

# An Analysis of Democrat and Republicans Voting Difficulty

Eric Wu, Gautam Sai Yarramreddy, Kevin Yi, and Xinran Zhang

2024-10-17

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data Description</b>	<b>1</b>
<b>3</b>	<b>Results and Discussion</b>	<b>2</b>
<b>4</b>	<b>Conclusion</b>	<b>3</b>

# 1 Introduction

In a well functioning republic, it is incredibly important that everyone has a voice and the ability to vote for policies and representatives. According to Pew Research, the past three elections have garnered the greatest voter turnout in US elections in decades with the 2020 election having the highest turnout since the 1900 election. Additionally, the margins between presidential race wins have become increasingly narrow and even involving the year 2000 and 2016 presidential election winners losing on the popular vote but winning the electoral college. These facts combined with many other factors make it prudent that people’s votes matter more than ever. The question of how to make voting more fair and equitable to everyone is crucial for election officials.

This analysis contributes to the discussion of voting rights, focusing on the voting difficulty for members of the Democrat and Republican parties. These two parties have had the biggest impact on the elections and have won the presidential election since 1852 and controlled the Congress since 1856. As a first step, this analysis aims to address the following research question:

*Do Democratic voters or Republican voters experience more difficulty voting?*

The answer to this question could provide valuable information to election officials on which groups need more resources and barriers to voting removed. This could also provide for NGOs and political activist groups to focus their attention on more marginalized groups. This could ultimately help further increase voter turnout and voting rights and allow previously disenfranchised groups to have a greater impact on policies and their representatives.

## 2 Data Description

Our analysis leverages data from the 2022 American National Election Studies (ANES). This is a survey dataset, based on a sample of respondents drawn from the YouGov platform. The survey period was from November 14-22, 2022 following midterm elections in the United States. Participants sign up on YouGov to complete online surveys in exchange for points that can be redeemed for gift cards. Due to these reasons, the YouGov panel is not nationally representative. This dataset consists of an opt-in online panel that consists of 1,585 total respondents where 1,500 were selected to produce the final dataset to be as representative of the United States voter population as possible.

A 2009 study by John Richard Petrocik in “Measuring party support: Leaners are not independent” concludes in his finding that there has been a rise in individuals that declare themselves independents (nearly 40%) but are actually partisan. However, we will use the self reported party affiliation to characterize who is a democrat or republican because roughly 30% of the survey respondent declared themselves independent or something else. Furthermore, voting history from 2020 and 2022, as provided in the survey, may not be an accurate way of categorizing who is a Democrat or Republican. This is because the Republican candidate, Donald Trump in 2022, and Trump supported candidates in 2022 did not fit neatly into the categorization of a Republican, thus leading to more drastic swing votes as suggested by the study.

Operationalizing the variables for our research question, we define voters as individuals that self-reported on the survey that they voted either by mail or in person for the 2022 November 8 election. We define participants as Democrat or Republican according to self-prescribed party affiliation. To operationalize the concept of a voter’s experience in voting difficulty, we used the “votehard” variable for which participants marked down their self-perceived level of difficulty they faced while voting. The question asked for the votehard information was: “How difficult was it for you to vote?”. Data is reported on a 5-point Likert scale that ranges from 1 (“Not difficult at all”) to 5 (“Extremely difficult”). It is important to note that there were no hard definitions for these categories. They are the self-perceived difficulty of voting. The votehard column additionally had another possible value labeled -1 representing an “inapplicable, legitimate skip” option.

The original data had 1500 rows with no null values. After keeping only the results of voters that voted, removing the -1 (“Inapplicable, Legitimate skip”) voting difficulty values, and filtering only for those who identified as Democrat or Republican, we are left with 820 total values. Of the 820 total values, 460 are Democrats, and 360 are Republicans.

We will be comparing two different groups (Democrats and Republicans) by evaluating difficulty of voting, measured by the likert scale. Therefore, we proceed with a Wilcoxon Rank-Sum (Hypothesis of Comparisons Version) to test our hypothesis.

The null hypothesis can be phrased as follows:

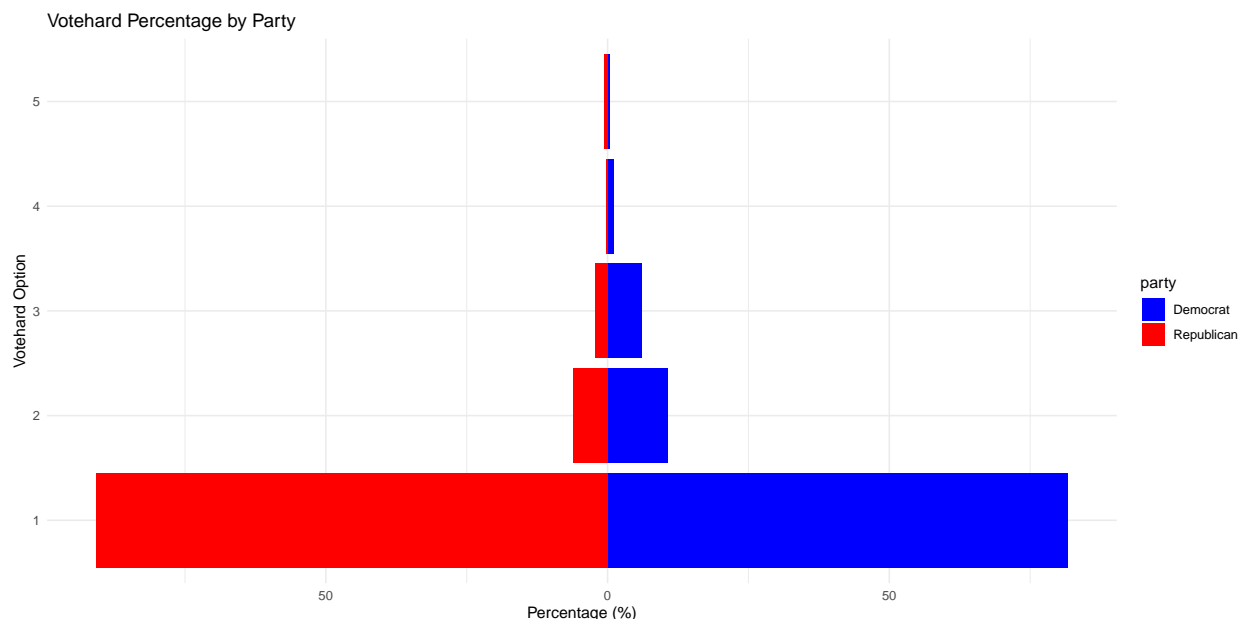
**Null Hypothesis:** *The probability that a Democratic voter's experiences more difficulty in voting than a Republican voter is equal to the probability that a Republican voter experiences more difficulty in voting than a Democratic voter.*

In order for a Wilcoxon Rank-Sum (Hypothesis of Comparisons Version) to produce reliable inference the following assumptions must be true:

1. Data must be measured on an ordinal scale.
2. Data must be IID (independent and identically distributed).
  - Each  $X_i$  is drawn from the same distribution, each  $Y_i$  is drawn from the same distribution, and all  $X_i$  and  $Y_i$  are mutually independent.
  - When samples are independent, the result of any one sample is not informative about the value of any of the other samples.
  - All samples are drawn from the same distribution.

First our outcome variable must be measured on an ordinal scale. In our case, voting difficulty is measured by a 5-point Likert scale that ranges from 1 (“Not difficult at all”) to 5 (“Extremely difficult”). Likert scale is a common form of ordinal value.

Second, the data must be generated via the IID process. The ANES 2022 pilot data used a panel of individuals from the YouGov platform. Users of this platform could tell friends and family members about YouGov and have them participate in surveys for rewards. This may create clusters of individuals that may give similar responses. However, this possibility is low since YouGov claims to have millions of users so links between users should be very low.



### 3 Results and Discussion

```
democrats <- anes2022_formatted$votehard[anes2022_formatted$party == 'Democrat']
republicans <- anes2022_formatted$votehard[anes2022_formatted$party == 'Republican']
```

```
test_result <- wilcox.test(democrats, republicans, paired = FALSE)
```

The p-value is 0.0002. Therefore, at the 5% significance level, we reject the null hypothesis. We conclude that the probability mean that a democrat voter experiences the same voting difficulties as a Republican voter are not equal.

There are some notes to mention about the results. First, the ANES data is not a nationally representative sample, meaning that the results of this test could be difficult to generalize to the entire voting population of the United States. Second, individuals have different scales of how difficult a task is. The dataset includes a subjective, self-perceived rating and not an objective metric. This could additionally make it more difficult to generalize the results to the wider population.

## 4 Conclusion

This study found evidence that Republican and Democrat voters on average experience different levels of difficulties when voting. The average Democrat voter experienced slightly higher levels of self perceived difficulty in voting than Republican voters.

Our results may be of key interest to election officials who work on breaking down barriers for voting and expanding voter rights. It additionally could be important to political campaigns and non-government organizations (NGOs) who want to increase voter turnout and increase representation of the constituency. Working towards these goals can increase the rates of first time voters.