American National Election Studies: The People and their Policies

Academic Track Edition
Summer 2024

1 Setting the Stage

Imagine you are a research assistant for a prominent professor at the University of Chicago. As the coming election looms, your advisor is interested in using historical data to paint a picture of the current landscape of American politics. Your advisor has used ANES data before, but his research has been focused in local politics for the past decade. As a refresher, he would like you to take a look at the data and do a preliminary analysis with another dataset of county level covariates compiled by Harvard professor Raj Chetty. He also knows that you have PhD applications coming up and could use some guidance on your writing sample. He suggests that you kill two birds with one stone and write a short academic style paper using the data he hopes to use in more depth later. He suggests you model your work on a research note and leaves the angle of your research to you. You're excited to pull together your disparate research experience in one place and get your advisor's feedback, so you decide to put your best foot forward for this project.

2 About American National Election Studies

The American National Election Studies were one of the first, if not the first, organizations to survey the American electorate back in 1948. In fact, the ANES was the only organization of the day that accurately predicted Harry Truman's defeat of Thomas Dewey. Since then, many more surveys have gained prominence, but ANES remains one of the longest running. Though the same individuals are not sampled every four years, a representative sample is captured each year alongside each individual's demographic characteristics and political beliefs and policy preferences. For more information, visit electionstudies.org.

3 Research Note Objectives

A research note is an academic paper, just a little bit shorter. You'll conduct a literature review, establish context, write out your empirical strategy mathematically, display your results, and both introduce and conclude your paper. The most challenging piece of this may well be generating your research question. Most of the analysis should mirror what you've already done in DPSS through homeworks, but this time you'll be guiding yourself. Below are some must haves for the research note to receive full points:

- At least five academic sources. If they haven't been published in a journal, these can be working papers published in the NBER Working Paper Series or similar reputable place. Because it can take years to write and publish an academic paper, working papers are often closest to the forefront of academic literature. Feel free to attend Alex's lecture on literature reviews on July 2nd, or review the slides/recording, for more information.

- One descriptive visual, ideally tying into the analysis. This is a graph or map that allows the reader to see and compare the data. Bar charts, violin diagrams, and histograms could all work for these purposes.

- An empirical strategy written mathematically. This is the formal equation that you are using for your analysis, in this case a regression. Below is an example of an empirical strategy from Mazzonna and Peracchi (2024):

> Our basic model for individual wealth changes is the following first-difference model:
>
> $$\Delta W_{it} = \beta_0 + \beta_1 \text{Aware}_{it} + \beta_2 \text{Unaware}_{it} + \beta_3 \text{Pessimist}_{it} \\ + \boldsymbol{\beta}_4^\top \boldsymbol{X}_i + \boldsymbol{\beta}_5^\top \boldsymbol{Z}_{it} + \psi_t + U_{it}, \tag{1}$$
>
> where $\Delta W_{it} = W_{it} - W_{i,t-1}$ is the change in real wealth (total, financial, or subcomponents; US\$1,000s in 2014 prices) of individual i between survey wave $t-1$ and t;[6] Aware_{it}, Unaware_{it}, and Pessimist_{it} are binary indicators for being aware, unaware, or pessimist in wave t (as defined in sec. III.B); $\boldsymbol{X}_i$ is a vector of time-invariant regressors that includes binary indicators for sex, race, and years of education; $\boldsymbol{Z}_{it}$ is a vector of time-varying regressors that includes a quadratic age term, lagged wealth and memory score, and a set of binary indicators for labor force status, marital status, and geographical region (census division); ψ_t is a survey wave effect common across individuals; U_{it} is an unobservable error term assumed to be mean independent of all included regressors; and the β_j are unknown coefficients to be estimated. We include lagged wealth and memory score because wealthier

- A copy of the code you used to clean and analyze the data.

- 3000-5000 words of text.

- APA style in-text citations and references section.

- LaTeX typesetting – this document was generated using LaTeX. A template will be provided and a workshop will be held on July 16th to introduce this tool.

- All of the components outlined in section 4.

4 Research Note Components

1. **Introduction with Context:** The introduction should not only include the research question, but a theoretical backing on why this is an important question. You are not expected to present a formal model, but you should be able to provide an intuitive argument for your hypothesis and explain the mechanisms of the phenomenon. Furthermore, you should substantiate why and to whom this question matters. Whether or not your question is context specific in time or place, you should be able to tell us how your findings are/would be informative beyond your specific application. Additionally, give any historical, environmental, cultural, or political context that will help us understand your research design. This will be a more in depth process for some of you that others depending on your research question.

2. **Literature Review:** We stand on the shoulders of our scientific forebearers – so make sure to cite them. You are not expected to do an expansive literature review of several dozen papers, but you should cite at least 5 academic papers. Even if your research idea is novel, there is always related literature be it a review of the region you're exploring, an econometrics paper, or an article that uses the same data in a different way. Two possible references are cited in the references section below

3. **Data and Empirical Strategy:** If you are using more than the data that has been provided, add a section describing this data source and how it's been adjusted. Make sure to note how the dimensions of the data may have been adjusted, i.e. if any merges were done. This section should include at least one visual describing the visual. In addition to data specifications, the empirical strategy should be outlined, similar to the example above. Explain how the methodological approach will help answer your research question. Be specific about which mechanisms of the social phenomenon the equation identifies and how it approaches causality.

4. **Results:** Results should be displayed in the form of a regression summary table and optionally other figures. In addition to the results themselves, you should include a discussion of how the results answer your research question. You may also briefly outline potential extensions to your work given these findings and explain how your findings suggest future avenues of research.

References

Gimpel, J. G., Lovin, N., Moy, B., and Reeves, A. (2020). The urban–rural gulf in american political behavior. *Political behavior*, 42:1343–1368.

Malhotra, N. and Krosnick, J. A. (2007). The effect of survey mode and sampling on inferences about political attitudes and behavior: Comparing the 2000 and 2004 anes to internet surveys with nonprobability samples. *Political Analysis*, 15(3):286–323.