

Writing Assignment 1

1 Instructions

- This writing assignment is due on September 29 by the beginning of class.
- Hopefully, you can have the computation (Part 1) done by September 22, so that we can spend a couple of class periods talking about the writing.
- Turn in paper and R code to Abhi (abhisekh@exchange.tamu.edu) and me (crainey@tamu.edu) by e-mail. This should be one .R file and one .pdf file. Also, submit the paper to through Turnitin. You can find the link on the eCampus sidebar. Lastly, bring a hard copy of the paper to class.
- Plagiarizing prose or code is not acceptable.
- Any sources that you rely on should be cited. See the APSA style guide ([pdf](#)) for the details on how citations typically work in political science. You can see this paper of mine ([pdf](#)) for an example. There is no need to include an abstract.
- Feel free to be creative with the style of the title, headers, etc, but make the paper look professional.
- We haven't talked much about graphs at this point, but feel free to include any tables that you feel help you make your point.
- Length is not important. You should, though, try to address the questions completely but compactly. Focus on the quality of the words, not the quantity. I think 750-1,000 words should be sufficient.

2 Problem

Surveys are frequently used to measure political behavior such as voter turnout, but some researchers are concerned about the accuracy of self-reports. In particular, they worry about possible social desirability bias where in post-election surveys, respondents who did not vote in an election lie because they may feel that they should have voted. Is such a bias present in the American National Election Studies (ANES)? The ANES is a nation-wide survey that has been conducted for every election since 1948. The ANES conducts face to face interviews with a nationally representative sample of adults. Table 1 displays the names and descriptions of variables in the turnout.csv data file.

| Variable | Description |
|-------------|---|
| year | election year |
| ANES | ANES estimated turnout rate |
| VEP | Voting Eligible Population (in thousands) |
| VAP | Voting Age Population (in thousands) |
| total | total ballots cast for highest office (in thousands) |
| felons | total ineligible felons (in thousands) |
| noncitizens | total non-citizens (in thousands) |
| overseas | total eligible overseas voters (in thousands) |
| osvoters | total ballots counted by overseas voters (in thousands) |

Table 1: Names and Descriptions of the United States Election Turnout Data.

2.1 Part 1: R Script

Using an R script, do the following:

1. Load the [data](#) into R and check the dimensions of the data. Also, obtain a summary of the data. How many observations are there? What is the range of years covered in this data set?
2. Calculate the turnout rate based on the voting age population or VAP. Note that for this data set, we must add the total number of eligible overseas voters since the VAP variable does not include these individuals in the count. Next, calculate the turnout rate using the voting eligible population or VEP. What difference do you observe?
3. Compute the difference between VAP and ANES estimates of turnout rate. How big is the difference on average? What is the range of the difference? Conduct the same comparison for the VEP and ANES estimates of voter turnout. Briefly comment on the results.
4. Compare the VEP turnout rate with the ANES turnout rate separately for presidential elections and midterm elections. Note that the data set excludes the year 2006. Does the bias of the ANES vary across election types?
5. Divide the data into half by election years such that you subset the data into two periods. Calculate the difference between the VEP turnout rate and the ANES turnout rate separately for each year within each period. (Edit 9-24 - Suggestion: you might also take the mean for each period as well.) Has the bias of the ANES increased over time?

2.2 Part 2: Essay

(Edit 9-24) Clearly divide the essay into five numbered sections (as this assignment is divided into three sections). Give each section a nice name. For example, don't call the first section "Section 1." Instead, call it something like "1 Difficulties in Measuring Voter Turnout." Section 1

should correspond to the first question, section 2 should correspond to the second question, and so on.

1. Clearly explain the problem and its importance. What is overreporting? Why should we care about it? Preview the idea of a voting rate based on the VAP and VEP. Which of these turnout rates is the more relevant theoretically and/or normatively?
2. Clearly explain the data. What variables do you use? What time frame do they span? Be sure to clearly explain what the variables mean.
3. Compare the turnout rate based on the VAP and the turnout rate based on the VEP. What are the differences? Are these difference increasing, decreasing, or staying about the same over time?
4. Compare the turnout rate based on the ANES and the turnout rate based on the VEP. What are the differences? Are these difference increasing, decreasing, or staying about the same over time? Should these two rates be about the same? Why might they be different? Are these difference smaller, larger, or about the same in midterm elections compared to presidential elections? Are the differences between the ANES and the turnout rate based on the VEP increasing, decreasing, or staying about the same over time?
5. Conclude with your ideas about the implications for researchers using self-reported turnout in the study of voter behavior.

3 Grading

- R Code: 30%
 - The code runs and computes the correct quantities: 15%
 - The code is neatly written and thoroughly commented: 15%
- Paper: 70%
 - Analysis: 15%
 - Organization: 15%
 - Punctuation: 10%
 - Grammar and usage: 10%
 - Spelling: 10%
 - Format: 10%