Influence of Party Affiliation on Reported Voting Difficulty

Kevin Lustig, Rebecca Nissan, Anuradha Passan, Giorgio Soggiu

2/29/2022

Contents

1 Importance and Context

The topic of obstacles to voting has been the focus of considerable discussion in the United States in recent years, including notably before, during, and since the 2020 federal election. The debate over the facility with which Americans are able to vote has, unsurprisingly, been divided along partisan lines. Elected officials and their supporters in both major parties use claims about difficulty to guide voting laws; Democrats tend to support efforts to increase access to the polls while Republicans tend to support restrictions on voting access.

This raises the question of whether the actual experience of difficulty voting is related to party affiliation. Perhaps attitudes toward measures to loosen or tighten voting regulations are driven by real or perceived differences in personal experience with voting difficulty. This analysis therefore seeks to address the following research question:

Did Democrats or Republicans have more difficulty voting in the 2020 election?

This question has the potential to illuminate important aspects of the debate about voting laws. What is the scale of actual voting difficulty for members of each party, and who would be most affected and in what way by changes meant to address the current ease or difficulty of voting?

2 Data and Methodology

This study uses data from the 2020 American National Election Studies (ANES) survey, specifically their 2020 Time Series Study, which includes survey responses for a random sample of US citizens from both before and after the 2020 election. The full dataset includes 8,280 observations. Our study focuses only on "voters" who are categorized as either Democrats or Republicans. We operationalize the concept of a "voter" and define party affiliation as follows:

We consider a *"voter"* to be anyone in the ANES survey sample who registered to vote.[1] Since this study examines **difficulty** voting, it is important to consider the subset of people who demonstrated some intention to vote by registering even though they may not have ultimately voted - perhaps because of some difficulty (see *Difficulty Category* below). Eliminating anyone for whom we did not have voter status info as well as those who were not registered by the time of the post-election survey leaves 7,053 voters. In order to determine *party affiliation*, we use a self-reported party identification variable.[2] Since this study focuses specifically on Democrats and Republicans, we classify anyone who reported a lean or stronger preference for the Democratic Party as a Democrat, and likewise for Republicans. Another possible way to define somebody's party is based on who they voted for, but there is no guarantee in 2020 or any other election that a voter will select the candidate from their party for any particular position. Therefore, we define party affiliation solely based on the self-reported variable. Of the original population, 6,316 respondents are identified as Democrat or Republican voters.

We next operationalize the concept of *"difficulty in voting."* Given the subjective nature of the term "difficulty", there are various ways difficulty in voting could be conceptualized. The ANES survey provides different information for those who cast a ballot and those who did not in the 2020 election. For **respondents who voted,** the ANES survey asks them both to enumerate specific difficulties in trying to vote (e.g. long lines, failing to provide proper registration papers, etc) and the degree of difficulty they encountered, if any. For **people who did not vote,** the ANES survey asks for the two main reasons they did not vote.[3] This inequality in the data for the two types of voters poses a challenge for analyzing the degree of difficulty faced, which we expand on in describing two variables that we elect to use to represent difficulty: difficulty ranking and difficulty category.

Difficulty ranking
Only respondents who did vote in the 2020 election were posed the question "How difficult was it to vote?" with the following possible responses: not difficult at all (1), a little difficult (2), moderately difficult (3),

[1] Survey question V202068x
[2] Survey question V201231x
[3] Survey questions: V202123 and V202124

very difficult (4), extremely difficult (5).[4] We employ these values as a simple variable indicator of degree of difficulty voting.

Difficulty Category
While the difficulty ranking variable appears to be a straightforward way to answer the research question premised on the idea of more or less difficulty voting, it's important to reiterate that it accounts only for respondents who were able to complete the act of voting. It does not provide insight on difficulty faced by voters (as per the definition mentioned above) who tried to vote, but encountered so much difficulty that in the end they were not able to vote. Because the ultimate concern about difficulty voting that motivates this research is that it may ultimately prevent eligible voters from casting a ballot, we propose that voters that are unable to vote due to difficulty are the *most severely* impacted by voting restrictions, and thus merit special focus in our analysis. Thus, we create a synthetic variable encompassing three categories: (1) voters who did not experience difficulty when trying to vote, (2) voters who self-reported any degree of difficulty and still managed to vote, and (3) voters who self-reported difficulty and were not able to vote.

In the ANES survey, respondents who did not end up voting were asked to attribute the main reason why.[5] It is important to note that these responses enable us to identify not simply voters who had difficulty and also didn't happen to vote; they have provided a specific difficulty as the reason they didn't vote. For this category, we include voters who provided the following responses: (6) "I did not have the correct form of identification", (8) "sick or disabled", (9) "transportation", (10) "bad weather", (11) "the line at the polls was too long", (12) "I was not allowed to vote at the polls, even though I tried", (13) "I requested but did not receive an absentee ballot", (14) "I did not know where to vote", (16) "Other". We disregard responses that indicate the voter simply forgot, wasn't interested in voting, etc. Those who provided one of these answers to this question were classified as not having experienced difficulty voting.

Proposed Testing Approach
Here, we also note another challenge with analyzing our research question: the distribution of the values for both the difficulty ranking and category variables are heavily skewed toward responses that indicate no difficulty voting (difficulty ranking = 1; difficulty category = 1 [Appendix: Figures 1-4]). Thus, we anticipate that measurement of any differences in voting difficulty between voters of the two major parties will be heavily influenced by these tendencies and unlikely to show a significant difference. We validate this supposition through statistical testing in the next section, and then propose an alternate analytical approach to overcome this limitation.

Both the difficulty variables are independent and identically distributed per the ANES survey methodology. Given that the difficulty ranking variable values seem to have an implicit equal distance between each value (no unexpected neutral values or jumps between values), this variable is metric in nature, and thus we employ a two sample t-test, with the null hypothesis stated as:

> **Null Hypothesis 1:** *The probability that a Democratic voter experienced increased difficulty voting is equal to the probability that a Republican voter experienced increased difficulty voting*

While the values of the difficulty ranking variable are not normally distributed (Appendix: Figures 1 and 2), due to the large sample size (n=6316) a t-test remains valid as the sampling distribution should be normal via the Central Limit Theorem. The difficulty category variable is ordinal, and thus we utilize the Wilcoxon Rank-Sum test, with the null hypothesis stated as:

> **Null Hypothesis 2:** *The probability that a Democratic voter experienced more difficulty voting than a Republican voter is equal to the probability that a Republican voter experienced more difficulty voting than a Democratic voter*

3 Analysis & Results

Ranked voting difficulty
Our analysis begins with an examination of the voting difficulty ranking. We test the difference, if any,

[4]Survey question V202119; responses such as "no post election data" or "no post election interview" have been filtered out
[5]Survey question: V202123

between the level of difficulty reported by Republican- and Democrat-identified voters:

```
t.test(diff_ranked~party,data = data_fin_reduced)
```

The result of this test provides no evidence that there is a significant difference in reported experience of voting difficulty based on political identification (t=1.2604,p=0.2076). This is unsurprising, as discussed previously; Democrats reported a mean level of voting difficulty of 1.176151, while Republicans reported a mean of 1.157778 (1 = "Not difficult at all"). That is, voters of both parties collectively experienced, on average, very little to no difficulty; in fact, 86.56% of voters reported no difficulty at all (Appendix: Table 1). We note again that the tested variable also does not incorporate the experience of those who were not able to cast a vote after experiencing difficulty, as these voters did not provide a ranked difficulty survey response.

Categorical voting difficulty and voting status
Therefore, we focus our analysis on the subset of voters who experienced difficulty in order to detect any difference in the experience of difficulty for Republicans and Democrats. To do so, we test our secondary variable categorizing voters by their report of having or not having difficulty and voting status in categories 2 (difficulty; voted) and 3 (difficulty; didn't vote).

```
wilcox.test(diff_categorical~party,data = data_fin_difficulty, conf.int=TRUE )
```

In this result, we find evidence of a difference in the distribution of difficulty and voting status between the two parties, with a small but significant lower shift in the category value for Republican voters as compared to Democratic voters (p=0.0005455). Because it excludes those who did not have difficulty voting, this test is not a true test of level of difficulty, but rather of the distribution of those that did or did not vote due to having difficulty. It therefore indicates that Democratic voters were significantly more likely to be unable to vote due to experiencing difficulty. Based on this result, we reject the null hypothesis that there is no difference in voting difficulty between the political groups.

It is important to note some key limitations of this approach. Factors that are not considered in this analysis are quite likely to have an impact on the experience of voting difficulty, such as method of voting and voting location. Moreover, it is critical to note that the method of questioning voters in the ANES survey asks only for self-reported and therefore subjective descriptors of difficulty. It does not attempt to validate or weight these incidents. Thus, this study does not attempt to analyze an objective measure of how difficult the voting experience was for any voter. In reality, two voters may regard the same experience as more or less difficult, and may choose to behave differently at the polls because of that perception. Different data than that found in the ANES dataset will be needed to form a clearer and less subjective picture of the nature of obstacles voters face when attempting to cast a vote.

4 Discussion

Our analysis finds, first, that the reported incidence of voting difficulty is quite low for all voters, and our analysis using the ranked level of self-reported difficulty showed no significant difference between Republican and Democratic voters. However, in examining the subset of voters who did self-report difficulty, we propose that it is meaningful to know how many were not able to ultimately cast a ballot due to that difficulty. With regard to this measurement, we do find that Democrats who experienced difficulty at the polls were slightly but significantly more likely to be able to vote than Republicans.

This study suggests a number of interesting questions that merit further research. Are there significant differences in the kind or amount of difficulty faced by just those voters who did experience difficulty? Do Democrat and Republican voters report or experience different kinds of difficulty at different rates? We believe our results are most interesting in the context of these and other answers in helping to form an evidence base for ongoing debates about voting access. The right mix of policy solutions for changing (or preserving) current voting practices depends on an understanding of the real ease or difficulty of voting for different groups. We hope that our findings can help to point the way toward additional answers that will contribute to effective and evidence-based policy-making that ensures fair access to the polls for all voters with a minimum of obstacles.

5 Appendix 1: Tables and Figures

Table 1: Self-Reported Difficulty and Party

	Difficulty	No difficulty
Democrat	0.075	0.457
Republican	0.059	0.408

Table 2: Self-Reported Difficulty and Voting Ability

	Difficulty	No difficulty
Didn't Vote	0.034	0.000
Voted	0.101	0.866

Figure 1: Difficulty Ranking of Voters Who Voted, Democrats

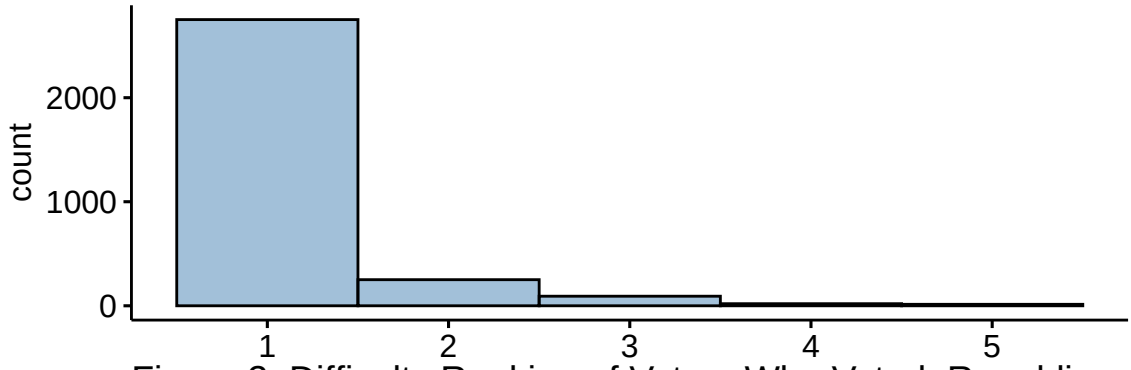

Figure 2: Difficulty Ranking of Voters Who Voted, Republicans

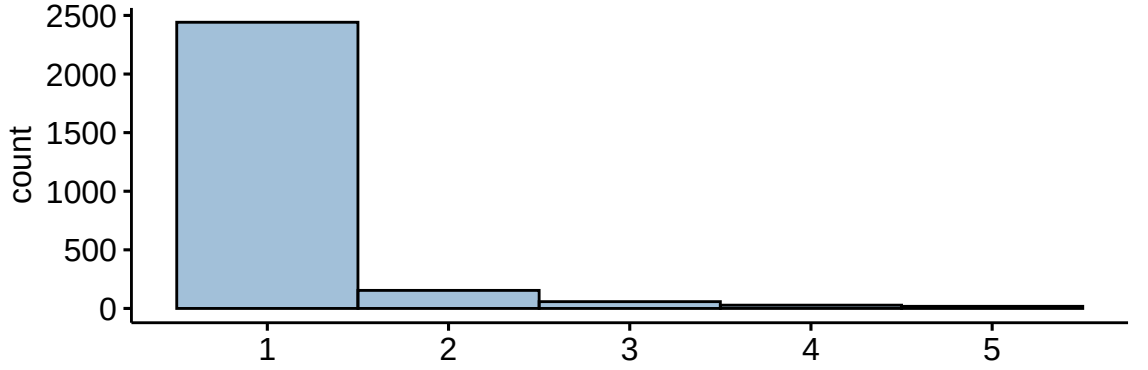

Figure 3: Difficulty Categories, Democrat Voters

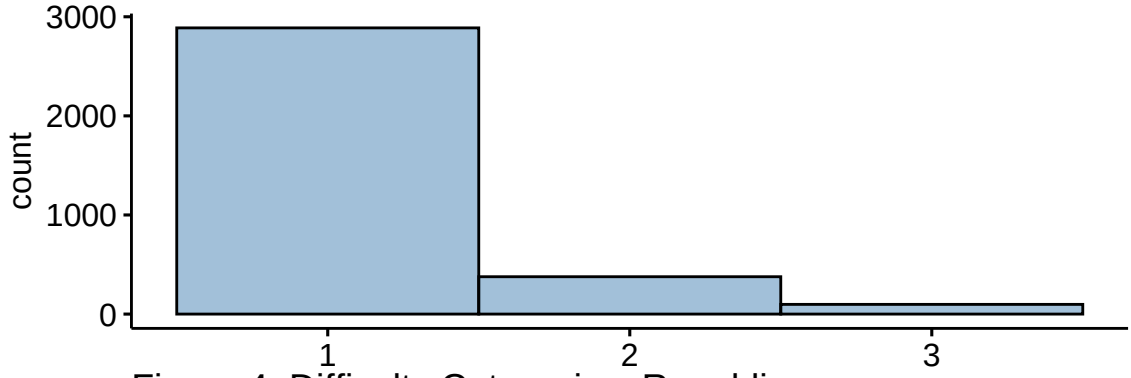

Figure 4: Difficulty Categories, Republican