Perspectives on Computational Research
Problem Set #2
Alice Mee Seon Chung

Boxell, Gentzkow and Shapiro (2017) studied how trends in political polarization relate to respondent's propensities to obtain news or information online or from social media. They combine nine previously proposed measures to construct an index of political polarization among US adults.

The researchers used 3 primary sources of data and those are the American National Election Studies' (ANES) 1948–2012 Time Series Cumulative, 2008 Time Series Study, and 2012 Time Series Study data sets. The ANES is 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. For clarification and maintaining consistency through out whole data set, they exclude a sample in the 2012 ANES survey since it contains respondents who completed the survey online. ANES does not have information of social media usage, the authors supplement the data with microdata from the Pew Research Center to plot trends in social media use from 2005 to 2012.

To interpret the data, the authors combine three different data sets and used proportions instead of simple count method to draw graphs across years. The author sliced the data from 1996 since the internet question was first asked in 1996. The 3 graphs in figure 1 showed the trends that 18-39-years-old had higher level of internet access and social media use and the elderly (75+) had lower level across years. The authors referenced few papers that has introduced nine measures of polarization to interpret the data. The nine measures are (1) Partisan affect polarization, (2) Ideological affect polarization, (3) Partisan sorting, (4) Straight-ticket, (5) Issue consistency, (6) Issue divergence, (7) Partisan-ideology polarization, (8) Perceived partisan-ideology polarization and (9) Religious polarization.

More specific, the authors denotes S_t the set of all face-to-face survey respondents in year t who have a valid response for their age. P_i denotes a 7-point measure of individual i's party from strong Democrat (-3) to Independent (0) to strong Republican (3), and B_i denotes a 7-point measure of individual i's ideological affiliation from strong liberal (-3) to moderate (0) to strong conservative (3). The authors denote $R_t := \{i: P_i > 1\}$ and $R_t^L := \{i: P_i > 0\}$ denote the sets of Republicans excluding and including leaners respectively in time t. Same methods applied to define D_t and D_t^L. The authors define $C_t := \{i: B_i > 0\}$ and $L_t := \{i: B_i < 0\}$ denotes the sets of conservatives and liberals.

The authors collected these nine measures of polarization that have been used in past research and commonly used in these days. When reconstructing measures, the authors tried to follow the process as they proposed in original research and only deviate where explicitly stated.

To measure the differences in warmth or coldness respondents feel toward their own or opposite party or ideology, the authors follow Iyengar et al. (2012) and Gentzkow (2016) in using ANES thermometer ratings of parties and ideologies.

(1) *Partisan affect polarization* M_t^{AP}:

$$M_t^{AP} = \frac{1}{\sum_{i \in D_t^L} w_i} \sum_{i \in D_t^L} w_i \left(A_i^D - A_i^R\right) + \frac{1}{\sum_{i \in R_t^L} w_i} \sum_{i \in R_t^L} w_i \left(A_i^R - A_i^D\right)$$

where A_i^D and A_i^R denote individual i's feeling toward Democrats and Republicans, respectively on 0-100 scale. The higher values mean higher favorable feelings.

(2) *Ideological affect polarization* M_t^{AI}:

$$M_t^{AI} = \frac{1}{\sum_{i \in L_t} w_i} \sum_{i \in L_t} w_i \left(A_i^L - A_i^C\right) + \frac{1}{\sum_{i \in C_t} w} \sum_{i \in C_t} w_i \left(A_i^C - A_i^L\right)$$

where A_i^C and A_i^L denote individual i's feeling toward conservatives and liberals, respectively on 0-100 scale. The higher values mean higher favorable feelings.

To measure and capture the degree of partisan sorting, the authors used the measure developed by Mason (2015) and Davis and Dunway (2016).

(3) *Partisan sorting* $M_t^{Sorting}$:

$$M_t^{Sorting} = \frac{1}{\sum_{i \in S_t} w_i} \sum_{i \in S_t} \frac{w_i}{105} \left[g(|P_i - I_i| + 1)(|P_i| + 1)(|B_i| + 1) - 7\right]$$

where $g(x) = max_{i \in \cup_t S_t}(|P_i - I_i| + 1) + min_{i \in \cup_t S_t}(|P_i - I_i| + 1) - x$.

The authors follow Hetherington (2001) in examining the frequency with which individuals split their votes across in an election.

(4) *Straight-ticket* M_t^{ST} is the survey-weighted proportion of respondents who report voting for the same party in both presidential and House elections.

The authors used the measure proposed by Abramowitz and Saunders (2008) to examine the similarity of ideological positions across issues by letting k_i denote a conservative(1), moderate(0), or liberal(-1) response to $k \in K$, where K is a set of seven different questions.

(5) *Issue consistency* $M_t^{Consistency}$:

$$M_t^{Consistency} = \frac{1}{\sum_{i \in S_t} w_i} \sum_{i \in S_t} w_i \left| \sum_{k \in K} k_i \right|$$

(6) *Issue divergence* $M_t^{Divergence}$:

$$M_t^{Divergence} = \frac{1}{|K|} \sum_{k \in K} Cor_{i \in R_t^L \cup D_t^L}(\mathbf{1}_{P_i > 0}, \boldsymbol{k})$$

where *cor* calculates Kendall's tau between the party indicator and the question coding.

Partisan-ideology polarization examines the extent which the self-reported ideological affiliation of Republicans and Democrats differ. To measure this, the authors followed a measure outlined in Abramowitz and Saunders (2008).

(7) *Partisan-ideology polarization* M_t^{PI}:

$$M_t^{PI} = \frac{1}{\sum_{i \in R_t} w_i} \sum_{i \in R_t} B_i w_i - \frac{1}{\sum_{i \in D_t} w_i} \sum_{i \in D_t} B_i w_i$$

To measure the extent to which individuals *perceive* there to be ideological differences between the Republican and Democrat parties.

(8) *Perceived partisan-ideology polarization* $M_t^{Perceived}$:

$$M_t^{Perceived} = \frac{1}{\sum_{i \in S_t} w_i} \sum_{i \in S_t} w_i \left(\tilde{R}_i - \tilde{D}_i \right)$$

where $\tilde{R}_i$ and $\tilde{D}_i$ denote individual i's perception on how conservative the Republican or Democratic party is, respectively, on the 7-point liberal-to-conservative scale.

The authors construct the measure of religious polarization similar to Abramowitz and Saunders (2008).

(9) *Religious polarization* M_t^{Relig}:

$$M_t^{Relig} = \frac{1}{\sum_{i \in CG_t} w_i} \sum_{i \in CG_t} P_i w_i - \frac{1}{\sum_{i \in NC_t} w_i} \sum_{i \in NC_t} P_i w_i$$

where CG_t is the set of white respondents who go to church no less than almost every week and NC_t is the set of white respondents who go to church no more than a few times a year.

This paper mainly uses identification exercise and simple descriptive study on trends in internet and social media usage with nine measures. The author also used overall index of polarization as identification exercise to interpret data. Overall index of polarization defined by authors is,

$$Index_t = \frac{1}{|M|} \sum_{m \in M} m_t / m_{1996}$$

where M is the set of all nine polarization measures. The authors compute this index for different groups of respondents. The authors normalize the group level values based on the overall 1996 value m_{1996}.

With the nine measures and one index the authors analyzed trends by age group. The figure 2 plots showed these nine polarization measures and the overall index across 1972 to 2012. The first panel of figure 3 shows trends polarization index by age group. Polarization for the elderly (75+) and older group (65+) grows fast after 2000, however for the younger group (18-39) decreases after 2004. Table 1 provides quantitative information and standard errors for inference using a nonparametric bootstrap with 100

replicates. Appendix figure 1 shows trends in polarization by demographic groups, individual measures. The authors discovered that polarization increased for the older American demographic group in eight out of nine polarization measures when compared to younger American.

In addition to nine measures and overall index, the authors used a broad index of predicted and actual internet access to see the trends in polarization. The authors supposed

$$\Pr(internet_i = 1|X_i) = X_i'\beta$$

where β is a vector of parameters and X_i is a vector of demographic characteristics including indicators for age group, gender, race, education, and whether an individual lives in the political South. The authors used sample in 1996 using weighted least squares with weights w_i. The second panel of figure 3 plots the polarization index for respondents in the first and fourth quantiles of predicted internet access for each respective survey year. The fourth quantile group, which having greater likelihood of accessing to the internet, growth slowly in polarization from 1996 to 2012 when compared to first quantile group. The third panel of figure 3 plots the trends in polarization by actual internet access. The indexes of respondents with internet access have highest polarization in 2012. However, the trends of both respondent groups with internet access and without internet access are parallel across years from 1996 to 2012.

This paper used all nine measures of political polarization, overall index and the broad index of predicted and actual internet access introduced above as the computation methods to answer the research question. As I already introduced detailed information of each element and relevant concrete results in the paper, I omitted explanation at this part.

From interpretation of data, the authors found that the growth in polarization in recent years is largest for the demographic groups is least likely due to the internet usage and social media. This result argues against the commonly accepted hypothesis that the internet is a primary driver of rising political polarization in contemporary politics. However, the author's conclusion is not to say that the rise of digital technologies is not important. They implied that internet usage and social media has influenced in some extent to the political polarization.

As the author focused on the age groups in the paper, I would suggest to control age variable and test the relationship between internet access and political polarization index or trends of polarization in a certain age group. It is commonly accepted that age affects to the internet access, but the impact of internet usage on political polarization will be clearly shown when we controlled age variable. This change will strengthen the results of this paper.

As the authors mentioned in the paper, their findings are difficult to account linkage between the recent rise of polarization and the internet. To improve the results, I would suggest to study more on the older group's propensities and root causes regarding what other factors, sources and forces can explain the large increase in polarization in the group. People can change their views through the policy or individual financial status. It is commonly believed that increasing income inequality and unemployment rates are one of main reasons of political polarization. For elderly people instead of using internet, they

prefer to use non-internet media like talk radio, newspaper and TV to get information. It is possible that they largely influenced by non-internet medias. These extended findings can explain why the internet is not a main driver for rising political polarization for older Americans.

Third suggestion to strengthen the evidence is to combine recent data set. The data that the authors used goes until 2012 and the social media and internet grows and transforms rapidly within short periods. The dataset only captures the first few years of the personalization trend (e.g. introduction of like button in Facebook (2009), introduction of Twitter (2006) and ranking algorithm (2016)). To cover more recent evidences and trends, recent related dataset is needed. In addition, Twitter's role in presidential election also has become increasingly important than the past. Since the statistics from 2016 presidential election showed that Trump won the election on strength with whites, older voters, if the authors can combine 2016 election surveys to the paper, it will amplify and back up the evidences.

References

Abramowitz, Alan I. and Kyle L. Saunders. 2008. Is polarization a myth? The Journal of Politics. 70(2): 542–555.

American National Election Study. 2016. The ANES 2012 time series study. Stanford Universityand the University of Michigan. Accessed at
http://www.electionstudies.org/studypages/download/datacenter all datasets.php
on December 22, 2016.
—. 2015a. The ANES 1948–2012 time series cumulative data file. Stanford University and the University of Michigan. Accessed at
http://www.electionstudies.org/studypages/download/datacenter all datasets.php
on December22, 2016.
—. 2015b. The ANES 2008 time series study. Stanford University and the University of Michigan. Accessed at
http://www.electionstudies.org/studypages/download/datacenter all datasets.php
on Februrary 18, 2017.

Boxell, Levi and Gentzkow, Matthew and Shapiro, Jesse M. 2017. Is the Internet Causing Political Polarization? Evidence from Demographics. NBER Working Paper No. w23258. Available at SSRN: https://ssrn.com/abstract=2937528

Davis, Nicholas T. and Johanna L. Dunaway. 2016. Party polarization, media choice, and mass partisan-ideological sorting. Public Opinion Quarterly. 80(S1): 272–297.

Gentzkow, Matthew. 2016. Polarization in 2016. Toulouse Network for Information Technology Whitepaper. Accessed at
http://web.stanford.edu/_gentzkow/research/PolarizationIn2016.pdf
on February 19, 2017.

Hetherington, Marc J. 2001. Resurgent mass partisanship: The role of elite polarization. The American Political Science Review. 95(3): 619–631.

Iyengar, Shanto, Gaurav Sood, and Ypthach Lelkes. 2012. Affect, not ideology: A social identity perspective on polarization. Public Opinion Quarterly. 76(3): 405–431.

Mason, Lilliana. 2013. Replication data for: "'I disrespectfully agree': The differential effects of partisan sorting on behavioral and issue polarization." Harvard Dataverse, V3. Accessed at
https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/22346
on January 20, 2017.

Pew Research Center. 2005. September 2005 — Online dating survey. Accessed at
http://www.pewinternet.org/datasets/september-2005-online-dating/
on February 16, 2017.