Kim Hancock

Voter Identification Laws and the Suppression of Minority Votes

The main takeaway of the paper by Hajnal, Lajevardi and Nielson (2017) is that strict voter ID laws disproportionately negatively affect minorities. This claim is portrayed by analyzing data from the 2006-2014 Cooperative Congressional Election Studies (CCES), a national stratified sample survey administered over the internet. Combined with data on the strictness of voter identification laws in place in each state, the paper's analysis purports that there is consistently lower voter turnout in states with strict voter ID laws. The paper also controls for other demographic and social characteristics that drive voter participation but ultimately land at the same result--that strict voter ID laws do cause voter suppression.

After the initial read, I was impressed by how the paper was able to present a thorough statistical analysis to support their claim; voter suppression is a slippery topic that often defies quantitative analysis. However, as a mathematics major interested in statistical analysis in social sciences, I wanted to learn more about their methodologies. In doing some additional research, I found a paper written in response to Hajnal, Lajevardi and Nielson. This paper boldly states right in the abstract that "the results of the (original) paper are a product of data inaccuracies" and that "the presented evidence does not support the stated conclusion" (Grimmer et al. 1).

Any good statistical analysis is only as good as its data set. That is where the root of the problem with the original paper lies, according to Grimmer, Hersh et al. In the response paper, they give three reasons for the CCES dataset being poorly suited for estimating state-level turnout. The first is that nationally representative surveys do not gather many responses from small states and minorities and thus these groups have a large error range (Grimmer et al. 2). Second, a person who lacks a voter ID is "unlikely to be accurately represented in the opt-in online CCES study" (Grimmer et al. 3). Third, there are inconsistencies "in the CCES vote validation process" that may "generate time-correlated measurement error in turnout estimates" (Grimmer et al. 3). The response paper then moves on to point out inconsistencies with the data in the original paper, the biggest being that their estimates of turnout by state often "deviate substantially from the truth" and are not properly corrected for since the error varies by state; some states were vast overestimates and some were vast underestimates (Grimmer et al. 4-5). It was at first surprising to me that the original paper missed these inconsistencies, but I realized that we are dealing with two potentially *huge* sources of error--the first is the error in the survey

data; the second is the error in the regression models used to compare ID laws to voter turnout. The original paper mentioned the first only briefly, and did not account for state-level variations.

Therefore, the goal of the response paper is to statistically show that the original paper extrapolated their results. The first thing they do is a placebo test meant to assess the plausibility that the relationship the original paper found between strict voter ID states and turnout is a result of other factors within the state more than the laws (Grimmer et al. 8). This placebo applies the original paper's regression models to turnout in the period before ID laws were enacted. The response paper finds that estimates from the placebo test are very similar to those in the original paper (Grimmer et al. 8). All in all, their conclusion is that the regressions "that comprise the central analysis in the study fail to adequately correct for omitted variable bias" (Grimmer et al. 12). Circling back to my original statement, an error-ridden data set is bound to produce error-ridden results.

Finally, this leaves me wondering how we can produce large, nationally representative surveys in a way that is actually representative of vast socioeconomic, racial, and geographic diversity (and more!). The problem is that getting more precise data requires more funding and resources. I do not blame the authors of the original paper for using this dataset--isn't it better to provide analytical results with the data we have rather than not produce it at all? Surely qualitative studies alone are not going to convince people that voter suppression is indeed a very serious and pressing issue! Thus, while strategies to produce better large data sets require more "financial investments and partnerships with governments," the stakes are high enough to justify the costs (Grimmer et al. 12). Sample bias can of course never be eliminated, but we can put the work in to reduce it.

So, given the low response rates in contemporary survey research, is it even possible to collect high-quality, nationally representative survey data? Without more accurate ways of portraying its effect on voter turnout, is voter suppression going to go left uncontrolled? These are questions I do not have any answers to right now, but simply are things I am thinking about. I have taken numerous anthropology courses and am currently applying to fellowships for next year, and have been thinking a lot about the differences between statistical research and ethnographic research in the social sciences. I think the two should serve to complement each other, with one method filling some gaps in the other. Data alone is not enough, but coupled with a solid qualitative analysis, even an error-ridden dataset can produce valuable results.

Work Cited:

- Justin Grimmer, Eitan Hersh, Marc Meredith, Jonathan Mummolo, Clayton Nallk, *"Comment on 'Voter Identification Laws and the Suppression of Minority Votes'"*, August 7, 2017, http://stanford.edu/~jgrimmer/comment_final.pdf
- Zoltan Hajnal, Nazita Lajevardi, Lindsay Nielson, *"Voter Identification Laws and the Suppression of Minority Votes,"* The Journal of Politics, volume 79, no. 2, January 5, 2017