PS1_Part2

Nora Nickels

4/15/2018

Problem Set 1

Part 2: Critique a computational research paper

State the research question of your assigned paper.

In this research paper, the authors are asking the following research questions: Are demographic differences in measures of political polarization consistent with the theory that the internet and social media play an important role in these measures? Specifically, how do trends in policital polarization relate to respondents' tendencies to obtain news or information online or from social media?

What data did the paper use?

The authors' primary sources of the data are from the American National Election Studies', or the ANES. Specifically, they include the 1948-2012 Time Series Cumulative, the 2008 Time Series Study, and the 2012 Time Series Study data sets. The authors describe the ANES as being a "nationally representative, face-to-face survey of the voting-age population that is conducted in both pre- and post- election rounds and contains numerous demographic variables and political measures" (Boxell et. al, 2017).

The authors also supplement the ANES data with microdata data from the Pew Research Center to include data on social media and internet use. The Pew surveys used are four surveys from 2005 - 2012 that focus on online dating, cloud computing, politics, adult social networking, apps and social media climate, and civic engagement.

What theory did the paper reference in order to interpret the data? (Note: it is possible that the paper has no reference to theory.)

This paper references two different topics of theory to interpret the data. First of all, they connect their research question to a focus of past theory and literature that attributes the trend of increasing polarization in America to the rise of social media and the internet in general. They reference authors what papers and figures, generally speaking, have connected this increasing political polarization to the internet based on the creation of echo chambers and the constant dissemination of both like-minded information and/or misinformation.

Further, the authors rely heavily upon theory and past work in order to compute their nine measures of political polarization that they use for their data exploration. The reconstruct measures of polarization that have been used in prior work, and reference all of the authors who have constructed these measures. The measures include information on measures of partisan affect polarization, ideological affect polarization, partisan sorting (extent to which partisan identity is correlated with self-reported ideology), straight ticket proportions, issue consistency, issue divergence, perceived partisan-ideology polarization, and religious polarization. These past measures allow the authors to create their overall index of polarization, and therefore are an important contribution from theory that allows the authors to interpret their data.

Was your assigned paper a descriptive study, an identification exercise, a numerical solution to system of equations study, or some combination of the three? (These are the three classifications we discussed in class.)

I would argue that that main classification of this paper would be a descriptive study. The paper combines data sets of polarization (AMES) to data sets of social media and internet use (Pew), and highlights interesting slices of the data set, using rigorous analysis by utilizing measures of political polarization to create a single polarization index. It highlights an unexplained relationship to focus on not whether polarization is increasing or how polarization should be conceptualized, but instead on demographics differences that exist in the measures used. The authors fully express that they are not seeking to measure or identify the increase of polarization or how it's conceptualized, but instead are focusing on dividing existing data information according to demographics predicting internet and social media use, and then to re-explore the data in that way. The index that is created from the models that define different aspects of polarization is certainly a strong numerical component of the study, but the solution itself involves just highlighting an interesting slice or demographic of a combined dataset.

What computational methods did this paper use to answer the research question? What was their result or answer to the question?

The computational methods used in this paper include the use of combining survey data, the use of numerical models to create an index of polarization, data visualization to explore trends in internet access and social media use by age group, trends in political polarization, and trends in polarization by demographic group, and statistical analysis (bootstrap standard errors) to show siginificant differences between age groups.

The result or answer to their question is that the increase in polarization is largest among groups that are least likely to use the internet and social media, with those groups being tied to older age. Specifically, the increase in polarization among those aged 18-39 is significantly less than the increase for those aged 65+ or for those aged 75+. They use these results to argue against the hypothesis that the internet or social media are the "main drivers of increasing polarization". They believe that their descriptive evidence rules out the most straight forward link between the internet and the increase in polarization.

Think of yourself as an academic referee. Give two suggestions to the author(s) of your assigned paper of things the authors might do to improve their results or strengthen their evidence for the answer to the question.

1. In the paper itself, could the authors discuss other aspects of demographics that have been tied to access to internet and social media use, and the defense for why any of these factors did not remain main predictors of interest? For example, are factors such as education or career tied to internet use or social media use in a similar way? It would benefit the paper if the authors would defend the connection of age as the focused demographic factor strongly tied to social media or internet use. Their appendix Table 1 lists all covariates that predict the dependent variable of "has internet", but the main predictor of focus of age is separate from the constructed index of predicted internet use based on the broader set of demographics.
2. Table 1 shows measures of polarization for all age groups, not just the focused ones of 18-39, 65+, and 75+. This table is essentially describing how most measures of polarization increase differently for different age groups. Could the authors visually show somehow that the differences between the youngest and oldest age groups are the most important, when compared to a comparison with 40 - 64 year olds? This table is important, as it basically leads to the significant difference in increase between the oldest and youngest groups (see results), but a figure such as a dot plot with error bars and each measures / age group on the axis, could help show the impotance of differences between the youngest and oldest age groups.

Reference

Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). Is the internet causing political polarization? Evidence from demographics (No. w23258). National Bureau of Economic Research.