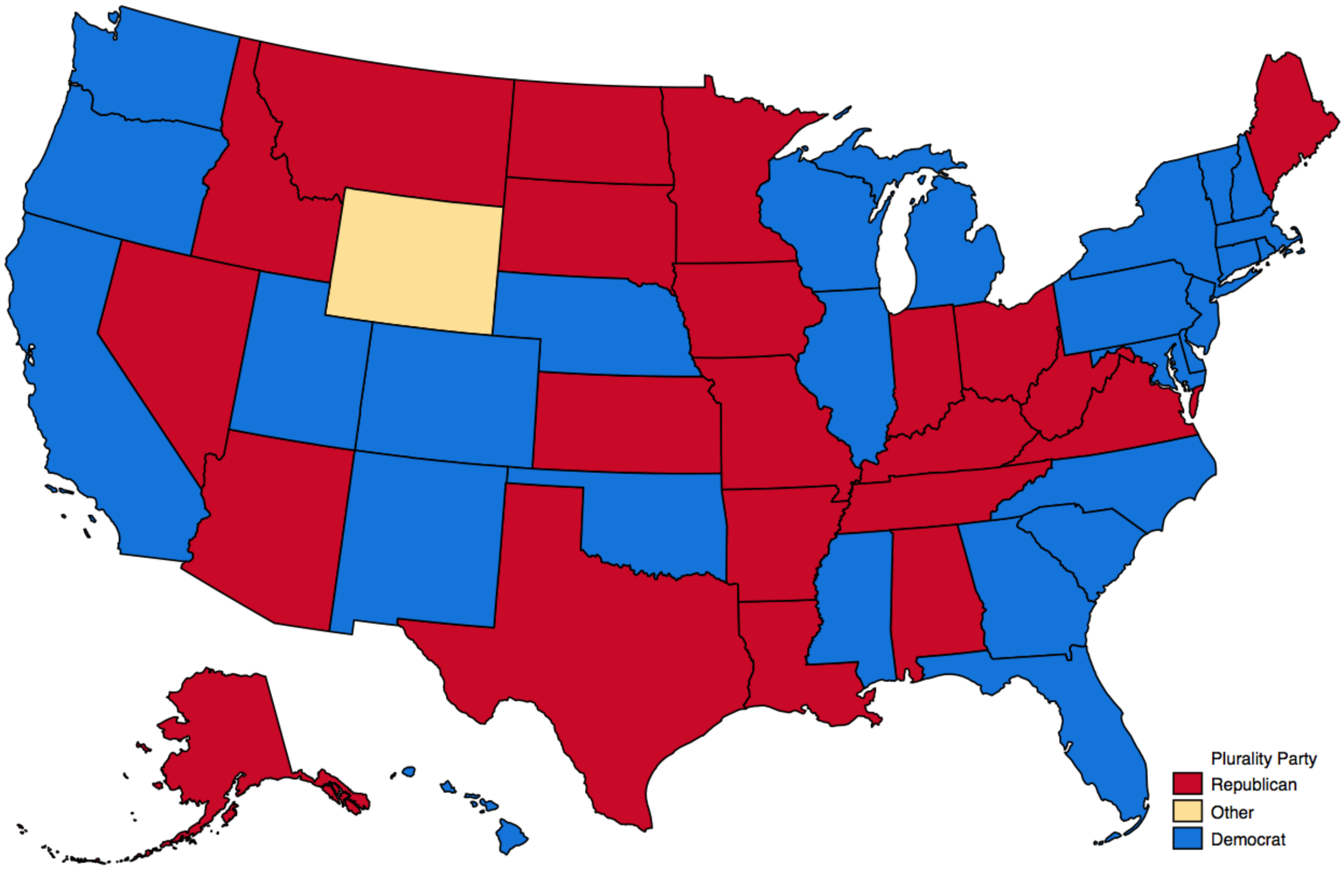


(b) Actual Outcome



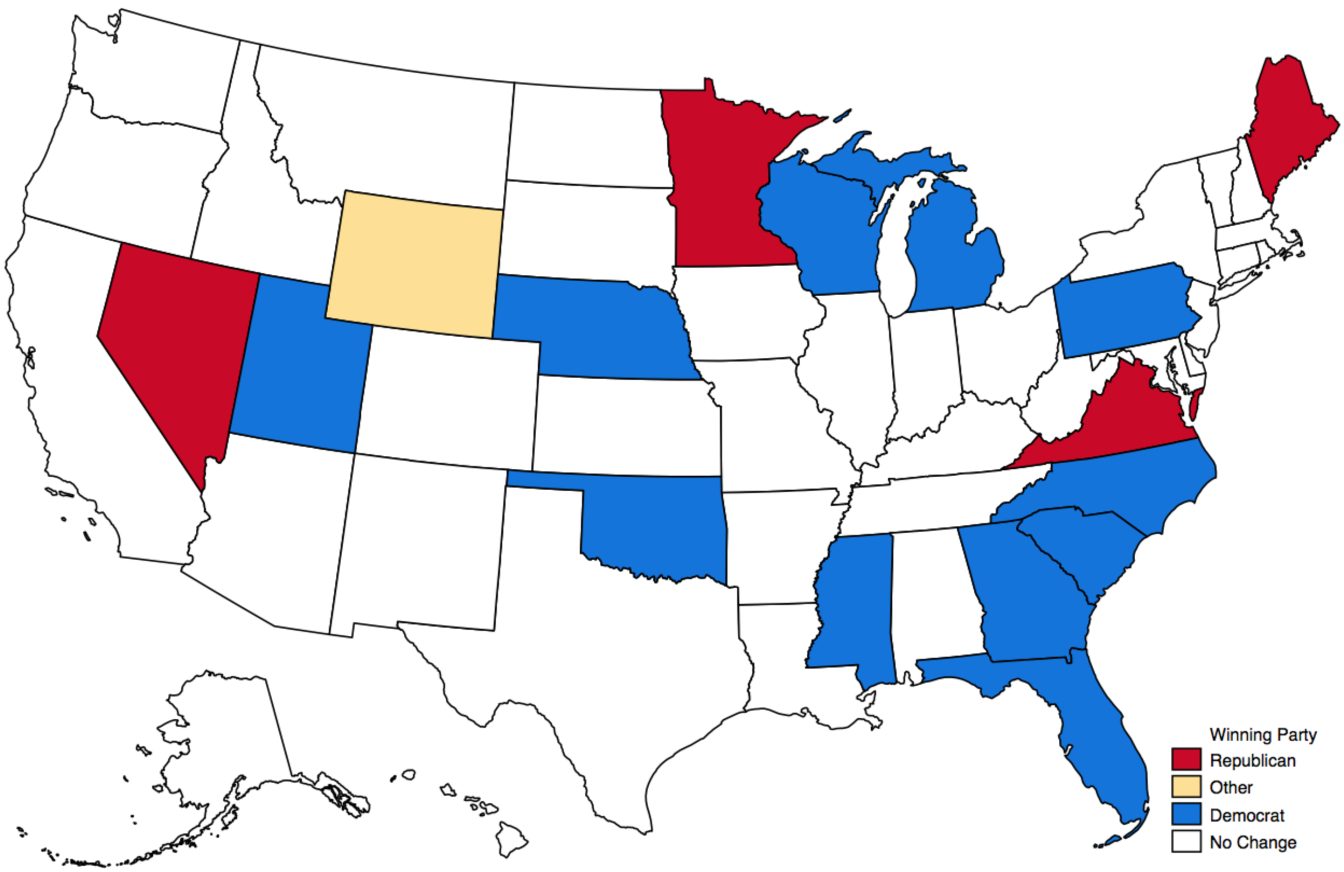
Notes: The figures show the counterfactual and actual distributions of electoral college votes for the 2016 presidential election, between Democrat party candidate Hilary Clinton and Republican party candidate Donald Trump. The yellow section represents the 3 votes for party candidates other than the major parties. I allocate electoral college votes to each candidate based on the "winner-takes-all" rule for 48 states and Washington D.C. I exclude the district-level electoral votes for Maine (2 votes) and Nebraska (3 votes), as I cannot estimate the plurality winner at the granularity of congressional districts using the ANES data. I also assume there are no "faithless electors," who do not vote for the candidate they pledged to vote for. I calculate the 95% confidence interval around the counterfactual votes for Clinton based on the standard error as calculated in Appendix A.2.

Figure 2: Counterfactual Plurality Party by State under Generational Vote Weighting



Notes: The map shows the counterfactual 2016 presidential election outcomes for each state or district with generational vote weighting. The "Others" category includes any Independent or other party candidate choice.

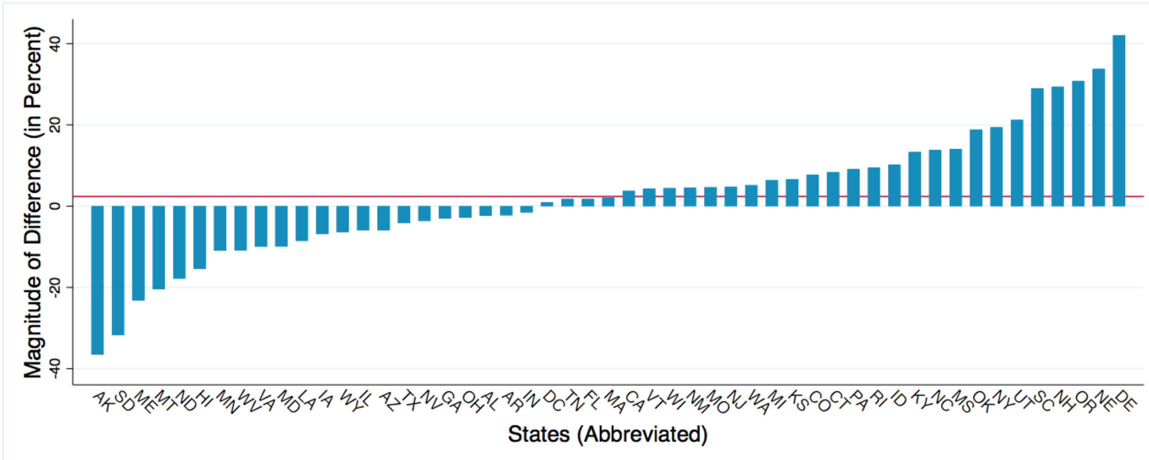
Figure 3: States where the Winner is “Flipped” by Generational Vote Weighting



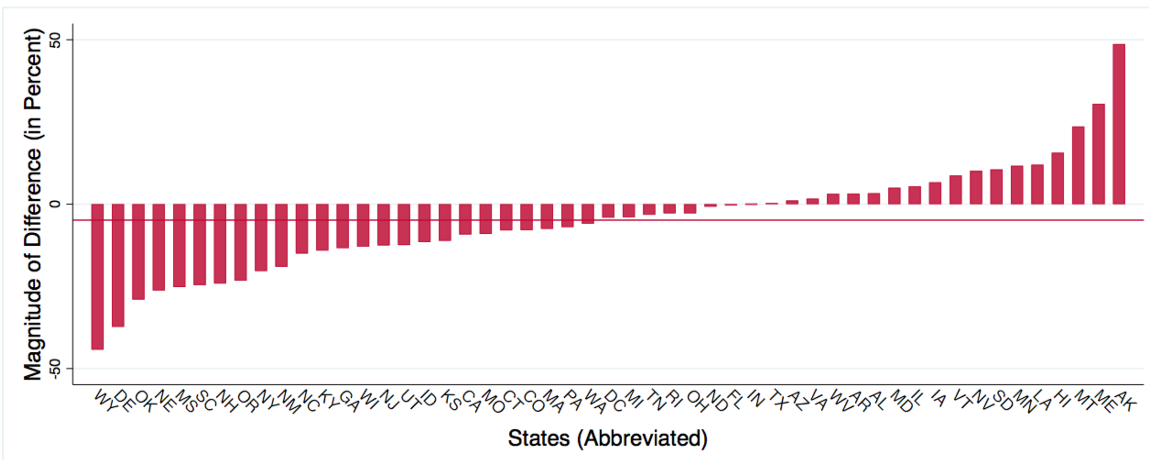
Notes: The map shows the counterfactual 2016 presidential election outcomes, highlighting only the states in which the plurality party is changed by generational vote weighting. There are 11 states that flipped from Republican to Democrat plurality, including: Utah, Nebraska, Oklahoma, Wisconsin, Michigan, Mississippi, Georgia, Florida, South Carolina, North Carolina, and Pennsylvania. There are 4 states that flipped to Republican from Democrat plurality, including: Nevada, Minnesota, Virginia, and Maine. The “Others” category includes any Independent or other party candidate choice.

Figure 4: Difference between Weighted and Actual Vote Percentage by State

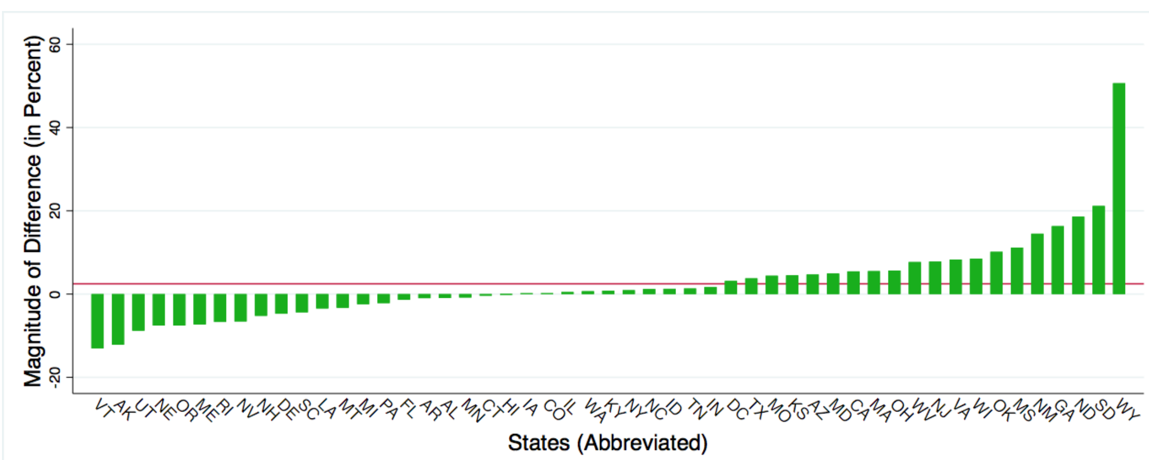
(a) Democrats



(b) Republicans



(c) Others



Notes: The figure shows the percentage point difference between the counterfactual (generationally weighted) and actual vote shares for Democrats, Republicans, and others. The "Others" category includes any Independent or other party candidate choice. The horizontal red line indicates the average difference across all states.

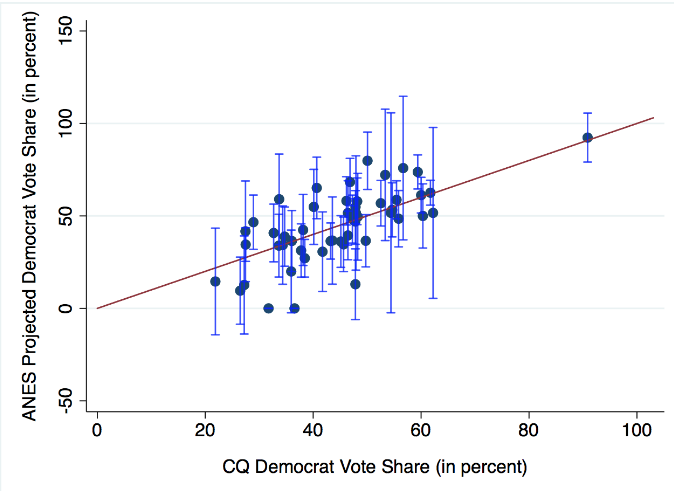
Table 1: Voting Behavior by Generations

Generations	Democrats	Republicans	Others
18-29	49.13	43.75	7.12
30-39	47.93	45.62	6.45
40-49	46.06	48.79	5.15
50-59	45.39	49.95	4.66
60-69	45.27	51.91	2.81
70-79	38.66	58.66	2.68
80-90	36.9	62.62	.48

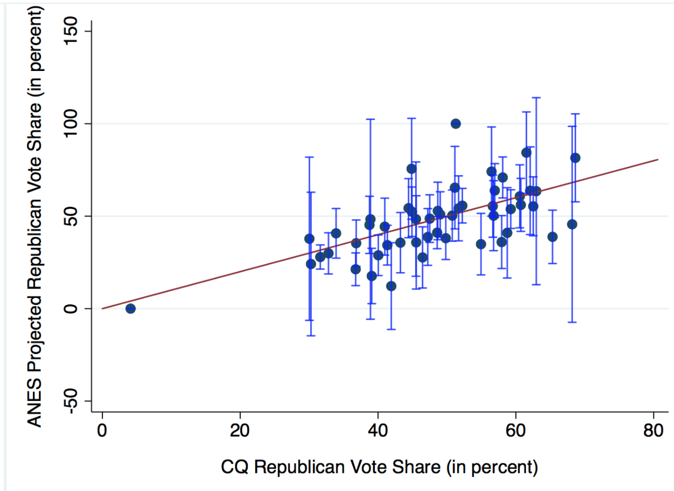
Notes: The table shows the vote proportion for each party by age group. The "Others" category includes any Independent or other party candidate choice. I estimate the vote proportions using vote counts from the ANES 2016 Time Series data set. I drop respondents who did not register to vote, cast a ballot, or correctly report their age. I also drop those who had inappropriate or missing answers for vote choice.

Figure 5: ANES Vote Shares vs. Actual Vote Shares

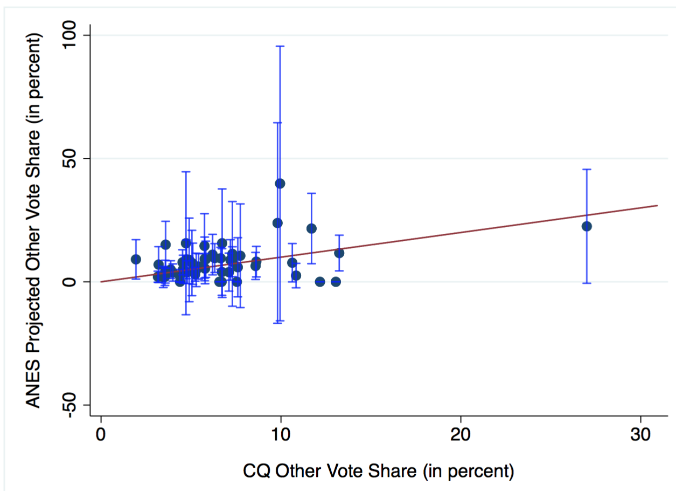
(a) Democrats



(b) Republicans



(c) Others



Notes: The data points represent vote shares for all parties, including Democrats, Republicans, and Others (all Independent and other parties). Each vertical bar represents the 95% confidence interval for an estimated vote share, as detailed in Appendix A.1. Some estimates (data points) are missing the 95% confidence interval, as the ANES counted no votes for a party in certain states. All 50 states and Washington D.C. are represented in the three figures above. I exclude any ANES individual who said they did not register to vote, did not vote in the presidential election, or did not correctly report their age.

Table 2: Sample Sizes for Each State of Party Registration

State of Party Registration	No.	%	State of Party Registration	No.	%
Alabama	36	1.1	Nebraska	17	0.5
Alaska	2	0.1	Nevada	23	0.7
Arizona	76	2.3	New Hampshire	28	0.8
Arkansas	38	1.1	New Jersey	89	2.7
California	327	9.8	New Mexico	32	1.0
Colorado	75	2.2	New York	131	3.9
Connecticut	46	1.4	North Carolina	135	4.0
Delaware	8	0.2	North Dakota	5	0.1
Florida	166	5.0	Ohio	135	4.0
Georgia	95	2.8	Oklahoma	44	1.3
Hawaii	8	0.2	Oregon	34	1.0
Idaho	46	1.4	Pennsylvania	143	4.3
Illinois	161	4.8	Rhode Island	6	0.2
Indiana	75	2.2	South Carolina	50	1.5
Iowa	26	0.8	South Dakota	9	0.3
Kansas	68	2.0	Tennessee	119	3.6
Kentucky	52	1.6	Texas	235	7.0
Louisiana	46	1.4	Utah	21	0.6
Maine	10	0.3	Vermont	9	0.3
Maryland	86	2.6	Virginia	72	2.2
Massachusetts	94	2.8	Washington	72	2.2
Michigan	111	3.3	Washington DC	22	0.7
Minnesota	68	2.0	West Virginia	14	0.4
Mississippi	37	1.1	Wisconsin	73	2.2
Missouri	51	1.5	Wyoming	4	0.1
Montana	10	0.3	Total	3340	100.0

Notes: The table displays the sample size, or the number of voters in each state, after I restrict the ANES 2016 Time Series sample to individuals who said they registered to vote. Here, I do not apply other sample restrictions that limit the sample to individuals who voted in specific elections (presidential, House, Senate, or governor) and individuals who correctly reported their age. The mean sample size is 62.27, with a standard deviation of 58.95.

References

Alesina, Alberto and Guido Tabellini, "Voting on the budget deficit," November 1988. National Bureau of Economic Research Working Paper 2759.

Aris, Elizabeth, Melonie P. Heron, and Jiaquan Xu, "United States life tables, 2014," Technical Report 4, National Center for Health Statistics (U.S.) 2017.

Bassetto, Marco and Thomas Sargent, "Politics and efficiency of separating capital and ordinary government budgets," *Quarterly Journal of Economics*, 2006, *121* (4), 1167–1210.

Bernstein, Robert, Anita Chadha, and Robert Montjoy, "Overreporting voting: Why it happens and why it matters.," *Public Opinion Quarterly*, 2001, *65* (1), 22–44.

Bisin, Alberto, Alessandro Lizzeri, and Leeat Yariv, "Government policy with time inconsistent voters," *American Economic Review*, 2015, *105* (6), 1711–37.

DeBell, Matthew, "How to analyze ANES survey data," Technical Report nes012492, ANES Technical Report Series, Palo Alto, CA, and Ann Arbor, MI: Stanford University and the University of Michigan 2010.

Demeny, Paul, "Pronatalist policies in low-fertility countries: Patterns, performance, and prospects," *Population and Development Review*, 1986, *12*, 335–358.

Enamorado, Ted and Kosuke Imai, "Validating self-reported turnout by linking public opinion surveys with administrative records," 2018.

_ , Benjamin Fifield, and Kosuke Imai, "Using a probabilistic model to assist merging of large-scale administrative records," *American Political Science Review*, 2018, pp. 1–19.

Halac, Marina and Pierre Yared, "Fiscal rules and discretion under persistent shocks," *Econometrica*, 2014, *82* (5), 1557–1614.

Ihori, Toshihiro and Takero Doi, *Economic analysis of Japanese politics*, BokutakuSha, 1998.

Kamijo, Yoshio, Yoichi Hizen, and Tatsuyoshi Saijo, "Hearing the voice of future generations: A laboratory experiment of "Demeny voting"," 2015. Kochi University of Technology, School of Economics and Management Working Paper SDES-2018-2.

Leighley, Jan E and Jonathan Nagler, *Who votes now? Demographics, issues, inequality, and turnout in the United States*, Princeton University Press, 2013.

Pew Research Center, "The generation gap in American politics.," Online March 2018.

Poterba, James M, "Demographic change, intergenerational linkages, and public education," *American Economic Review*, 1998, *88* (2), 315–320.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Rotten parents and disciplined children: A politico-economic theory of public expenditure and debt," *Econometrica*, 2012, *80* (6), 2785–2803.

Tabellini, Guido, "The politics of intergenerational redistribution," *Journal of Political Economy*, 1991, *99* (2), 335–357.

Wolfinger, Raymond E and Steven J Rosenstone, *Who votes?*, Vol. 22, Yale University Press, 1980.

A Appendix

A.1 Estimating Standard Errors of State-level Vote Shares

To understand how confident I can be of the difference between the actual and counterfactual election outcomes, I estimate the standard errors for the state-level vote shares. For each counterfactual vote share for party p and voter age a within state or district s, recall that I denote the within-state or within-district vote share for each party and age group by $\hat{X}_{asp} = \frac{\hat{N}_{asp}}{\hat{N}_s}$. I estimate the variance of $\hat{X}_{asp}$ by

$$Var(\hat{X}_{asp}) = \frac{1}{\hat{N}_s^2}\Big\{\hat{Var}(\hat{N}_{asp}) - 2\hat{X}_{asp}\hat{Cov}(\hat{N}_s, \hat{N}_{asp}) + \hat{X}_{asp}^2\hat{Var}(\hat{N}_s)\Big\}. \tag{1}$$

Equation (1) follows the formula for the variance of sample ratio discussed in Rice (2007). I describe the steps to calculate $\hat{N}_{asp}$, $\hat{N}_s$, and $\hat{X}_{asp}$ in Section 3.

Below, I describe how to calculate the remaining components of equation (1): $\hat{Var}(\hat{N}_{asp})$, $\hat{Cov}(\hat{N}_s, \hat{N}_{asp})$, and $\hat{Var}(\hat{N}_s)$. I do so by incorporating ANES's nonrandom, stratified sampling design into consideration. The ANES organizers construct strata and their sampling weights within the sample that reflect their original sampling design. Specifically, the ANES sample consists of 132 strata, or independent subsamples. Each stratum contains several primary sampling units (PSUs), which are clusters of individuals (survey respondents). Finally, ANES produces respondent-specific weights to reflect the sampling probability of each individual, based on their geographical location and demographics. For more details on how ANES constructs the strata and PSUs, see DeBell et al. (2018).

I calculate $\hat{Var}(\hat{N}_{asp})$, the variance of the vote count for party p and age a within state or district s as follows. For each stratum $h = 1, \cdots, L$ and each PSU $i = 1, \cdots, n_h$ in stratum h,

$$\hat{Var}(\hat{N}_{asp}) = \sum_{h=1}^{L} \frac{n_h}{n_h - 1}\Big\{\sum_{i=1}^{n_h} {y_{(a,p)shi}}^2 - \frac{(y_{(a,p)sh})^2}{n_h}\Big\}, \tag{2}$$

where $y_{(a,p)shi}$ is the vote count for party p and age a in each PSU i of stratum h:

$$y_{(a,p)shi} = \sum_{j=1}^{m_{hi}} w_j \times I_{s,hij} \times y_{(a,p)hij},$$

where each respondent j is indexed by $j = 1, \cdots, m_{hi}$. $y_{(a,p)hij}$ is 1 if respondent j is age a and voted for party p, and 0 otherwise. The value w_j is the ANES sampling probability weight for each respondent j. $I_{s,hij}$ is 1 if the j-th respondent in state or district s and 0 if not. I calculate $y_{(a,p)sh}$, the total vote count for a and p across all PSUs of stratum h by

$$y_{(a,p)sh} = \sum_{i=1}^{n_h} y_{(a,p)shi}.$$

Intuitively, equation (2) finds the variance of vote counts for each party p in age a across PSUs contained in the same stratum, and then sums the within-stratum variance across all strata. For a more in-depth discussion on variance estimation for survey subpopulation totals, see West et al. (2008).

My estimation of $\hat{Var}(\hat{N}_s)$ is similar to equation (2). The key difference is, I replace $y_{(a,p)shi}$ in equation (2) with y_{shi}, the vote count for each PSU i of stratum h across parties and ages:

$$y_{shi} = \sum_{j=1}^{m_{hi}} w_j \times I_{s,hij} \times y_{shij},$$

where y_{hij} is the vote count for each individual j in PSU i and stratum h. All respondents who said they voted have $y_{hij} = 1$, regardless of their age or vote choice. I also replace $y_{(a,p)sh}$ in equation (2) with $y_{sh} = \sum_{i=1}^{n_h} y_{shi}$, the vote count for each stratum h across parties and ages.

Finally, I estimate $\hat{Cov}(\hat{N}_s, \hat{N}_{asp})$ as follows. For each stratum $h = 1, \cdots, L$ and each PSU $i = 1, \cdots, n_h$,

$$\hat{Cov}(\hat{N}_s, \hat{N}_{asp}) = \sum_{h=1}^{L} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (y_{(a,p)shi} - \bar{y}_{(a,p)sh})(y_{shi} - \bar{y}_{sh}), \tag{3}$$

where $\bar{y}_{(a,p)sh}$ is the average vote count for party p and age a across all PSUs in stratum h:

$$\bar{y}_{(a,p)sh} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{(a,p)shi}.$$

$\bar{y}_{sh}$ is the average vote count across all PSUs in stratum h:

$$\bar{y}_{sh} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{shi}.$$

Intuitively, equation (3) finds the covariance between $y_{(a,p)shi}$ and y_{shi} across all PSUs $i = 1, \cdots, n_h$ in a given stratum h. It then sums the stratum-specific covariances across all strata. For further details on estimating covariance for a stratified sample, see Heeringa et al. (2010).

I use $\hat{Var}(\hat{X}_{asp})$ to calculate the standard error of state-level vote shares $\hat{X}_{sp}^{CF} = \dfrac{\sum\limits_a w_a \hat{X}_{asp}}{\sum\limits_p \sum\limits_a w_a \hat{X}_{asp}}$.

To do that, I multiply both the vote share estimate and standard error values by age-specific weights, and sum across age-levels:

$$SE(\hat{X}_{sp}^{CF}) = \frac{1}{\sum\limits_p \hat{X}_{sp}^{CF}} \times \sqrt{\sum_a {w_a}^2 \times \hat{Var}(\hat{X}_{asp})}.$$

I assume that if there is no vote share estimate for a given age (when the ANES sample size is 0 for a given age), the standard error value is also 0. I also assume independence between the vote shares for a given party across different ages. To relax this assumption in the aforementioned step, I can calculate $SE(\hat{X}^{CF}_{sp})$ by

$$SE(\hat{X}^{CF}_{sp}) = \frac{1}{\sum_p \hat{X}^{CF}_{sp}} \times \sqrt{\sum_a w_a^2 \times \hat{Var}(\hat{X}_{asp}) + \sum_a \sum_{a' \neq a} w_a w_{a'} \times \hat{Cov}(\hat{X}_{asp}, \hat{X}_{a'sp})},$$

which follows from this rule: $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$. See Heeringa et al. (2010) for further details on the method.

Testing Significant Differences between Party Vote Shares

In Table 4, I use a one sample t-test to gauge at the significance of the differences between the counterfactual Democrat and Republican vote shares. For Democrats p_1 and Republicans p_2, my null hypothesis is $\hat{X}^{CF}_{sp_1} - \hat{X}^{CF}_{sp_2} = 0$ and my alternative hypothesis is $\hat{X}^{CF}_{sp_1} - \hat{X}^{CF}_{sp_2} \neq 0$. Treating each state as an individual sample, I calculate the standard error of the estimated difference between $\hat{X}^{CF}_{sp_1}$ and $\hat{X}^{CF}_{sp_2}$ by

$$SE(\hat{X}^{CF}_{sp_1} - \hat{X}^{CF}_{sp_2}) = \sqrt{SE(\hat{X}^{CF}_{sp_1})^2 + SE(\hat{X}^{CF}_{sp_2})^2 + 2 \times Cov(\hat{X}^{CF}_{sp_1}, \hat{X}^{CF}_{sp_2})}, \tag{4}$$

where I estimate $Cov(\hat{X}^{CF}_{sp_1}, \hat{X}^{CF}_{sp_2})$ as follows. Note that for $Y_1 = \sum_{j=1}^{p} c_j X_j$ and $Y_2 = \sum_{k=1}^{p} d_k X_k$, I have

$$Cov(Y_1, Y_2) = \sum_{j=1}^{p} \sum_{k=1}^{p} c_j d_k Cov(X_j, X_k).$$

I apply this formula to estimate $Cov(\hat{X}^{CF}_{sp_1}, \hat{X}^{CF}_{sp_2})$. Given $\hat{X}^{CF}_{sp_1} = \sum_a w_a \hat{X}_{asp_1}$ and $\hat{X}^{CF}_{sp_2} = \sum_{a'} w_{a'} \hat{X}_{a'sp_2}$, I calculate the covariance of the two linear combinations using

$$Cov(\hat{X}^{CF}_{sp_1}, \hat{X}^{CF}_{sp_2}) = \sum_a \sum_{a'} w_a w_{a'} Cov(\hat{X}_{asp_1}, \hat{X}_{asp_2}).$$

I follow the same steps to calculate the covariance between counterfactual vote shares when all parties are included (with all non-major parties grouped togther as "Other"). In some election levels in the 2016 ANES Time Series data set, there were no reported votes for the "Other" category. In those states, I assume a covariance value of 0 between the counterfactual "Other" vote share and the counterfactual Democrat/Republican vote share.

I also test if the winning party vote share is significantly different from the runner up party vote share (i.e. received second most votes). For winning party p_1 and runner up party p_2, my null hypothesis is $\hat{X}^{CF}_{sp_1} - \hat{X}^{CF}_{sp_2} = 0$ and my alternative hypothesis is $\hat{X}^{CF}_{sp_1} - \hat{X}^{CF}_{sp_2} \neq 0$. I use equation (4) to calculate the standard error of $\hat{X}^{CF}_{sp_1} - \hat{X}^{CF}_{sp_2}$.

Confidence Intervals

I calculate confidence intervals in Figure 5 as follows. According to the "svy: tabulate twoway" documentation, the confidence intervals are calculated as follows. I first find $f(\hat{X_{sp}})$, the logit transformation of $\hat{X_{sp}}$ by

$$f(\hat{X_{sp}}) = \ln(\frac{\hat{X_{sp}}}{1-\hat{X_{sp}}}).$$

Applying the logit transformation means the values would be contained between 0 and 1. For $\hat{s} = SE(\hat{X_{sp}})$, the standard error estimate is given by

$$\hat{SE}\{f(\hat{X_{sp}})\} = f'(\hat{X_{sp}})\hat{s} = \frac{\hat{s}}{\hat{X_{sp}}(1-\hat{X_{sp}})}.$$

I then can find the $100(1-\alpha)\%$ confidence interval using

$$\ln(\frac{\hat{X_{sp}}}{1-\hat{X_{sp}}}) \pm \frac{t_{1-\alpha/2,v}\hat{s}}{\hat{X_{sp}}(1-\hat{X_{sp}})},$$

where $t_{1-\alpha/2,v}$ is the critical value at the $(1-\alpha/2)$th quantile of the t distribution with v degrees of freedom. Finally, suppose that $y = \ln(\frac{\hat{X_{sp}}}{1-\hat{X_{sp}}}) \pm \frac{t_{1-\alpha/2,v}\hat{s}}{\hat{X_{sp}}(1-\hat{X_{sp}})}$, I use the formula below to find the inverse of the logit transform and the final confidence interval values

$$f^{-1}(y) = \frac{e^y}{1+e^y}.$$

A.2 Estimating Standard Errors of Electoral College Vote Shares

In Figure 1, I assess the reliability of the final counterfactual electoral college outcome by estimating its standard error. For each state s, I have estimated the counterfactual vote shares for Democrats, p_1, Republicans, p_2, and Others, p_3, denoted as $\hat{X}^{CF}_{sp_1}$, $\hat{X}^{CF}_{sp_2}$, and $\hat{X}^{CF}_{sp_3}$, respectively. As detailed in Appendix A.1, I also estimated $\hat{SE}(\hat{X}^{CF}_{sp_1})$, $\hat{SE}(\hat{X}^{CF}_{sp_2})$, and $\hat{SE}(\hat{X}^{CF}_{sp_3})$. I use them to compute the standard error of the final counterfactual election college outcome as follows. For $t = 1, \cdots, 500$ simulations,

(1) **Draw vote shares:** For states with 10 or more observations in the ANES sample, I draw $\hat{X}^{CF}_{sp_1}(t)$ from the normal distribution with mean $\hat{X}^{CF}_{sp_1}$ and standard deviation $\hat{SE}(\hat{X}^{CF}_{sp_1})$. I also draw $\hat{X}^{CF}_{sp_3}(t)$ from the normal distribution with mean $\hat{X}^{CF}_{sp_3}$ and standard deviation $\hat{SE}(\hat{X}^{CF}_{sp_3})$. I define $\hat{X}^{CF}_{sp_2}(t) \equiv 1 - \hat{X}^{CF}_{sp_1}(t) - \hat{X}^{CF}_{sp_3}(t)$.

(2) **Allocate electoral college votes:** For states with less than 10 observations in the ANES sample, I assign electoral college votes based on the actual outcomes. For the other states, I allocate the electoral college votes based on the plurality winner. For each state s, I denote the electoral college votes for Democrats and for Republicans by $Dem_s(t)$ and $Rep_s(t)$, respectively.

(3) **Sum electoral college votes:** I sum the electoral college votes to find $Dem(t) \equiv \sum_s Dem_s(t)$ and $Rep(t) \equiv \sum_s Rep_s(t)$.

Finally, I compute the standard deviation of electoral college votes for Democrats across 500 simulations. I use this as the standard error of my counterfactual electoral college votes for Democrats (Hilary Clinton).

Results. The standard error of the electoral college votes for Hilary Clinton turns out to be 33 votes. Using this standard error, I run a t-test to check whether the counterfactual electoral college votes (336 votes found in Figure 1) is significantly different from the minimum threshold of electoral college votes needed for a candidate to win the election (270 votes). My null hypothesis is $h_0 : \hat{Dem} = 270$, and my alternate hypothesis is $h_a : \hat{Dem} \neq 270$. I find the counterfactual electoral college votes to be significantly different from 270 at the 95% confidence level (p-value = 0.046). I also test whether the estimated counterfactual electoral college votes for Democrats is significantly greater than 270, with the null hypothesis $h_0 : \hat{Dem} \geq 270$, and the alternate hypothesis, $h_a : \hat{Dem} < 270$. I find the counterfactual electoral college votes for Democrats to be significantly greater than 270 at the 95% confidence level (p-value = 0.023).

Table 3: Original Vote Shares

State	Dem. Vote %	Dem. SE	Dem. p-val	Rep. Vote %	Rep. SE	Rep. p-val	Other Vote %	Other SE	Other p-val
Alabama	34.2	10.7	.942	63.8	12	.42	2	1.8	.077
Alaska	0	0	N/A	100	0	N/A	0	0	N/A
Arizona	36.1	7.1	0	52.8	7.9	0	11	4.2	0
Arkansas	33.9	8.6	.85	60.7	8.6	.922	5.4	3.1	0
California	62.5	3.4	0	27.9	3.3	0	9.5	2.2	0
Colorado	57.9	6.4	0	35.7	8.2	0	6.4	2.8	0
Connecticut	53.3	7.4	.251	44.3	7.8	.01	2.4	1.3	0
Delaware	72.2	18	.032	12.2	11.9	.001	15.6	14.7	.042
District of Columbia	92.4	6.7	.342	0	0		7.6	6.7	.014
Florida	47	6.2	.117	50.9	6.2	0	2.1	1.1	0
Georgia	34.6	7.5	0	50.3	7	.567	15.1	4.8	0
Hawaii	51.6	23.4	.369	37.8	22.3	.481	10.6	10.6	.24
Idaho	34.5	4.6	0	53.8	5.3	0	11.7	3.7	0
Illinois	48.5	7.7	0	45.3	7.8	0	6.2	2.6	0
Indiana	31.3	7.3	0	63.9	7.4	0	4.8	3.4	0
Iowa	30.7	10.9	0	65.4	11.3	0	3.9	3.9	.485
Kansas	36.4	8.3	.725	55.4	8.5	.282	8.1	3.1	0
Kentucky	40.8	7.9	0	55.3	8.1	0	3.9	2.3	0
Louisiana	27.1	5.1	0	70.9	5.6	0	2	2.2	.27
Maine	13.1	9.7	0	75.6	13.8	0	11.3	10.7	.025
Maryland	50	8.8	0	40.7	6.8	0	9.3	4.5	0
Massachusetts	61.2	4.9	.026	29.9	5.6	0	8.9	4.2	0
Michigan	48.3	6.6	.14	48.7	6.5	.084	3.1	1.4	0
Minnesota	39.4	6.6	0	52.4	6.8	0	8.2	3.1	0
Mississippi	54.9	10.3	0	36	7.2	0	9.1	4	0
Missouri	42.4	9.7	.008	50.3	9.6	0	7.3	4.2	0

State	Dem. Vote %	Dem. SE	Dem. p-val	Rep. Vote %	Rep. SE	Rep. p-val	Other Vote %	Other SE	Other p-val
Montana	20	11.3	.003	74.1	12.2	.002	5.9	6.1	.085
Nebraska	59	12.3	0	41	12.3	0	0	0	N/A
Nevada	51.6	15.6	.294	48.4	15.6	.404	0	0	N/A
New Hampshire	68.3	6.5	0	27.7	8.4	0	4	4.8	.174
New Jersey	58.7	5.2	0	34.3	5.4	0	7	3.7	0
New Mexico	49.5	11.9	.571	28.9	5.5	0	21.6	7.2	0
New York	73.8	4.7	0	21.3	4.5	0	4.9	1.8	0
North Carolina	58.2	6.6	0	38.1	5.8	0	3.8	1.8	0
North Dakota	12.6	13.4	.072	63.5	25.6	.964	23.8	20.6	.092
Ohio	36.7	11.9	0	54.2	8.9	.003	9.1	4	0
Oklahoma	46.6	7.4	0	38.8	7.3	0	14.6	6.6	0
Oregon	79.8	7.8	0	17.6	7.6	0	2.5	2.5	0
Pennsylvania	54.8	4.5	0	41.1	4.4	0	4.1	2.3	0
Rhode Island	51.7	27.3	.833	48.3	27.3	.483	0	0	N/A
South Carolina	65.1	8.4	0	34.9	8.4	0	0	0	N/A
South Dakota	0	0	N/A	84.4	11.1	.002	15.6	11.1	.013
Tennessee	38.8	8.1	0	56.1	7.2	0	5.1	2.9	0
Texas	36.4	4.9	0	55.7	4.7	0	7.9	2.6	0
Utah	41.7	13.8	.001	35.8	12.8	.008	22.5	11.7	.738
Vermont	75.9	19.6	.019	24.1	19.6	.375	0	0	N/A
Virginia	36.6	7.1	0	54.3	8.1	0	9.1	3.7	0
Washington	56.9	6.2	0	35.4	6.3	.078	7.7	3.9	0
West Virginia	9.6	9.2	0	81.5	12	.008	8.9	8.6	.026
Wisconsin	51.6	8.8	0	38.8	7.8	0	9.6	3	0
Wyoming	14.5	14.6	.389	45.6	26.8	.191	39.9	28.2	.088

Notes: The table above shows all party vote shares and standard errors estimated using 2016 ANES data. The "Others" include all Independent and other party candidates. For each party and state, I compute p-values using one sample t-tests comparing the estimated ANES vote shares to the actual CQ vote shares, as detailed in Appendix A.1. With exception of Alaska, some p-values are so small that they are displayed as 0 when I round them to 3 digits after the decimal point. Some p-values are missing for states in which one or two parties received no votes. I excude any ANES respondents who did not register as voters, did note vote in the presidential election, or did not correctly report their age.

Table 4: Counterfactual Party Vote Shares

State	Dem. vote %	Dem. SE	Rep. vote %	Rep. SE	Other vote %	Other SE	p-value
Alabama	32	15.9	65.4	18.3	2.6	2.3	0
Alaska	0	0	100	46.6	0	0	.371
Arizona	39.2	11.5	49.8	13.3	10.9	4.7	.924
Arkansas	31.4	13.1	63.8	17.1	4.8	3.1	0
California	65.5	8	22.4	3.2	12.1	3.6	0
Colorado	55.9	12.9	35.3	11.4	8.8	4.6	0
Connecticut	63	21	33	9.2	4	2.2	0
Delaware	95.4	48.6	4.6	4.5	0	0	.048
District of Columbia	91.8	12.5	0	0	8.2	9	0
Florida	49.6	8.6	48.6	8.7	1.8	1	.549
Georgia	42.6	14	37.4	7.6	19.9	6.2	0
Hawaii	46.8	28	45.7	33.6	7.6	7.6	.81
Idaho	37.7	5.2	47.7	11	14.5	4	.21
Illinois	49.9	8.6	44.2	8.3	5.9	2.3	.89
Indiana	36.2	8.3	56.9	11.5	7	5.6	0
Iowa	34.9	14.4	57.8	21	7.3	7.3	.098
Kansas	42.7	13.2	45.5	11.2	11.8	5.5	.004
Kentucky	46	13.9	48.4	11.8	5.6	3.4	.364
Louisiana	29.9	8.3	70.1	13.5	0	0	0
Maine	24.6	17.9	75.4	29	0	0	.022
Maryland	50.4	11.2	38.9	9.6	10.7	5.2	.749
Massachusetts	62.1	13.8	25.3	7.4	12.7	6.8	0
Michigan	53.7	10.7	43.5	8.2	2.8	1.6	.001
Minnesota	35.5	8.7	56.6	16.1	7.8	3.6	.003
Mississippi	54.2	20.1	32.7	8.3	13.1	3.3	.262
Missouri	42.8	14.2	47.7	12.1	9.5	5.7	.225

State	Dem. vote %	Dem. Std. Error	Rep. vote %	Rep. Std. Error	Other vote %	Std. Error	p-value
Montana	15.5	13.4	80.1	31.9	4.3	4.5	.022
Nebraska	67.5	27.6	32.5	15.8	0	0	.041
Nevada	44.3	18.2	55.7	26.1	0	0	.328
New Hampshire	76.2	16.4	22.3	3.6	1.5	1.8	0
New Jersey	60.2	10	28.8	5.9	11	6.6	N/A
New Mexico	52.8	13.9	21	4.4	26.2	3.9	N/A
New York	78.8	14.8	16.4	4.1	4.8	2.2	0
North Carolina	60	10.5	34.8	6.1	5.2	2.7	0
North Dakota	9.4	10	62.2	29.1	28.4	23.6	.402
Ohio	40.7	8.9	48.9	9.8	10.4	4.1	.261
Oklahoma	47.8	15.1	36.3	9.3	15.9	7.7	
Oregon	80.9	19.5	15.8	10	3.3	3.3	0
Pennsylvania	57	9.5	41.6	6.8	1.4	.9	0
Rhode Island	63.9	41.9	36.1	24	0	0	.499
South Carolina	69.7	17.7	30.3	9.5	0	0	0
South Dakota	0	0	72.1	35.7	27.9	20.9	.153
Tennessee	36.5	8.2	57.5	13.3	5.9	3.4	0
Texas	39.1	5.6	52.6	7.9	8.3	3.1	0
Utah	48.7	22.9	33.1	14.8	18.2	10.6	.821
Vermont	61	23.9	39	31.7	0	0	.205
Virginia	39.8	10	46.1	9.9	14.1	6.2	.003
Washington	57.7	13.3	30.9	7.9	11.3	5.8	0
West Virginia	15.6	15	71.8	27.7	12.6	12.1	.034
Wisconsin	50.9	10.5	34.3	7.6	14.8	5.6	N/A
Wyoming	15.5	15.6	23.9	15.8	60.6	42.8	.655

Notes: The table above shows the counterfactual party vote shares and standard errors estimated using the 2016 ANES data set. Other party vote shares includes all independent and other party candidates. The p-values are from t-tests on the significance of the difference between the counterfactual plurality winner vote share and the runner-up vote share (h_0: Plurality winner vote share - Runner-up vote share = 0), as detailed in Appendix A.1. With the exception of California and Oregon, some p-values are so small that they are displayed as 0 when I round them to 3 digits after the decimal point. We exclude any ANES respondents who reportedly did not registered as voters, did not vote in the presidential election, or did not correctly report their age.

References

DeBell, Matthew, Michelle Amsbary, Vanessa Meldener, Shelley Brock, and Natalya Maisel, "Methodology report for the ANES 2016 time series study," Technical Report, Stanford University and the University of Michigan., Palo Alto, CA, and Ann Arbor, MI February 2018.

Heeringa, Steven G., Brady T. West, and Patricia A. Berglund, *Applied survey data anlysis* Statistics in the Social and Behavioral Sciences Series, Boca Raton, FL: CRC Press, 2010.

Rice, John A., *Mathematrical statistics and data analysis*, third ed., Duxbury, 2007.

West, Brady, Patricia A. Berglund, and Steven G. Heeringa, "A closer examination of subpopulation analysis of complex-sample survey data," *The Stata Journal*, 2008, *8* (4), 520–531.