**Cosette L. Hampton**

Assignment #9

MACS 30000, Dr. Evans

Due Friday, Dec. 14 at 5:30pm

# Referee report with extension (10 points; 3-5 pages long)

Paper Choice:

Gentzkow, Matthew, Jesse M. Shapiro, and Matt Taddy. “Measuring Group

Differences in High-dimensional Data: Method and Application to Congressional

Speech," NBER Working Paper 22423, National Bureau of Economic Research

<http://web.stanford.edu/gentzkow/research/politext.pdf>. August 2018.

The following paper is what MACS 30000 would consider a “referee report.” It is designed to comment on the validity, accuracy, relevance and quality of the paper, and also present possible options for an extension of the methods or topics that could lead to the formulation of another research question or topic in the future. In the “push” of this assignment in the folder “A9” I included the .pdf of the paper that contains a set of highlights:

* Yellow – general comments and points that help answer the questions (a)-(f),
* Red – misspellings and points that either are not necessary or not well-explained
* Green – points that help show that the paper is answering the necessary questions and providing the necessary statistics to either imply causation or statistical significance.

Please see the red highlights in the paper for answers to part (e) and citation issues to answer part of (d).

Gentzkow, et al. discover in “Measuring Group Differences in High-Dimensional Data” different norms and methodologies that make it easier for researchers to, “…characterize differences in behavior between groups [when] the space of possible choices is high-dimensional” (Gentzkow, 2018). Substantively, they found that in contrast with previous studies on political polarization, by 1994, members of Congress had begun to take on much more polarized political language. They state that, “…partisanship was low and roughly constant until the early 1990s, then exploded, reaching unprecedented heights in the recent years of our sample” (Gentzkow, 2018).

These findings come from their primary, compelling research question: Has the “average partisanship of speech” (Gentzkow, 2018) using penalized estimates changed over time? This may not have been explicitly stated in the text, but it is clearly inferred as the above quote comes from the title of Figure 3 which they consider their “main result.” Using both the visualization of Figure 3 and in the results section, they answer the research question in rigorous terms and with simpler terms in the introduction and repeating and summarizing the results section in the discussion section.

Other questions that come up throughout the paper that the authors attempt to answer either analytically or with empirical support are the following (Gentzkow, 2018):

* “Is today’s partisan language a new phenomenon?”
* “What caused the dramatic increase in the partisanship of speech?”
* “Does the growing partisanship of language matter?”

The first question is answered with the ongoing discussion of 1994 being an inflection point in the magnitude partisanship of language—namely, during this time “…partisanship turned sharply upward… [which] coincides precisely with the Republican takeover of Congress… Consistent with this, we show that phrases from the text of the Contract with America see a spike in usage in 1994, and then exhibit a particularly strong upward trend in partisanship” (Gentzkow, 2018). The remaining questions are analytically answered using some other coincidental findings and related texts, but they are not substantively proven to have results that imply causation or statistical significance. This is addressed in the document as some of those red highlights.

The methods of answering the research question are appropriate and sufficient. In the introduction section, they introduce various methods that relate to “structural estimation and machine learning” (Gentzkow, 2018) for exploring information around the partisanship of speech by members of Congress. They are:

1. Using all (the total set) of Congressional speeches during a Congressional session (instead of a selection of them, for example) to have a constant scene and source.
2. Estimating partisanship using a probability model that looks at the amount of speech by the speaker in the session, the probability of the speaker speaking a certain phrase, and speaker characteristics.
3. Estimating partisanship using a choice model that looks at the utility each speaker receives as they attempt to choose phrases that garner more support—these phrase probabilities “depend on speaker characteristics” (Gentzkow, 2018).
4. Using said estimates of partisanship of speech to look at “isolation,” which Gentzkow, et al. call “an affline transformation of partisanship” because basically they see how far away phrases stated by Speakers are from one another or like, polarization.

Some methods they discuss to improve their model in comparison to similar studies of segregation or polarization are:

1. Controlling bias of the maximum likelihood estimator (MLE) due to finite samples of phrases using “penalization,” or more specifically, a Poisson approximation
2. And, an “L1” penalty that helps to minimize a “Bayesian Information Criterion.”
3. Finally, they also use a placebo of phrases where parties are reassigned “…to speakers at random and then [the measure of partisanship] is re-estimate[d]” (Gentzkow, 2018).

These methods all help to account for biases that may have been present in Jensen’s previous iteration of similar research and similar research they cite on residential segregation, and lead to definitive, statistical, reproduceable and measurable statistical significance using coefficients, p-values and other qualitative measures common in sociological and experimental study. According to the lecture “Analysis with Experiments” of MACS Fall 2018, “Experiments intervene in the world to see if data changes and are ideal for answering cause and effect questions.”

The citations both at the end and throughout the paper show that it engages well with broader literature. The research itself uses Congressional speeches similar to residential inhabitance, work of (White 1886; Cutler et al. 1999) and Jensen et al.’s (2012) work on characterizing “…the partisanship of language from the late nineteenth century to the present” (Gentzkow, 2018). They say explicitly what papers relate to their work methodologically, analytically, and substantively, and they use approaches that they say other researchers would say, “performs well in related settings” (Gentzkow, 2018) like Taddy (2015). Mostly all of the citations they use seem fitting and necessary, though there are some spaces where those cited in the References section are only used in the footnotes and not in the actual text of the paper itself, so having the source in the footnote only may have removed some redundancy.

One important issue that I found within the paper is the lack of discussion around the selection and affirmation that an observer is “neutral.” At no point in the paper was the definition of a “neutral observer” given, though this is one of the main points of importance as this is the person who is set to determine whether or not a selected phrase is partisan or not. Readers are just expected to trust that there are no counterfactuals relative to this that may bias the findings. Another issue is that they state that, “To our knowledge, ours is the first paper to estimate group differences based on preference parameters in a structural model. It is also the first to use a penalization scheme to address the finite-sample bias arising in segregation measurement,” (Gentzkow, 2018). It is nice that the paper gives fresh, reproduceable methods for the research but I wonder if this is true, and if it is true, how another researcher can effectively evaluate the methodology.

Nevertheless, overall the paper has sufficient background research, relevant findings to issues today, and it both presents, and effectively answers the research question with methodologies that have previously proven to be successful and that are improved upon with new methods. These factors make the research valid, accurate, and of high quality, and in times where the political environment in the U.S. is very polarized, relevant as well.

In subsequent research, Gentzkow, et al. may want to define neutral observers and provide further qualitative findings around the reason for the inflection at 1994 and beyond. The analytical discussion at the end of the paper is persuasive, but “innovation in political persuasion,” “political marketing” and “changes in the media environment” (Gentzkow, 2018) could be proven to have causal effect with more rigorous research and applied methods as used to find fluctuation in Congressional speech partisanship. The research questions for these purposes would be:

1. What are innovations in political persuasion and how have they influenced partisanship in Congress?
2. How successful are political marketing strategies at polarizing discussions on a political topic? – This could be done as an experiment where you have a control group, and a treatment group that gets exposure to political marketing and observing changes in how they are polarized over time. And
3. How have changes in the media environment influenced partisanship? For this question, I may look at the data used in this study, and data from viewership and associated polls on issues, and look at similar groups takes on issues before Congressional hearings became public and Congresspeople spoke outside of sessions. I may also use an experimental approach where a treatment group has media exposure and maybe in a country where political discussion in the media is censored and looking at partisanship.