Comparing Voter Difficulty in the 2020 General Election

w203 2022 sec-02, Victoria Hollingshead, Scott Kao, and Brian Tung

6/25/2022

Contents

1 Importance and Context

Over the last several decades, United States legislators have instituted barriers to ballot accessibility, resulting in rising concerns around the topic of abandoned ballots[1]. Tighter voting regulations, which take place in forms such as strict voter ID laws, registration restrictions, and restricted absentee voting, place disproportionate difficulties on varying demographics and result in tens of millions of abandoned ballots[2]. With the consideration that five out of the last six elections since 2000 had a popular vote difference of under five percentage points between the top Democrat and Republican candidate, voter difficulties leading to lost votes at any scale has the potential to influence the results of the presidential election.

This study aims to contribute to the discussion of voter difficulty by identifying the existence of disproportionate levels of difficulty experienced by self-claimed Democrat and Republican voters. In our attempt to better understand this issue, this analysis is driven by the following research question:

Did Democrat voters or Republican voters experience more difficulty voting in the 2020 election?

Discovering an answer to this research question is the first step toward improving voter infrastructure. An improved understanding of voter difficulty discrepancies could provide insight for political campaign organizers and grassroot organizations on crucial improvements that need to be made to their campaign strategies. Additionally, findings from this study could serve as a significant foundation for future studies on ballot abandonment.

2 Data and Methodology

Our analysis was conducted using the American National Election Studies (ANES) Time Series Study, centered around the U.S. Presidential Election of 2020. This is a cross-sectional study of 8,280 pre-election interviews and 7,449 post-election re-interviews collected from General Social Survey (GSS) respondents. Below are the key concepts we will use throughout our report.

Voter: A voter is a respondent who registered, voted, and provided voter difficulty information in the post-survey. The best available measurement for a voter is a self-reported confirmation from the respondent in the pre-survey or post-survey. If affirmative, these fields are used to define the universe for variable *V202119 POST: HOW DIFFICULT WAS IT FOR R TO VOTE*, which is used in our analysis. A system-reported feature that records each confirmed vote could be the best possible measurement of a voter, as this approach would eliminate potential discrepancies found in the self-reported post-survey. We identified 5,828 total voters.

Voter Difficulty: To measure voter difficulty, we used the Likert Scale response from variable *V202119 POST: HOW DIFFICULT WAS IT FOR R TO VOTE* in the post-survey. The valid responses used in our study were not, a little, moderately, very, and extremely difficult. All other responses were omitted. The best possible measurement of voter difficulty would include a set of verified difficulty options based on third party evidence. The post-survey would also allow respondents to provide multiple responses, well beyond the current primary and secondary response limitation. This would require additional research and data gathering, but may be possible in an ideal study. Because of the limited options and quantity of options voters could choose from when expressing voter difficulty, the best available measurement is the Likert Scale measurement.

Democrat/Republican: We further demarcated the voter sample according to the respondent's political leaning and affiliation. If a respondent indicated in the pre-survey, for *V201231x PRE: SUMMARY: PARTY ID*, that they leaned toward or identified with the Democrat (i.e., Strong Democrat, Not very strong Democrat, Independent-Democrat) or Republican party (i.e., Independent-Republican, Not very strong Republican, Strong Republican), they were categorized accordingly. This is the best available measurement for

[1] https://www.brennancenter.org/issues/ensure-every-american-can-vote/vote-suppression
[2] https://www.americanprogress.org/article/increasing-voter-participation-america/

party affiliation. We did not include the following in our study: respondents who identified as Independent, refused to answer, or did not know how they identified in the pre-election survey.

Because voters are required to state their political preferences upon registration and re-register if their political preferences change, the best possible measurement for party affiliation can be found in voter registration data. This would require additional data gathering, but would be possible in an ideal study. We opted to use the pre-survey variable, *V201231x PRE: SUMMARY: PARTY ID*, instead of inferring partisan leaning from voter behavior (i.e. labeling a respondent a Democrat if they voted for Joe Biden), because we deemed our third party inference to be less accurate than the respondent's self-identification. Following our definitions, we identified 3,128 Democrat and 2,700 Republican voters.

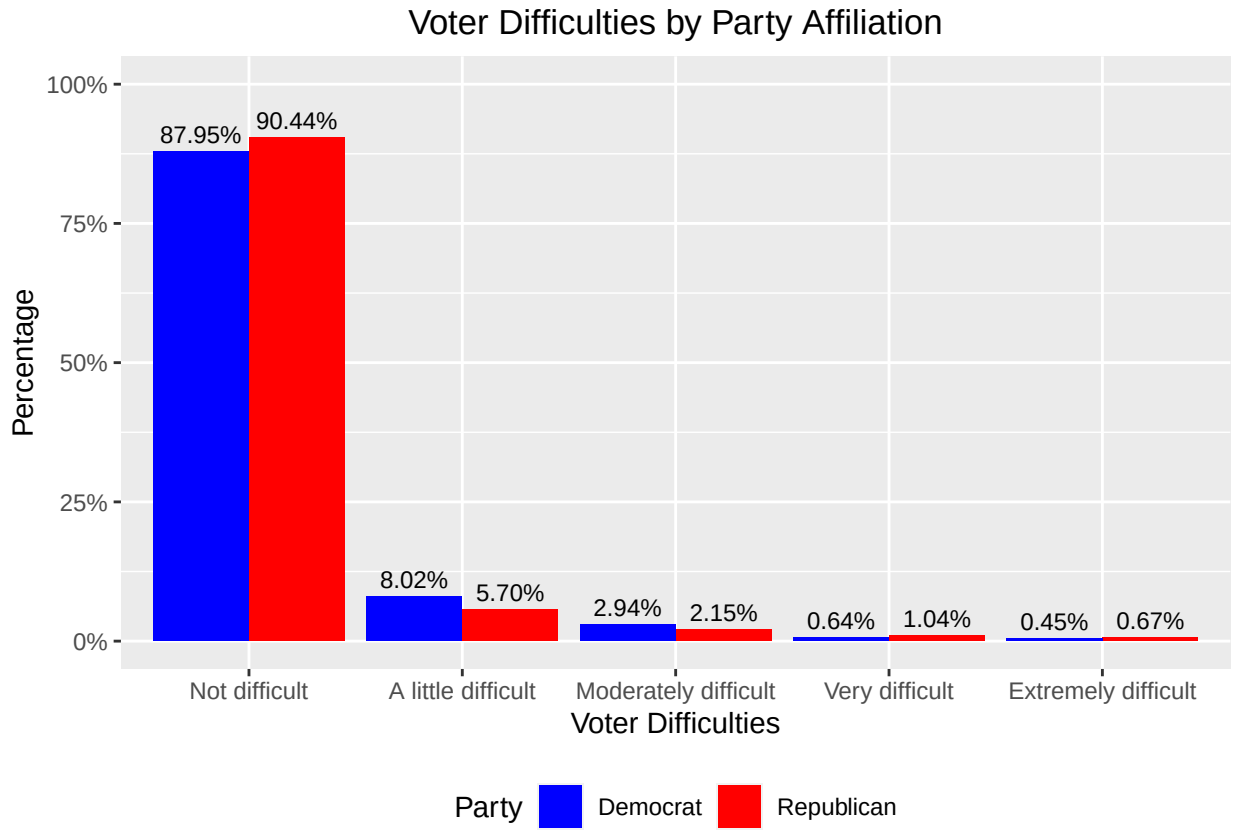

Figure 1: 10% ~ 12% of Voters experienced some form of difficulties

Figure 1 highlights that approximately 12% of Democrat voters and 10% of Republican voters experienced some form of difficulty in the 2020 General Election. On the scale of 168 million registered voters in 2020[3], these differences can amount to considerable voter disparities. The percentage of difficulty declines steadily from a little difficulty to extremely difficult, with less than 1% of Democrats and Republicans experiencing high amounts of voter difficulty.

Using the concept definitions above, along with the respective variable groupings, the null hypothesis was defined as follows:

> **Null Hypothesis:** *The probability that Democrat voters experienced more difficulty than Republican voters is the same as the probability that Republican voters experienced more difficulty than Democrat voters.*

[3]https://www.statista.com/statistics/273743/number-of-registered-voters-in-the-united-states/

To test our null hypothesis, we used a nonparametric test of two independent variables, the Hypothesis of Comparisons Version of the Wilcoxon Rank Sum Test. The Hypothesis of Comparisons Version requires two assumptions of the data to be fulfilled: 1) the data must be at least ordinal in scale, and 2) the data must be independent and individually distributed.

First, by virtue of using the Likert Scale as our output variable, the sample is ordinal in scale. The difference between each difficulty level is not quantifiable or consistent. For this reason, a nonparametric test was required. Second, the independent and individually distributed assumption states that each sample is from the same distribution and are mutually exclusive. Because our sample is randomly drawn from the USPS computerized delivery sequence file (C-DSF), and each residential address across the 50 states and Washington DC has an equal probability of being selected, we can confirm that each sample meets this criteria.

3 Results

```
test <- wilcox.test(anes_filtered$V202119 ~ anes_filtered$Party)
```

This test indicates that there is a statistically significant difference between Democrat and Republican voting difficulties ($w = 4323774$, p-value $= 0.003541$), and as a result, we reject our null hypothesis. However, further exploration of practical significance using rank biserial correlation yielded a value of 0.04. Due to the poor practical significance, we determined that the results of this study may not be reliable and representative in the real world.

4 Discussion

Given 87.95 - 90.44% of voters self-reported no difficulty in the voting process, the limited practical significance of this study was expected. Therefore, we conclude that the actions of political campaign organizers and grassroot organizations to improve voter turnout would be limited for this focus group. For the highest impact, future analysis and resources may be best allocated to respondents who had intended to vote, but were unable due to voting difficulties. Early in our research design, we had considered analyzing the voting experience of this demographic. While we did not pursue this analysis due to the limited sample size and necessity of several test assumptions, we encourage future research for this population. The results of this future study has the potential to remove voting barriers from disenfranchised populations. Overall, we hope our study contributes to the discussion of alleviating voter difficulty.