Analysis Reflection Paper

I. Introduction: The effect of gerrymandering on the political ideology of the House of Representatives Members

The political ideology among members of the House of Representatives (HOR) is becoming increasingly polarised. In my research paper, I explore the correlation between gerrymandering and the political polarisation of the House of Representatives (HOR) members between 2010-2020 (using the 2010 redistricting cycle). Extreme ideological polarisation could threaten the very foundations of democracy, causing a deficiency in the representation of more neutral and minority views. I contribute to the existing literature on polarisation by contending that gerrymandering - the re-drawing of district lines for the protection and unfair advantage of incumbents – is a catalyst of political polarisation in the HOR. The correlation between gerrymandering and political polarisation of the HOR has been disregarded in the past due to a lack of outcome analysis. However, I hypothesise that in gerrymandered districts, members of the HOR obtain more extreme views to please their most vociferous voters, moving away from the median political ideology of their district to the median of their party. A comprehensive analysis of these theories should shed light on whether there is a correlation between gerrymandering and the political polarisation of HOR members.

II. Method 1: Statistical Analysis of district-wide Data

One way to approach this analysis is through a quantitative analysis of nominal data for numerous districts. Using two binary variables (gerrymandering in the 2010 cycle on the X axis and polarisation of the HOR members on the Y axis), I will carry out this analysis on multiple districts, half of which have been gerrymandered in the 2010 redistricting cycle, and

half which have not. To gauge whether a state has been gerrymandered, I deem an efficiency gap of over 7% to indicate a gerrymandered state (Stephanopoulos 2014). I will be choosing these districts by searching for states that have been subject to legal challenges, for example North Carolina and Maryland. To assess whether the HOR member holds polarised political ideology, I use the DW-Nominate score, in which a score of 0.3 or above indicates polarised ideology (Poole and Rosenthal 1985). With this statistical approach, I hope to observe a positive correlation for both the 'gerrymandered-polarised' and 'not gerrymandered-not polarised' outcomes, so that a significant correlation can be drawn between gerrymandering and the polarisation of political ideology of the HOR. The ideal type of data would be districts that not only suffer from an efficiency gap but other signs of gerrymandering such as compactness and packing and cracking. Real-world data sources that approximate these ideal features individually should be in abundance but finding data that embodies numerous aspects of gerrymandering might be more difficult. The simplicity of this method, with only four outcomes, makes it easy to interpret.

III. Method 2: Sentiment Analysis on District-wide Data

A further way to approach this analysis is through a modern natural language processing approach to count the occurrences per year of political terms related to either Republicans or Democrats (e.g., "Trump" for red and "Biden" for blue) between 2010 (after the redistricting cycle) - 2020. This data will be found on the internet, for example political news articles and video campaigns. Although this is a quantitative approach, the data I am modelling is qualitative. I will choose two districts, one that has been gerrymandered and one that has not and observe if over ten years the frequency of either side's words increase. I hope to observe an increase in frequency of the words in the gerrymandered district, to show a correlation between gerrymandering and political polarisation of the HOR members to please their more

extreme voters. In an ideal data set, there would be a predetermined list of political terms that are related to either Republicans or Democrats which serve as buzz words. Availability of resources is vast, from Government data sources to news archives. Large amounts of text data can be studied efficiently. This method would also be easy to interpret, and with the temporal nature of the study shows a change over time which would more strongly support my hypotheses.

IV. Limitations

Both methodological approaches to this analysis pose limitations. The statistical analysis can only produce four outcomes because of its binary structure, and so will not elucidate the extent of the connection between gerrymandering and political polarisation of the HOR member. Furthermore, there are significant external validity issues. Although there might be a correlation in the tested districts, the limited scope of my method means it is difficult to generalise my analysis. As I am choosing districts that are specifically significantly gerrymandered and districts that are not, there is an issue in representativeness in that it might not be reflective of other districts. Finally, although this data is easy to interpret, it oversimplifies both the terms 'gerrymandering' and 'polarisation,' by using only one measure. Both gerrymandering and political polarisation are complex issues with several dimensions. Simplifying them is assuming that one metric embodies both definitions.

There are also limitations for the sentiment analysis approach. Namely, there are internal validity issues such as the misinterpretation of sources, or inaccurately contextualising key words. If key words are used sarcastically, or in the context of opposition, they do not accurately represent the incumbent's ideology. In addition, there are limitations as to generalisability of the results, as this test only compares between two

districts. The time-consuming nature of this study makes it unfeasible to do on a larger scale, limiting the veracity of the results it yields. Finally, it is likely that there will be bias in the keywords chosen, making for inaccurate results. Because of the subjective nature of this study, it is impossible to neutrally carry out the study.

Overall, I think the quantitative statistical analysis is favourable to the sentiment analysis approach for multiple reasons. First, the data is easy to find and readily available. In comparison, it is difficult to find sources for the sentiment analysis approach, and even harder to process the data. Although setting the variables of gerrymandering and polarisation as binaries oversimplifies the complex nature of this issue, the results, regardless of metric would prove a correlation between the two.

Works Cited

Stephanopoulos, N. O (2014). The Measure of a Metric: The Debate over Quantifying Partisan Gerrymandering. Columbia Law Review, 114 (7), P1683-1742.

Poole, K. T., & Rosenthal, H (1985). A spatial model for legislative roll call analysis. American Journal of Political Science, 29 (2), 357-384.