Internet Use and Political Polarization

Sebastian Deri

July, 1945

“Consider a future device for individual use, which is a sort of mechanized private file and library …

… a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility”

Vanevvar Bush

March, 1960

“The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought …

… it seems likely that the contributions of human operators and their equipment will blend together so completely in many operations, that, it will be difficult to separate them neatly in analysis”

J.R. Licklider

August 30, 1965

"In our view, it is possible to build a large common-user communication network able to withstand heavy damage …

… the system is a marked departure from the existing communication system …"

F.R. Collbohm, President of RAND Corporation, in letter to Deputy Chief of Staff, Research and Development of U.S. Air Force, about the possibility of developing a distributed, digital communication system (as outlined by Paul Baran)

January 4, 1978

“The largest single surprise of the ARPANET program has been the incredible popularity and success of network mail. There is little doubt that the techniques of network mail developed in connection with the ARPANET program are going to sweep the country and drastically change the techniques used for intercommunication in the public and private sectors”

The Completion Report for ARPANET, Defense Funded Effort to Create Resilient Network for Sharing Computer Resources, which eventually formed the foundation of the Internet

1995

“Computing is not about computers anymore. It's about living.”

Nicholas Negroponte.

1999

“Similarly with the Net. A force of unimaginable power – a Leviathan, to use a Biblical (and Hobbesian phrase) – is loose in our world, and we are as yet barely aware of it. It is already changing the way we communicate, work, trade, entertain and learn; soon it will transform the ways we live and earn. Perhaps one day it will even change the way we think.”

John Naughton, in *A Brief History of the* Future, a book written on the history of the internet as things stoof at the turn of the millennium

October 10, 2017

“Today, the Chairman of the Congressional Black Caucus, Congressman Cedric L. Richmond (D-LA-02), released the following statement on recent reports that Russia purchased Facebook ads that targeted Black Lives Matter.

… We can’t conclusively say these actions impacted the outcome of the election. But we can say that these ads caused harm and additional resentment to young people who unselfishly fight for justice and equality for African Americans and other marginalized communities.”

Congressman Cedric L. Richmond, Chairman of the Congressional Black Caucus

October 27, 2017

“Facebook is what propelled Breitbart to a massive audience. We know its power.”

Steve Bannon, Chief Executive Officer for Trump 2016 Presidential Campaign and Former White House Chief Strategist for President Donald Trump

April 21, 2018

“The germs are ours, but Facebook is the wind.”

Harindra Dissanayake, Presidential Adviser in Sri Lanka

November 5, 2018

“We want Facebook to be a place where people can express themselves freely and safely around the world.

As part of that commitment, we commissioned an independent human rights impact assessment on the role of our services in Myanmar and today we are publishing the findings ...

The report concludes that, prior to this year, we weren’t doing enough to help prevent our platform from being used to foment division and incite offline violence. We agree that we can and should do more.”

Alex Warofka, Product Policy Manager, Facebook

Spring, 2018

There has been much discussion recently in both the popular press and in academic psychology about increasing belief polarization, especially political polarization, in the modern world. A sometimes-explicit and sometimes-implicit claim is that this increase has been aided and abetted by modern communication technologies. Has it? That is, what is the evidence that the beliefs and opinions a person encounters on the internet and through social media is more polarized and more “bubbly” than the beliefs and opinions a person encounters in daily, non-mediated social life? Furthermore, what is the evidence that (1) the U.S. population, as opposed to U.S. politicians, are more polarized now than they were previously, and (2) any such trend is due to modern communication technologies? (You can feel free to examine the same questions with respect to other countries if that is helpful, but you needn’t feel compelled to.) Finally, is there any reason to believe that information received through modern IT at a given level of “bubbliness” might be more or less influential in influencing people’s beliefs than the same information received through day-to-day social life?

Dr. Tom Gilovich

OVERVIEW

To answer this question, this response is organized into three major parts. In the first part, I examine whether there has indeed been a rise in political polarization in the US in the last several decades. The remaining second and third parts deal with the relationship between polarization and internet use. I begin, in the second part, by reviewing evidence pertaining to the question of whether internet use plays a causal role in bringing about polarization. I then move, in the third part, to exploring the possible means by which internet use might bring about polarization. By analogy to cigarettes and cancer, the second part examines *whether* cigarette smoking causes cancer, while the third part examines *how* cigarette smoking causes (or might cause) cancer. One focus, in the third section, is on the most often discussed mechanism of internet-caused polarization: segregated information exposure, which corresponds to claims that polarization is been driven by an internet ecosystem characterized by “echo chambers”, “filter bubbles”, and otherwise partisan information consumption and dissemination.

Word Count (starting after this page, and ending before Appendix): 23,313

OUTLINE

PART 1: HAS POLITICAL POLARIZATION BEEN RISING IN THE US?

In order to examine whether internet use contributes to increasing polarization—namely, political polarization among citizens in the US—it is first necessarily to establish whether political polarization in the US has indeed being increasing. This is where I will begin. I start by distinguishing the different ways polarization has been defined and measured. I then examine the evidence that polarization has been rising, according to each of these measures.

Definitions

As the conversation on political polarization broadens, what exactly is under discussion can get muddled. The single most useful and comprehensive resource for wading through that muddle is a white paper on political polarization by Stanford economist Matthew Gentzkow (2016). In it, Gentzkow delineates several ways in which political polarization has been measured and studied, which I rely heavily throughout, adding other notable metrics in the process.

A reasonable way of categorizing how polarization has been studied is to delineate the following groups: polarization as measured by (1) party identification, (2) political ideology, (3) voting behavior, (4) opinions on policy issues, (5) alignment between policy positions and political identification ("party sorting"), or (6) attitudes towards those who have the opposite political affiliation (Gentzkow, 2016). I now unpack these in more detail.

party identification (definition)

As Gentzkow (2016) points out, party identification is perhaps the simplest and most straightforward way that one might assess political polarization. In studies that focus on this measure (e.g. Miller, 1991), researchers analyze changes in the extent to which people identify with the two major U.S. political parties—the Republican Party and the Democratic Party. Party identification is measured along one-dimensional continuum, usually in the form of a 5-point scale, where people are asked to place themselves into one of five boxes, for example: Republican, Lean Republican, Independent, Lean Democrat, and Democrat. An increase in polarization would be evinced by a greater proportion of people shifting to the endpoints ("poles") of the scale.

political ideology (definition)

Closely related to party identification, but distinguishable from it (e.g. Schiffer, 2000), is political ideology. Studies of this kind (e.g. Abramowitz & Saunders, 2008) focus not on people's party loyalties but on their social beliefs systems (Jost, Federico, & Napier, 2009). Commonly and most simply, this also often measured along one-dimensional continuum in the form of 5 or 7-point scale, where people are asked to categorize the extent to which they are "conservative" or "liberal", for example: Very Conservative, Conservative, Moderate/Centrist, Liberal, Very Liberal. Again, an increase in polarization would be evinced in a shift to the poles of the scale.

voting behavior (definition)

Gentzkow (2016) also points to less psychological measure of polarization which rely on actual voting behavior. Voting behavior can be analyzed to assess the extent of polarization in at least two main ways. One focuses on geographic units, such as counties, and then examines, the changes in the distribution of votes in these areas. (Sometimes this is referred to as “geographic sorting”, e.g. Fiorina & Abrams, 2008.) In these analyses, the proportion of highly skewed areas is examined for evidence of polarization. For example, some studies look at how the number of “landslide counties” has changed over time (that is, the number of counties where the winning candidate receives a huge majority of the vote) (Bishop, 2009). Evidence of polarization here would come in the form of an increase in the number of areas with disproportionate vote totals (e.g. landslide countries). The other way of assessing polarization through voting behavior examines patterns of voting across a ballot, across time, and across levels of voting (e.g. votes in congressional elections v. votes in presidential elections). Studies that look across the ballot, commonly examine how likely voters are to “split tickets” (that is, for one person to vote for candidates of different parties on the same ballot) (Hetherington, 2001; Mayer, 1998). Studies that look across time, examine how likely voters or areas are to switch their party allegiances across elections. Studies that look at patterns across levels of voting examine, for example, the correlation between the percentage of votes that presidential candidates receive in a county and the percentage of votes that congressional candidates receive in a county (Bartels, 1998; Fleisher & Bond, 2004; Jacobson, 2000, 2003). Evidence of polarization in these cases would come in the form of a decrease in the proportion of split ballots, a decrease in the extent that voters vote for different candidates across time, and an increase in the correlation between voting at different levels (e.g. increase in the correlation between presidential and congressional votes).

One important concern with these measures that must immediately be pointed out is that they entangle voters’ political preferences with candidate options (that is, they conflate “popular polarization” with “elite polarization”; Fiorina & Abrams, 2008). Assume that some proportion of the population is “moderate”; that is, they lean towards one party, but may vote for a member of the other party in certain circumstances. Without any change in the underlying distribution of people’s beliefs, a change in the candidates offered by each party could result in more polarized voting behavior, under any of the measures defined above. Imagine that one or both parties start selecting and nominating candidates who are more ideologically extreme. Such a selection magnifies the leanings of voters, even if the voters’ underlying leanings don’t change at all. Democrat leaning moderates, who may have voted for a Republican in certain circumstances, are now less likely to cross over if only extremely conservative Republican candidates are ever offered. The same applies in the other direction as well. This may increase the proportion of people living in landslide counties, and it could increase the correlation between voting at two different levels (e.g. between President and Congress) as long as the trend is taking place at both levels.

policy positions (definition)

Polarization may also be revealed by measuring attitudes on key policy issues, such as health care, immigration, and gun control. Under this method, as (Gentzkow, 2016) rightly points out, there are at least two ways to assess whether polarization has taken place. First, as Fiorina & Abrams (2008) contend, polarization may be evinced by a change in the shape of the overall, aggregated distribution of attitudes on policy issues, from one that is unimodal to one that is more bimodal (as illustrated in Figure 1). This looks at the views of all partisans together. However, another way that one may assess polarization with regard to policy issues, is to first separate people into political camps (i.e. Republicans and Democrats) and assess whether the positions of members of these parties diverge over time. These two measures of polarization need not yield the same results (because of party sorting, which is explained just below).

Figure 1: Polarization as defined by Fiorina & Abrams (2008)

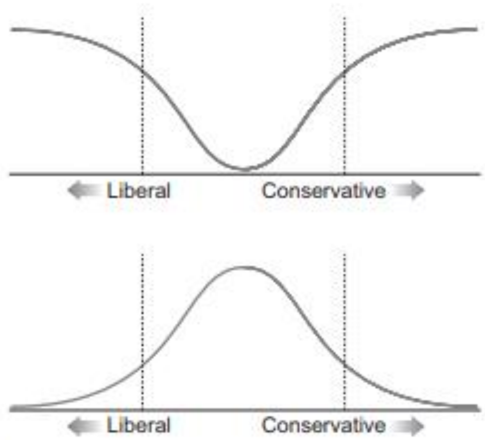


Figure 1
Polarized and nonpolarized distributions.

party sorting

Polarization can also be identified through the presence of party sorting (also sometimes called "social sorting"), which looks at the alignment between policy positions (e.g. views on issues like taxation, abortion, and immigration) and one's party identification (i.e. Republican-Democratic) or political ideology (i.e. conservative-liberal), as well as the general alignment between party identification and political ideology. Studies on party sorting**,** examine things like the size of the correlation between policy positions and political ideology over time (e.g. Abramowitz & Saunders, 2008). Under this type of metric, an increase in polarization would manifest as an increase in the size of the correlation between policy positions and political ideology over time. That is, policy views become more strongly linked to political identity.

An increase in the alignment between policy positions and party identity could occur in such a way that the overall distribution of policy views is unchanged, but their apportionment among the parties changes. As a simplified case, imagine 60 out of 100 Republicans support Immigration Policy X, while only 40 out of 100 Democrats do. An increase in party alignment could make it so that 80 out of Republicans support Immigrant Policy X and only 20 out of the 100 Democrats do. In both cases, the overall distribution is the same (100 of the 200 people support Immigration Policy X), yet the increase in party sorting has made it so that the parties have moved farther apart on the issues. It is in this way (alluded to at the end of the previous section on policy position measures of polarization) that an increase in party sorting could reveal evidence on some of the policy position measures of polarization (e.g. when looking at the divergence of views on policies between parties), but not others (e.g. when looking at the shape of the overall distribution of views).[1]

affective polarization (definition)

Yet another way polarization can be assessed is by looking at the attitudes and feelings that members of different political parties or camps have towards themselves and each other. Studies on affective polarization (e.g. Iyengar, Lelkes, Levendusky, Malhotra, & Westwood, 2019) commonly examine items like the "feeling thermometer" from the American National Election Study ("American National Election Studies," 2018). This question asks Republicans and Democrats to rate how they feel about members of their own party and members of the other party from "very cold" to "very warm." Other studies ask people to make other interpersonal evaluations of party members—such as how intelligent or selfish they seem, or how much they would trust them with various decisions. Additional studies even assess people's implicit evaluations of in-party and out-party members, (e.g. Iyengar & Westwood, 2015, who use a political IAT). Evidence of polarization on such metrics would come in the form of increasingly hostile feelings towards members of the other party, and increasingly large gap between the feelings that party members have towards members of their party and the feeling they have towards members of the opposite party.

Evidence of Polarization

So what does the evidence show? Having laid out some of the most common methods of assessing polarization, I now examine the evidence regarding whether polarization has been increasing or decreasing under each of them.

party identification (evidence of polarization)

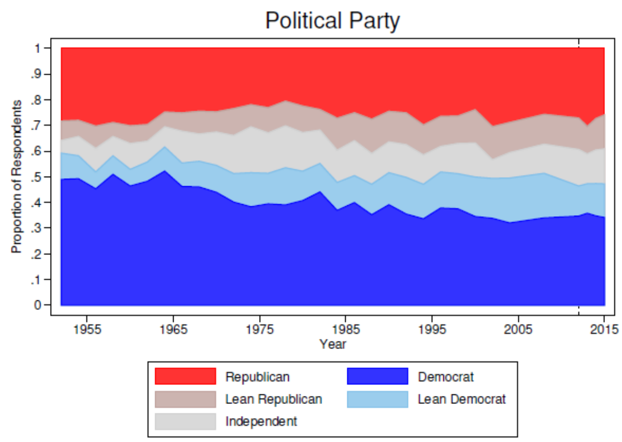
On measures of party identification and ideology, there is not much evidence of polarization. Gentzkow (2016) summarizes data from the American National Election Survey (see Figure 2), which has asked political identity questions in a consistent way since 1948. As can be seen, there is little evidence that of a rise in the proportion of people occupying the scale's poles (if anything there is an appearance of the widening of the middle).

That these questions have stayed consistent over time allows for more meaningful comparisons over time. Although it may be worth noting that caveat that while the literal questions have stayed constant over time, the meaning of them may not have. For example, if the policies of the parties have become more extreme, it might mean something more extreme to say that one is "Republican" in 2018 than to have said the same in 1960. So, if the meaning of the ends of the continuum have come to have a more extreme connotation over time, the fact that the

[1] In this very simple example, I relied on a binomial distribution (i.e. people either agree or disagree with the policy), which can't "really" be bimodal. But the core of the intuition remains the same in other cases. One need just imagine that case of a policy where support can be expressed in some continuous manner (e.g. ratings on a 1-7 scale) and the same phenomenon can emerge (overall distribution remains unchanged, but distributions among parties becomes more extreme because of great alignment between party and policy).

percentage in each has stayed the same, could indeed evidence of polarization. This is not to say that this has in fact happened, but just to note that such a nuance is possible, complicating the interpretation of even such a simple measure of polarization.

Figure 2: Polarization as measured by stated party identification (from Gentzkow, 2016)

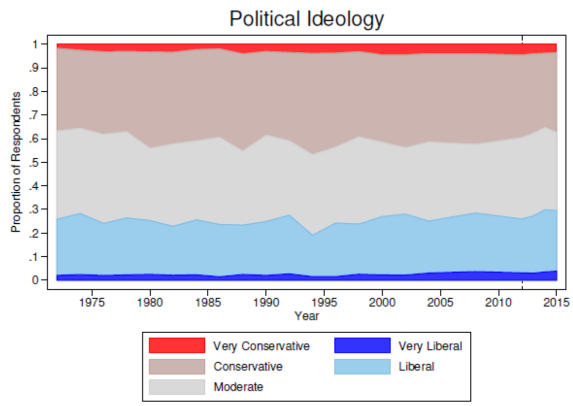


Notes: Figure shows the proportion of respondents to the American National Election Study survey who identify as Republican, lean Republican, identify as Independent, lean Democrat, or identify as Democrat. The post-2012 data comes from a separate survey conducted by the Pew Research Center and is rescaled in such a way that the overlapping time periods have the same mean.

political ideology (evidence of polarization)

The results for polarization as measured by political ideology look much the same as the results for party identification. We can again see this in data from the American National Election Survey, summarized by Gentzkow (2016). As can be seen in Figure 3, there is little evidence of a major shift to the poles in reported ideology over time. The same caveat as for stated party identification should be noted—that the meaning of each identity *may* have shifted over time.

Figure 3: Polarization as measured by stated ideology (from Gentzkow, 2016)



Notes: Figure shows the proportion of respondents to the American National Election Study survey who identify as very conservative, conservative, moderate, liberal, and very liberal. The post-2012 data comes from a separate survey conducted by the Pew Research Center and is rescaled in such a way that the overlapping time periods have the same mean.

voting behavior (evidence of polarization)

Analysis of voting behavior starts to reveal more evidence of polarization. In Gentzkow's (2016) summary, there is evidence in both the geographic and voting pattern data of polarization. Geographically, the data suggest the number of voters living in landslide counties has increased from 1976 to 2004 (Bishop, 2009). And patterns of voting reveal evidence that (1) voters have become less likely to split tickets (Hetherington, 2001; Mayer, 1998), (2) vote shares in counties have become more correlated over time, and (3) at the county level, the percentage of votes that presidential candidates receive and congressional candidates receive has become more correlated over time (Bartels, 1998; Fleisher & Bond, 2004; Jacobson, 2000, 2003).

Although, as Gentzkow (2016) also points out, there is some debate about the size and strength of these patterns. Glaeser & Ward (2006), for example, present evidence that the current correlation between vote shares is actually pretty close to historical averages over the last hundred years—and that there was more cross party voting during the 60s and 70s when the South switched from being Democratic to Republican. And perhaps now we are simply returning to baseline.

Further, as mentioned before, Fiorina & Abrams (2008) make clear the necessity of distinguishing between “mass” and “elite” polarization, when looking at polarization in voting behavior. Indeed, there is much evidence of “elite” polarization—which has been primarily studied at the U.S. Congressional level—although there is some debate in the literature about exact extent. Thus, we cannot unequivocally conclude from evidence of polarization in voting behavior that the US citizens, in mass, are polarized, as such polarization in voting patterns might simply be the result of elite polarization. We must look elsewhere for more clear-cut evidence.

policy positions (evidence of polarization)

Whether there is evidence of rising polarization in people's policy positions depends on how you look.

evidence against polarization

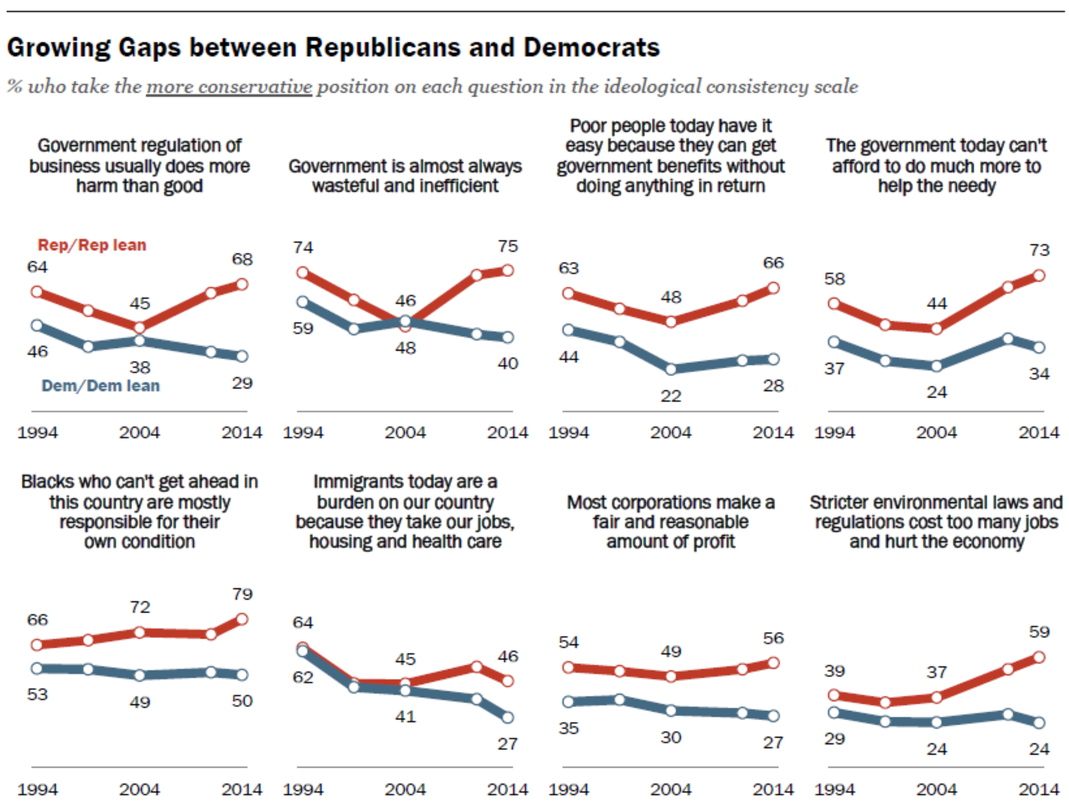
Under the method that examines the shape and nature of the overall, aggregated distribution of views on policy positions, there is little evidence of a shift from unimodal distribution to a bimodal one. Fiorina & Abrams (2008) find no strong evidence for polarization in their review of research—aggregated distributions of views remain unimodal. And as Gentzkow (2016) notes, an analysis from Ansolabehere, Rodden, & Snyder (2006) that looks at relevant data up until the 1990s similarly finds no evidence of polarization on economic issues, and only a slight hint of evidence in favor of polarization on moral issues[2].

evidence for polarization

However, we see quite another pattern when separating and comparing the policy opinions of Democrats and Republicans. This can be seen in nationally representative survey data from Pew, that Gentzkow (2016) compiles. As we can see in Figure 4A, members of both parties have been moving away from each other on key policy issues like immigration, corporate taxation, and protecting the environment (since 1994, in the data shown). And the same pattern is evidence when looking at aggregated differences between political values

[2] While economic issues may be less polarized than other issues, note also the time period of study of the economic issues, which goes up until the 1990s. Polarization may have simply increased after that time. For example, the term "culture war" first rose to national prominence with James Davis Hunter's 1991 publication of *Culture Wars: The Struggle to Define America*, which was followed then by Pat Buchanan's use of the term in his speech at the 1992 Republican convention. This was followed by the 1994 Congressional elections which gave Republicans a majority in the House of Representatives for the first time in four decades, in a campaign that ostensibly focused on economic issues in the Newt Gingrich's "Contract with America" (although some argue this victory was also aided in part by increased focus on cultural "wedge issues" like abortion ("How 1994 Gave Us Today's Politics," 2018)).

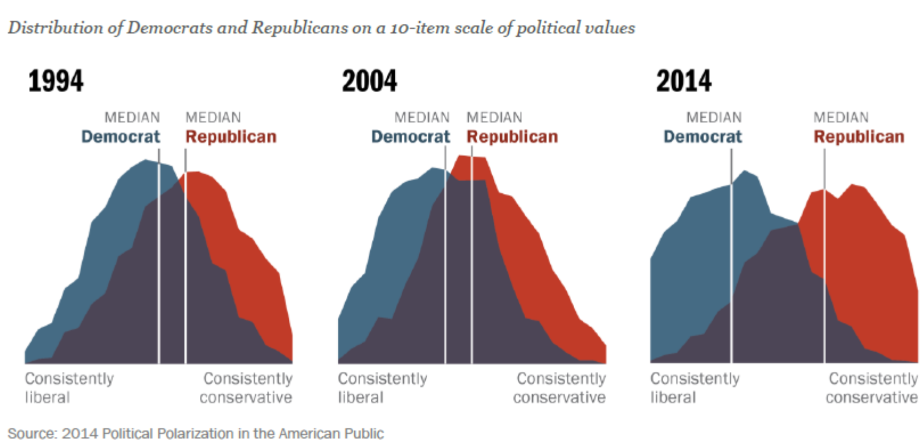
Figure 4A: Evidence of Policy Polarization from Pew



Source: Pew Research Center (2014).

In addition to the tighter linking between party identification and policy preferences, we can also see (in Figure 4B) a tighter linking over time between party identification and general political values, in Pew data compiled by Gentzkow (2016). views.

Figure 4B: Further Evidence of Policy Polarization, from PEW



Source: Pew Research Center (2014).

How can it be that there is no evidence of the overall distribution on policy positions becoming more bimodal, yet Democrats and Republicans appear to be moving further apart on political issues? The most compelling explanation, the possibility alluded to earlier, seems to be that there has been an increase in party sorting—the correlation between political party and policy positions. While the overall distribution on issues (e.g. abortion) may not have changed, the parties have moved farther part on these issues—because policy positions have become increasingly associated with party identification.

party sorting (evidence of polarization)

There is indeed evidence that party sorting has increased over time. Abramowitz & Saunders (2008), for example, show that policy positions and party identification have become more linked over time (Figure 5), as have party identification and political ideology (Figure 6).

Figure 5: Correlation Between Party Identification and Policy Preferences over Time

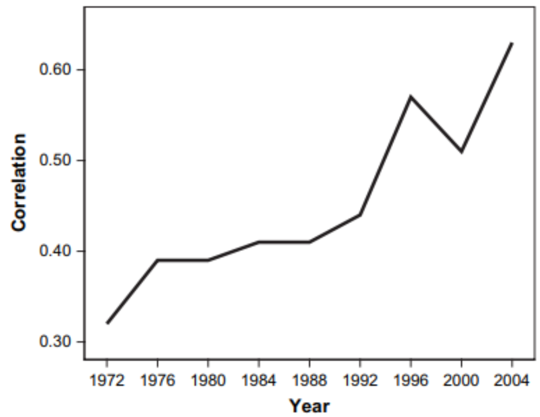
TABLE 3 Trends in Partisan Polarization on Issues, 1972-2004

Issue	1972-1980	1984-1992	1996-2004
Aid to Blacks	.20	.27	.35
Abortion	−.03	.08	.18
Jobs/Living Standards	.28	.34	.40
Health Insurance	.25	.31	.39
Lib/Con Id	.42	.49	.62
Presidential Approval	.42	.56	.61
Average	.26	.34	.43

Note: Entries shown are average correlations (Kendall's tau) between issues and party identification (strong, weak, and independent Democrats vs. strong weak and independent Republicans).
Source: American National Election Studies

Figure 6: Correlation Between Party Identification and Political Ideology Over Time

FIGURE 3 Correlation of Party Identification with Liberal-Conservative Identification, 1972-2004



Note: Correlation coefficient is Pearson's r based on 7-point party Identification scale and 7-point liberal-conservative identification scale.

Source: American National Election Studies

Reviewing evidence on party sorting, Fiorina & Abrams (2008) also conclude that the evidence suggests party sorting has increased over time—and contend that the debate is simply over the matter of degree.

Others go even further, arguing that political ideology and party identification are becoming bound with an increasingly wide array of social and demographic aspects of identity. A strong proponent of this claim is Lilliana Mason (2018), who puts forward the following claim.

> "The American political parties are growing socially polarized. Religion and race, as well as class, geography, and culture, are dividing the parties in such a way that the effect of party identity is magnified. The competition is no longer between only Democrats and Republicans. A single vote can now indicate a person's partisan preference as well as his or her religion, race, ethnicity, gender, neighborhood, and favorite grocery store. This is no longer a single social identity. Partisanship can now be thought of as a mega-identity, with all the psychological and behavioral magnifications that implies. American citizens currently believe they are in a partisan competition against a socially homogenous group of outsiders." (p. 14)

As evidence, she cites ANES data showing that over time party identification has been become more linked not only to political ideology but also other aspects of identity like race and economic status.

Figure 7: Links Between Party Identification and Aspects of Social Identity

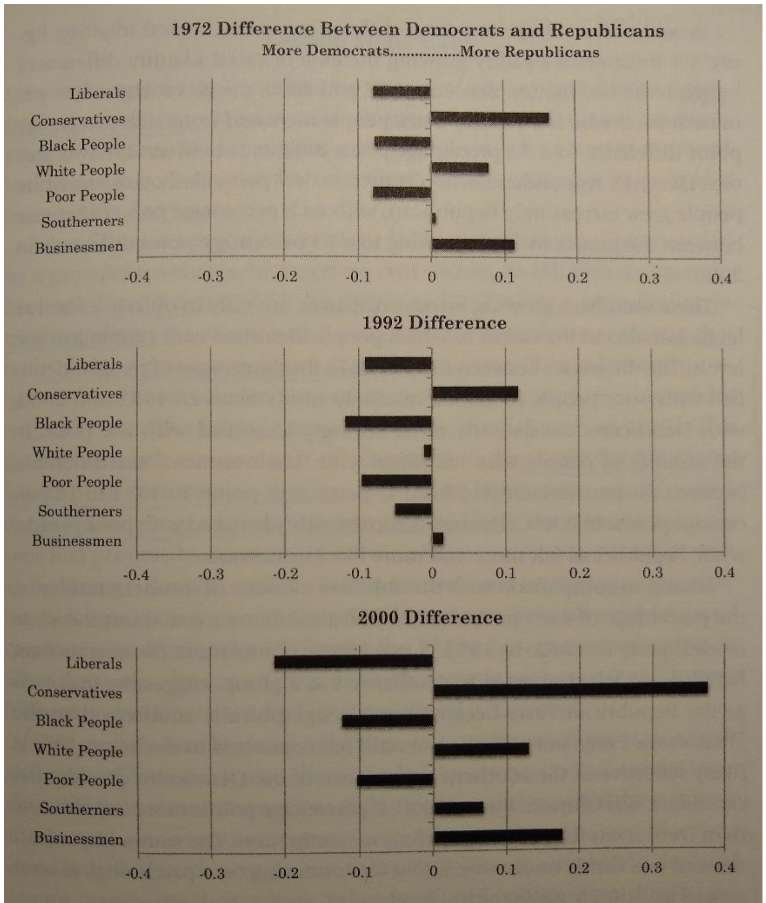


Figure 3.5. Identity-based social sorting
Note: Bars represent the difference between the percentage of Republicans who claimed to feel identified with each group minus the percentage of Democrats who did so (including partisan leaners). Data drawn from the ANES data files of 1972, 1992, and 2000.

The net effect of this, she claims is ever growing negative sentiments and feelings of competition between members of the two parties

affective polarization (evidence of polarization)

And indeed, the starkest evidence for polarization comes may come from an examination of the attitudes and feelings that members of the two parties have toward each other. We can see this, for example, in ANES feeling thermometer ratings. It is not a surprise that party members might feel more warmly towards members of their own party than towards members of the other party. However, as Iyengar et al. (2019) show, this gap has been growing since the 1990s, and is driven in particular by increasingly negative feelings towards out-party members (Figure 8).

Figure 8: Feeling Thermometer Ratings, Aggregated Across Parties

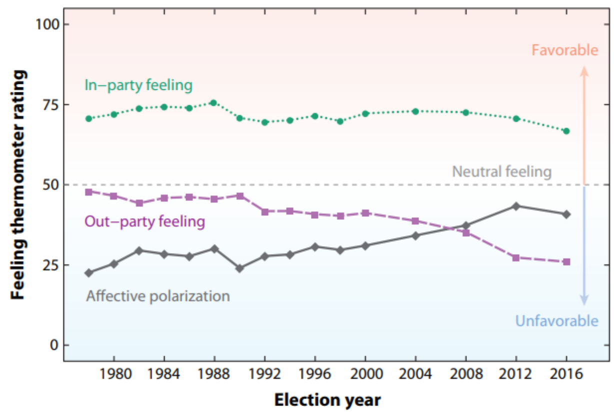
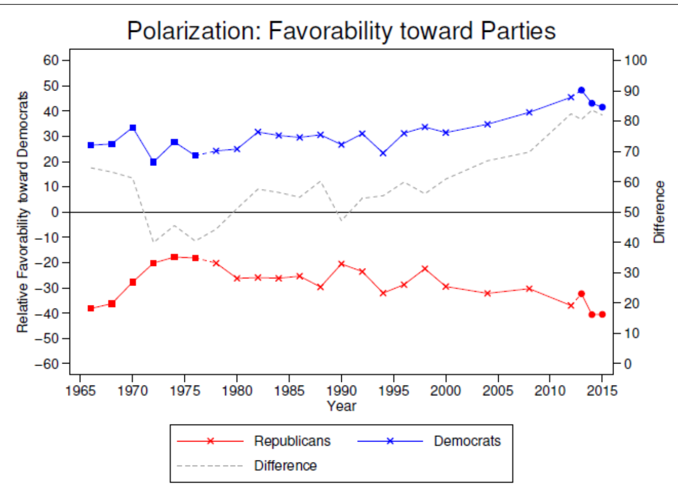


Figure 1

Using data from the American National Election Study (ANES), the figure shows trends in average feeling for the party participants identify with (in-party) and for the opposing party (out-party). In-party feeling (*green line*) has remained high over the period plotted, though it has decreased slightly in recent years. Out-party feeling (*purple line*) has decreased dramatically, especially after 1990. We also plot affective polarization (*gray line*)—the difference between mean in-party feeling and mean out-party feeling—which shows a significant increase over time (Iyengar et al. 2012), mainly due to an increase in animus against the out-party. However, it is worth noting that affective polarization actually decreased between 2012 and 2016 due to decreases in feeling toward the in-party.

In Gentzkow's (2016) analysis of this data, we can see that this trend holds for members of both political parties. Democrats have been increasing in the extent to which they feel warmer towards Democrats than Republicans, and likewise Republicans have been increasing in the extent to which they feel warmer towards Republicans than Democrats (Figure 9).

Figure 9: Feeling Thermometer Rating Difference, Separated by Party



Notes: Using data from the American National Election Study, the red line shows the average favorability of Republicans towards Democrats minus the average favorabilty of Republicans towards Republicans on a scale from 0-100. The blue line shows an analogous time series for Democrats. The grey line plots the diference between the blue and red lines over time. The National Election Study changed the phrasing of their question over time. The square points represent the original phrasing, the x points represent the revised version of the question, and the circle points represent a similar question used in monthly political surveys by the Pew Research Center. The different time series are then rescaled so that the average values of the overlapping time periods are equivalent.

There is also evidence that Democrats and Republicans make more negative trait evaluations of each other than they once did. Gentzkow (2016) compares intelligence and competence ratings made by Democrats and Republicans of both their own party members and opposite party members from a study by Almond & Verba (1960) to the same ratings made in a 2008 YouGov poll. As he shows, across both parties, party members now rate their own party members as more intelligent than they used to, and rate members of the opposite party as less intelligent than they used to—thus increasing the gap between how intelligent partisans view in-party members and out-party members. Another comparison that can be made from the Almond & Verba (1960) and YouGov data concerns attitudes towards intermingling between party members. In both studies, Democrat and Republicans were asked if they would be displeased if their child married a member of the opposite party. For both parties, the portion of people saying they would be displeased moved up from below 5% in the 1960s, to at or above 20% in 2008. Although it does not allow for a temporal comparison, another statistic that speaks to the extent to which party members currently evaluate each other as incompetent comes from a study by (Pew Research Center, 2014, p. 100), which finds that around 27% of Democrats and 36% of Republicans believe that members of the opposing party "are so misguided that they threaten the nation's well-being."

Summary

So has polarization increased or not? As we can see, it depends what you measure. While there is little evidence of a moving towards the "poles" in terms of party identification or political ideology, there is evidence of polarization when it is measured in other ways. Voting has become more partisan and homogenous. Differences in Republican and Democrat policy preferences have widened. Links between party and policy preferences (and perhaps social identity more broadly) have tightened. And bad feelings between parties have intensified.

Regardless of how any definitional debate could be decided, in the realm of political life in the US, there does seem to be something to be worried about. Widening policy disagreement, rising animus, and increasing ideological division between Republicans and Democrats, or conservatives and liberals, are ample causes for concern.

PART 2: DOES THE INTERNET CAUSE POLITICAL POLARIZATION?

Our next question is about the role that the internet might play in bringing about the negative political outcomes previously documented (which we might as well continue to refer to as "political polarization"). To answer this question, it seems like we need to get a handle on two things. *Does* the internet cause political polarization? And *how* does (or *how* could) the internet cause political polarization? The section will be solely concerned with answering this *does* question.

Of course, the "internet" is not one thing. It consists, physically, of myriad laptops, desktops, cell phones, routers, data centers, underground, above-ground, and undersea fiber optic cables, as well as all other kinds of computing and communications machinery, not to mention the people who maintain that machinery. It is used by many people and many institutions for many things—buying scarves, keeping up with relatives, watching news videos, and indeed organizing events aimed at reducing political polarization. Asking whether the internet causes polarization may sound a little like asking whether telephones cause racism. Things become clearer, however, when we recognize the internet as technological innovation which can and has altered both societal and individual level behavior. It makes sense to wonder about the net effect of this behavior change, brought about by internet use, on social outcomes, like political polarization. Our focus then is really on *internet use*—which, for a given person over a given time scale, can be thought of as the accumulated differences in their behavior when using the internet, considered against the counterfactual of what they would otherwise be doing in that same situation if they were not using the internet.[3] Our question, more precisely, is about aggregated internet use and whether *more* internet use—averaging across "all" people and "all" uses—has, on net, the effect of resulting in more political polarization.

Empirically, there are two main ways to broach this question. One is through retrospective, longitudinal, or correlation-type analyses, that examine how rises in internet use over time track rises in polarization over time, while attempting to rule out confounds. This can be thought, in a sense, as asking whether there is historical evidence of a dose-response

[3] Of course, its hard to grasp the exact nature of what such a counterfactual encompasses. For example, imagine someone who is browsing articles on www.nytimes.com while on a line to get into a Dave Chapelle comedy show they found out about earlier in the week through a post in their Facebook Newsfeed from Dave Chapelle's Facebook page, which they started following earlier in the year after binging on YouTube videos of Chapelle's earlier standup acts. One possible counterfactual is to imagine what that person would be doing in that same situation on the line for the comedy show if they weren't using the internet to browse news articles on their phone. Bu we can also enlarge the counterfactual and imagine what the person might be doing if they hadn't used the internet all week. Would they even be at the comedy show? And larger yet, if they hadn't used the internet in the last year would they even know about or have become interested in Dave Chapelle at all? In this paper, the implicit assumption that I make throughout is that we are considering "smaller" counterfactuals—which can be thought of as the absence of the internet on smaller time scales, and applied to cases of individual people, not accounting for inevitable network effects. "Larger" counterfactuals, especially on the societal scale, would seem to have more drastic and unpredictable consequences. Consider the most drastic of all counterfactuals for our purposes—that the internet doesn't exist or never come about at all. You've never skyped, emailed, or been enraged by a tweet. The New York Times never reported on Cambridge Analytica scandal because Facebook never existed. There were no protests about the location of Amazon's headquarters because Amazon never existed. It seems likely that the course of social and political history would be unrecognizably altered. Quite possibly things like Donald Trump's election to the presidency and Barack Obama's election would have never come about, with very likely differences in political polarization trends over time.

relationship between amount of internet use and amount of political polarization. These studies examine internet use over larger time scales (i.e. years), but have a harder time providing definitive conclusions about causality. The other empirical strategy is direct intervention—that is, manipulating people's amount of internet use and observing subsequent changes in political polarization. These studies examine internet use over smaller time scales (e.g. weeks), but are much more definitive in terms of causality.

Retrospective Analyses

The most clear and informative paper examining the causal relationship between internet use and political polarization in a retrospective, correlational fashion comes from Boxell, Gentzkow, & Shapiro (2017). Their idea is quite simple. Some people spend more time on the internet than others. If internet use in some way contributes to polarization, then we should see a steeper increase in polarization among those people who spend more time on the internet.

To examine this question, the authors use various survey data from the American National Election Studies (ANES) from 1982 to 2016, as well as survey data from Pew Research Center from 1996 (when Pew first started asking about internet use) to 2016. To account for the various different ways of operationalizing polarization, the authors compute eight different measures of polarization from their survey data.[4] They then normalize and average these into one overall polarization metric.

Most of the analysis focuses on the demographic variable most strongly associated with different amounts of use of the internet—age. Internet use meanwhile can be measured in several ways. But as we can see in Figure 10, across the three primary measures of internet use available in the authors data[5], younger people use the internet more than older people on average. Across all age groups, internet use has been rising over time and at seemingly similar rates.

Figure 10: Internet Use by Age Groups

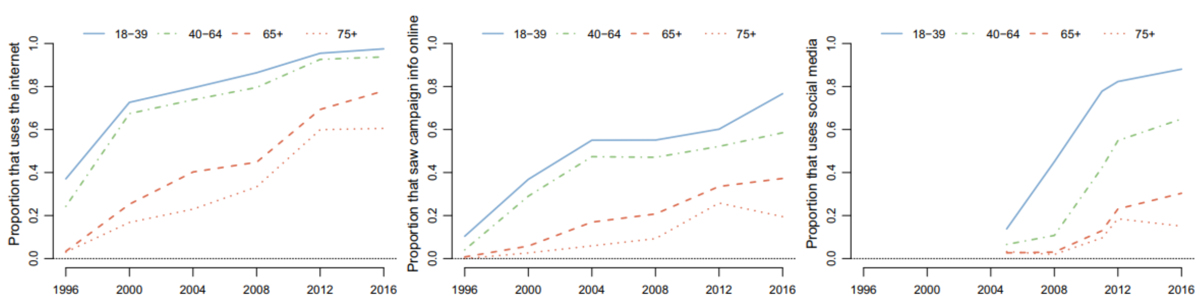


Fig. 2. Trends in Internet and social media use by age group. Each plot shows trends in Internet or social media use by age group. *Left* shows the weighted proportion of respondents that use the Internet by age group, using data from the ANES. *Center* shows the weighted proportion of respondents that obtained campaign information online by age group, using data from the ANES. *Right* shows the weighted proportion of respondents that use social media by age group, using data from the Pew Research Center. See *SI Appendix*, section 1 for details on variable construction.

[4] These are: partisan affect polarization, ideological affect polarization, partisan sorting, partisan-ideology polarization, perceived partisan-ideolog polarization, issue consistency and issue divergence, straight ticket voting. Across these measures, the several major categories of polarization, described Part 1, are accounted for.

[5] These are detailed in the panels of the related figure: proportion that use the internet, proportion that saw campaign information online, and proportion that use social media.

While there is little evidence that internet use is growing differentially among the different age groups, as we can see in Figure 11, polarization has increased most drastically among the age group least like to use the internet (those who are 75+ years old). These are not the first-order pattern of results we would expect if internet use were a strong driver of polarization.

Figure 11: Changes in Polarization by Age Group

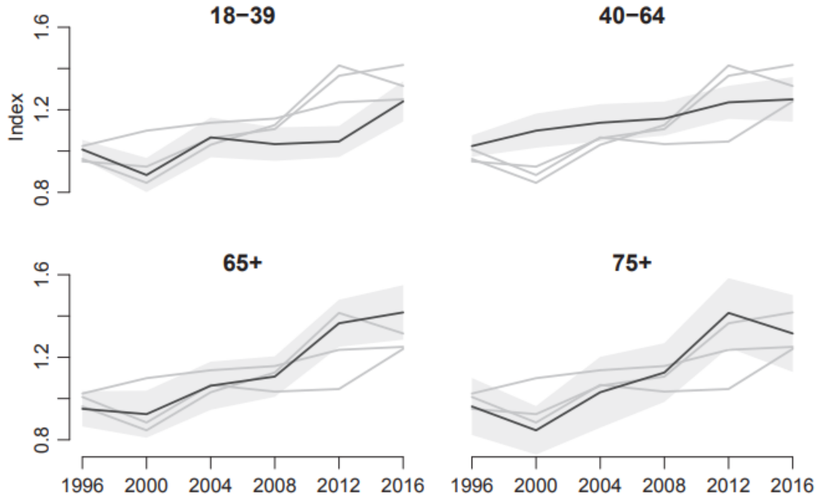


Fig. 3. Trends in polarization by age group. Each plot shows the polarization index for each of four age groups. Each plot highlights the series for one age group in bold. Shaded regions represent 95% pointwise CIs for the bold series constructed from a nonparametric bootstrap with 100 replicates. See main text for definitions and *SI Appendix*, section 3 for details on the bootstrap procedure.

Next, Boxell et al. (2017) move beyond this zero-order correlational analyses, to a more sophisticated analytic strategy. To do this, they build a model to predict likelihood of internet use based on demographic characteristics and other relevant factors available from their survey data from 1996 to 2016. Those variables are: age group, gender, race, education, geographical area, and survey year. Using this model, Boxell et al. (2017) can generate a prediction for each participant of their likelihood of being an internet user, based on the totality of their demographic and other information. This quantitative assessment allows them to group participants by their likelihood of being an internet user. And finally, high likelihood internet users can be compared to low likelihood internet users in terms of their actual polarization scores. If internet use contributed strongly to rises in polarization, we would see a steeper increase in polarization among those participants predicted to be likely internet users. This is exactly the opposite of what we see in Figure 12, which displays the results of this analysis. In fact, what we see is that those in bottom quartile of predicted internet use have seen a steeper increase in polarization than those in the top quartile of predicted internet use (especially from 2008 to 2012). (As we can also see, those in the top quartile of predicted internet use are more polarized overall than those in the lowest quartile of predicted internet use. However, this is likely because such people differ from low internet users from the outset—e.g. more educated, more politically engaged to begin with.

This does not explain the rise in polarization over time, it merely corresponds to a static difference in polarization between groups.)

Figure 12: Polarization by Predicted Internet Use

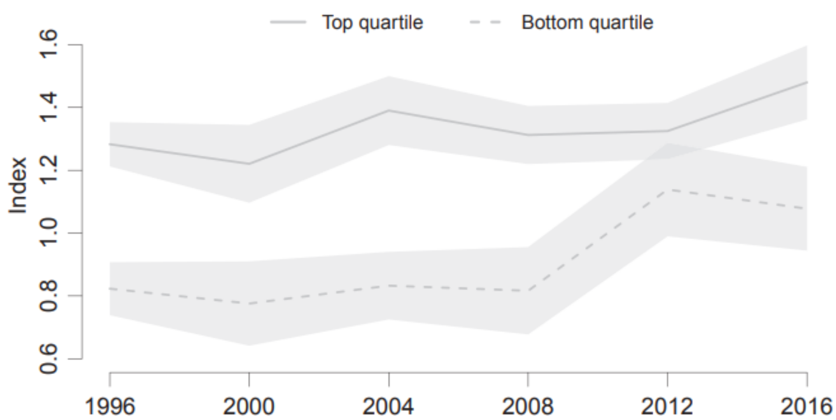


Fig. 4. Trends in polarization by predicted Internet use. The plot shows the polarization index broken out by quartile of predicted Internet use within each survey year. The bottom quartile includes values that are at or below the 25th percentile, while the top quartile includes values greater than the 75th percentile. Shaded regions represent 95% pointwise CIs constructed from a nonparametric bootstrap with 100 replicates. See main text for definitions and *SI Appendix*, section 3 for details on the bootstrap procedure.

The authors conclude with another analysis aimed at estimating the effect of internet use on polarization. They do this by constructing two models to predict the increase in polarization over time. In one model polarization is predicted by the year of the survey and other demographic characteristics. The other model is exactly the same, except they add a parameter which measures the amount of internet and social media use for each participant. The predictions from the second model can then be subtracted from the predictions of the first, to obtain an estimate of the extent to which internet use contributes to the rise polarization, as everything else is equal between the models. Boxell et al. (2017) run this analysis on four datasets (two from Pew, two from ANES). In three of the four datasets, the result is that internet use actually has a negative predicted effect on polarization—that is, increases in internet use are predicted to be associated with decreased polarization. In the one model that predicts a positive effect of internet use on polarization, the coefficient indicates that internet use accounts for 5.6% of the linear trend in increasing polarization—a meager effect.

If internet use accelerates political polarization, one uncontroversial implication should be that those who spend more time using the internet will exhibit larger rises in polarization. This is not what the data bears out in Boxell et al.'s (2017) analysis. Nevertheless, this alone does not rule out the possibility that internet use increases polarization.

For one, even within Boxell et al.'s (2017) analysis there is at least some evidence that internet use contributes to polarization (in that one "outlying" of the four datasets). This raises possibility that the "true" net effects of internet use on polarization may be obscured because of measurement issues—both in the form of how exactly internet use is operationalized and how much noise it is measured with. Other analyses, using different proxies for internet use, indeed arrive at different estimates. For example, Lelkes, Sood, & Iyengar (2017) try to identify an instrumental variable which causes differential amounts of internet use but does not affect polarization "through any other means" (p. 8). To do this, they focus on public right-of-way regulations (ROW). Right-of-way laws concern the rules and costs that can be imposed on internet service providers, in setting up infrastructure needed for internet access. The authors point out that Section 253 of the 1996 Telecommunications Act gave municipalities the authority the determine their own right-of-way regulations, leading to significant regional variation. They use this to instrument for variation in broadband internet access, demonstrating that ROW regulations predict broadband internet access and are unrelated to other regional differences which might affect polarization (e.g. the political orientation of regional governing bodies). Focusing on the years 2004 to 2008 and on affective polarization in particular, Lelkes et al. (2017) estimate that changes in broadband internet access do in fact cause statistically "significant" increases in affective polarization. Although, their estimated effects are very (perhaps even inappreciably) small; a 10% increase in broadband access being associated with a 0.03% increase in affective polarization. In sum then, these results, while different in exact magnitude from those of Boxell et al. (2017), might still point to the same conclusion. There is no strong evidence that changes in internet use account for a substantial portion of the increase in political polarization in the US in recent decades.

Still, none of this conclusively demonstrates that internet use plays no causal role in bringing about political polarization at all. Many factors possibly cause and contribute to polarization, among them: increasing lack of cross-cutting political interactions (e.g. Levendusky, 2009; Mason, 2018) and the rise of higher choice and less balance in news environment (e.g. Iyengar et al., 2019; Levendusky & Malhotra, 2016; Prior, 2013). These other factors might simply overwhelm internet use, making it harder to detect the unique effects of internet use in bringing about political polarization. More importantly, all of these factors intermingle. For example, McGregor & Molyneux (2018) demonstrate that many news journalists use information and discussions among users on internet platforms (Twitter in their study) to make coverage decisions. If the subsequent coverage of these journalists results in an increase in polarization is it properly characterized as caused by news consumption or internet use? The best way to answer question of whether internet use plays any causal role is to directly intervene in internet use.

Direct Intervention

The most thorough, most recent, and really only study directly intervening on internet use and measuring the effect on subsequent polarization comes from Allcott, Braghieri, Eichmeyer, & Gentzkow (2019). In their study, they paid Facebook users to quit the platform (Facebook) in the four weeks prior to the 2018 U.S. midterm election and then compared polarization in this group to a control group.

While Facebook is only a subset of the "internet", it accounts for a very sizable chunk of all US internet users and all internet usage. In terms of number of *users*, the authors report that 172 million adults in the US use Facebook, which accords with Pew estimates that, in 2018, 68%

of US adults were on Facebook[6] (Gramlich, 2019). In terms of internet *uses*, people spend a good portion of their time on the internet on Facebook. According to the company itself, users spent an average of 50 minutes per day on the platform in 2016 (Stewart, 2016)[7], which accords with the authors' own estimates that their participants used Facebook for just over 59 minutes per day on average. Additionally, the variety of activities that one can engage in on Facebook encompasses the main activities one can participate in on the internet in general—following news, messaging friends, reading or writing on public message boards, watching videos, or looking at pictures for entertainment. If internet use causes political polarization, there is a reasonable chance it flows through Facebook, or the type of activities that one can engage in on Facebook. A study examining the causal impact of Facebook use on polarization lends insight into the causal effect of internet use in general on polarization.

To recruit Facebook users, the authors placed an advertisement on the platform, asking for participants in a study about "internet browsing". In the end, 1,690,076 Facebook users saw the ad, 30,064 clicked on it, and 2,897 qualified to participate, by virtue of passing a prescreen, providing high enough quality initial data, consenting to participate and indicating they would be willing de-activate their account[8].

Once this sample was collected, there were three critical periods of data collection. First, during the initial recruiting (September 24th to October 3rd, 2018), participants completed a "baseline" survey consisting of measures of polarization, political beliefs, and news engagement (detailed later). Second, on October 11th, 2018, all participants completed a "midline survey". The purpose of this midline survey was to enact the de-activation procedure. *All* participants were asked to de-activate their Facebook for a 24 hour period—to make sure any subsequent long term differences in de-activation were not due to the upfront cost (annoyance) of merely going through the process of de-activating one's Facebook account. Next, the authors asked participants to report their "willingness to accept (WTA)" to keep their Facebook de-activated for four more weeks (a period which would end *after* the 2018 midterm election). WTA measures the minimum amount of money participants would be willing to accept to de-activate Facebook[9]. Randomization to condition took place after this. With approximately 35%

---

[6] The only possible content-based platform (i.e. not a search engine like Google) which might be considered comparable in number of users and frequency of use is YouTube, which according to Pew in 2018 was used by 73% of adults (Gramlich, 2019).

[7] Note that this does not need to come in the form of one continuous 50 minute long session, but could for example be composed of several short length sessions (e.g. 10 five-minute sessions).

[8] The authors tried to recruit a representative sample of Facebook users, by recruiting a sample stratified along 48 demographic dimensions. In the end, they managed to recruit a diverse sample, which relative to the US population and estimates of the average Facebook users, were somewhat more wealthy, well-educated, female, young, Democratic, and tended to use Facebook more. The authors examine their results with and without re-weighting for demographic, finding no major differences.

[9] WTAs were elicited using what is known as the "Becker-DeGroot-Marschak (BDM)" procedure. In this procedure, participants are informed that they will be offered a randomly generated amount of money to stay deactivated for four weeks. Before being given that offer, they are asked to indicate the lowest amount of money they would be willing to accept to stay de-activated for four weeks (WTA). Finally, they are told that if the randomly generated amount of money is above their indicated minimum, then they will be offered this amount to stay de-activated for four weeks. Participants were not given any information about the distribution from which the offer would be generated. I believe that the authors use this particular procedure because they believe it is the most effective available way to elicit honest willingness to accept responses.

probability, participants were offered $102 to de-activate for four weeks. With approximately 65% probability, participants were assigned to a condition where they were offered $0 to de-activate for four weeks. The subset of participants who were made the $102 offer and indeed had a willingness to pay that was below $102, were considered part of the "treatment" condition. They were asked to de-active their accounts for four weeks. The subset of participants who were made the $0 offer and had a willingness to pay that was below $102 were considered the "control" condition. They were not asked to de-activate their accounts for four weeks. Participants in either the "treatment" and "control" conditions were considered part of the "impact evaluation sample" (N=1,661)[10]. The third critical period occurred on November 8th, 2018, (two days after the midterm election). In this session, participants were administered the "endline survey", which asked the same questions about polarization, political beliefs, and news engagement as the "baseline survey", allowing for a pre-post analysis of changes on these measures. Several other measures were also collected between these periods[11] as well as after them. Attrition in the study was low and compliance was high[12].

The authors measured a host of variables relating either directly or indirectly to polarization[13]. I focus here on the seven measures which directly assessed polarization, and return to the variables that indirectly relate to polarization in Part 3, as they are more relevant to mechanistic questions of *how* internet use might bring about polarization, rather than *whether* it does so.

The first direct measure of polarization is what the authors refer to as (1) *party affective polarization*, and corresponds to the now familiar difference between in-party and out-party warmth ratings on the "feelings thermometer". Next, the authors measure (2) *Trump affective polarization*, which corresponds to ratings of warmth toward Donald Trump on the feeling thermometer. The also authors also measure (3) *issue polarization*, which corresponds to a metric of how close participants' opinions are to the mean opinions of their party over nine political issues (e.g. "On the whole, do you think immigration is a good thing or a bad thing for this country today?"). The authors then record (4) *vote polarization* wherein participants were asked to rate the strength of their preference for the candidate they would vote for (and then did

---

[10] In addition to the ~35% of participants who were given a $102 offer to de-activate and the ~65% of participants who were given a $0 offer, another 0.2% of the sample were made an offer *p* to de-activate their Facebook where *p* was an amount randomly sampled from the uniform distribution [$0, $170]. I believe this was done so that the authors would not be engaging in explicit lying, after having told participants, in the Becker-DeGroot-Marschak procedure, that they would be offered a randomly generated amount of money. If the only possible payouts were $0 or $102, then it might be considered a lie to say that a they would be offered a randomly generated amount of money; however, with at least the possibility of being assigned to a condition where their payout would really be drawn from a random distribution, this same statement might be considered less misleading. Although for practical purposes, the two setups are the same.

[11] such as daily text-messages asking for subjective well-being ratings

[12] Only 1.2% of those originally assigned to the "treatment" condition and 1.6% originally assigned to the control condition did not complete the endline survey. Participants in the "treatment condition" were de-activated on 90% of checks between the midline and endline survey period (compared to only 2% of participants in control, who for whatever reason, also deactivated their accounts during this period).

[13] The authors actually measure a wealth of variables, which go beyond the topics of polarization and politics. The paper intends to examine the general "welfare effects" of Facebook use. Thus, in additional to political outcomes, they measure effects on social outcomes (e.g. social time spent with others offline) and subjective well-being (e.g. life satisfaction, loneliness).

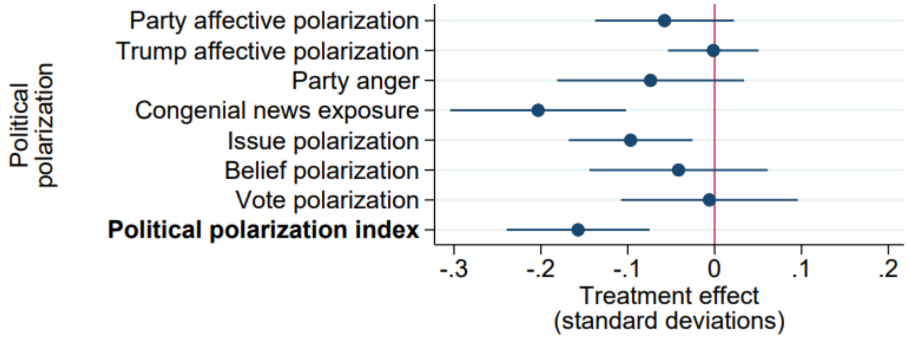
vote for or would have voted for) in the midterm election. A measure of (5) *belief polarization* recorded the extent to which participants' beliefs about the accuracy of news headlines were near the mean beliefs of their own party, aggregated from a 15-items news knowledge quiz, where participants evaluated real and made-up news story headlines (detailed more in the Part 3). An metric intending to measure differential (6) *party anger* asked participants to list news events that made them angry at the Republic party and news events that made them angry at the Democratic party. The metric itself was computed as the difference between the length of the angry responses wrote about the out-party and the length of the angry responses they wrote about their own in-party. Finally, there was an item assessing (7) *congenial news exposure*, where participants were asked to rate both "how often they saw news that made them better understand the point of view of the Republican party" and the same thing with regard to the Democratic party. The measure was computed as the difference between this response for one's own party and the response for the out-party.

So what were the effects of quitting Facebook on polarization?

The authors found that reducing Facebook usage did reduce political polarization, although both the extent and significance of this reduction varied across the seven different polarization measures[14]. These results are shown in in Figure 13. As can be seen, de-activating Facebook for four weeks did not result in a significant reduction in either (1) *party affective polarization* nor (2) *Trump affective polarization*. With regard to (3) *issue*, (4) *vote*, and (5) *belief polarization*, there was a significant decrease in *issue polarization*, although not *vote* or *belief* polarization. De-activating Facebook for four weeks appeared to reduce *issue polarization* by about a tenth of a standard deviation—bringing Democrats and Republicans 8% closer on the *issue polarization* index. In Figure 14, these results are broken down further, examining individual effects on each of nine issue items which comprised the scale. Here we can see that the treatment effect is particularly pronounced (and indeed only significant) for the issues of police racial bias and the fairness of the Mueller investigation.

[14] All primary results that the authors present are done on the participants in the "impact evaluation sample" (i.e. the participants who made it into the "treatment" and "control" conditions, by indicated a WTA lower than $102). All treatment effects are "local average treatment effects (LATE)" calculated as the value of τ in the regression equation $Y_i = (\tau * D_i) + (\rho * Y^b_i) + \nu_s + \varepsilon_i$ where Y_i is the value of a given metric on the endline survey, Y^b_i is the value of that metric in the baseline survey (if measured at baseline), D_i is the percent of times that the participant *i* was observed to be de-activated, and ν_s as a vector corresponding to a vector of the 48 stratification variables. The authors submitted a pre-analysis plan on October 12th, which they modified on November 7th (just before the endline survey). Based on their analysis of some of the non-compliance with the de-activation, they switched their analysis plan from an intent-to-treat analysis to an instrumental variable (IV) estimate plan. Thus, in the regression equation above, D_i is "instrumented" for T_i which is an indicator variable ($T_i \in \{0, 1\}$) for whether participant *i* was in the treatment condition or control condition.

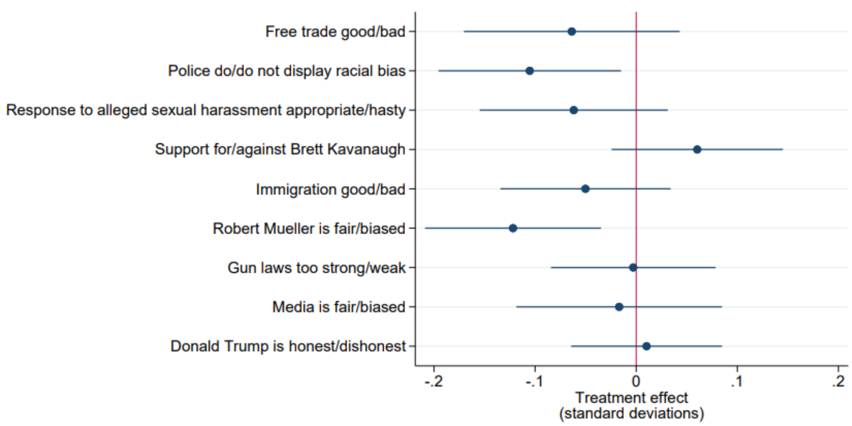
Figure 13: Effect of Facebook De-Activation on Polarization and other Political Outcomes



Notes: This figure presents local average treatment effects of Facebook deactivation estimated using Equation (1). All variables are normalized so that the Control group endline distribution has a standard deviation of one. Error bars reflect 95 percent confidence intervals. See Section 2.3 for variable definitions.

Figure 14: Effect of Facebook De-Activation on Issue Polarization, by Each Issue

Figure A30: **Effects on Issue Polarization**



Notes: This figure presents local average treatment effects of Facebook deactivation estimated using Equation (1). All variables are normalized so that the Control group endline distribution has a standard deviation of one. Error bars reflect 95 percent confidence intervals. See Section 2.3 for variable definitions.

Finally, there was no significant treatment effect on the (6) *party anger* measure, which aimed to measure differential anger at news covering one's own party relative to the other party. However, the largest and most significant effect of quitting Facebook was on the (7) *congenial news exposure* measure—which aimed to measure the difference in the extent to which participants stated they saw news which allowed them to better understand the perspective of the other party relative to news that allowed them to better understand their own party. De-activating Facebook for four weeks reduced that difference by about a fifth of a standard deviation. When these seven measures are all aggregated together (*political polarization index*, bolded in Figure 13), we also see a significant reduction in polarization accompanying Facebook de-activation.

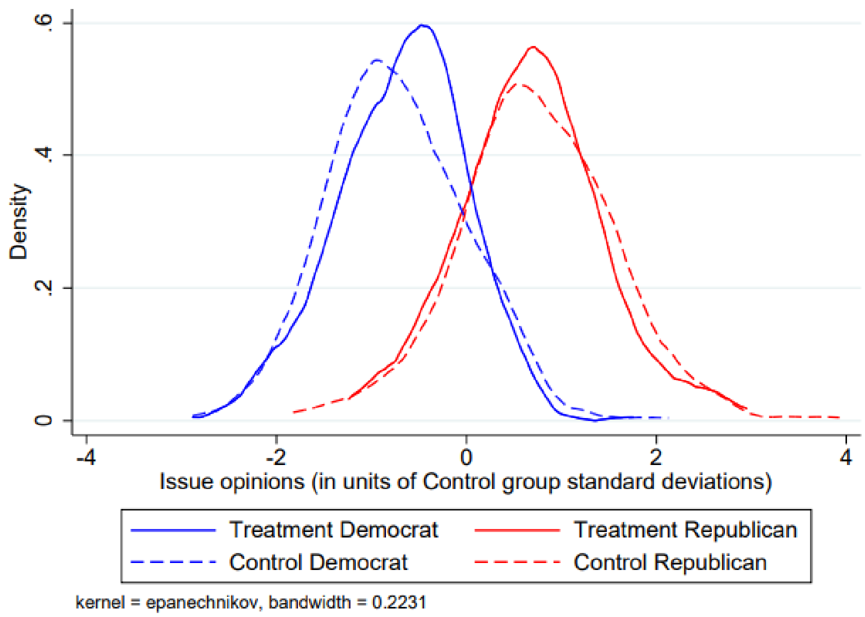
In sum, it does indeed appear that internet use, or at least Facebook use, can cause political polarization. To contextualize the *size* of their effects, the authors observe that the effect size of the decrease in the aggregate *political polarization index* that came from quitting Facebook (0.16 standard deviation reduction) was about 42% of the size of the rise in the polarization between 1996 and 2016, as measured by Boxell et al. (2017). Although it's not clear how meaningful this comparison is, which occurs across different operational definitions and time scales.

A few caveats might be noted. For one, variation in the treatment effect depending on which measure of polarization is focused on makes it hard even get a sense of the *type* of polarization that Facebook might contribute to. There was no detectable effect on either of the *affective polarization* measures, which correspond to one of the types of polarization that shows the strongest and most concerning evidence of an increase in over time in US and which has been most strongly suspected of being fomented by internet use. Meanwhile, the measure which showed the strongest evidence of polarization—*congenial news exposure*—is an unusual one. It does not correspond to any of the major typical categories of polarization. And further, this question, which asks participants to make judgments about differences in the extent to which they saw news that made them understand the perspective of in and out parties, requires a fair deal of accurate introspection and memory from participants. Such measures could easily be distorted by pre-existing expectations, as the participants were surely aware of the condition to which they have been assigned, likely had heard about the society-wide debate that is occurring with regard to Facebook's effects on political acrimony, and quite likely may have had their own theories on such effects. To their credit though, the authors do try to assess concerns of demand effects, and find no major evidence that their results were driven by them[15]. The most solid evidence for an increase in polarization as standardly measured comes from the significant treatment effect on the *issue polarization*, which more clearly corresponds to measures of polarization referred to in Part 1 as polarization on *policy preferences*. In almost a mirror image to Figure 5 in Part 1, where we saw Republicans and Democrats moving farther apart on policy preferences over time, we see Republicans and Democrats moving closer together as a result of reducing their Facebook use in this study (Figure 15). This is also perhaps more convincing evidence of a causal effect of internet use on polarization, as unlike the *congenial news exposure* measure, the question which comprise the *issue polarization* measure do not make deep introspective or memory-contingent demands of participants. There is less room for participants' expectations about the effects of Facebook use to creep into their responses. They can simply report the strength and direction of their opinions on political issues in a straightforward manner. And these opinions seem to be less drastically divergent after a reduction in Facebook use.

[15] Participants were asked whether they thought the researchers had a particular "agenda". A plurality (around 45%) thought that the researchers didn't have an agenda, many thought the researchers had an agenda to "show that Facebook is bad for people" (about 35%), while very few (less than 5%) thought they had an agenda to "show Facebook is good for people", and the rest (more than 15%) were unsure. However, these proportions did not differ between conditions. Furthermore, those who thought the researchers had an agenda to show that Facebook is bad for people did not show any significant differences on the major polarization indices between treatment and control conditions. Nevertheless, this doesn't rule out the possibility that some of the results were driven by, not demand effects, but the participants own pre-existing expectations about the effects of Facebook use on socio-political outcomes.

Figure 15: Issue Polarization Distributions by Party and Condition

Figure 4: **Issue Opinions by Party at Endline**



Summary

Taken together, the retrospective analyses and direct manipulation, suggest that there is at least some causal contribution of internet use to political polarization in the US. The retrospective studies marshal against a version of reality where the rise in the political polarization in recent decades is *mostly* accounted for by the dramatic rise in internet use over this same time period. However, portions of these same results hint at the possibility of at least some contribution of internet use. And Allcott et al.'s (2019) direct intervention study confirms this possibility, at least over short time scales of personal internet use. At the individual level, not using Facebook for four weeks, results in significant, and non-trivial sized, decreases in polarization surrounding political opinions and policy preferences. Thus, internet use can contribute to polarization.

But how?

PART 3: HOW MIGHT THE INTERNET CAUSE POLARIZATION?

Determining *how* internet use might increase polarization is an even more difficult question to adjudicate than *whether*, on net¸ internet use increases polarization. This is because of the enormity of the internet, ubiquity of its use, and diversity in its uses. As of 2016, the "size" of the internet, measured in the amount of information transferred per year, was estimated to exceed one thousand exabytes, the equivalent of 36 millennia worth of HD video (Pappas, 2016). In terms of ubiquity, according to Pew Research, as of 2018, 77% percent of adults in the US said they go online daily, with 26% saying they are online "almost constantly", (Perrin & Jian, 2018). In terms of diversity, as of 2015, there were over 935 million "live" websites (LaFrance, 2015). And putting statistics aside, the cliché must be noted that never before has communication been possible between so many people on such a large scale. Not only can any pair of people hypothetically communicate in the form of video, audio, picture, or text messaging over the internet, but information from one person or groups of people of any size can be broadcast in real time to millions and millions of other people, without reliance on major societal institutions like the media or government—allowing us, for example, to witness first hand accounts of sarin gas attacks in Douma, Syria, or police violence in Ferguson, Missouri.

It is possible, nevertheless, to cut through this enormous and dense jungle of internet usage and make sense of our more focused question about how internet use could affect polarization in the US. This can be done by taking note of a few observations. First, while there are an enormous array of websites on the internet, some are visited much more commonly than others. Second, while there an enormous amount of uses of the internet, some are again much more typical than others. And finally, some types of internet use, prima facia, have much plausible pathways to polarization, which also helps direct our attention.

According to Alexa, a private internet research firm, the top ten most "popular"[16] websites in the US are Google, Youtube. Facebook, Amazon, Wikipedia, Reddit, Yahoo, Twitter, Linkedin, and Instagram ("Top Sites in United States," 2019). And attempts at estimating suggest that Facebook (which also owns Instagram) and Google (which also owns YouTube) alone directly influence over 70% of web traffic (Cuthbertson, 2017; Staltz, 2017)[17]. The dominance of these websites in terms of individual people's web experience is reaffirmed by results from representative surveys of US adults, which show 73% of all US adults use Youtube, 68% use Facebook, 35% use Instagram, and 24% use Twitter (Gramlich, 2019). While many of these websites (e.g. Facebook, Twitter) are often link out to other website (e.g. to news articles), if polarization increases through internet use, there is a good chance that it goes through these websites in some form or another, helping to curtail the scope of our inquiry.

Indeed, this power law distribution of internet use—whereby a few very popular websites account for a large portion of web traffic—can be seen on smaller scales as well. In a study

[16] as measured by page views and number of users

[17] Although the exact significance and interpretation of such results is unclear. Most notably, the amount of web traffic generated by a website does not directly map on to the experiences of time spend on the internet from an average or median person. A small portion of people or firms may generate a large portion of traffic (e.g. one person might watch 5 GB of YouTube videos on astrology), while a larger portion of people may generate a smaller portion of web traffic (consuming only MBs of text, by reading news articles). Focusing on aggregate web traffic can give us a distorted image of the experience of the average web user.

examining online news consumption specifically (reviewed in more depth later), Gentzkow & Shapiro (2011) show that a large majority of online news traffic goes to a just a few websites. As can be seen in Figure 16, the top 10 news websites account for more than 60% of online news consumption, and the top 20 news websites account for nearly 80% of online news consumption.

Figure 16: Cumulative Online News Traffic by Website Popularity

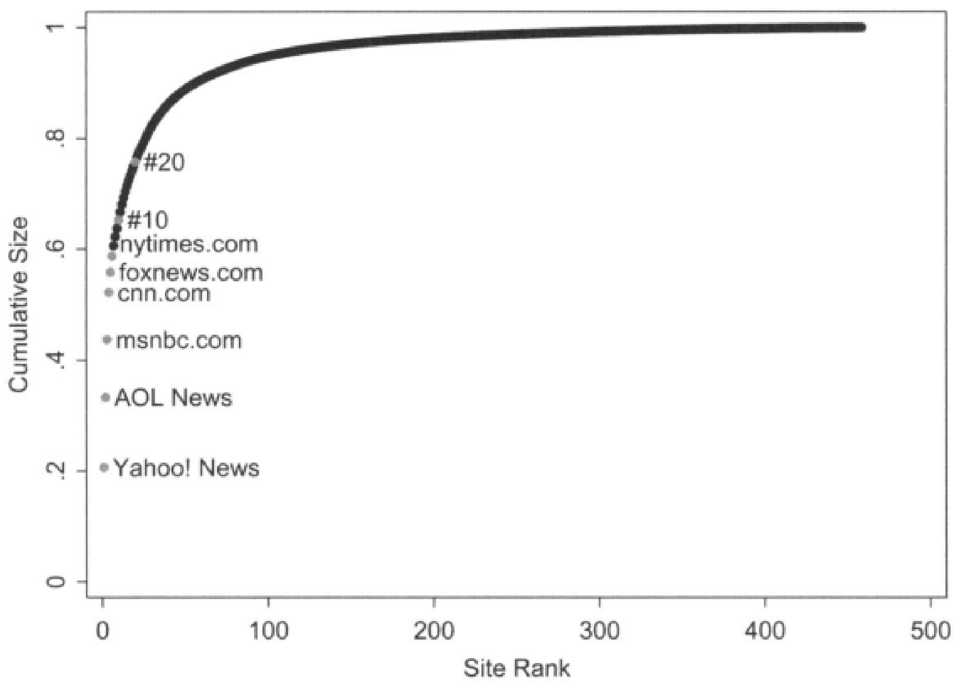


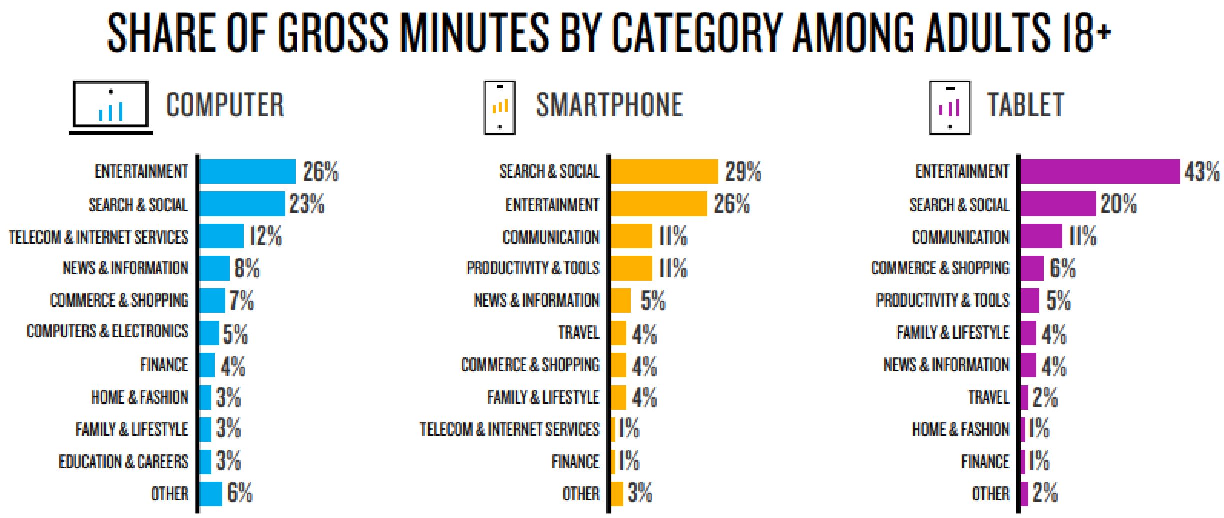
FIGURE V

Cumulative Distribution of Internet Unique Visits

Data are from comScore. Size is measured by average daily unique visitors.

We can further narrow and define our focus by looking at the *types* of activities people use the internet for. According to self-reported internet time use, collected by Statista, a market research firm, US adults spend the largest portion of their internet time on social networking sites (37 minutes per day), followed by 29 minutes on email, 23 minutes on video, 23 minutes on search, and 19 minutes on online games (Fitzgerald, 2014). Somewhat similar results can been seen in a 2017 Nielson report—which claims to be based on analysis of actual (i.e. not self-reported) internet browsing behavior of a representative panel of 200,000 internet users, whose internet use was tracked across three major surfaces (computer, smartphone, and tablet). As can be seen in Figure 17, more than 50% of time on the internet is spent on use involving “entertainment”, “search”, and “social” media. While the remaining portion of time use is accounted for by a “long tail” of other activities, including usages such as “news and information” consumption (8% of time on computer, 5% on mobile) and “commerce and shopping” (7% on computer, 4% on mobile).

Figure 17: Categories of Internet Time Use, By Surface



From this data on the dominant websites and predominant uses of the internet, which captures a wide portrait of internet use, we can begin culling those websites and uses which are unlikely to affect political polarization as they have no obvious mechanism of affecting polarization. From the top websites, this eliminates sites which have little ostensible connection to anything political, like Linkedin, which is for finding jobs, and Amazon, which is for shopping[18].

This leaves us mostly with websites that are used for information or news aggregation (e.g. Reddit, Wikipedia) or varied forms of social interaction (e.g. Facebook, Twitter)[19]. If internet use increases polarization, it seems a reasonable bet that it might occur through one, or both, of these two major avenues. Though the first major avenue, internet use increases polarization due to a biased exposure to political information online. Broadly, the argument would be that the political information that people encounter online is somehow distorted, in ways that exacerbate polarization—which may lead to more negative opinions of out party members (i.e. affective polarization), or stronger and less mutable opinions on policies (i.e. policy preference polarization), both of which, as we have seen, have some evidence of causal linkage to internet use. I will call this mechanism of internet-based polarization *differential exposure*. Through the second major avenue of internet-based polarization, the types of social interactions that are promoted online—both among those who agree politically and between those who disagree politically—bring about greater polarization, perhaps by strengthening party

[18] See, however, Shi, Shi, Dokshin, Evans, & Macy (2017) on politically divided book buying on Amazon. Nevertheless, the dominant activity on this platform is not political, even though evidence of political division may be evidence on the website.

[19] For the time being, I have left out Youtube from this analysis. A great deal of political content and debate is present on the platform, which some have tied into political "radicalization" (e.g. Roose, 2019). Researchers have begun to explore this, developing their own theories of the political dynamics of the platform (e.g. Munger's, 2019 "supply theory" of political Youtube content). However, because the topic is so new but also so large, I leave it to a future version of myself to include a more serious and lengthy discussion.

allegiance (i.e. party sorting polarization), in addition to possibly increasing affective and policy preference polarization. I will call this *polarizing social interactions*. I now examine the evidence internet use increasing polarization through these two routes.

Differential Exposure

The most popular accounts of how internet use might increase polarization in the US all have something to do with the information environment of the internet. In one way or another, Democrats and Republicans, conservatives and liberals, encounter different political information online—different news articles, different hot topic discussions, and so on. And somehow this *differential exposure* exacerbates political polarization. The most common mechanism by which differential exposure is purported to increase polarization argues people are disproportionally[20] exposed to information which *reaffirms* and is *consistent* with their already existing political beliefs and attitudes (Bakshy, Messing, & Adamic, 2015; Stroud, 2008). A highly-touted claim is that much of internet use occurs in "echo chambers", wherein users of different political persuasions are largely walled off from each, disproportionately talking to and sharing information with only like-minded individuals, further entrenching their political views and identities (Sunstein, 2007). For example, Democrats might predominantly click on news articles shared by other Democrats on Facebook and discuss them predominantly with other Democrats, while Republicans do the same, both on Facebook and elsewhere, for example largely visiting politically-aligned news sites, such as, in their case, Breitbart and Fox News. Another claim is that this differential exposure is caused by "filter bubbles", whereby website algorithms, designed to tailor content to existing preferences, overwhelmingly expose people to information which is consistent with their existing attitudes, including their political attitudes (Pariser, 2011). For example, Facebook's NewsFeed algorithm might be more likely to recommend news articles form political like-minded friends in the first-place, leading this type of information to be disproportionally consumed. Implicit in these accounts is that biased exposure might occur through deliberate means or non-deliberate means. The internet may enable greater deliberate "selective exposure", whereby people consciously prefer and thus choose to consume information that is consistent with their existing views (Stroud, 2008). And it may also result in differential exposure through non-deliberate means, because of biases in the proportion and types of information available for consumption in the first place (Sears & Freedman, 1967). These are not always neatly separated, nor separable, in analysis. Their final manifestation, *differential exposure*, is ultimately the same however. And thus, I examine whether there is evidence for *differential exposure* generally. Later, I direct some attention to sussing out the extent to which any differential exposure comes from deliberate and non-deliberate means.

differential news exposure online

One obvious place to look for differential exposure is in political news consumption. The ideal study would track all the online behavior of a representative sample of internet users in the U.S. We could then examine whether there is evidence of differential exposure by political affiliation. Fortunately, there is one gold standard study which does exactly this. In a month long

[20] What exactly is meant by "disproportionally" is not always clear and not always constant. One relevant baseline is some sort of comparison to what people would be exposed to otherwise in their "everyday life" during comparable activities.

longitudinal study, Guess (2018) tracks the internet activity of broad sample of internet users in the U.S., examining how political news consumption varies by political ideology.

Guess (2018) tracks the internet activity of a sample 1,392 people, recruited from YouGov's participant pool, from February 27 to March 19, 2015. Participants were asked to download specialized software ("Wakoopa"), which was able to monitor the entirety of their internet activity any time that person used their computer, collecting everything aside from financial transactions and passwords[21]. In this period of time, the users engaged in a total of 6,319,441 website visits, allowing the author to examine not only news consumption that flows from major websites like Facebook or Reddit, but almost any news consumption that might occur at all (e.g. direct visits to major news sites like the page for the *New York Times* as well as visits to more obscure or smaller regional sites like *The Oklahoman*). Because participants were recruited from YouGov's participant pool, each participants' political ideology was also known and could be tied to their internet browsing activity[22]. A second analysis is also conducted on a dataset collected using similar methods, by Guess, Nyhan, & Reifler (2018), tracking the internet use of 2,512 people from October 7th to October 31st, 2016 (with a total of 16,984,969 site visits).

To identify differences between people in political news consumption, there must first be some way of classifying the political bent of the news consumed by those people. The standard way to do this is at the source level. News sources are classified according to their "partisan slant"—that is, the extent to which they tilt toward one political perspective or another. Any given news item (e.g. a news article) can then be classified according to the partisan slant of the source from which it comes.

There are several ways to infer "partisan slant" of a news source, which can be grouped into two major categories. In one category, slant is assessed by examining the typical content of a new source. For example, one method records the number of times that a news source cites different policy groups and think tanks and then associates that with the number of times Congressmembers of different ideologies reference those same policy groups and think tanks (Groseclose & Milyo, 2005)—making the inference, for example, that if a news source cites policy groups and think tanks that more conservative congressional representatives cite, then that news source is more likely to be conservative. The other major method involves examining the readership of a particular news source—and characterizing its political orientation or slant according to the distribution of that source's readership (e.g. if 95% of readers are liberal, then that source is assumed to be more ideologically liberal than a source where 60% of readers are liberal). Such methods are imperfect as they rest on the assumption that reader partisanship is a valid proxy for political leanings of the news source and the content of their news pieces (Sood

[21] The only exception being that participants were allowed to pause the monitoring for 15 minute breaks. One concern with this might be that people who would agree to download such software are in some important way unrepresentative of the population to which we wish to generalize. Guess (2018) reports in Appendix B, an analysis to examine whether the type of people who would download this tracking software are unrepresentative. He finds that, on only one of four measures asking about concern with privacy and sharing personal information online, members of the online tracking survey indicate slightly less concern. In the main analysis, the author examined the results under various weighting of the data, adjusting in various ways for demographic differences.

[22] with their knowledge and consent, to be clear

& Lelkes, 2016). In practice, this is often a reasonable assumption[23]. Such a method is used by Bakshy et al. (2015) to examine the news consumption of 10.1 million Facebook users. They are able to assign a partisan slant score ("alignment score") to hundreds of news websites, because they have access to the political ideology of the Facebook users who visit those sites. Guess (2018) opts to use this latter method to infer the partisan slant of a news source—specifically, he literally uses the partisan slant scores already generated and made available by Bakshy et al. (2015). This is primarily for the pragmatic reason that the news consumption he is looking at are online news sources (e.g. websites); for which content-based alignment scores are simply not available for most websites. As Guess (2018) further points out, these readership based scores are face valid, as well-known publications fall where we would expect along the continuum (see Figure 18, where we see for example news sources like Breitbart receive high conservative partisan slant scores, and publications like the New York Times receiving partisan slant scores in the liberal direction).

Figure 18: Alignment Scores for Top News Outlets in Guess (2018)

Domain	Slant	# Visits	Domain	Slant	# Visits
msn.com	-0.08	36,263	msn.com	-0.08	146,302
en.wikipedia.org	-0.21	21,496	en.wikipedia.org	-0.21	56,520
aol.com	0.01	9,110	huffingtonpost.com	-0.62	45,652
finance.yahoo.com	0.08	5,949	aol.com	0.01	38,666
news.yahoo.com	0.05	5,000	washingtonpost.com	-0.26	32,067
foxnews.com	0.78	4,777	nytimes.com	-0.55	29,454
townhall.com	0.93	4,372	foxnews.com	0.78	22,450
buzzfeed.com	-0.52	4,077	cnn.com	-0.27	18,055
huffingtonpost.com	-0.62	3,898	dailykos.com	-0.90	17,838
nytimes.com	-0.55	2,532	fivethirtyeight.com	-0.52	11,335
dailykos.com	-0.90	2,148	politico.com	-0.13	11,042
washingtonpost.com	-0.26	2,025	realclearpolitics.com	0.66	10,795
bbc.com	-0.26	1,976	finance.yahoo.com	0.08	9,859
theblaze.com	0.89	1,640	buzzfeed.com	-0.52	7,918
cnn.com	-0.27	1,565	nbcnews.com	-0.27	7,777
dailymail.co.uk	0.29	1,551	slate.com	-0.68	7,166
businessinsider.com	-0.06	1,544	rawstory.com	-0.85	6,826
nation.foxnews.com	0.90	1,469	cbsnews.com	-0.13	6,786
hotair.com	0.92	1,373	dailymail.co.uk	0.29	6,740
breitbart.com	0.91	1,317	breitbart.com	0.91	6,688

Table E1: Alignment scores and the raw number of visits to the 20 top hard news domains in the 2015 (left) and 2016 (right) YouGov Pulse samples.

With this, Guess (2018) can proceed with his analysis. The main goal of the analysis is to compare political slant in news consumption between Democrats and Republicans. To do this, Guess (2018) first computes an ideological alignment score for each participant, aimed to measure the skew in that individual's political news consumption. For each participant, this is the average partisan slant of the news sources they visited, averaged across all their web visits. So, for example, if a participant viewed 7 articles on foxnews.com, 3 on breitbart.com, and 2 on news.yahoo.com during the study, their ideological alignment score would be 0.73 (i.e. (7*0.91 +

[23] For example, Bakshy et al., 2015 report correlations between 0.5 and 0.9 between their audience-based measures and content-based measures, for those sources on which there is information for both

3*0.78 + 2*0.05)/12). Guess (2018) then aggregates to the party level and compares these ideological scores between Democrats and Republicans, as well as Independents[24].

Guess (2018) does an admirable job examining whether his findings hold across a range of specifications. Specifically, he examines his results under the following conditions.

- **Raw**: This is the most rudimentary analysis and is essentially exactly what I described in the paragraph just above. Republicans, Democrats and Independents are compared by looking at their alignment scores across all their web visits, to all websites which have an alignment score.[25]
- **Raw (weighted)**: This is the same as the "raw" analysis just above, except all data in the analysis have been weighted to be representative of the U.S. population—in terms of age, gender, race, and region.
- **News/politics only (weighted)**: This is the same as "raw (weighted)" analysis just above, except only looking at the subset of web visits where participants viewed an article that was explicitly classified as a "hard news" article. The aim here is to remove visits to news website like the New York Times when those visits are to sections like the Sports section or Travel section, which are unlikely to be about political or polarizing topics.[26]
- **News/politics only, portals excluded (weighted)**: This is the same as the "news/politics only (weighted)" analysis just above, with a further restriction. Visits to articles that are appear on general purpose portal websites (i.e. aol.com, msn.com, google.com) are excluded. These portals have fairly non-extreme alignment (partisan) scores, and this analysis is done to see how news consumption looks like, when visits to these large, common, and fairly neutral portal websites are excluded.
- **News/politics only, portals excluded (weighted, de-duped)**: This is the same as the analysis just above, but attempts to account for some technical issues that might overinflate visits to some websites (e.g. when a user refreshes a page, this counts as a

---

[24] On a superficial glance, because the alignment scores for website are based on the partisan distribution of readership, this analysis might appear circular. It is not. By way of example, the news source alignment scores can be thought of as asking for the probability: **p(Democrat | News Source)**, i.e. how likely is it that someone is a Democrat if they are reading this particular new source. And here we are interested in figuring out another probability: **p(News Source | Democrat)**, i.e. how likely is someone to read this particular news source if they are a Democrat. If all the data were drawn from the same underlying population, and we knew the base rates of p(Democrat) and p(New Source), then theoretically either of the conditional probabilities could be derived from the other, using Bayes Rules. But, of course, the actual populations from which the two conditional probabilities are drawn aren't actually the same (one is a sample of Facebook users, and the other is a sample of people in the YouGov tracking survey). And further, the important point, which is made clear by Bayes rule, is that $p(A|B) \neq p(B|A)$.

[25] A few websites, which are not exactly news websites, yet still have partisan alignment scores from the Bakshy et al. (2015) study, are excluded. These are Twitter, YouTube and Amazon. Both Twitter and YouTube host content from people all across the political spectrum determining by the ideology of the individual posters to these websites, not some overall editorial bent, while Amazon has an alignment score, but is clearly not a news source.

[26] Guess (2018) describes the process he uses to classify articles as "hard news" in detail in Appendix C, but the basic idea is that he builds a classifier that analyzes the actual text of an article and uses very basic features (e.g. unigrams) to classify the article as "hard news" or not. He reports that his performance of the model is high. Although some decisions are idiosyncratic and performance is not evaluated on the full set of the data (as the ground truth for the full set of articles is not known). However, the fact that he shows his results both with and without applying this "hard news" classification ultimately shows that the results look pretty similar either way.

new visit, over-inflating the number of times that that web page is counted; such cases are now "de-duplicated").

- **News/politics only, portals excluded (weighted, de-duped)**: This is the same as the analysis in the bullet point above, except each participant's alignment score is computed by weighting visits by the amount of time spent on each web visit (e.g. in the previous methods, if a participant had 3 visits to foxnews.com and 3 visits to nytimes.com, those visits would contribute the same amount to a participant's overall alignment score; but now, if one spent 50 minutes across those foxnews.com visits and only 4 minutes on the nytimes.com visits, the foxnews.com visits will get more weight).

So what does he find? Across the six different analyses outlined above, there are three key results, which remain consistent throughout (see Figure 19).

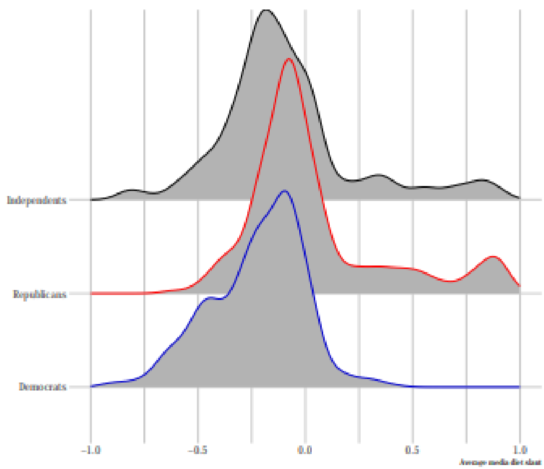
First, overall, there is a large degree of overlap in news consumption between partisans. Specifically, Guess (2018) computes a number called the "overlapping coefficient", which measures the extent to which two distributions overlap. This is a number that can vary from 0% to 100%, where 100% would mean the distributions entirely overlap, and 0% would mean that there is no overlap at all. When the distributions of alignment scores are compared for Republicans and Democrats, they range between 61% and 69% depending on which of the six analyses is employed. This can also be seen visually in Figure 19, where the Democrat, Republican, and Independent's distributions are presented beside each other, under the six different analyses plans—and where we see that the distributions overlap to a fair degree in each case.

Second, the overall distributions are of course slightly different from each other, as would be expected. On average, Democrats consume more liberal media diets than Independents, who consume more liberal media diets than Republicans. In other words, the means of the ideology scores of the groups do significantly differ. It would be surprising if this weren't the case—that is, if there were no difference at all in the ideological alignment of the news sources that Republicans and Democrats consume.

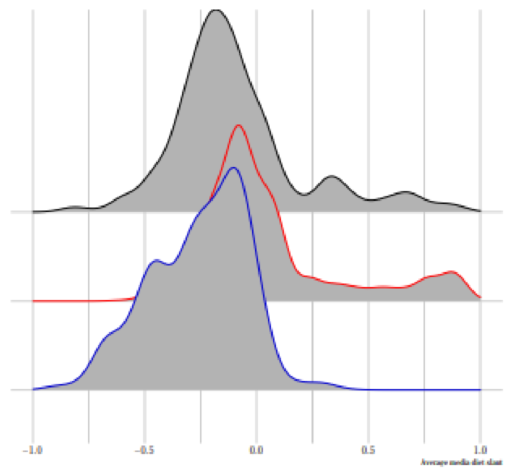
Third, there is greater variability in the distribution of media consumption of Republicans. This is evidenced by two empirical facts. For one, the standard deviation of the Republican distribution is higher. But, more importantly, when visually examining the distributions, we can see a smaller second mode ("bump") emerge in far right of the Republican distribution. This represents the subset of Republicans who do in fact consume a fairly extreme media diet (the mode of this second bump is in the 0.8 range, which is around where foxnews.com and breitbart.com fall). This bump seems most pronounced when looking specifically at hard news consumption, and weighting visits by time spent (panel f in Figure 19).

And finally, the three patterns noted above also emerge in the second data set that Guess (2018) analyzes (form October 7th to 16th, 2016)—the only difference being perhaps some slight indication that the smaller second mode ("bump") in the Republican media consumption becomes larger (i.e. the subset of Republicans consuming extreme media diets may have increased slightly).

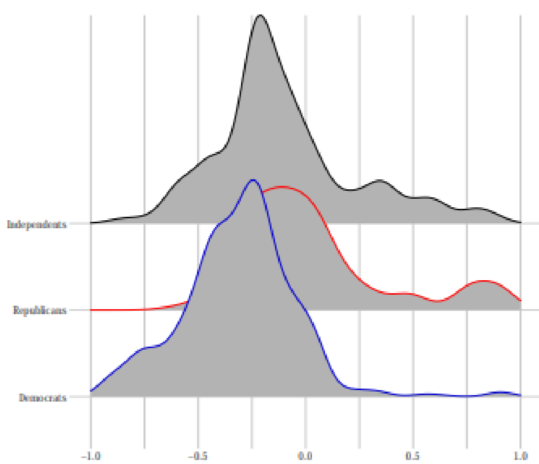
Figure 19: Overlap in Partisan Online Media Diets, Across 6 Different Analysis Types



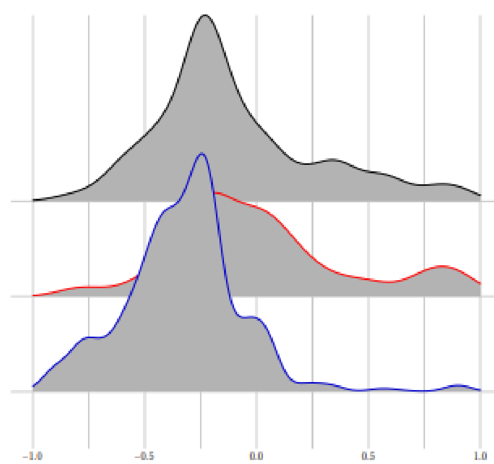
(a) Raw



(b) Raw (weighted)

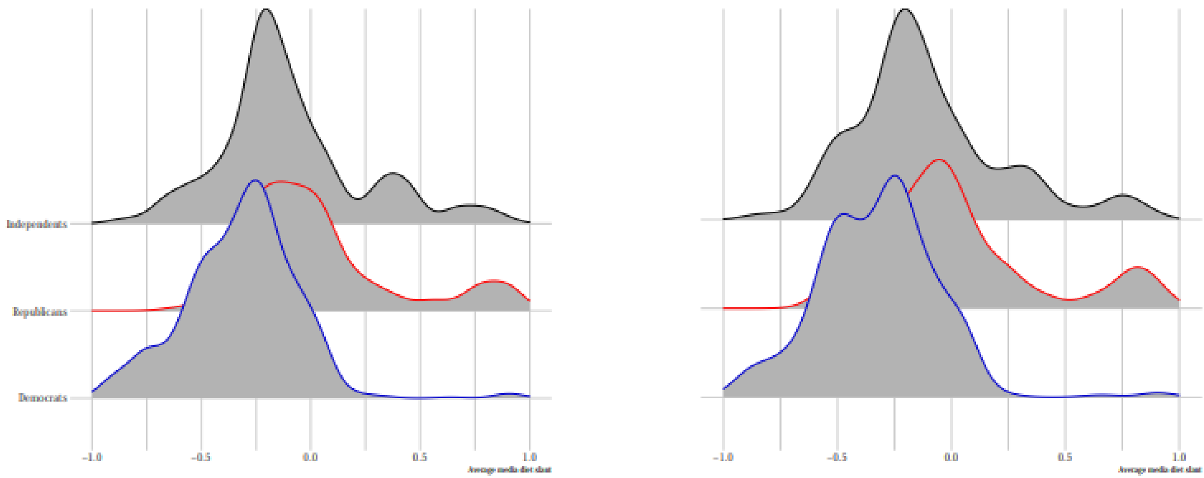


(c) News/politics only (weighted)



(d) News/politics only, portals excluded (weighted)

(Figure 19 continued)



(e) News/politics only (weighted, de-duped) (f) News/politics only (weighted, de-duped, by duration)

These results offer us the most precise evidence on differential news consumption habits between Republicans and Democrats. And the most central finding is the quantification of the overlap in media consumption between the two sides—which appears to be between 61-69%. While the results are clear, how to interpret them is another matter. Guess (2018) interprets 61-69% as a good deal of overlap, which might imply that *differential exposure* is not necessarily a huge factor driving the polarization that results from internet use. But how is this to be judged? The best way a determination could be made is if we knew how much varying amount of overlap contribute to varying amount of political polarization. Of course, this is not known.

Another way to make sense of this data is by comparison to meaningful reference points. For example, using publicly available data on the means and standard deviations for men's and women's heights, I computed the overlap coefficient for the distribution of men's and women's heights—which is approximately 52.6%. Thus, by comparison, there is more overlap in ideology of news consumption between Democrats and Republicans than between the heights of men and women. A more meaningful comparison, however, is perhaps to the degree of segregated exposure encountered across other mediums in everyday life.

differential exposure across mediums

If differential exposure is one of the ways that internet use contributes to polarization, one expectation we *may* have is that the level of differential exposure to information by political ideology might be higher on the internet than in other mediums in everyday life. This need not be the case, as other mediums (e.g. cable news) might also contribute to political polarization even more so than internet use, and they may also do so through partisan segregation in information exposure (e.g. Prior, 2013) which may be even greater than it is online. Such a comparison, where differential exposure is higher in other mediums, would not necessarily indicate differential exposure isn't a mechanism by which internet use can increase polarization. Nevertheless, a comparison is useful as it *could* reveal that differential exposure is much higher on the internet than on other mediums, at least suggesting that something is highly unusual about the medium in this regard. Even if such a huge difference is not revealed, comparisons to long-standing mediums, with which we are more familiar, help contextualize the amount of differential exposure we see online.

Thus, we can turn our attention to a study by Gentzkow & Shapiro (2011) which aims to exactly to compare "ideological segregation" across a host of mediums through which people are exposed to political information. Specifically, the authors compare ideological segregation across 12 different mediums, which they group into the following four categories:

- *internet*: (1) internet news;
- *offline media*: (2) cable news, (3) broadcast news, and (4) national newspapers
- *acquaintances (i.e. face-to-face interpersonal interactions)*: (5) family, (6) work, (7) people who are "trusted", (8) mutual membership in a voluntary association group, (9) people named as part of one's neighborhood, (10) people with whom politics are discussed;
- *communities (i.e. local physical communities)*: (11) county and (12) zip code level communities.

To enable comparison of political segregation across mediums, Gentzkow & Shapiro (2011) rely on a simple measure of segregation called the "isolation index," previously used to quantify racial segregation (White, 1986; Cutler, Glaeser, & Vigdor, 1999). This measure is cleverly repurposed to quantify the extent to which conservative and liberals are segregated from each other across various mediums through which political information may travel.

Here's an example of how the "isolation index" is computed for the *internet*, where the focus is on partisan segregation in internet news consumption. For any online news website, a certain percent of its readers are conservative and a certain percent are liberal. Thus, each information source (website, here) can be characterized by the percent of its audience that is conservative, referred to as its "share conservative".[27] For example, if 90% of visits to the Fox News website comes from conservatives, 50% of the visits to CNN's website come from conservatives, and 10% of visits to the New York Times website comes from conservatives, these sources would have "share conservatives" of 90%, 50%, and 10% respectively. All people's media diets can then be characterized by a number called "conservative exposure," which is the weighted average of the "share conservative" across the sources they are exposed to. For example, in this hypothetical, a person who visits Fox News nine times a day and CNN once a day would have a conservative exposure of 86% (i.e. (90%*9 + 50*1) / 10), and a person who only visits only the CNN and the New York Times websites in equal proportion would have a conservative exposure of 30% (i.e. (50% + 10%) / 2). We can then group all conservatives together and compute their average "conservative exposure" and group all liberals together and compute their average "conservative exposure". The average liberal's conservative exposure score can then be subtracted from the average conservative's conservative exposure score. And this number is the author's "isolation index", for this medium. The larger the difference in

[27] We could also compute the "share liberal" of course. The results are the same either way, as one is just 100% minus the other. In the paper, the authors choose to focus on the percent conservative for their analysis. It does not make any difference, but one must be picked and then kept constant.

conservative exposure between the groups, the larger the isolation index, and the more ideological segregation there is in this medium. For any medium, this isolation index can range from 0% to 100%. At one extreme, an isolation index of 100% would mean that conservatives and liberals only visit websites that are only ever visited by people of their own party. At the other extreme, an isolation index of 0% is consistent with all websites receiving an equal number of visits from conservatives and liberals.[28]

Crucially, with some adjustment for context, the authors can compute an isolation score for *any* given medium (internet news, cable news, interactions with family etc.) through which political information can travel between partisans. Take the medium loosely defined as "interactions with family." If, for the average conservative, 80% of their family members are also conservative, then we can say the say the average conservative exposure for conservatives on this medium is 80%. And, if the average conservative exposure for liberals is 20% on this same medium (i.e. on average 20% of liberals have a conservative family member), then the isolation index score in this medium of face-to-face interactions with family would be 60% (80% - 20%).[29] Various different adjustments and weightings are made for each medium (e.g. in online news consumption, people visit different websites with different frequencies and thus we want to weight conservative exposure by the frequency of visits). And of course what constitutes a "unit" of political information in each medium is always up for debate (e.g. ideally we could weight family interactions by the frequency with which we have political conversations with each family member). Nevertheless, the overarching point is that such a metric can be computed across mediums, which can be used to at least get somewhat of a sense of differential exposure in each.

The authors then compile data from a wide array of different sources in order to compute a political isolation index score for the 12 different mediums outlined earlier. Like Guess (2018), Gentzkow & Shapiro (2011) rely on actual tracked internet browsing behavior (restricted to websites categorized as related to "general news" or "politics") to compute an isolation index score for partisans on the *internet*. Information to compute political isolation index scores for mediums that fall in the other three categories of *offline media*, *face-to-face interpersonal*

[28] Note that the average conservative share among conservatives and liberals does not need to be 50% for the isolation index to be 0%. So long as conservatives and liberals have the same conservative exposure, there is no "segregation" or difference between the two groups in information consumption for that medium. For example, if conservatives had an average conservative exposure score of 70% and liberals also had an average exposure score of 70%, the isolation index would still be 0%. There would be no difference in website exposure between these groups. Such an outcome is, at least hypothetically possible, because exposure is weighted by visits. For example, a case of 70% of conservative exposure among both conservatives and liberals is possible if conservatives simply engage in more web visits (e.g. in the simplest case, imagine a universe of one liberal and one conservative who only ever visit one website, but the conservative visits it 7 times a day and the liberal visits it 3 times a day).

[29] Note that, for families, share conservative for liberals and share conservative for conservatives does not need to sum to 100%. For a simple counterexample, imagine two families: in one family both members are conservative, in the other family one member is conservative and the other is liberal. For liberals, 100% of family members (not including themselves) would be conservatives. For conservatives, 67% of their family members (not including themselves) would be conservatives. This counterexample works because there are an uneven number of conservatives and liberals in the total population, but conservative exposure for liberals and conservatives need not sum to 100% even when there are equal numbers of conservatives and liberals (for a counter example of this variety imagine 3 families: one family with two conservatives, and two families with two liberals and one conservative; conservative exposure for liberals would be 50%, but for conservatives it would be 75%.)

interactions, and *local physical communities* come from a huge variety of high-quality sources, which the author leverage in ingenious ways to infer political isolation between conservatives and liberals. (See Appendix, for more detailed review of methods and data sources).

The authors main results are nicely summarized in Figure 20, where we can see how political segregation varies between partisans on the *internet*, compared to political segregation between partisans on media that fall into other categories: *offline media* (e.g. partisan segregation in national newspaper, or broadcast news consumption), *acquaintances* (e.g. political segregation in face-to-face interactions, like with family, friends and co-workers), and *local communities* (e.g. political segregation at the county level).

Figure 20: Partisan Segregation Across a Host of Mediums

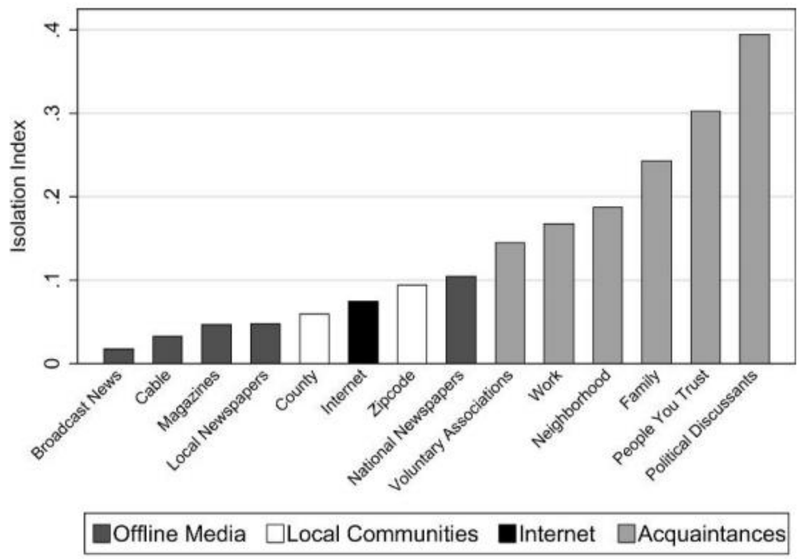


FIGURE II

Ideological Segregation by Medium and Type of Interaction

Internet data are from comScore. County, ZIP code and offline media data are from MRI. Voluntary associations, work, neighborhood, family, and "people you trust" data are from the GSS. Political discussants data are from the CNES. See Section III for details on the construction of the isolation index.

This allows us to contextualize *differential exposure* on the internet, by comparison to differential exposure elsewhere. Among the various "news" mediums (i.e. the set of mediums consisting of: broadcast news, cable news, magazines, local newspapers, and national newspapers), the *internet* is one of the mediums with the highest degree of partisan segregation—only surpassed by, although roughly comparable to, partisan segregation in national newspaper readership (e.g. partisan segregation in readership of the New York Times v. Wall Street Journal). On the other hand, there is a greater degree of partisan segregation in every single of the six mediums of *face-to-face interactions*. Some of these show dramatically higher levels of partisan segregation than the internet (e.g. partisan segregation in in-person political discussion networks, which has an isolation index score of 39%). This may be due to a core feature of many human social networks—homophily, the tendency of like people to associate

(McPherson, Smith-Lovin, & Cook, 2001), which is perhaps most dramatically the case in person.[30]

In absolute terms, the amount of partisan segregation on the internet, as recorded by the authors isolation index, is not all too high. Figure 21 displays isolation index scores (right-most column) as well as conservative exposure scores, of liberals and conservatives (from which the isolation index is computer; middle and left-most columns). The precise isolation index score that the authors compute for the internet is 7.5%—meaning that the average difference in conservative exposure between conservatives (which is 60.6%) and liberals (which is 53.1%) is 7.5%, which might be characterized as appreciable but not enormous. The authors point out that this difference is comparable to difference in conservative exposure between those who live in Minnesota or Iowa (61% share conservative) and those who live in Massachusetts (52% share conservative).[31]

Figure 21: Conservative Exposure by Political Ideology (and Resulting Isolation Index Scores)

TABLE IV

IDEOLOGICAL SEGREGATION BY MEDIUM AND TYPE OF INTERACTION

	Conservative exposure of		
	Conservatives	Liberals	Isolation index
Internet	.606	.531	.075
Offline media			
Broadcast news	.677	.660	.018
Cable	.712	.679	.033
Magazines	.587	.540	.047
Local newspapers	.695	.647	.048
National newspapers	.612	.508	.104
Face-to-face interactions			
County	.682	.622	.059
ZIP code	.637	.543	.094
Voluntary associations	.625	.480	.145
Work	.596	.428	.168
Neighborhood	.627	.439	.187
Family	.690	.447	.243
People you trust	.675	.372	.303
Political discussants	.796	.402	.394

Notes: Internet data are from comScore. County, ZIP code, and offline media data are from MRI. Voluntary associations, work, neighborhood, family, and "people you trust" data are from the GSS. Political discussants data are from the CNES. See Section III for details on the construction of exposure and isolation measures.

It should be noted that the degree of partisan segregation on the internet does seem to vary depending on which "parts" of the internet one is focusing on. The authors present evidence that larger, more popular websites tend to be less ideologically extreme than smaller, less popular ones. The isolation index for the 10 most popular news websites is 6.2%, and for the next 11-25 websites it is similarly 5.8%, but jumps to 8.6% for the next 25-50, and spikes to 21.3% for the all the remaining websites after the 50th most popular.

[30] In a later part of their analyses, the authors also compute an isolation index measure across these various mediums for several non-political demographic variables (gender, race, income and education). For the most part, across all mediums, there is less segregation for these demographics than for political ideology. The only demographic that clearly surpasses political segregation is racial segregation in geographic and face to face interactions.

[31] This is ideological identification, not the same as presidential vote share—which of course, for states like Massachusetts is more solidly non-Republican (32.8% for Trump in 2016, 37.5% for Romney in 2012, 36.0% for McCain in 2008, 36.8% for Bush in 2004, etc).

The possibility that partisan segregation varies by the "part" of the internet that one is focusing is also supported by a study Flaxman, Goel, & Rao (2016)—where, in one part of their analysis, the authors examine how partisan segregation varies by what can be considered "type" of internet use, or the effects of different internet platforms on *differential exposure*. In their study, Flaxman, Goel, & Rao (2016) also examined online news consumption. They did so by tracking the internet browsing habits of 50,000 people from March to May 2013, via a Bing Toolbar plugin on Internet Explorer. Like Gentzkow & Shapiro (2011), they also examined political segregation in internet news consumption, and did so using a measure that is conceptually similar to Gentzkow & Shapiro's (2011) isolation index measure. In part of their analysis, they examine, how much partisan separation there is between individuals depending on how they arrive at online news articles and news opinion articles. That is, they sought to examine whether there are differences in partisan segregation in online news consumption depending on whether users: (1) arrived at an article through an "aggregator site" (e.g. Google News, Reddit), (2) arrived at an article "directly" (e.g. by going straight to NYTimes.com from your browser), (3) arrived at an article through social media (e.g. by following a link from Facebook or Twitter), and arrived at an article through (4) search (e.g. Google). They find, indeed, that different avenues to online news lead to different degrees of partisan segregation in what people read. For both news articles and opinion articles, the authors find that *differential exposure* (i.e. as measured by partisan segregation) is highest when news articles are accessed either through social media and search (Figure 22).

Figure 22: Political Segregation by Internet Platform

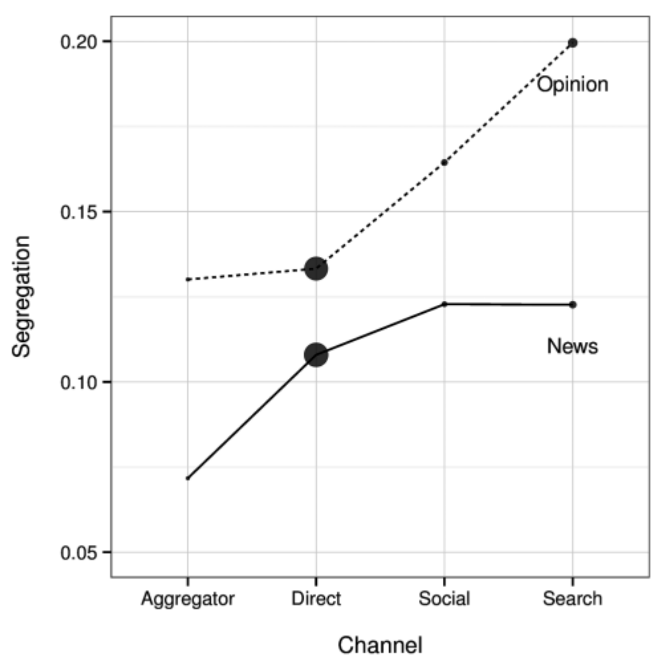


Figure 3. Estimates of Ideological Segregation across Consumption Channels. Point sizes indicate traffic fraction, normalized separately within the news and opinion lines.

In sum, Gentzkow & Shapiro's (2011) results suggest that *differential news exposure* on the internet is comparable to differential news exposure in other news media (e.g. national newspapers), but not astronomically high—with the caveat, from their own analyses and those of Flaxman et al. (2016) that there may be considerably variation in *differential exposure* on the internet, depending on the websites and types of internet use one is focusing on. None of this comparison, however, can actually determine whether *differential exposure* is one of the mechanisms by which internet use increases polarization.

differential news exposure and direct intervention

The clearest evidence that *differential exposure* drives the internet-based polarization would come from showing a tight linkage between internet use and differential exposure, whereby increases in internet usage only result in increases in political polarization when also accompanied by increases in *differential exposure*. I'm not aware of any studies that have exactly this design, but some come pretty close. Together, they provide mixed and conflicting evidence regarding *differential exposure* as a mechanism driving polarization.

First, Allcott et al.'s (2019) direct intervention on internet use (reviewed at the end of Part 2) provides some at least suggestive of a linkage between internet use and *differential exposure*. Some of the steepest reductions that the authors observed among those who quit Facebook for a month were on their measures of *congenial news exposure*. Although they classified this as a measure of political polarization (grouping it in with more classic measures of polarization like *issue polarization*), it could also rightly be considered a self-report measure of *differential exposure*. The measure indeed directly asked participants to reflect on the extent to which they saw news which allowed them to better understand the out-party compared that to the extent to which they saw news which allowed them to better understand their own in-party. Quitting Facebook resulted in almost a fifth of a standard deviation reduction in *congenial news exposure*, evidence that internet use and *differential exposure* might be coupled.

If internet use increases polarization through *differential exposure*, we might also suspect that de-activating Facebook for four weeks would result in larger reductions in polarization among those who most heavily use Facebook for news consumption. Those who use Facebook most heavily for news should be subject to the largest amount of *differential exposure*. However, the authors do not find any moderation of their affects by whether participants were heavy" or "light" news users[32], marshalling slightly against a differential exposure account.

Further complicating this already inconsistent picture is a study from Bail et al. (2018). These authors directly manipulate exposure to cross-cutting views, among active[33] Twitter users and examine subsequent effects on political polarization. This can be thought of as a direct manipulation of *differential exposure*, examining the effects of directly intervening and reducing *differential exposure*, by increasing cross-cutting content.

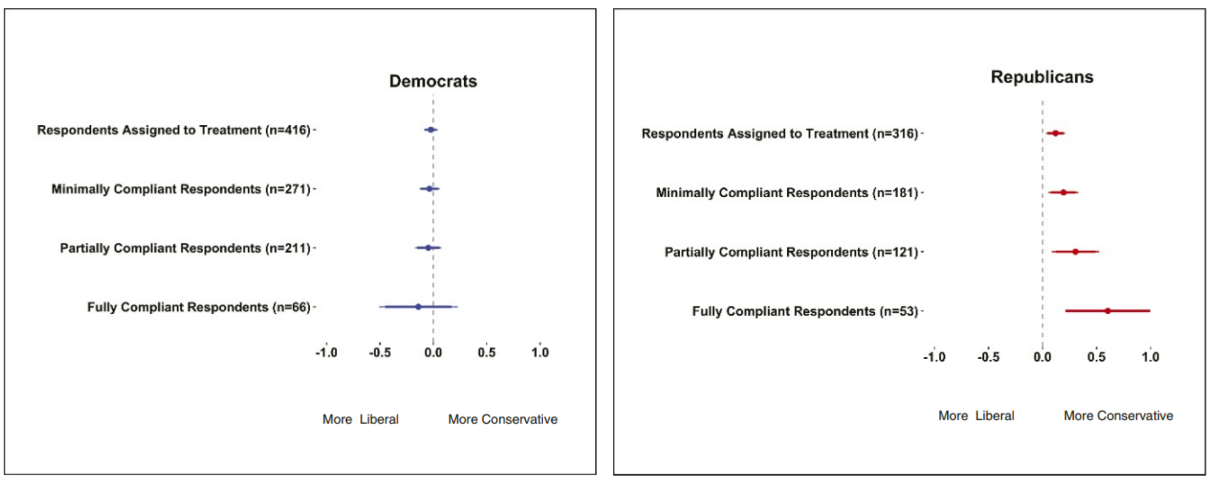
---

[32] Defined as those who report getting their news from Facebook very or fairly often v. sometimes, hardly ever, or never

[33] "active" means they self-report visiting Twitter at least three times per week

The authors study took place over 1-and-a-half-month period. On October, 2017, the authors recruited 1,652 Republicans and Democrats and asked them to report their positions on 10 policy issues. Changes in the extremity of these views constituted their main dependent measure (and correspond to “policy preference” polarization, in the terms used in Part 1, or *issue polarization* in Allcott et al.'s 2019 terminology). These participants were then randomly assigned to either follow a Twitter “bot” or to not do so. Republicans in the treatment condition were assigned to follow a bot that retweeted 24 messages per day, sampled from the Twitter accounts of elected politicians, political organizations, and political “opinion leaders”, who were liberal[34]. Democrats in the treatment condition were assigned to follow a bot that retweeted 24 messages per day, sampled form the Twitter accounts of elected politicians, political organizations, and political “opinion leaders”, who were conservative. No interventions were done on those in the control condition. Compliance, in the treatment condition, was measured throughout the study—by periodically quizzing participants about the content tweeted by the bot they had been assigned to follow. Finally, at the conclusion of the study, the authors asked participants to again record their opinion on the same 10 policy issues from beginning of the study. As can be seen in Figure 23, the end result of greater exposure to cross-cutting messages from popular Twitter users of the opposite part was an *increase* political polarization. This increase in issue polarization was significant among Republicans (at all levels of compliance), but not significant among Democrats (at any levels of compliance).

Figure 23: Issue Polarization as a Result of "Treatment"
(following Twitter-Bot Tweeting Cross-cutting Political Content)



This might seem to marshal against an account whereby internet use increases polarization through greater *differential exposure*. However, some important caveats should be noted. Most importantly, the authors could not definitively rule out the possibility that their treatment increased overall exposure to political information on Twitter, compared to the control

[34] The authors describe a methodology to infer the ideology of Twitter users, based on a principal components analysis of a followership adjacency matrix among their network of political leaders, political organization and opinion leaders. This leads to a continuous measure of ideology; allowing the authors to sample liberals by selecting users who fall on liberal portion distribution and conservatives by selecting users who fall on the conservative portion of the distribution.

condition. That is, those in the treatment condition may have spent more time on Twitter overall—consuming not only more cross-cutting content from the authors' bot, but also consuming more content from like-minded political Twitter users, who they already follow[35]. It is at least possible then that the net effect of their treatment was to actually increase *differential exposure*. If this were the case, then their intervention would have increased political polarization precisely through the expected means, of greater exposure to messages from one's own political "side".

Furthermore, the authors here are examining the effects of a very particular type of (ostensible) reduction in *differential exposure*. They exposed partisans to cross-cutting political content from highly followed political Twitter users (e.g. Hilary Clinton was one account Republicans were exposed to). The messages that they were exposed to as a result of this may have consisted of a large portion of emotionally provocative opinions aimed by political leaders to signal strong stances to their base. Such an intervention might increase political polarization, as members of one political leaning recoil in self-defense at particularly aggressive cross-cutting content. This is not the same as demonstrating that regular internet use increases political polarization by through systematic, long-term differences in exposure to, for example, fact-based news accounts of political events. If nothing else, however, these results show that all differential exposure is not equal. If internet use does increase polarization through *differential exposure*, it must be through a particular type of different exposure—not the type the authors reduced in their study.

In sum, there is no "slam dunk" direct intervention evidence linking *differential exposure* to political polarization; however, none of the evidence clearly rules out this mechanism of polarization either. The results from Bail et al. (2018) suggest that decreasing *differential exposure* does not automatically reduce political polarization. While these results from Allcott et al. (2019) present some mixed, partly supportive, evidence that internet use contributes to political polarization through *differential exposure*.

deliberate and non-deliberate differential exposure

If differential exposure *is* a mechanism of internet-based polarization, one question we may have is the extent to which internet use brings about differential exposure through deliberate or non-deliberate means. Through deliberate means, internet use brings about *differential exposure* by encouraging and increasing conscious choices to consume congenial information, often referred to as "selective exposure" (Sears & Freedman, 1967; Stroud, 2008). Through non-deliberate means, internet use brings about *differential exposure* because of underlying

---

[35] The authors acknowledge they cannot rule out this alternative explanation in the main text of the article. In supplemental analyses, they do conduct one analysis that tries to get at this alternative. They examine whether Twitter usage increases with increasing compliance to the treatment (among only those in the treatment condition). Because increasing compliance with the treatment is associated with larger reduction in polarization, this might suggest that—if the alternative explanation completely accounts for these observed effects—we should see corresponding changes in Twitter usage, according to compliance level. The authors do not find a relationship between *self-reported* Twitter usage and degree compliance with the treatment. However, this says nothing about overall differences between those in the treatment condition and those in the control condition. The authors, for example, do not examine the amount of Twitter activity of those in the treatment condition compared to those in the control condition.

differences in the availability of congenial and non-congenial information that is present on the internet in the first place, due to forces like website algorithms which create "filter bubbles" (Pariser, 2011).

Overall, there is evidence that internet use enables *differential exposure* through both deliberate and non-deliberate means.

Evidence for, at least, the *potential* of differential exposure due to deliberate selective exposure in online information consumption comes from a study by Garrett (2009). In this study, 727 known partisan internet news consumers were recruited from a liberal leaning (AlterNet) and a conservative leaning (WorldNetDaily) website. In an experiment, these users were then presented with real news articles, taken from Google News, on hot-button political issues at the time (gay marriage, social security, civil liberties). Participants were presented with five articles at a time, for which they were given the headline, source, and a two-sentence excerpt. They made a binary choice for each article, ticking a checkbox if they were interested in reading it. They were then asked to rate the extent to which the articles reinforced their existing opinions and the extent to which they challenged their existing opinions. Finally, the participants were given time to actually read any of the articles they had selected. Overall, the authors found that participants were more likely to read an article if it was opinion reinforcing, and were less likely to read an article if it was opinion challenging (although this latter effect was only marginal)[36]. That people deliberately choose to read congenial news content is suggestive that *differential exposure* to political and news information on the internet might also be driven by deliberate, conscious choices. However, these results do not allow us to know whether this is how things actually play out at scale, in people's real internet environments where both deliberate and non-deliberate forces are at play.

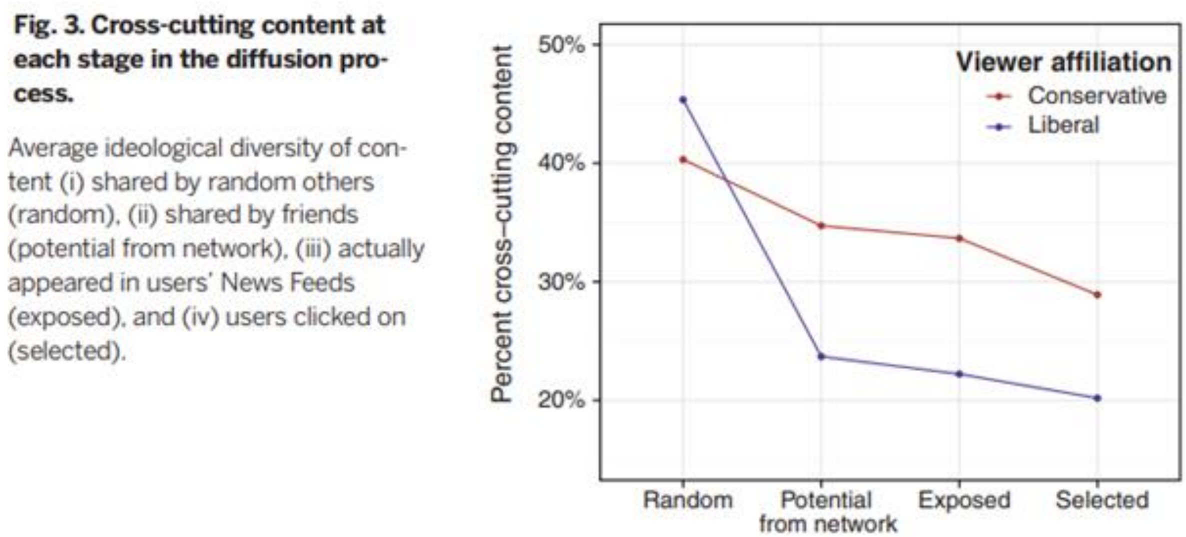
A study which examines the relative contribution of deliberate and non-deliberate sources of *differential exposure* at scale comes from Bakshy et al. (2015), who analyze the news consumption habits of 10.1 million Facebook users from July 7, 2014 to January 7, 2015, using proprietary, internal data from Facebook. The authors examine the degree of exposure to "cross-cutting" news content (i.e. the extent to which liberals were exposed to and clicked on articles shared by conservatives, and vice versa). As can be seen in Figure 24 (on the point along the x-axis labeled "random"), the authors find that if users were exposed to posts selected at random from users across the entire website, 45% of the news content that liberals would see would be "cross-cutting", and 40% of the content conservatives would see would be cross-cutting—revealing the possibility of a fairly low degree of *differential exposure* from the outset. However, various deliberate and non-deliberative forces result in greater and greater *differential exposure*, i.e. reductions in cross-cutting content.

First and most importantly, users are exposed to a much smaller degree of cross-cutting content due homophily in their friendship networks. The median percent of conservatives that liberals are friends with on Facebook is 20%, and the median percent of liberals that conservatives are friends with on Facebook is 18%. In Figure 24 (on the point along the x-axis

[36] The authors also conduct additional analyses examining the time spend on articles once they are clicked (whether they are opinion reinforcing or opinion challenging). Some of these findings are more nuanced, but at a basic level people prefer and select congenial news articles with greater frequency.

labeled "potential from network"), we can see that this reduces the potential cross-cutting content that liberals could be exposed to down to 24% for liberals, and 35% for conservatives. Whether this source of *differential exposure* is exactly deliberate or non-deliberate in nature may be a matter of debate. To the extent that people curate their Facebook friends consciously, this may be rightly classified as a source of deliberate differential exposure, as users are choosing not to attend to those who may be unlike them politically. However, friendship ties on the platform might simply reflect people's offline social connections, many of which are either non-deliberately chosen (e.g. family) or are not a function of deliberate choice online (e.g. college classmates and co-workers are determined by offline school and career decisions). In any case, *differential exposure* due to network homophily is not a source of deliberate *differential exposure* in the form of direct decisions made on political news content itself.

Figure 24: Exposure to Cross-Cutting Content,
Through Deliberate & Non-Deliberate Mechanisms



After this, there are two more sources of *differential exposure*, more clear in their division along deliberate and non-deliberate lines. One is the Facebook News Feed algorithm, which may further limit exposure to cross-cutting content by preferentially promoting certain content, in ways that are clearly outside users deliberate control (i.e. a "filter bubble" effect). The authors are able to quantify the extent of this affect by comparing the amount of cross-cutting content users would be exposed to with and without algorithmic curation[37]. As can be seen in Figure 24 (on the point along the x-axis labeled "exposed"), this algorithmic curation has some, but very little effect on *differential exposure*—reducing cross-cutting content among liberals to 22% (from 24%), and to 34% among conservatives (from 35%). Finally, among the content already filtered by homophilic and algorithmic forces, there is the opportunity for further *differential exposure* due to deliberative decision about which articles users consciously choose to click on and consume. Again, as can be seen in Figure 24 (on the point along the x-axis

[37] Which these particular authors can examine, because they are Facebook employees and have access to internal data and content generating mechanisms. We will have to take them at their word that they estimated their effects reliably.

labeled “selected”), this results in another small reduction in cross-cutting content—bringing cross-cutting content among liberals to 20% (from 22%), and just under 30% (from 34%) among conservatives. Thus, in sum, at least on Facebook, forces of deliberate and non-deliberative filtration bring about *differential exposure*, although non-deliberative “filter bubbles” and deliberative consumption choices due not appear to exert overwhelming effects, at scale, in people’s real internet browsing environments.

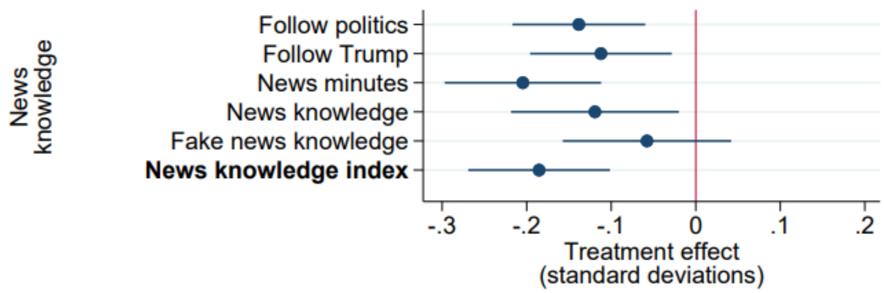
Even among platforms of the same type (e.g. social media), there may be differences in the mechanisms of *differential exposure*. For example, Bakshy et al. (2015) suggests that connections on Facebook are determined by offline social connection to a large extent, while connections on Twitter, for example, are determined more by common content-based interests. These may lead to difference in *differential exposure* by way of different amounts of network homophily, a source of segregated political content. For example, one analysis by Conover et al. (2011) of a set of 250,000 retweets prior to the 2010 midterm elections in the US, suggests that political homophily may be even higher on Twitter than Facebook. The authors identify two major network clusters in their dataset, which they then manually code a random subset of for political orientation. At least in their particular dataset, they find that between 85.7% and 93.4% of retweets occur from users of the same political affiliation. This is even higher than the degree of political homophily that Bakshy et al. (2015) observe on Facebook.

These results suggest that if internet use brings about political polarization through *differential exposure*, that *differential exposure* likely comes about through several different means—some of which are deliberate, like consciously choosing to read reading politically congenial articles, some of which are non-deliberate, like being shown articles highly ranked by a website’s algorithm, and some of which are in between deliberate and non-deliberate, like the extent to which the information available for consumption comes from a network of people who are similar to the user.

amount of exposure

Finally, although not exactly *differential exposure*, internet use might affect polarization simply by reducing the overall *amount* of political news that users consume—which may affect polarization simply by bringing political topics to mind more frequently, possibly ossifying existing political opinions and identities. Indeed, as can be seen in Figure 25, from Allcott et al.'s (2019) study, de-activating Facebook significantly reduced the extent to which participants (1) followed politics and (2) followed news about President Trump. It also reduced the number of (3) minutes participants reported watching, listening to, or reading the news (by 15 percent, or about eight minutes per day). Objective new knowledge was also reduced among participants who de-activated their Facebook accounts. Participants in the treatment condition did significantly worse on a (4) “news knowledge” quiz (0.12 standard deviations lower than in the control condition), although there was no difference in (5) “fake news knowledge” (ability to correctly identify popular fake news stories as false). These news exposure and knowledge reductions occurred specifically through reduction in *online* news consumption. The authors find that their intervention led to participants getting significantly less of their news both from Facebook as well as other (non-Facebook) social media outlets. However, it did not lead to any reductions (nor increases) in news consumption from any other mediums: print, radio, local TV, network TV, Cable TV.

Figure 25: Changes in News Consumption After Quitting Facebook for Four Weeks



Notes: This figure presents local average treatment effects of Facebook deactivation estimated using Equation (1). All variables are normalized so that the Control group endline distribution has a standard deviation of one. Error bars reflect 95 percent confidence intervals. See Section 2.3 for variable definitions.

These results suggest that internet use may affect polarization through changes in the sheer volume of news consumption generally, in addition to any changes in the *differential exposure*.

Polarizing Social Interactions

People's social relationships and social interactions have a profound effect on their political beliefs and attitudes, throughout their lifetimes (Neundorf & Smets, 2017). Up until now, we have been focusing on a mechanism of polarization, which is not very social in nature. *Differential exposure* implicitly characterizes people as information-processors, and accounts for internet-based polarization largely as a result of some sort of distortion to the information fed to these information-processors. However, another major plausible avenue by which internet use might affect political polarization is through a systematic change to social interactions. The internet may contribute to political polarization by enabling *polarizing social interactions*.

This is a plausible mechanism of polarization for a number of reasons. For one, one of the most popular uses of the internet is for various types of socialization or social interaction. Indeed, on ARAPENET, a predecessor to the internet, one of the uses that emerged as most popular was "network mail", i.e. email—which of course allows for virtual communication between users, a form of social interacting (Heart, McKenzie, McQuillian, & Walden, 1978; Naughton, 2001). More relevant, as we saw in the results of a 2017 Nielson survey on internet use (Figure 17), "social" media (combined with search) [38] is either the second or first most popular internet use depending on the surface being examined (computer, smart phone, or tablet), while "communication" is the third most popular use of the internet on both tablets and smart phones. Further, four of the ten most popular websites in the US are social media sites (Facebook, Twitter, Instagram, as well as Reddit, which can be considered both an aggregator

[38] For some odd reason, the authors of that study lump together social media use and search use as one common internet activity, even though they are quite obviously disparate in nature.

site and a social media site). A great deal of internet use thus involves some sort of communication or social interactions between people.

Importantly, these social interactions on the internet differ in systematic ways from their offline counterparts. Social interactions online—whether on Facebook, Reddit, Instagram or the comment sections of news articles—are often characterized by conditions of anonymity, diminished identification, or depersonalization. This, along with several other systematic differences (e.g. less rich representation of others, group immersion) are relevant because, in many circumstances, they can increase aggression and hostility in online social interactions involving politics. In the context of an already heated political environment, these forces may lead to social interactions that exacerbate political polarization.

anonymity

Anonymity or degrees of anonymity, through the veiling of identity, are thought to result in disinhibition. Hirsh, Galinsky, & Zhong (2011) outline a specific pathway by which this might happen. They argue that the behavioral inhibition system (BIS)—a brain network which forms the neural basis of anxiety (Carver & White, 1994; Gray, 1982; Gray & McNaughton, 2000)—serves to halt immediate behavioral reactions, especially during situations in which there are competing possible responses. By halting any immediate response, the BIS allows for a more careful selection among possible behavioral responses. Disinhibition in turn is characterized by situations in which BIS activity is diminished, leading to less careful selection among conflicting alternatives of action, which often leads to the selection of a “dominant response” (one that is the most salient and immediate). Anonymity diminishes BIS activity by reducing a concern that frequently interferes with other goals—social evaluation concerns. Thus, in situations where our dominant response is antisocial in nature, anonymity may increase antisocial behavior. More simply, we might just say that anonymity can lead to more antisocial behavior[39] because it reduces many of the personal consequences of bad behavior (Hoffman, McCabe, & Smith, 1996).

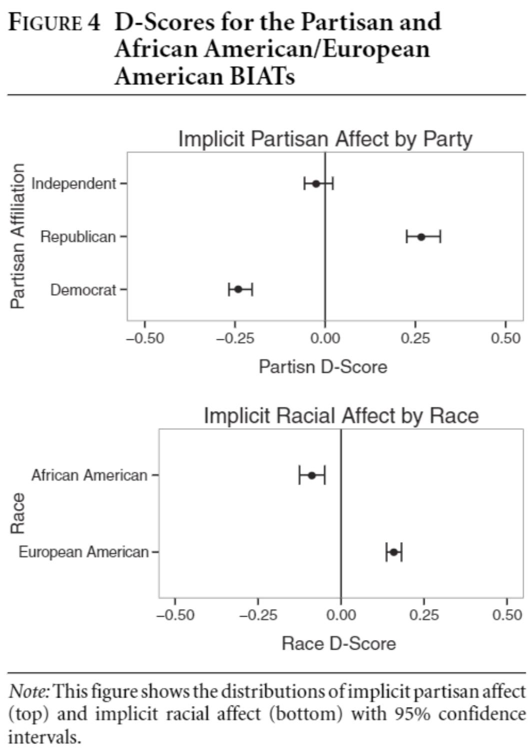
For the anonymity that accompanies internet usage to result in more aggressive and antisocial *political behavior*, our dominant response in this realm must be of an antisocial nature. And there is evidence to support this possibility. Of course, classic “minimal group” paradigm demonstrations (Tajfel, 1970; Tajfel, Turner, Austin, & Worchel, 1979) show just how easily we are willing to favor in group members at the expense of out-group members, even when the in-group and out-group distinctions are much less meaningful than the political distinctions which color our contemporary world. And there is indeed evidence of our strong favoritism towards our own political party and antagonism towards opposing political parties,. A notable demonstration comes from Iyengar & Westwood (2015), who had Democrats and Republicans complete a

---

[39] Outside the realm of politics, there are many demonstrations of how anonymity can lead to more antisocial behavior. For example, people who have the opportunity and may be tempted to steal or cheat are more likely to do so when their visibility is obscured by darkness or lack of personal identification (Diener, Fraser, Beaman, & Kelem, 1976; Zhong, Bohns, & Gino, 2010). Likewise, people act more on their impulses to drive aggressively in driving simulators when they are more visibly obscured to other drivers (e.g. because of tinted windows) (Ellison-Potter, Bell, & Deffenbacher, 2001). People engage in more physical aggression the less they attend to themselves (Scheier, Fenigstein, & Buss, 1974).

modified (partisan) version of the brief implicit associations test (BIAT), where reactions times are compared between the speed of categorizing various stimuli as either “Republicans or Good” to the speed of categorizing various stimuli as either “Democrats or Good”. These authors also had participants complete a racial version of the BIAT. Their sample included members not only of different political orientations, but also different racial groups (i.e. African Americans and European Americans). As can be seen in Figure 26, the difference in valanced reaction times was larger between Democrats and Republicans on the partisan BIAT than between African Americans and European Americans on the racial IAT. This suggests that there may strong latent negative associations between these parties. (And indeed, as reviewed when discussing affective polarization in Part 1, there is evidence of very conscious bad feelings as well.)

Figure 26: Implicit Associations, Between Party and Between Race



In another experiment, Iyengar & Westwood (2015) had Republicans and Democrats engage in a hypothetical task where they selected between various candidates for a $30,000 academic scholarship—some of which were identifiably Republicans or Democrats (e.g. leader of campus Republican group). They found huge partisan differences in selection rates, that were again larger than differences when the same task was repeated among European Americans and African Americans selecting between identifiably European American and African American candidates. Unlike with the racial version of the task, this effect persisted not only when candidates were equally qualified but also when the out-group candidate was more qualified (e.g. when the identifiably Republican candidate was more qualified, there was still only around 30% chance that a Democrat would pick them for the scholarship). Further, in both dictator and trust games, Iyengar & Westwood (2015) found significant smaller allocations of money when partisans of opposite parties were paired (e.g. a Democrats and Republican) than when partisans of the same party were paired (e.g. a Democrat and another Democrat). All of this also seems comport with daily experience. We are quite willing to express strong favoritism of our political party and hostility towards members of the other "side". And while most of us would be mortified to express racial animus, we are quite willing to openly, actively, and publicly denigrate members of the opposite political party as stupid, ill-intentioned or worse.

These factors combine into the exact right ingredients that would be needed for internet use to lead to more *polarizing social interactions*. The greater veil of anonymity that accompanies online social interactions[40], combined with strong partisan differences in affect and in-group favoritism, may lead people to act on their "dominant responses" to exclude, alienate and aggress against those of the opposite political orientation. The net result of this would seem to be to increase political polarization, for example increase *affective polarization*. There is no demonstration of this entire causal pathway. However, there is highly suggestive evidence from demonstrations of the effects on online anonymity on other social interactions.

In a highly-cited theoretical paper in the field of "cyberpsychology", Suler (2004) posits a general *online disinhibition effect*, whereby the anonymity of internet use may contribute to "toxic disinhibition" and antisocial behavior. This is accompanied by empirical support. For example, in a longitudinal study on 146 college students over a year, Barlett, Gentile, & Chew (2016) found that especially among those who have an inclination towards bulling others, greater initial feelings of anonymity online[41] are linked to higher incidences of reported "cyberbullying" over time.

Better evidence on the disinhibiting effects of anonymity in internet use comes from Lapidot-Lefler & Barak (2012). They begin by carefully delineated that there are actually many sources of what they term an *online sense of unidentifiability*. They highlight three particularly notable sources of unidentifiability: anonymity, visibility, and eye-contact. Anonymity, in their definition, concerns not having one's name (and other personally identifying information) known. Visibility concerns seeing some visual representation of a person (you may see someone in a picture or video chat, but still not know their name). And eye contact, of course, involves looking someone in the eyes.

They then examine the effects of each source of unidentifiability on aggressive behavior online, in an experiment, where they orthogonally manipulate each of these three factors

[40] Internet use often takes place behind a total or partial veil of anonymity. Least anonymous are websites like Facebook, where all users post under their real names. Nevertheless, there are still features of the website that hide identity. There is no face-to-face communication; users do not look each other in the eye, nor are they in each others physical company at all when communications occur. Furthermore, many communications are extended over time; unlike real conversation, back and forth exchanges may be separate by many hours. On other websites, like Twitter, people may choose between posting under their real names or under an online alias. And on other websites, like Reddit or 4chan, all users post through completely anonymous accounts—where there are indeed strong rules and norms against "doxxing" (i.e. revealing a user's identity or other personal information).

[41] Note, however, that the authors method of assessing perceived anonymity is not the most clear cut measure one could imagine. We might want to author to simply measure how anonymous users feel online, with questions like "when I'm interacting with others online, I feel that I can't be seen" (or something of this sort). The author's assessment of anonymity, instead includes questions like "Sending mean e-mails or text messages is easy to do because I am not face-toface with the other person", which seem to lump together anonymity with their predicted effects. So, subsequent correlations with cyberbullying are perhaps artificially inflated.

(resulting in a 2x2x2 design: anonymity [yes/no] X visibility [yes/no] X eye-context [yes/no]). To do this, they have participants take part in an online chat with one other person, where the pair engage in a discussion about the allocation of a limited resource (who between the two of them should be allocated a life-saving drug). Anonymity was manipulated by either presenting participants names to each other or not doing do. Visibility was manipulating by setting up a web camera, which makes the upper portions of participants bodies visible to each other (those in the "invisibility" condition simply had no web camera). And finally, eye-contact was manipulated by setting up a second camera focused on participants faces specifically and explicitly instructing participants to engage in eye-contact throughout the course of their communication.

Subsequent negative behavior was then assessed in three different ways. (1) A panel of experts read the transcripts of the communications between dyads and rated the extent to which there was the presence of "threats" and a "negative atmosphere". (2) Coders examined the transcripts, counting the number of instances of pre-specified words and expression, indicating hostility (e.g. swear words, symbolic aggression through means like capitalization and punctuation). (3) Finally, participants themselves rated the extent to which they believed they expressed hostility and aggression in the exchange (e.g. "I had trouble controlling my temper").

In their analysis, the authors found evidence that each of the three types of online unidentifiability (anonymity, lack of visibility, and lack of eye-contact) contributed to online hostility, on at least some of the various measures. The factor which they found to account the most to online hostility was the lack of eye-contact.

All of these sources of *online unidentifiability*—anonymity, lack of visibility, and lack of eye-contact—are present in online social interactions writ large. It is reasonable to assume that if they have the effect of making dyads of participants more hostile and aggressive towards each other in a debate about the hypothetical allocation of resources, it could easily make actual partisans more hostile and aggressive towards each other in very non-hypothetical online political discussions (e.g. discussions of how to allocate federal budget resources among options like military spending or funding for research and the arts).

Indeed, a study by Santana (2014) presents relevant (but not definitively causal) evidence of the role that anonymity might play in reducing civility in online political discussions. By way of background, the authors point out that among the top 100 most popular newspapers, as of 2010, 92% allowed online comments (up from 75% in 2008, and just 33% in 2007) (Johnson, 2008; Santana, 2011). And among the top 137 newspaper (classified as having a circulation of 50,000 or more), 42% allow anonymous commenting, while 49% have commenting that is not anonymous (the remaining 9% don't have commenting of any type). In their study, the authors examine the civility of comments left on online newspaper articles, comparing newspapers that allow anonymous commenting to newspapers that allow commenting but require it be non-anonymous. Because it is a topic that has the potential to invoke incivility, the authors examine comments left on news articles focusing on immigration. Ideally, they would be able to compare newspaper that are exactly alike, but only differ in whether they allow anonymous commenting or not. This is not possible however. And the authors instead aim to compare mostly newspapers in similar geographic areas (focusing on California, Texas, and Arizona), where immigration is a hot button issue. They compare the civility of comments left on online newspapers that allow anonymous commenting (e.g. the Los Angeles Times, the Houston Chronicle), with the civility of comments left on newspapers that allow commenting but require it not be anonymous (e.g. San Jose Mercury News, El Paso Times)[42]. The newspaper that don't allow anonymous commenting most typically use a Facebook plugin[43], which requires users to login in through Facebook and post comments under their Facebook account information (less commonly, the non-anonymous newspaper require users to make an account, under the name, to leave comments; e.g. the Wall Street Journal does this).

To examine the civility of comments, the authors had a group of coders read through a randomly select subset of 450 comments from both the anonymous and non-anonymous newspaper and code them as either civil or not civil. A comment was judged uncivil if it contained "personal or inflammatory attacks, threats, vulgarities, abusive or foul language, xenophobic or other hateful language or expressions, epithets or ethnic slurs, sentiments that are racist or bigoted, disparaging on the basis of race/ethnicity or that assign stereotypes". Ultimately, the authors found that while 29% of the comments left on newspaper which forced non-anonymous commenting were uncivil, 53% of comments on the newspapers that allow anonymous commenting were uncivil. While lacking in the causal force that accompanies a true randomized experiment, these results do strongly suggest that anonymity may bring about more hostile political discussions online—which may contribute to greater political polarization, by further entrenching partisan in their policy opinions (e.g. *issue polarization*) and leading to more negative sentiments between partisans (e.g. *affective polarization*).

everything else (shallow representations, online mobs, lower social costs)

Diminished identification is only one way in which the nature of online social interaction may bring about political polarization. I will now briefly review a host of other very plausible (but less empirically verified) differences in online social interaction that may also exacerbate political polarization. These are: shallower representations of others, immersion in online crowds, and lower social costs.

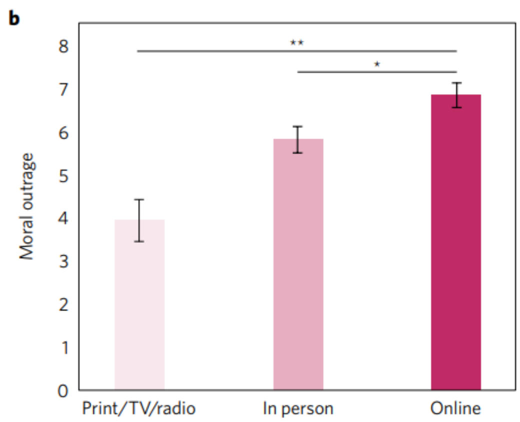
As noted, online anonymity might embolden possible aggressors because their own identity is diminished. However, conditions of anonymity might also increase online aggression because of the diminished identity of those who might be aggressed against. While we are often averse to harming others (Crockett, Kurth-Nelson, Siegel, Dayan, & Dolan, 2014), the greater physical separation we have from others and the less richly we represent others in our minds, the more we are likely to dehumanize them and aggress against them. We know, for example, from Milgram's classic obedience studies that "teachers" deliver more severe shocks to "learners" when there is more of a physical separation between the two (Milgram, 1963, 1974). More severe shocks were delivered as teachers moved from having to place the "learners" hand on a plate, to only being in the same room as the learner, to being in a different room and only hearing

[42] Although they try to bring some geographic balance between the newspaper that allow and don't allow anonymous comments, they certainly do not exactly achieve this goal. Their sample of newspaper which don't allow anonymous comments includes more national newspaper (e.g. the Wall Street Journal) and newspaper from other regions (e.g. the Detroit Free Press). In another attempt to adjust for this, they also sample a large portion of their comments from those newspaper in the focal geographic regions: California, Texas, and Arizona.
[43] Facebook Comments plugin (https://developers.facebook.com/docs/plugins/comments/)

the learner, to neither seeing nor hearing the learner. It is easy to imagine how similar dynamics might play out online, in the context of a heated political discussion. The physical separation of political opponents might increase the "severity" (in terms of words and insults) with which each side is willing to treat the other. Further, the lack of interpersonal contact in online social interactions is likely to lead discussants to mentally represent each other with diminished richness and greater shallowness. A relevant finding comes from Schroeder & Epley (2015) who find that hearing someone's voice during a social exchange (e.g. interview) has a humanizing effect. Interviewees whose voices are heard are evaluated as more competent and intelligent. It is again easy to imagine the effects this might have on online social interactions. Social interactions online often take the form of textual exchanges (e.g. arguments in online news comments, or Twitter feuds), which of course lack humanizing elements like the sound of another's voice. This may lead us to represent our political "opponents" less richly and less humanly than if we were interacting with them in person. This too might result in less tempering of bad feelings and more aggression against political "opponents", which are exactly the symptoms of *affective polarization*.

Another notable aspect of online interactions is that they often take place with some sort of online "crowd" in the background. Whether sharing a post on Facebook, making a point on Twitter, or commenting on a news article, there is in the backdrop a host of others, from personal contacts (e.g. friends, family, coworkers) to quasi-acquaintances, public figures (e.g. celebrities, journalists, politicians, professionals in one's field), and online strangers. Offline, we know that immersion in a groups, crowds, or "mobs" is often accompanied by more disinhibition, "deindividuation", and antisocial behavior (Ed Diener, Lusk, DeFour, & Flax, 1980; Festinger, Pepitone, & Newcomb, 1952; Le Bon, 1895; Singer, Brush, & Lublin, 1965; Zimbardo, 1969)—everything from more Halloween-candy stealing (Edward Diener, Fraser, Beaman, & Kelem, 1976) to greater suicide baiting (Mann, 1981). However, we also know that groups may come to more extreme decisions and extreme attitudes—a phenomenon know, in fact, as "group polarization" (Isenberg, 1986; Myers & Lamm, 1976). There is evidence for two mechanisms which bring this group polarization. One posits that such polarization happens through exposure to a greater number of ideas supporting one's pre-existing convictions during group discussion (Burnstein & Vinokur, 1977; Isenberg, 1986)—which, online, might be similar to *differential exposure* mechanism of internet-based political polarization that we examined. The other mechanism posits that group polarization happens as everyone in the group tries to outdo each other and come across as the most righteous or correct (Isenberg, 1986; Sanders & Baron, 1977). The online manifestation of this latter mechanism may be things like "Twitter mobs" and online moral outrage cascades, where thousands and thousands of users pile on in increasing moral indignation, anger, and disgust at a moral or political offense (Boot, 2019; M. Miller, 2019; Ronson, 2016; Williamson, 2018). Because everyone wants to demonstrate their righteous to members of their own political group, the entire emotional atmosphere of a discussion becomes more extreme and hostile. Some very indirect empirical support for this possibly can be found in a few places. Hofmann, Wisneski, Brandt, & Skitka (2014) conducted an experience sampling study in which people were messaged five times a day and asked to record various moral acts which they had witnessed, experienced or were a part of. A re-analysis of this data by Molly Crockett (2017) finds that moral outrage is higher in response to offenses heard about online than in person or through other media sources (print, TV, and radio) (see Figure 27). This may be due to the background presence of online crowds, crowds which are not usually present in person or when reading print news or watching news at home. Further, a study by Brady, Wills, Burkart, Jost, & Van Bavel (2018) finds that tweets from presidential candidates, senators, and congressional representative receive more retweets if they contain more moralizing language. Again, one possible explanation for this is that the backdrop of hundreds or thousands of others users on Twitter heightens each users concern with appeasing and pleasing the group, leading to greater moralizing and political righteousness. All this may contribute to political polarization by making people's political views more emotional intense and extreme.

Figure 17: Moral Outrage by Medium



Finally, as some have pointed out (e.g. Crockett, 2017), negative, hostile, or aggressive social interactions over the internet might be less costly than offline. Even without the veil of anonymity and even when parties represent each other "richly", the physical separation that social interactions over the internet entail necessarily eliminate certain costly social consequences of bad behavior. Calling someone you are having a political argument with a "m@therfucker!!" in person might get you punched in the face; no such possibility exists online.

Summary

So what can we say overall about *how* internet use might bring about political polarization? Social interactions over the internet clearly differ from their offline counterparts. There is fair evidence that greater anonymity can lead to more hostility and incivility in computer and internet mediated interactions. And it is reasonable to suspect that this could contribute to political polarization, by making partisans more hostile and aggressive towards each other. Likewise, shallower representations of others online, immersion in moralizing online "crowds", and the lower costs of online social interactions are also factors which may plausibly contribute to social interactions that are more politically polarizing. However, strong empirical demonstrations are lacking. And it is also unclear how such factors play out at scale. Better empirical evidence comes with regard to the other major mechanism of internet-based political polarization—*differential exposure*. Here the evidence suggests that claims of completely segregated online "echo chambers" and "filter bubbles", with partisans living in completely different online worlds, are likely overstated. Nevertheless, there are significant and meaningful differences in the political content that partisans of different political orientations consume online, comparable to the degree of segregation in national print newspaper readership. The causal evidence linking this actual degree of *differential exposure* to political polarization, however, is not as strong as the evidence that this degree of *differential exposure* exists.

One final possibility should be noted. Up until now, quietly in the background, we have assumed that political polarization is rising because something has “gone wrong”. And the internet might have contributed by enabling or increasing certain “bad” forces—*differential exposure*¸ or hostile anonymous interactions. However, the great hope during the early rise of the internet was that it would be a communication technology that would open up the world and lead to wider, freer communication and exchange between all people (Naughton, 2001). Perhaps it has done so, but we are surprised by what we see. A study by Jensen (2007) shows the revolutionary effects that the introduction and wide-spread adoption of mobile phones had on the fishing industry in Kerala, India—leading to dramatic reductions in price dispersion and market asymmetries between different regions, as lighting fast and ubiquitous information exchange was enabled by these mobile phones (Figure 28). What if the internet has had a similar effect on the political realm? Perhaps the widespread adoption of Twitter, Facebook and smartphones has allowed for more accurate perceptions of the political conflict that is rife in our society—from racial conflict between African American communities and police forces to deep seated differences in cultural values between urban coastal centers and more rural conservative areas that have simmered since the founding of our country. Perhaps previous levels of polarization were artificially low; the semblance of political unity maintained by a media echo system which was unnaturally restricted and homogenized. There is no objective “baseline” of political polarization. Just because polarization was lower in the past does not mean that political sentiment was somehow more calibrated to reality than it is now. Perhaps all the internet has done is enable us to see more clearly just how much latent political conflict there is in our country.

Figure 28: Price Dispersion Before and After Mobile Phone Introduction

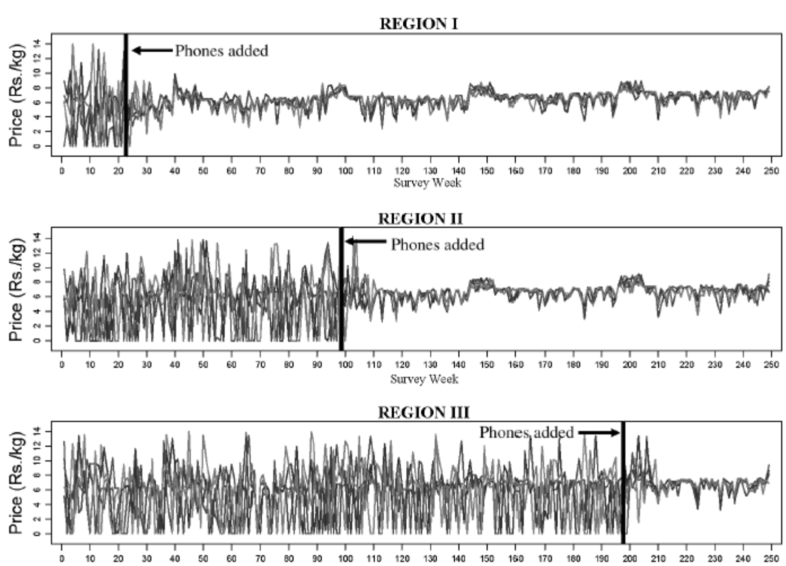


FIGURE IV
Prices and Mobile Phone Service in Kerala
Data from the Kerala Fisherman Survey conducted by the author. The price series represent the average 7:30–8:00 A.M. beach price for average sardines. All prices in 2001 Rs.

FINAL SUMMARY

In sum, it seems that political polarization among citizens in the US has been on the rise—especially when measured by differences between Republicans and Democrats on policy preferences or the emotions and attitudes they have toward each other. The conclusions we can reach about the role of internet are more tempered. Almost certainly, the large increases in internet use over the last few decades are not the primary cause for the increases in political polarization over that same time period. However, internet use does appear to play some causal role in bringing about political polarization—as evinced by the fact that quitting the most popular social media platform on the planet for a month results in significant reductions in forms of polarization, like polarization over policy preferences. The case for “echo chambers”, “filter bubbles”, and completely separate online worlds for partisans is overstated, yet differences do exist, and it seems reasonable that they might contribute to some extent of internet-based political polarization. Other ways that internet use alters social interaction—more anonymity, more moral “mob-ishness”, lower interpersonal costliness to social infractions, and shallower representations of others—may also play at least some role in bringing about political polarization by making political interactions on the internet more hostile. Although strong empirical confirmation awaits.

These tempered conclusions about the role of internet use in bringing about polarization might seem out of line with the apparently large amounts of vitriol we see online. How can there be such meager evidence that internet use contributes to polarization, when our daily online experience is often so politically contentious? One possibility is simply that the internet is a communication technology which allows people to express and exhibit high degrees of polarization, which nevertheless have their origins elsewhere. Like any other vector of polarization—print news or face to face interactions—it has its own idiosyncrasies and personal contribution. But primarily, the internet might simply be a new theater for an already existing conflict. A war that has moved from land to the sea is not caused by the sea.

APPENDIX

Data Sources for Gentzkow & Shapiro (2011)

- To collect the information needed to infer ideological segregation on *online media*, the authors get their data from comScore, a popular private media research firm. Specifically, comScore assesses web traffic across the internet by having a panel (of a million people) download software which tracks their browsing behavior. A subset of 12,000 people in that group have also been administered a survey asking about their political affiliation. This allows the authors to compute an isolation index score across the websites visited by these participants for whom political affiliation is also know. The authors restrict their analysis visits to websites which comScore labels as “General News” or “Politics.” Ultimately, this ends up being a set of 119 high traffic news websites. This allows them to compute an isolation index for the medium of (1) internet news.
- To collect the information needed to infer ideological segregation across various form of *offline media*, the authors rely on data from Mediamark Research and Intelligence (MRI), which is a media research firm that conducts national surveys to collect data on people’s media consumption habits. Because MRI asks about political affiliation in some of their surveys, the authors are able to compute an isolation index score for the *offline mediums*: (2) cable news, (3) broadcast TV, (4) national newspapers, and (5) local newspapers. (To give an intuition behind how the isolation index is computed in each of these cases, take the example of Cable News. From the dataset, the authors can compute the “share conservative” for channels like MSNBC, CNBC, Fox News and so on, then compute the average conservative exposure among liberals and conservatives, and use the difference to compute an isolation index.)
- To make inferences about the ideological segregations in *face-to-face interactions*, the authors use data from the General Social Survey and Cross-National Election Survey (CNES). The GSS asks respondents to describe the demographic characteristics (including political affiliations) of people in their personal social networks. Specifically, they ask respondents to describe the demographic characteristics of these groups: the respondent’s family, the respondent’s co-workers, people the respondent considered part of their network of people who they “trust,” people, the respondent knows through mutual membership in voluntary associations (e.g. school, church, clubs), and members of the respondent’s neighborhood. Specifically, for each group, respondents are asked to estimate the number in each group that are conservative and the number in each group that are liberal. This allows the authors to compute an isolation index for each of the following mediums of *face-to-face interaction*: (5) family, (6) work, (7) people who are “trusted”, (8) mutual membership in a voluntary association group, (9) people named as part of respondent’s neighborhood. The CNES meanwhile asks people to list out several people with whom the respondents discuss politics. These listed people are then actually contacted by the CNES staff and asked for their political ideology. This allows the authors to compute an isolation index for the *face-to-face interaction* medium of (10) people with whom politics are discussed
- Finally, the MRI data also retains information about survey participants’ geographic location, allowing the authors to assess ideological segregation in the “medium” of *local physical communities*. Specifically, the data allows the authors to compute an isolation index score of two geographic local physical community units: (11) one’s county and (12) one’s zip code (“mediums” of interaction and information, under a broader meaning of “medium”).

REFERENCES

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics*, *70*(2), 542–555.

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2019). The Welfare Effects of Social Media. *National Bureau of Economic Research*, 114.

Almond, G., & Verba, S. (1960). *Civic Culture Study, 1959–1960* (No. ICPSR07201-v2. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2009-02-12.).

American National Election Studies. (2018). Retrieved April 18, 2019, from ANES | American National Election Studies website: https://electionstudies.org/data-center/

Ansolabehere, S., Rodden, J., & Snyder, J. M. (2006). Purple america. *Journal of Economic Perspectives*, *20*(2), 97–118.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., … Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Barlett, C. P., Gentile, D. A., & Chew, C. (2016). Predicting cyberbullying from anonymity. *Psychology of Popular Media Culture*, *5*(2), 171.

Bartels, L. M. (1998). Electoral continuity and change, 1868–1996. *Electoral Studies*, *17*(3), 301–326.

Bishop, B. (2009). *The big sort: Why the clustering of like-minded America is tearing us apart*. Houghton Mifflin Harcourt.

Boot, M. (2019, January 31). On Wednesday, the Twitter mob came for me. *Washington Post*. Retrieved from https://www.washingtonpost.com/opinions/2019/01/31/wednesday-twitter-mob-came-me/

Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences*, 201706588.

Brady, W. J., Wills, J. A., Burkart, D., Jost, J. T., & Van Bavel, J. J. (2018). An ideological asymmetry in the diffusion of moralized content on social media among political leaders. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000532

Burnstein, E., & Vinokur, A. (1977). Persuasive argumentation and social comparison as determinants of attitude polarization. *Journal of Experimental Social Psychology*, *13*(4), 315–332.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*(2), 319.

Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *Fifth International AAAI Conference on Weblogs and Social Media*.

Crockett, M. J. (2017). Moral outrage in the digital age. *Nature Human Behaviour*, *1*(11), 769.

Crockett, Molly J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, *111*(48), 17320–17325.

Cuthbertson, A. (2017, November 2). Google and Facebook's dominance could lead to the "death of the internet," warns programmer. Retrieved April 23, 2019, from Newsweek website: https://www.newsweek.com/facebook-google-internet-traffic-net-neutrality-monopoly-699286

Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (1999). The rise and decline of the American ghetto. *Journal of Political Economy*, *107*(3), 455–506.

Diener, Ed, Lusk, R., DeFour, D., & Flax, R. (1980). Deindividuation: Effects of group size, density, number of observers, and group member similarity on self-consciousness and disinhibited behavior. *Journal of Personality and Social Psychology*, *39*(3), 449.

Diener, Edward, Fraser, S. C., Beaman, A. L., & Kelem, R. T. (1976). Effects of deindividuation variables on stealing among Halloween trick-or-treaters. *Journal of Personality and Social Psychology*, *33*(2), 178.

Ellison-Potter, P., Bell, P., & Deffenbacher, J. (2001). The effects of trait driving anger, anonymity, and aggressive stimuli on aggressive driving behavior. *Journal of Applied Social Psychology*, *31*(2), 431–443.

Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of de-individuation in a group. *The Journal of Abnormal and Social Psychology*, *47*(2S), 382.

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annu. Rev. Polit. Sci.*, *11*, 563–588.

Fitzgerald, B. R. (2014, May 20). Data Point: Yes, Email Still Eats Up a Good Chunk of Your Day. Retrieved April 23, 2019, from WSJ website: https://blogs.wsj.com/digits/2014/05/20/data-point-yes-email-still-eats-up-a-good-chunk-of-your-day/

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*(S1), 298–320.

Fleisher, R., & Bond, J. R. (2004). The shrinking middle in the US Congress. *British Journal of Political Science*, *34*(3), 429–451.

Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of Computer-Mediated Communication*, *14*(2), 265–285.

Gentzkow, M. (2016). Polarization in 2016. *Toulouse Network of Information Technology White Paper*.

Gentzkow, M., & Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, *126*(4), 1799–1839.

Glaeser, E. L., & Ward, B. A. (2006). Myths and realities of American political geography. *Journal of Economic Perspectives*, *20*(2), 119–144.

Gramlich, J. (2019, February 1). 10 facts about Americans and Facebook. Retrieved April 19, 2019, from Pew Research Center website: https://www.pewresearch.org/fact-tank/2019/02/01/facts-about-americans-and-facebook/

Gray, J. A. (1982). *The neuropsychology of anxiety*. New York, NY: Oxford University Press.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford, England: Oxford University Press.

Guess. (2018). (Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets. *Unpublished Manuscript*, 56.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council*.

Heart, F., McKenzie, A., McQuillian, J., & Walden, D. (1978). *The Completion Report* (No. BBN Report 4799). Retrieved from http://walden-family.com/bbn/arpanet-completion-report.pdf

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

Hirsh, J. B., Galinsky, A. D., & Zhong, C.-B. (2011). Drunk, powerful, and in the dark: How general processes of disinhibition produce both prosocial and antisocial behavior. *Perspectives on Psychological Science*, *6*(5), 415–427.

Hoffman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, *86*(3), 653–660.

How 1994 Gave Us Today's Politics. (2018, October 25). *The New York Times*. Retrieved from https://www.nytimes.com/2018/10/25/podcasts/the-daily/midterm-elections-trump-clinton-gingrich.html

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, *50*(6), 1141.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science*, *22*(1), null. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690–707.

Jacobson, G. C. (2000). Party polarization in national politics: The electoral connection. *Polarized Politics: Congress and the President in a Partisan Era*, *5*, 17–18.

Jacobson, G. C. (2003). Partisan polarization in presidential support: The electoral connection. *Congress & the Presidency: A Journal of Capital Studies*, *30*, 1–36. Taylor & Francis.

Jensen, R. (2007). The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics*, *122*(3), 879–924. https://doi.org/10.1162/qjec.122.3.879

Johnson, J. (2008). The use of the Internet by America's largest newspapers (2008 Edition). *Bivingsreport. Com, December*, *18*.

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, *60*, 307–337.

LaFrance, A. (2015, September 30). How Many Websites Are There? Retrieved April 22, 2019, from The Atlantic website: https://www.theatlantic.com/technology/archive/2015/09/how-many-websites-are-there/408151/

Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, *28*(2), 434–443.

Le Bon, G. (1895). *The Crowd: A Study of the Popular Mind*.

Lelkes, Y., Sood, G., & Iyengar, S. (2017). The hostile audience: The effect of access to broadband internet on partisan affect. *American Journal of Political Science*, *61*(1), 5–20.

Levendusky, M. (2009). *The partisan sort: How liberals became Democrats and conservatives became Republicans*. University of Chicago Press.

Levendusky, M., & Malhotra, N. (2016). Does media coverage of partisan polarization affect political attitudes? *Political Communication*, *33*(2), 283–301.

Mann, L. (1981). The baiting crowd in episodes of threatened suicide. *Journal of Personality and Social Psychology*, *41*(4), 703.

Mason, L. (2018). *Uncivil agreement: How politics became our identity*. University of Chicago Press.

Mayer, W. G. (1998). Mass Partisanship, 1946-1996. *Partisan Approcahes to Postwar American Politics*, 186–219.

McGregor, S. C., & Molyneux, L. (2018). Twitter's influence on news judgment: An experiment among journalists. *Journalism*, 1464884918802975.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444.

Milgram, S. (1963). Behavioral Study of Obedience. *Journal of Abnormal and Social Psychology*.

Milgram, S. (1974). *Obedience to Authority: An Experimental View*. Harper & Row.

Miller, M. (2019, January 22). Viral standoff between a tribal elder and a high schooler is more complicated than it first seemed. Retrieved May 3, 2019, from Washington Post website: https://www.washingtonpost.com/local/social-issues/picture-of-the-conflict-on-the-mall-comes-into-clearer-focus/2019/01/20/c078f092-1ceb-11e9-9145-3f74070bbdb9_story.html

Miller, W. E. (1991). Party identification, realignment, and party voting: Back to the basics. *American Political Science Review*, *85*(2), 557–568.

Munger. (2019). *Mapping YouTube*. https://doi.org/None

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*(4), 602.

Naughton, J. (2001). *A brief history of the future: The origins of the internet* (4. impr). London: Phoenix.

Neundorf, A., & Smets, K. (2017). *Political socialization and the making of citizens*.

Pappas, S. (2016, March 18). How Big Is the Internet, Really? Retrieved April 23, 2019, from Live Science website: https://www.livescience.com/54094-how-big-is-the-internet.html

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Perrin, A., & Jian, J. (2018, March 14). A quarter of Americans are online almost constantly. Retrieved April 22, 2019, from Pew Research Center website:

https://www.pewresearch.org/fact-tank/2018/03/14/about-a-quarter-of-americans-report-going-online-almost-constantly/

Pew Research Center. (2014). *Political Polarization in the American Public.* Retrieved from http://www.people-press.org/2014/06/12/political-polarization-in-the-american-public/

Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, *16*, 101–127.

Ronson, J. (2016). *So you've been publicly shamed*. Riverhead Books.

Roose, K. (2019, June 8). The Making of a YouTube Radical. *The New York Times*. Retrieved from https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html, https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html

Sanders, G. S., & Baron, R. S. (1977). Is social comparison irrelevant for producing choice shifts? *Journal of Experimental Social Psychology*, *13*(4), 303–314.

Santana, A. D. (2011). Online readers' comments represent new opinion pipeline. *Newspaper Research Journal*, *32*(3), 66–81.

Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice*, *8*(1), 18–33.

Scheier, M. F., Fenigstein, A., & Buss, A. H. (1974). Self-awareness and physical aggression. *Journal of Experimental Social Psychology*, *10*(3), 264–273.

Schiffer, A. J. (2000). I'm not that liberal: Explaining conservative democratic identification. *Political Behavior*, *22*(4), 293–310.

Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological Science*, *26*(6), 877–891.

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, *31*(2), 194–213.

Shi, F., Shi, Y., Dokshin, F. A., Evans, J. A., & Macy, M. W. (2017). Millions of online book co-purchases reveal partisan differences in the consumption of science. *Nature Human Behaviour*, *1*(4), 0079.

Singer, J. E., Brush, C. A., & Lublin, S. C. (1965). Some aspects of deindividuation: Identification and conformity. *Journal of Experimental Social Psychology*.

Sood, G., & Lelkes, Y. (2016). *Dont expose yourself: Discretionary exposure to political information*. Working paper.

Staltz, A. (2017, October 30). The Web began dying in 2014, here’s how. Retrieved April 23, 2019, from https://staltz.com/the-web-began-dying-in-2014-heres-how.html

Stewart, J. B. (2016, May 5). Facebook Has 50 Minutes of Your Time Each Day. It Wants More. *The New York Times*. Retrieved from https://www.nytimes.com/2016/05/06/business/facebook-bends-the-rules-of-audience-engagement-to-its-advantage.html

Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, *30*(3), 341–366.

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, *7*(3), 321–326.

Sunstein, C. (2007). *Republic.com 2.0*. Princeton, New Jersey: Princeton University Press.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96–103.

Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational Identity: A Reader*, 56–65.

Top Sites in United States. (2019). Retrieved April 23, 2019, from Alexa website: https://www.alexa.com/topsites/countries/US

White, M. J. (1986). Segregation and diversity measures in population distribution. *Population Index*, 198–221.

Williamson, K. D. (2018, April 20). When the Twitter Mob Came for Me. *Wall Street Journal*. Retrieved from https://www.wsj.com/articles/when-the-twitter-mob-came-for-me-1524234850

Zhong, C.-B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science*, *21*(3), 311–314.

Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. *Nebraska Symposium on Motivation*. University of Nebraska press.