The Political Landscape of the U.S. Twitterverse

Subhayan Mukerjee, Ph.D.¹, Kokil Jaidka, Ph.D.¹, Yphtach Lelkes, Ph.D.²

¹National University of Singapore, ² University of Pennsylvania

forthcoming in *Political Communication*

Abstract

Prior research suggests that Twitter users in the United States are more politically engaged and more partisan than the American citizenry, who are generally characterized by low levels of political knowledge and disinterest in political affairs. This study seeks to understand this disconnect by conducting an observational analysis of the most popular accounts on American Twitter. We identify opinion leaders by drawing random samples of ordinary American Twitter users and observing whom they follow. We estimate the ideological leaning and political relevance of these opinion leaders and crowdsource estimates of perceived ideology. We find little evidence that American Twitter is as politicized as it is made out to be, with politics and hard news outlets constituting a small subset of these opinion leaders. Ordinary Americans are significantly more likely to follow non-political opinion leaders on Twitter than political opinion leaders. We find no evidence of polarization among these opinion leaders either. While a few political professional categories are more polarized than others, the overall polarization dissipates when we factor in the rate at which the opinion leaders tweet: a large number of vocal non-partisan opinion leaders drowns out the partisan voices on the platform. Our results suggest that the degree to which Twitter is political has likely been overstated in the past. Our findings have implications about how we use Twitter and social media, in general, to represent public opinion in the United States.

The study of social media platforms has become prominent in political communication for at least three reasons: one, they increasingly mediate our access to news (Matsa & Shearer, 2018; Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2017; Scharkow, Mangold, Stier, & Breuer, 2020); two, they are increasingly crucial as messaging tools for political elites (Gulati & Williams, 2010; Larsson, 2015; Lee & Oh, 2012; McGregor, Mourão, & Molyneux, 2017); and three, they are used by journalists, politicians, academics to gauge public opinion (McGregor, 2019) and even predict elections (Tsakalidis, Papadopoulos, Cristea, & Kompatsiaris, 2015). These factors have led to speculation regarding the extent to which the social media experiences of Americans on Twitter deviate from their quotidian media experiences. As a result, a large body of literature has emerged that seeks to better understand the relationship between social media and the political process, particularly in the context of the United States.

Much of this research is fueled by speculation about how political polarization can reinforce selective exposure dynamics: that despite an increase in the quantity and diversity of political information available online, people may only be exposed to information that is consistent with their existing beliefs (Lazarsfeld, Berelson, & Gaudet, 1968; Lee, Choi, Kim, & Kim, 2014; Sears & Freedman, 1967). This can give rise to echo chambers, which, in turn, can lead to a more polarized, less nuanced, and less deliberative citizenry (Sunstein, 2017).

While these concerns have long troubled social media researchers, they stand in tension with a key feature of the American public: most Americans are not politically engaged. In the US presidential election conducted in 2016, a few years before the data for this study was collected, only 54.8% of the eligible voting population of the United States cast their vote. The United States ranks only 31 out of 35 developed nations in terms of voter turnout. Prior work has also reported on the low levels of political knowledge among the American public (Delli Carpini & Keeter, 1996). These trends would imply that while most Americans use social media and roughly a quarter of the public use Twitter (Matsa & Shearer, 2018), only some may self-select into using

Twitter for politics. Furthermore, among those who are interested in politics, the increasingly polarized American political landscape may have turned some off from discussing politics online (Jurkowitz & Mitchell, 2020). These findings undermine many expectations regarding online polarization made in prior studies and raise questions regarding the prominence of politics in the social media experiences of American citizens.

Scholars investigating polarization on social media, often focus on *politically engaged* audiences and *political* discourses, by using politically relevant keywords (e.g. Conover, Ratkiewicz, & Francisco, 2011), or the behavior of (the followers of) political opinion leaders (e.g. Barberá, 2015) to collect data. Such datasets typically reflect the attitudes and behavior of a specific subset of the larger national population – that of those who actively follow politics – which may be far removed from the attitudes and behaviors of the average online citizen. These studies, therefore, rely on non-representative samples of social media users. However, they implicitly contribute to advancing a general narrative regarding the prevalence of polarization on these digital platforms, which gets further amplified by mainstream media coverage (Bump, 2017; Mims, 2020).

Our study examines these assumptions and offers an understanding of the *politicization* and *polarization* of the American Twitter landscape, based on a representative sample of Twitter users. Twitter is one of the most prominent social media platforms today, with over 330 million monthly active users. It has also grown in importance as a platform used by political elites for communicating and connecting with users (McGregor, 2019; McGregor et al., 2017) (for instance, almost every prominent politician in the US uses Twitter). A Twitter-based observational study can thus help understand its users (a steadily growing chunk of the American citizenry) by allowing us to appraise their relationship with these political opinion leaders vis-a-vis non-political opinion leaders. Notably, it sheds light on whom Twitter users choose to follow, the primacy of politics manifested in these choices, and the resulting polarization of these opinion leaders that potentially reflects the partisan divides in their follower

base and, consequently, American society. It can also address whether politically disengaged people do use social media platforms, such as Twitter, to discuss or at least be exposed to politics (Gil de Zuniga, Weeks, & Ardevol-Abreu, 2017).

The objective of the present study is to contextualize political polarization on Twitter relative to the politicization of the platform. We use rich relational data obtained from Twitter to identify the opinion leaders, both political and non-political, on the platform based on a set of accounts that a random sample of American Twitter users (from the Twitter 1% firehose) follow. These opinion leaders are manually classified based on their professional categories (e.g., entertainment, sports, media outlet, public figure, political figure etc.), following which their ideological slants and political relevance levels are estimated. We further crowdsource their *perceived* partisan slants using an online survey. Our analyses reveal that politics is essentially a sideshow on Twitter. The most followed American opinion leaders on the platform are primarily non-political, as well as largely non-partisan. Opinion leaders across all the professional categories are similarly popular on average. However, an ordinary Twitter user is far more likely to follow non-political opinion leaders than political opinion leaders. In short, ordinary Americans do not use Twitter for politics. Non-political opinion leaders also rarely tweet about politics. There is also little evidence that these opinion leaders are polarized, although the level of polarization varies by professional category. Finally, we find that while ordinary Americans *perceive* opinion leaders to be more liberal than they are according to network-based measures of ideology, the overall distribution of perceived ideologies remains indisputably unimodal as well. In the following two sections, we develop a conceptual argument as to why ordinary Americans choose to follow political versus non-political opinion leaders before introducing the hypotheses we set out to test.

Selective exposure and polarization on social media

Selective exposure theory suggests that people prefer to consume like-minded content (Sears & Freedman, 1967; Zillman & Bryant, 1985), especially in high-choice

information environments that characterize online social media platforms. Traditional application of this theory in the context of the United States has been restricted to the realm of politics, often being used to understand selective exposure to partisan information. A large body of literature has studied the relationship between these dynamics and the creation of online echo chambers and the rise of group polarization (Lee et al., 2014; Sunstein, 2017; Tucker et al., 2018). When Twitter users selectively follow a few partisan stakeholders – both politicians and traditional media alike – it can alter their social media feeds and expose them to more political content, which platform algorithms can further exacerbate. Relatedly, some scholars have described social media to be “hyperpartisan” (Tucker et al., 2018). Selective exposure may not simply affect ordinary users but also journalists. Usher and Ng (2020), for instance, found political journalists on Twitter to be operating within insular political “microbubbles” far removed from the political ground reality of the country.

However, prior research has shown that people use social media for a variety of purposes, including “social interaction, information seeking, pass time, entertainment, relaxation, communicatory utility, convenience utility, expression of opinion, information sharing, and surveillance/knowledge about others” (Whiting & Williams, 2013). Twitter is no exception, with the platform’s use being largely dictated by social and informational motives (Johnson & Yang, 2009; Quan-Haase, Martin, & McCay-Peet, 2015). Therefore, in a high-choice media environment such as Twitter, people can also choose to disengage from politics altogether (Prior, 2009). Decisions to follow opinion leaders on the platform instead of others incur a ‘cost’ on individual attention, which is in short supply given the fragmented media environment. Even though Twitter “recommends” users to follow, the choice to follow remains a conscious choice that the user makes and can be assumed to be a reasonably accurate representation of the pay-off they seek out of using the platform.

To summarize, the choice of whom to follow on Twitter has an attention cost, which Twitter users can choose to incur or not by deciding to engage or disengage with political topics altogether. Given the fragmented media landscape on Twitter and the

proliferation of entertainment content on the platform (Barthel & Shearer, 2015), it is necessary to consider whether selective exposure would indeed be linked to a polarized social media landscape or, instead, to a centrist and mostly apolitical landscape. Grounded in evidence about the political apathy of American citizens, we discuss these possibilities in the following section.

The apolitical American life

Selective exposure may inadvertently polarize social media news feeds along partisan lines. However, the notion of characterizing the entirety of the Twitterverse as a politically charged platform stands in contrast with what we already know about American citizens more generally: that they are usually low on political knowledge (Delli Carpini & Keeter, 1996) and as evidenced by low voter turnout (Desilver, 2020), largely indifferent to politics compared to other developed democracies. Compared to those who get their news through traditional platforms, Americans who turn to social media for election news are less likely to follow the coverage closely. They are also less likely to be aware of current events, as well as less knowledgeable and less engaged with the news (Mitchell, Jurkowitz, Oliphant, & Shearer, 2020). Furthermore, counter ideological polarization among political elites, most Americans are ideologically moderate (Hill & Tausanovitch, 2015).

One reason for any potential disconnect between the general political landscape in the United States, and the political landscape of American social media users, may be because the latter is not a representative population of the United States. American Twitter users skew younger, more educated, and more liberal than the general American public (Wojcik & Hughs, 2019). Other studies have found that “Twitter users who write about politics tend to be male, to live in urban areas, and to have extreme ideological preferences” (Barberá & Rivero, 2015), and that “online discussions feel more hostile, in part, because the behavior of such individuals is more visible than offline” (Bor & Petersen, 2019). This view is reinforced when journalists actively incorporate the use of social media data in their media reporting, often using Twitter sentiment as a barometer

of public opinion, even as it may misrepresent the larger electorate (McGregor, 2019; McGregor & Molyneux, 2020). Moreover, these practices effectively allow Twitter content to set news agendas as well (Molyneux & McGregor, 2021).

However, there is some reason for skepticism towards the claim that Twitter is substantially different from the political landscape at large. While the majority of Americans use social media, 70% of social media users report never posting or rarely posting or sharing about political issues (McClain, 2021). Hence, studies that characterize Twitter by content analyzing only political posts may vastly overestimate polarization on the platform. Previous surveys have reported that rather than all Twitter users having a general political inclination, a small proportion of active Twitter users are responsible for producing most of the political tweets on Twitter (Wojcik & Hughs, 2019). Users who tweet about politics are also more likely to be strong partisans, with higher levels of political interest than the average citizen (Bekafigo & McBride, 2013). The high levels of political interest found in survey data of Twitter users may also be inflated due to the very high level of political engagement of longitudinal survey panelists, in general, (Karp & Lühiste, 2016).¹ Furthermore, a study of individual Twitter accounts conducted by Pew suggested that 23% of all the accounts followed by users are non-public accounts, such as friends and family (Barthel & Shearer, 2015). While only 9% of the accounts followed by them were by general and political news outlets, these accounts produced 23% of the total tweets on their timeline. On the other hand, 35% of the accounts they followed were related to entertainment, sports, and lifestyle news, which produced 32% of all total tweets on their timeline. However, this study was based on an analysis of only 176 Twitter users' handles, which casts doubts on its representativeness for an ordinary American Twitter user.

Hence, while some literature on echo chambers and selective exposure suggest a polarized and hence a politicized Twitterverse, we expect that, given the low levels of political knowledge and interest in politics among Americans, most Twitter users are more interested in non-political information. In other words, we expect, within the set

¹ Respondents to the Pew Survey were, for instance, 2-3 more likely to be registered voters than the population at large.

of opinion leaders, non-political opinion leaders such as celebrities and sportspersons to be more followed than politicians or hard news accounts. This line of reasoning yields our first hypothesis:

H1: Individual American users are more likely to follow non-political opinion leaders than political opinion leaders on Twitter.

In order to test H1, we will first appraise the degree of *politicization* of Twitter – that is, the extent to which politics dominate Twitter in the first place. Doing this will allow us to formally compare the number of non-political opinion leaders an average user follows with the number of political opinion leaders, which in turn will let us test H1.

Measurements of social media polarization

Related to the question of Twitter’s politicization is that of Twitter’s *partisan polarization*. Partisan polarization on Twitter matters because a polarized Twitter has implications for the manner in which ordinary American Twitter users are exposed to partisan political information. This has become increasingly relevant as Twitter has emerged as a prominent medium for official communication among America’s political elites (McGregor et al., 2017). Partisan polarization on such an influential platform can have downstream effects for the political process at large, as it can further increase the partisan divide in the US via echo-chamber effects (Sunstein, 2017) precluding the possibility of “the common ground” needed for nuanced, deliberative, and democratic debate.

Partisan polarization is typically defined as the degree to which Twitter users cluster around ideologically extreme positions. Therefore, the first step towards the appraisal of polarization among users on a platform is the estimation of their ideological positions. Most methodological approaches used for this purpose are based on preferential attachment – that similar people are more likely to be connected than to dissimilar people. Scholars adopting these approaches have operationalized ideological position as a structural phenomenon that can be understood by analyzing users’ connections on a platform. Such approaches have been applied to Twitter (Barberá, Jost,

Nagler, Tucker, & Bonneau, 2015; Halberstam & Knight, 2016; Wong, Tan, Sen, & Chiang, 2016), Facebook (Bakshy, Messing, & Adamic, 2015), and Reddit (Soliman, Hafer, & Lemmerich, 2019), among other social media platforms.

Other approaches have focused on the content being shared by users to identify polarization. Conover et al. (2011), for example, analyzed a large corpus of partisan tweets to understand whether polarization manifested itself in the semantic structure of how people of different political orientations talk on the platform. Similarly, Cinelli, Morales, Galeazzi, Quattrociocchi, and Starnini (2021) did a comparative analysis of 100 million posts on political topics, collected from Facebook, Twitter, and Reddit to understand how partisans interacted with those posts and shared those posts. Perhaps as an artifact of their approach, their findings suggest that online dynamics are typically composed of users in homophilic clusters, with Facebook appearing to be more polarized than Reddit.

Yet other studies have corroborated the importance of social network relationships in identifying individuals' partisan- and issue-focused leanings (Pennacchiotti & Popescu, 2011; Yardi & Boyd, 2010; Zamal, Liu, & Ruths, 2012) across partisan issues (Lai, Bosco, Patti, & Virone, 2015; Yardi & Boyd, 2010), and across contexts, such as in Egypt (Weber, Garimella, & Batayneh, 2013), France (Lai et al., 2015), Venezuela (Morales, Borondo, Losada, & Benito, 2015), and Canada (Gruzd & Roy, 2014). Similar top-down approaches to polarization have assumed ideological concordance between elites (such as media outlet accounts (Tewksbury, 2005)) and their digital audiences through the lens of selective exposure theory. These approaches create and assign audience-based metrics of partisan bias to elite accounts on the relevant platforms (Bakshy et al., 2015; Dilliplane, 2011; Flaxman, Goel, & Rao, 2016; Messing, van Kessel, & Hughes, 2017).

Prior work is invaluable to understand the polarization processes on social media; however, they may not offer a complete picture of the overall political landscape of digital platforms due to sampling biases. For instance Barberá (2015) looked at the polarization of *political elites* on Twitter – a group of Twitter users that are certainly

not representative of the overall Twittersphere. Other studies handpick only those users who talk about politics on social media – again, a patently non-representative section of the online population who are more politically engaged than other users. Our study circumvents this problem by instead focusing on those users who organically emerge as the most popularly followed accounts on American Twitter, based on multiple random samples of ordinary American Twitter users. While we still estimate their ideologies using Barberá’s algorithm, the political landscape we paint with our analysis is not one of the handpicked political elites – but of a sample that is much more representative of American Twitter.

Our hypothesis that non-political actors are more popular than political actors on Twitter leads us to expect that any political polarization in *political circles* on Twitter may be drowned out by the simultaneous presence of these prominent apolitical actors on the platform, as evidenced by the statistical distribution of their ideology. Thus, while it is possible that the political opinion leaders alone are driving the polarization in the overall Twittersphere, it is also possible that their influence matters little in the larger, noisier picture. This leads us to our second hypothesis and an exploratory research question:

H2: The distribution of ideology among opinion leaders is unimodal.

RQ 1: How does the extent of polarization differ across the different professional categories of American opinion leaders on Twitter?

Data and Methods

This study focuses on the “opinion leaders” followed by a majority of a random sample of American Twitter users. By focusing on the people that users follow, we gain insight into their preferences. First, we manually annotated each opinion leader, classifying them into one or more genres based on their professional backgrounds. Next, we applied a lexicon-based algorithm to their tweets to examine how often they tweet about politics. Then, we measured their partisan slant by scraping their Twitter networks and analyzing whom they follow. In order to distinguish between partisanship

and *perceived* partisanship, we further crowdsourced the opinions of Americans on these opinion leaders using an online survey. Finally, we examined the distribution of the partisan slants. The following subsections describe the data collection and analytical procedures in further detail.

Data collection

Our first improvement over existing studies that look at elite polarization on Twitter began with our choice of opinion leaders. We did not decide on the elites or opinion leaders based on who they are or their position in real life. Instead, we began with a sample of tweets from three months (January - March 2019) from the Twitter 1% firehose, which was collected and made available for public use as the County Lexical Bank by Giorgi et al. (2018). We used the geo-location information provided with the tweets and information from user bios to geo-tag 25% of the random sample and identify those posted from the United States. We then sampled 10,000 Twitter accounts from this list of users and discarded any accounts which were no longer available (either suspended, deleted, or changed to a private setting). Finally, we were left with 9959 Twitter accounts who followed a total of 393,919 individual accounts.

Analysis

Identifying American opinion leaders. To protect against potential sampling bias, we sampled ten sub-samples of 1000 users each without replacement from the geo-tagged 1% filtered in the last step. Next, we obtained the *following* networks (i.e., the list of accounts they follow) of the users in each sub-sample. Then, we identified the most followed one thousand accounts in each sub-sample. Subsequently, we pooled the most followed Twitter accounts across all the sub-samples into one list. We denoted this set as the set of ‘American opinion leaders’ that we used for subsequent analyses. Thus, if any Twitter account was among the top thousand most-followed accounts in any of the ten sub-samples of American Twitter users, we deemed it an opinion leader. The final number of elites that we ended up with was 1,822. After removing accounts that no longer existed or were suspended, we were left with 1,761

opinion leaders. Any tweet posted, retweeted, or liked by this group of users gains massive exposure through a readership totaling hundreds of millions. The popularity of these opinion leaders were also highly correlated across the ten sub-samples (pairwise Spearman’s ρ between two samples ranged from 0.72 to 0.78; all p-values ≈ 0) showing that the sub-samples were similar in which opinion leaders they followed.

Opinion leaders were categorized into groups or ‘genres’ based on their professional backgrounds in the next step. The opinion leaders comprised Twitter handles representing both organizations and individuals. Some of the non-individual opinion leaders were “media outlets” such as CNN or Fox News, while others were more specifically “hard news” programs such as NBC Nightly News and Anderson Cooper 360°. Other non-media-related non-individual opinion leaders included “brands” such as Twitter, Nike, and Wendys, and “organizations” such as NASA and ACLU. Among the individual opinion leaders, we observed “political figures” such as Hillary Clinton and Donald Trump, who hold/have held public offices, and “political pundits” such as Nate Silver and Ben Shapiro, who provide political analyses and forecasts. The list of opinion leaders also comprised many celebrities from the fields of “entertainment” and “sports.” The “entertainment” genre also included writers such as J. K. Rowling and YouTubers such as Jenna Marbles. All other opinion leaders who did not fit into any of these above categories were classified as “public figures.” This genre included lawyers such as Preet Bharara, CEOs of tech companies such as Elon Musk, and religious figures like the Dalai Lama. Based on these general observations, all three authors classified each opinion leader individually into the genres mentioned above. Any disagreements in the classification were related to individuals who could fit into more than one category. Therefore, some opinion leaders were assigned to multiple genres. For instance, John Oliver was classified as both “entertainment” and “political pundit”. @FLOTUS (the account then representing Melania Trump), similarly, was classified as both a political figure owing to her tenure as former First Lady and a public figure owing to her prominence outside politics.

Inferring Ideologies. The identification of opinion leaders was followed by the use of the ideal-point estimation algorithm (Barberá, 2015) to infer their ideologies. This method works under the assumption that “[the] decision to follow [on Twitter] is considered a costly signal that provides information about Twitter users’ perceptions of both their ideological location and that of political accounts” (p. 77). This approach assumes two types of “cost.” The first cost is the cognitive dissonance that a Twitter user can incur by following politicians with whom they disagree. The second cost is an opportunity cost that the user can incur by *not* following politicians, as it reduces their likelihood of being exposed to their messages. In other words, decisions to follow or not follow political elites “provide information about how social media users decide to allocate a scarce resource – their attention.” (p. 78) The ideal point estimation method then works by first identifying Twitter handles of well-known liberal and conservative actors at two ends of the latent ideological spectrum and then interpolating the latent ideal points of a random user based on which of these known actors they follow. Estimates of user ideology using this method are strongly correlated with other estimates of ideology, e.g., DW-Nominate (Barberá, 2015) and survey measures (Eady, Nagler, Guess, Zilinsky, & Tucker, 2019).

Note that an ‘apolitical’ ideology is not the same as a ‘neutral’ ideology, as the latter implies that an individual may still express a centrist perspective and follow political accounts. Ideological estimation can capture neutral ideology as a small score left or right of 0. It is also possible that ideology estimates based on follow networks and calculated using the ideal-point estimation algorithm do not reflect how ordinary Americans perceive the ideological stance of an opinion leader. Another challenge was that the method relies on a user following at least one political elite from a small set ($N = 599$) selected as anchors. For 568 out of 1761 opinion leaders in our sample, this is why ideological estimation returned an ‘NA,’ i.e., an invalid ideological score. We made a working assumption that a lack of any valid ideological score indicated neutrality, although we show results with and without this assumption.²

² We validate this assumption in the Appendix using crowdsourced ideological scores for those

Validating ideology scores through crowdsourcing:

We compared the ideological estimates based on follower lists against the perceived ideology based on crowdsourced human judgments. We set up a pre-registered annotation task on Amazon Mechanical Turk to collect multiple perceived ideological scores for each opinion leader. Only MTurkers, based out of the United States, who had a track record of completing at least 5000 approved tasks, with an approval rate of at least 98% were eligible to participate in our task. The task showed MTurkers the name and Twitter handle of an opinion leader (hyperlinked so they could navigate to their Twitter profile if they wanted to) and asked them to rate them on a 7-point Likert scale where lower values indicated liberal while higher values indicated conservative. As an attention check (that was pre-registered before the data were collected), we required every MTurker to rate two specific opinion leaders from our set, Donald Trump (@realDonaldTrump) and Bernie Sanders (@BernieSanders)³. Any MTurker who rated Donald Trump to be more liberal than or equal to Bernie Sanders was considered to have failed the attention check, and their responses were consequently removed from the analysis. The MTurkers also had the option to answer “Don’t know/Can’t say” if they could not rate an opinion leader. In the end, each opinion leader was rated by 13.98 MTurkers on average ($sd = 2.68$). Additionally, we also asked the MTurkers whether they had a Twitter account or not. While we used the responses of all MTurkers who passed the attention check, dropping those who claimed not to have a Twitter account had no qualitative effect on our subsequent findings.

As an additional step, we estimated the political relevance of each opinion leader by calculating the number of a predetermined set of political keywords they included in their tweets as a fraction of all the words they tweeted. This list of keywords was curated and validated by political scientists in a study of tweets by Preotiuc-Pietro, Liu, Hopkins, and Ungar (2017).

To test H1, we conducted pairwise (i.e. between two categories) paired

“neutral” opinion leaders.

³ Donald Trump has since been banned from the platform.

non-parametric tests of differences in medians of the number of opinion leaders followed by the ordinary individuals in our samples.

Estimating polarization using ideology scores weighted by tweet frequency: H2 required all the opinion leaders on Twitter to be characterized according to their political ideology. However, because we sought to understand Twitter’s information environment and its user base, we created a measure of polarization that accounts for each user’s contribution to the information environment by weighting a user’s ideal point by their tweeting frequency (between January and March 2019). Weighting ensured that a highly partisan account (with an extreme ideal point or perceived ideological score) that did not tweet at all during the data collection period would not contribute to the polarization on the platform. In this manner, we obtained the ideological distribution of the Twittersphere both before and after weighting. We also included a subgroup analysis of the unweighted ideological distribution within each decile of tweeting frequency for added nuance. This analysis gave a clearer picture of the relationship between polarization and tweeting frequency, allowing us to distinguish between the polarization patterns of opinion leaders who are more vocal on Twitter and those who are not, a picture that a simple weighted distribution potentially obfuscates. We conducted statistical tests to appraise the deviation of each of these distributions from unimodality (that is, the lack of polarization) to formally test H2 and answer RQ 1.

Results

Our strategy for the identification of opinion leaders on American Twitter yielded a set of 1,822 Twitter accounts. After inferring their ideal points and crowdsourcing their perceived ideological slant, we were left with 1,761 accounts, as the remaining accounts were either private or had been suspended or deactivated. When reporting the ideological distributions, we drop the two accounts (@realDonaldTrump and @BernieSanders) that we used as attention checks in our crowdsourcing tasks and show the distributions of the remaining 1,759 accounts. This is because these two

accounts were rated by every Mechanical Turk worker, and their mean response would not be directly comparable to the mean response for all the other accounts that were rated by an average of 14 workers.⁴

The twenty most popular opinion leaders, as followed by our sample of American Twitter users, are listed in Table 1. This list includes popular politicians, mainstream media outlets, and celebrities (such as sportspersons and Hollywood celebrities). Crucially, only six of the top twenty accounts belonged to one of the political categories (i.e., a political pundit or a political figure), and only two of the twenty were hard news outlets. Four of the political accounts were alternate accounts for Donald Trump and Barack Obama.

Panel A in Figure 1 shows the reach of each genre (measured by the number of unique Twitter users in our dataset who follow at least one opinion leader in that genre) and the tweeting activity of the genre (total number of tweets sent out by all opinion leaders in that genre between January and March 2019). The size of the bubble indicates the number of opinion leaders. Political figures and political pundits comprise less than 10% of the opinion leaders. The percentage rises to less than 20% when combined with hard news and media outlets. Entertainment (44.9%) is by far the most popular category, with sports personalities (14.7%) and public figures (9.74%) following far behind. As panel B in Figure 1 shows, the genres that we decided during the manual classification to be non-political were indeed less politically relevant than the genres we had decided to be political (political pundits and political figures) or hard news outlets.

The results illustrate the differences between the popularity of the different kinds of opinion leaders. However, they do not speak to the Twitter experiences of ordinary individuals. We now turn to this and formally test H1. For each ordinary user in our sample who follows at least one account ($N = 9378$), we make non-parametric statistical comparisons of the distribution of opinion leaders (by category) in their following lists. The results are shown in Figure 2 (panel A). We find that the median Twitter user in our sample followed 10 entertainment accounts, 1 brand account, 1

⁴ Including these two accounts in the ideal points distribution did not substantively change our results.

Table 1

The list of the 20 most popular opinion leaders, that are most followed by our sample of ordinary Americans

Rank	Opinion Leader Twitter Handle	Opinion Leader Name	Political Relevance [†]
1	BarackObama	Barack Obama	0.062
2	RealDonaldTrump	Donald Trump	0.039
3	TheEllenShow	Ellen DeGeneres	0.0026
4	Drake	Drake	0.0069
5	rihanna	Rihanna	0.0045
6	kanyewest	Kanye West	0.0047
7	SportsCenter	SportsCenter	0.0010
8	jimmyfallon	Jimmy Fallon	0.0022
9	POTUS44	President Obama	0.022
10	KingJames	LeBron James	0
11	POTUS	President Trump	NA*
12	cnnbrk	CNN Breaking News	0.060
13	nytimes	The New York Times	0.015
14	espn	ESPN	0.0012
15	ArianaGrande	Ariana Grande	0.0014
16	HillaryClinton	Hillary Clinton	0.050
17	elonmusk	Elon Musk	0.00037
18	AOC	Alexandria Ocasio-Cortez	0.020
19	JColeNC	J. Cole	0.0016
20	KimKardashian	Kim Kardashian	NA**

[†]This score is calculated following prior literature (Preotiuc-Pietro et al., 2017) by identifying the number political keywords the user tweets as a fraction of the total number of words they tweeted.

*The score for POTUS was not estimated because the Twitter handle had changed to represent Joe Biden’s administration by the time the calculations were done.

** The score for Kim Kardashian was not estimated as her tweets were not available through the Twitter API at the time of data collection. We expect that this may have been because her account was temporarily deactivated.

sports account, 1 political figure, 1 public figure, 0 political pundits, 0 hard news, 0 meme accounts and 0 organization accounts. Pairwise paired one-sided Wilcoxon’s signed-rank tests with Holm’s p-value correction show that of all possible 45 combinations of genres, 44 differences in medians are statistically significant at the one-sided $p < 0.001$ level. The only two genres that are not significantly different are “organization” and “meme”. The table showing all the test results is provided in the Appendix (Table A1). For comparison, the relative distribution of opinion leaders and non-opinion leaders in the users’ following lists is depicted in panel B. The difference in

Figure 1. Panel A shows a scatterplot illustrating the relative prominence of the different types of opinion leaders in our dataset. The x-axis shows the total number of unique followers of all the opinion leaders in a given type. The y-axis indicates the total number of tweets posted by all the opinion leaders in a given type between January - March 2019. The size of the circles indicates the number of opinion leaders in each genre, while the color indicates their political relevance as captured by the percentage proportion of their entire vocabulary that was relevant to politics. Panel B shows the distribution of political relevance scores by genre.

the median, in this case, is also significant. (paired samples Wilcoxon’s signed-rank test; $V = 37302544$, $p \approx 0$.) The low prominence of politics is also manifested in the fact that of the 8,649 users who followed at least one opinion leader, 27.8% of them did not follow a single political opinion leader (including media outlets and hard news). Only 6,246 ordinary users followed at least one political and one non-political leader. Among them, the average ratio of non-political opinion leaders followed to political opinion leaders followed was 8.4 (median = 4.5). H1 is therefore, supported, by these analyses.

These results already suggest that both, at the macro level, as well at the micro level, politics is but a sideshow in the Twitter landscape. Collectively, the political opinion leaders are overshadowed in reach and popularity by the non-political opinion leaders. At the individual level, we find that users tend to follow accounts that are not political, in line, with our expectations, given that the American citizenry is primarily not interested in politics.

Figure 2. Panel A shows a boxplot illustrating the relative prominence of the different categories (genres) of accounts as followed by individual Twitter users in our sample.

All *paired* pairwise differences in medians except that between “organization” and “meme” are significantly different (one-sided $p < 0.001$). Panel B shows that ordinary users are more likely to follow accounts that are not in our set of opinion leaders than they are to follow accounts that are.

Analysis of Ideologies of Opinion Leaders

The ideal point estimation algorithm (Barberá, 2015) returns a value between -3 and $+3$, depending on the balance of predetermined liberal and conservative accounts that the opinion leader follows. A negative value implies a liberal bias, while a positive value implies a conservative bias. If the user does not follow any of the predetermined liberal and conservative actors, the algorithm cannot calculate an ideal point. For our analyses, we impute the ideal points of such users to be 0 – implying a lack of partisan bias in any specific direction. We validate this assumption with our crowdsourced scores of perceived ideology in the Appendix.

If this set of opinion leaders – who constitute the most popular Twitter accounts in the US and who subsequently post content that is viewed the most on the platform –

is polarized, the distribution of their ideal points would, by definition, be bi-modal. An inspection of Figure 3, however, shows that this is not the case. The distribution of ideal points is multimodal with four peaks. Three are local maxima, implying a high density of opinion leaders around three points on the latent ideological spectrum. However, the global maximum at the mid-point of the spectrum was composed of opinion leaders whose ideal points could not be inferred (and thus imputed to be 0) towers over these local peaks. This indicates an overall lack of bimodality in the distribution. If opinion leaders with an imputed “neutral” are removed, the distribution is more polarized.⁵

The distribution of the mean perceived ideological slants tells a similar story. In order to make the perceived ideological slant scores (ranging from 1 [most liberal] to 7 [most conservative]) directly comparable to the inferred ideal points (ranging from -3 [most liberal] to 3 [most conservative]), we subtracted 4 from the former, before graphing their distribution. The lack of any evidence of bi-modality indicates the lack of any perceived polarization among these opinion leaders. However, the distribution is shifted to the left, which implies that, in general, our annotators perceived the opinion leaders to be more liberal than what their ideal points would indicate. A non-parametric paired Wilcoxon signed-rank sum test between the perceived ideologies and their corresponding ideal points reveals that the former is significantly lesser than the latter ($V = 429162$; $p \approx 0$).

In order to test H2, we inspected the distribution of ideologies weighted by frequency (See Figure 4). As is evident, weighting further attenuates polarization. This implies that even while partisan opinion leaders on Twitter exist, their content is more likely to come from less partisan and centrist voices. However, the left-bias in perceived slant (compared to the ideal points) persists even after weighting them both; a non-parametric paired Wilcoxon signed-rank sum test between the *weighted* perceived ideologies and their corresponding *weighted* ideal points reveals that the former is significantly lesser than the latter ($V = 397483$; $p \approx 0$).

⁵ We also estimated the ideologies of the ordinary American users as manifested in their following decisions using the ideal point algorithm. In line with our expectations, most ordinary users do not follow enough political elites to estimate an ideal point. If we impute these respondents’ ideal point as 0, the distribution of ideology on Twitter is decidedly unimodal (see Figure A2 in the Appendix.)

Figure 3. Comparison of the distribution of ideal points with the distribution of perceived ideologies with higher numbers indicating more conservative. While the latter is multimodal, with a global peak at 0, neither distribution shows bimodality. There is a high density of opinion leaders in the middle of the spectrum. However, the perceived ideologies bear a distinctive liberal slant, with non-parametric tests confirming that they are significantly lower than the corresponding ideal points

Next, we tested these observations by conducting Hartigan’s dip tests (Hartigan & Hartigan, 1985) on each of the distributions. Hartigan’s dip test uses a simulation-based method to test the null hypothesis that a given set of values is unimodally distributed. Greater multimodality of the distribution is associated with a higher Hartigan’s Dip Statistic (HDS) produced by the test, and the associated p-value is used to test its deviation from the unimodal null model. A significant p-value rejects the null hypothesis that the distribution is unimodal. The test results indicated that while the unweighted ideologies exhibit significant (in terms of p-value) deviation from a unimodal distribution (ideal points: $HDS = 0.027$; $p \approx 0$; perceived ideologies: $HDS = 0.022$; $p \approx 0$), the weighted ideologies are indisputably unimodal (ideal points:

$HDS = 0.0036; p = 1$; perceived ideologies: $HDS = 0.0047; p = 1$). This implies that after factoring in the tweeting frequency of the opinion leaders, the platform clearly does not appear to be polarized. In the robustness section in the Appendix, we redo these analyses after removing the opinion leaders whose ideal points could not be inferred and had been assumed to neutral. We find that upon weighting, these partisan opinion leaders do not exhibit any polarization neither (Figure A1).

Figure 4. Comparison of the distribution of weighted ideal points with the distribution of weighted perceived ideologies with higher numbers indicating more conservative. Weighting was done by multiplying the ideology scores by the scaled tweeting frequency of the corresponding opinion leader. Both distributions become patently unimodal after weighting, showing that partisan opinion leaders do not tweet much - and therefore do not contribute to the polarization of the content on the platform. However, despite weighting, pairwise non-parametric tests indicate that the perceived ideologies are significantly lower than the corresponding ideal points.

To further unpack the relationship between polarization and tweeting activity, we plot the distribution of ideal points for opinion leaders within each decile of tweeting activity (figure 5). To do this, we divide the set of opinion leaders into ten roughly

equal-sized groups after sorting them by tweeting activity. The first decile consists of opinion leaders who tweeted the least, while the tenth decile comprises opinion leaders who tweeted the most. By plotting and comparing the (now unweighted, since we are already controlling for tweeting frequency by dividing into deciles) ideal point distributions, we find that the density of liberals roughly remains constant at all levels of tweeting activity, the density of conservatives gradually increases. In the tenth (or the most active) decile, conservatives dominate the distribution, rising above even the neutrals.

Interestingly, the perceived ideological scores do not tell the same story. Across all the deciles, their distribution within each decile remains unimodal and barely shifts. However, the perceived left bias persists across all the deciles, with our annotators agreeing that opinion leaders are more liberal than their ideal points suggest at levels of tweeting activity. Table A2 in the Appendix lists the V-statistic and p-value for each decile’s corresponding Wilcoxon signed-rank test.

Hartigan’s Dip test on each of these distributions show that while the perceived ideologies in each decile are unimodal, the ideal points in some of the deciles, tend to show some evidence of multimodality. The test results are provided in the Appendix in Table A3.

To answer RQ1, our next set of findings reveals the level of polarization within various professional categories.⁶ We report the weighted ideal points and perceived ideological slant scores in the main text. The corresponding unweighted versions are provided in Figure A3 in the Robustness section in the Appendix. Here again, we find the lack of any stark polarization in the distribution of opinion leaders within each category (Figure 6). The ideal points indicate some level of polarization within the category of “media outlets”, with a local peak at the conservative end of the spectrum and a fat tail for the “hard news” and “political pundit” categories. This implies that there are more media outlets and political pundit Twitter accounts on the conservative

⁶ In our classification of opinion leaders, each opinion leader may belong to more than one professional category, as described in the Data section. In the Robustness section in the Appendix, we show that the results presented here are very similar, even if we only use the primary category of each opinion leader (Figure A4).

Figure 5. Comparison of the distribution of *unweighted* ideal points with the distribution of *unweighted* perceived ideologies within each decile of tweeting frequency (higher numbers indicating more conservative). Decile 1 comprises the opinion leaders who tweeted the least, while decile 10 comprises those who tweeted the most. While the distributions of ideal points indicate there are more conservative opinion leaders in the higher deciles (as evidenced by progressively thicker tails on the right), the distributions of perceived ideologies, however, do not change qualitatively. The significant left-bias in the perceived ideologies (compared to the ideal points) is also preserved in every decile.

end of the spectrum than on the liberal end. Interestingly, the distribution of the *perceived* ideological slant of media outlets and political pundits does not corroborate the distribution of the corresponding ideal points. Non-parametric paired Wilcoxon signed-rank tests between the *weighted* perceived ideologies and their corresponding *weighted* ideal points reveal that the former is indeed significantly lesser (at the $p < 0.05$ level) than the latter within most categories. Only in one category (“memes”), this is not the case. Table A4 in the Appendix lists the V-statistic and p-value for the corresponding test for each category.

We ran Hartigan’s dip test on the *weighted* ideal points as well as perceived

Figure 6. Comparison of the distribution of *weighted* ideal points with the distribution of *weighted* perceived ideologies, within each genre of opinion leaders. The distribution of ideal points of “media outlet” accounts show some evidence of polarization, and a fat tail in the conservative end of the spectrum. The latter is also visible within the “hard news” and “political pundit” category. Moreover, in all except one categories (“memes”) are the perceived ideologies significantly ($p < 0.05$) lesser than the corresponding ideal points.

ideologies within each professional category and found none of the distributions to be not significantly deviated from the unimodal null model. The results are shown in Table A5 in the Appendix.

Discussion

Commentators regularly caution that the politicking, vitriol, and partisanship they see on Twitter “is not real life.” However, our results suggest that Twitter, in many ways, does reflect real life in America. Just as most Americans are not interested in politics outside of Twitter, neither are most people on Twitter. Similarly, just as the average American is fairly centrist (Fiorina, Abrams, & Pope, 2006), so are their

Twitter feeds. More importantly, these findings force us to question some of the fundamental assumptions about social media – particularly the supposed pre-eminence of politics – that many researchers take for granted. It also could imply that researchers and journalists assume that their timelines, which might, in fact, be polarized are the same as that of the modal twitter user.

Through the four main analytical choices made in this paper, we offer a methodological contribution to the study of online polarization. First, a bottom-up approach using the follow networks of American Twitter users helps to identify Twitter opinion leaders based on prominence and *then* political influence, rather than the other way around. Second, considering the variation in their tweeting activity accounts for the dynamic variation in the polarized messages on the Twitter timeline at any given moment. Third, subgroup analyses help to contextualize the normative expectations of polarization within smaller factions of influence. Finally, triangulating network-based ideology scores with human judgments offers a way to validate the results.

Our analyses suggest that the hyper-partisan and hyper-political Twitter feed that scholars and pundits use to generalize social media is not representative of the modal user’s experience. Instead, Twitter is essentially a platform of centrists and the politically disinterested. This set of findings has important implications for understanding the relationship between digital media platforms and the political process in the United States.

First, we find evidence that undermines the supposed primacy of politics on Twitter. Political and hard-news opinion leaders comprise only a small subset of the most popular Twitter accounts in the US. Instead, we find that there is far more to Twitter than just politics, at least for the average American user, and it is the presence of a large number of non-political actors on the platform that considerably diminishes the level of political polarization by drowning out the political voices. This indicates that most Americans use Twitter for entertainment rather than politics or news in so far as the following choices reveal. While we do not make any causal claims regarding the motivations behind Twitter use, it is possible that the lack of interest in politics,

coupled with the simultaneous presence of non-political information sources, translates to ordinary Americans being more likely to follow non-political opinion leaders than political ones. In other words, the selective exposure patterns that dominate Twitter are in line with what some scholars have called the “real divide in America”: that between “political junkies and everyone else” (Krupnikov & Ryan, 2020, 2022) as opposed to one that is partisan. Nearly 30% of our sample of ordinary Americans did not follow a single political opinion leader (including hard news and media outlets), and those who did, followed far more non-political opinion leaders in comparison. This underscores the need for a theoretical pivot to focus more on selective exposure to political vis-a-vis non-political content as opposed to selective exposure to left vis-a-vis right leaning content in order to accurately characterize the social media habits of ordinary Americans (Prior, 2007).

Although the apolitical may virtually constitute the ‘periphery’ on political topics, their association with other users with political leanings may be critical in understanding how information is disseminated and circulated, based on studies of social media protests (Barberá, Wang, et al., 2015). Politically disinterested voters are usually the swing voters (Fowler et al., 2021) and play a critical role in close elections. Therefore, it is especially crucial to understand how they receive and process political information in their social environment. Furthermore, as per studies on Facebook, apolitical users with weak ties to other users who are polarized, are more likely to also be polarized (Settle, 2018, pp. 244-245).

Findings from our crowdsourcing task also corroborate this narrative: a large fraction of the annotators are unable to assess the political slant of these opinion leaders, implying that either these opinion leaders are not perceived to be political, or that our annotators do not know who they are. This result questions certain assumptions that prior researchers have made by focusing solely on political actors or political conversations. In assuming the politicization of Twitter, they have perhaps overstated the level of political polarization on the platform.

Second, our study makes the distinction between the ideological distribution of

opinion leaders (which has traditionally been used to characterize polarization on online platforms) and the distribution of actual information production on Twitter. An information environment like Twitter is political or polarized only to the extent to which the information on it is political or polarized. We introduce a weighting method that accounts for the fact that many partisan users who are rarely active would not make Twitter a polarized platform. By weighting the estimated ideologies of the Twitter opinion leaders by their tweeting activity, we report a unimodal distribution, implying that possible polarization in the information produced on Twitter is less than what the distribution of partisan opinion leaders would suggest. The lack of polarization also has theoretical implications for how we understand echo-chamber effects, and parallels recent scholarship that finds little evidence that they at all exist online (Dubois & Blank, 2018; Eady et al., 2019; Guess, Lyons, Nyhan, & Reifler, 2018).

Finally, we find differences between the distribution of ideal points and the perceived political slant of the opinion leaders. While both distributions agree on the lack of polarization, the perceived slant of the opinion leaders is significantly more liberal than what their ideal points suggest. In the absence of self-reported ideology from these individuals, the difference in measured versus perceived ideology could indicate either of two measurement issues. The first is that an individual's follow network is not as accurate an approximation of their ideologies as the ideal point estimation would suggest. Alternatively, it could also reflect annotator bias. Most MTurk workers are from liberal, urban areas (Hitlin, 2016). They may have thus perceived opinion leaders to be more like them, i.e., more liberal than the ideological estimates based on their follow networks suggest. The (lack of) perceived polarization across all professional genres, including those that are non-political, also raises interesting questions regarding what polarization could mean when applied to non-political contexts. Since traditional measures of polarization pre-suppose an ideological stance and thereby ignore those whose stances cannot be inferred or do not exist, more theoretical research is needed to understand how one can grapple with polarization beyond the left-right spectrum.

While we use validated and standard methods to estimate ideology, these methods are not without their limitations. First, we assumed that American Twitter users posting geo-coded tweets are no different from those who do not post geo-coded tweets. Second, we assume a single left-right ideological spectrum that precludes any opportunity to compare with and generalize to more pluralistic political contexts. Future researchers can seek to adopt more nuanced strategies to unpack *within-party* differences, including in the context of the United States. For instance, the unweighted distribution of our opinion leaders does reveal two peaks on the liberal end of the spectrum, potentially indicating a cleavage between the progressives and the liberals in the Democratic party. Another avenue for future research lies in quantifying longitudinal trends in politicization and polarization – something that our cross-sectional design is unable to illuminate. Finally, for a complete picture of the landscape of social media, future studies should compare elites with non-elites on Twitter and other platforms and countries.

Why do our findings conflict with the traditional narrative of polarization on social media? Social media news feeds “fosters people’s recognition of social and identity differences that align with political views.” (Settle, 2018, p.18). This is perhaps why a study of Facebook users found that people are able to infer political inclinations of seemingly apolitical content, such as posts about Chick-Fil-A and hybrid cars, with reasonable accuracy (Settle, 2018). Settle argues that social media affordances prime personal and social identity characteristics, so that even apolitical content is interpreted to be political by the social media audience.

Another reason for the dissonance may be that journalists and politicians often use social media to gauge public sentiment (McGregor, 2019, 2020). However, they may be responding to a distorted perception of public sentiment if the information they see on their feeds is more partisan and political than the modal piece of information on Twitter. It would, therefore, potentially bode well for the democratic process if media practitioners and public office holders step outside their *political* (and not just their ideological) echo-chambers and take a step back in assessing the primacy of politics in

ordinary Americans' lives in the first place. For instance, the mainstream media is increasingly covering the social media posts of political elites. While these political elites have large numbers of followers, their tweets and posts are still only organically visible to a small subset of the users on the platform. Their impact, then, is potentially overstated when they are used by media outlets to set the public agenda. It is perhaps imperative that media coverage of political elites is based more on potentially substantive issues that affect more people, such as their policies or their stances on issues, as opposed to their social media activity.

References

- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132. Retrieved from <http://www.sciencemag.org/cgi/doi/10.1126/science.aaa1160> doi: 10.1126/science.aaa1160
- Barberá, P. (2015). Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis*, 23(1), 76–91. Retrieved from <http://pan.oxfordjournals.org/cgi/doi/10.1093/pan/mpu011> doi: 10.1093/pan/mpu011
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10), 1531–1542. doi: 10.1177/0956797615594620
- Barberá, P., & Rivero, G. (2015). Understanding the Political Representativeness of Twitter Users. *Social Science Computer Review*, 33(6). doi: 10.1177/0894439314558836
- Barberá, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J., & González-Bailón, S. (2015). The critical periphery in the growth of social protests. *PloS one*, 10(11), e0143611.
- Barthel, M., & Shearer, E. (2015). How do americans use twitter for news. *Pew Research Center*. Retrieved from <https://www.pewresearch.org/fact-tank/2015/08/19/how-do-americans-use-twitter-for-news/>
- Bekafigo, M. A., & McBride, A. (2013, October). Who Tweets About Politics?: Political Participation of Twitter Users During the 2011Gubernatorial Elections. *Social Science Computer Review*, 31(5), 625–643. (Publisher: SAGE Publications Inc) doi: 10.1177/0894439313490405
- Bor, A., & Petersen, M. B. (2019, December). The Psychology of Online Political Hostility: A Comprehensive, Cross-National Test of the Mismatch Hypothesis. *American Political Science Review*, forthcoming. Retrieved 2021-08-16, from <https://psyarxiv.com/hwb83/> (type: article) doi: 10.31234/osf.io/hwb83
- Bump, P. (2017). Analysis | How politicians’ use of social media is reinforcing a partisan media divide. *Washington Post*. Retrieved 2021-04-09, from <https://www.washingtonpost.com/news/politics/wp/2017/12/18/how-politicians-use-of-social-media-is-reinforcing-a-partisan-media-divide/>
- Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*,

118(9).

- Conover, M., Ratkiewicz, J., & Francisco, M. (2011). Political polarization on twitter. In *ICWSM 2011 International Conference on Weblogs and Social Media*. doi: 10.1021/ja202932e
- Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Retrieved from <https://www.jstor.org/stable/j.ctt1cc2kv1>
- Desilver, D. (2020). *In past elections, U.S. trailed most developed countries in voter turnout* (Tech. Rep.). Pew Research Center. Retrieved from <https://www.pewresearch.org/fact-tank/2020/11/03/in-past-elections-u-s-trailed-most-developed-countries-in-voter-turnout/>
- Dilliplane, S. (2011). All the news you want to hear: The impact of partisan news exposure on political participation. *Public Opinion Quarterly*, 75(2), 287–316. doi: 10.1093/poq/nfr006
- Dubois, E., & Blank, G. (2018, May). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society*, 21(5). Retrieved 2021-12-22, from <https://doi.org/10.1080/1369118X.2018.1428656> doi: 10.1080/1369118X.2018.1428656
- Eady, G., Nagler, J., Guess, A., Zilinsky, J., & Tucker, J. A. (2019). How many people live in political bubbles on social media? evidence from linked survey and twitter data. *Sage Open*, 9(1), 2158244019832705.
- Fiorina, M. P., Abrams, S. J., & Pope, J. C. (2006). *Culture War? The Myth of a Polarized America*. doi: 10.2139/ssrn.586463
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(Special Issue), 298–320. doi: 10.1093/poq/nfw006
- Fowler, A., Hill, S., Lewis, J., Tausanovitch, C., Vavreck, L., & Warshaw, C. (2021). Moderates. *Unpublished manuscript*, August, 3, 2021.
- Gil de Zuniga, H., Weeks, B., & Ardevol-Abreu, A. (2017). Effects of the news-finds-me perception in communication: Social media use implications for news seeking and learning about politics. *Journal of Computer-Mediated Communication*, 22(3), 105–123. doi: 10.1111/jcc4.12185
- Giorgi, S., Preotiuc-Pietro, D., Buffone, A., Rieman, D., Ungar, L. H., & Schwartz, H. A. (2018). The remarkable benefit of user-level aggregation for lexical-based population-level predictions. In *Proceedings of the 2018 conference on empirical methods*

in natural language processing.

- Gruzd, A., & Roy, J. (2014). Investigating Political Polarization on Twitter: A Canadian Perspective. *Policy & Internet*, 6(1), 28–45. Retrieved from <http://doi.wiley.com/10.1002/1944-2866.POI354> doi: 10.1002/1944-2866.POI354
- Guess, A., Lyons, B., Nyhan, B., & Reifler, J. (2018). *Avoiding the Echo Chamber about Echo Chambers: Why selective exposure to like-minded political news is less prevalent than you think*. Retrieved from https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/133/original/Topos_KF_White-Paper_Nyhan_V1.pdf
- Gulati, G. J., & Williams, C. B. (2010). Congressional candidates' use of youtube in 2008: Its frequency and rationale. *Journal of Information Technology & Politics*, 7(2-3), 93–109.
- Halberstam, Y., & Knight, B. (2016). Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of Public Economics*, 143, 73–88. Retrieved from <http://dx.doi.org/10.1016/j.jpubeco.2016.08.011> doi: 10.1016/j.jpubeco.2016.08.011
- Hartigan, J. A., & Hartigan, P. M. (1985). The Dip Test of Unimodality. *The Annals of Statistics*, 13(1), 70–84. Retrieved 2021-08-11, from <https://www.jstor.org/stable/2241144> (Publisher: Institute of Mathematical Statistics)
- Hill, S. J., & Tausanovitch, C. (2015). A disconnect in representation? comparison of trends in congressional and public polarization. *The Journal of Politics*, 77(4), 1058–1075.
- Hitlin, P. (2016). 4. turkers in this canvassing: Young, well-educated and frequent users. *Pew Research Center*, 437.
- Johnson, P. R., & Yang, S.-U. (2009). Uses and Gratifications of Twitter. In (Vol. 54, p. 33).
- Jurkowitz, M., & Mitchell, A. (2020). *Almost half of Americans have stopped talking politics with someone* (Tech. Rep.). Pew Research Center. Retrieved from <https://www.pewresearch.org/journalism/2020/02/05/a-sore-subject-almost-half-of-americans-have-stopped-talking-politics-with-someone/>
- Karp, J. A., & Lühiste, M. (2016). Explaining political engagement with online panels: Comparing the british and american election studies. *Public Opinion Quarterly*, 80(3), 666–693.
- Krupnikov, Y., & Ryan, J. B. (2020, October). Opinion | The Real Divide in America Is Between Political Junkies and Everyone Else. *The New York Times*. Retrieved 2021-08-23, from <https://www.nytimes.com/2020/10/20/opinion/>

polarization-politics-americans.html

- Krupnikov, Y., & Ryan, J. B. (2022). *The Other Divide*. Cambridge: Cambridge University Press. Retrieved from <https://www.cambridge.org/core/books/other-divide/BAEDF5146C86E57F26696EF7AAB800FE> (DOI:)
- Lai, M., Bosco, C., Patti, V., & Virone, D. (2015). Debate on political reforms in Twitter: A hashtag-driven analysis of political polarization. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 1–9). IEEE. Retrieved from <http://ieeexplore.ieee.org/document/7344884/> doi: 10.1109/DSAA.2015.7344884
- Larsson, A. O. (2015). The eu parliament on twitter—assessing the permanent online practices of parliamentarians. *Journal of Information Technology & Politics*, 12(2), 149–166.
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1968). *The people's choice; how the voter makes up his mind in a presidential campaign*. Columbia University Press.
- Lee, Choi, J., Kim, C., & Kim, Y. (2014, August). Social Media, Network Heterogeneity, and Opinion Polarization. *Journal of Communication*, 64(4), 702–722. doi: 10.1111/jcom.12077
- Lee, & Oh, S. Y. (2012). To Personalize or Depersonalize? When and How Politicians' Personalized Tweets Affect the Public's Reactions. *Journal of Communication*, 62(6), 932–949. Retrieved 2021-09-08, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-2466.2012.01681.x> (_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-2466.2012.01681.x>) doi: 10.1111/j.1460-2466.2012.01681.x
- Matsa, K. E., & Shearer, E. (2018). *News Use Across Social Media Platforms 2018 / Pew Research Center* (Tech. Rep.). Pew Research Center. Retrieved from <https://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/>
- McClain, C. (2021, May). *70% of u.s. social media users never or rarely post or share about political, social issues* (Tech. Rep.). Pew Research Center. Retrieved from <https://www.pewresearch.org/fact-tank/2021/05/04/70-of-u-s-social-media-users-never-or-rarely-post-or-share-about-political-social-issues/>
- McGregor, S. C. (2019). Social media as public opinion: How journalists use social media to represent public opinion. *Journalism*, 20(8), 1070–1086.
- McGregor, S. C. (2020). “taking the temperature of the room” how political campaigns use social media to understand and represent public opinion. *Public Opinion Quarterly*,

84(S1), 236–256.

- McGregor, S. C., & Molyneux, L. (2020, May). Twitter's influence on news judgment: An experiment among journalists. *Journalism*, 21(5), 597–613. (Publisher: SAGE Publications) doi: 10.1177/1464884918802975
- McGregor, S. C., Mourão, R. R., & Molyneux, L. (2017). Twitter as a tool for and object of political and electoral activity: Considering electoral context and variance among actors. *Journal of Information Technology & Politics*, 14(2), 154–167.
- Messing, S., van Kessel, P., & Hughes, A. (2017). *Sharing the News in a Polarized Congress* (Tech. Rep.). Pew Research Center. Retrieved from <https://www.people-press.org/2017/12/18/sharing-the-news-in-a-polarized-congress/>
- Mims, C. (2020, October). Why social media is so good at polarizing us. *Wall Street Journal*.
- Mitchell, A., Jurkowitz, M., Oliphant, J. B., & Shearer, E. (2020, July). *Americans Who Mainly Get Their News on social media Are Less engaged, less knowledgeable*. Retrieved 2021-05-04, from <https://www.journalism.org/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/>
- Molyneux, L., & McGregor, S. C. (2021, January). Legitimizing a platform: evidence of journalists' role in transferring authority to Twitter. *Information, Communication & Society*, 0(0), 1–19. doi: 10.1080/1369118X.2021.1874037
- Morales, A. J., Borondo, J., Losada, J. C., & Benito, R. M. (2015). Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3), 033114. Retrieved from <http://aip.scitation.org/doi/10.1063/1.4913758> doi: 10.1063/1.4913758
- Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. a. L., & Nielsen, R. (2017). Reuters Institute Digital News Report 2017. *Reuters Institute for the Study of Journalism*, 1–108. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/21670811.2012.744561%0Ahttp://www.digitalnewsreport.org/survey/2017/malaysia-2017/#fn-6162-2> doi: 10.1080/21670811.2012.744561
- Pennacchiotti, M., & Popescu, A.-M. (2011). A Machine Learning Approach to Twitter User Classification. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM)*. doi: 10.1145/2542214.2542215
- Preoțiuc-Pietro, D., Liu, Y., Hopkins, D., & Ungar, L. (2017). Beyond binary labels: political ideology prediction of twitter users. In *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 1: Long papers)*.

- Prior, M. (2007). *Post-broadcast democracy : how media choice increases inequality in political involvement and polarizes elections*. Cambridge University Press. Retrieved from https://books.google.com/books/about/Post_broadcast_Democracy.html?id=imT0jwEACAAJ
- Prior, M. (2009). The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure. *Public Opinion Quarterly*, 73(1), 130–143.
- Quan-Haase, A., Martin, K., & McCay-Peet, L. (2015, May). Networks of digital humanities scholars: The informational and social uses and gratifications of Twitter. *Big Data & Society*, 2(1), 2053951715589417. Retrieved 2021-08-11, from <https://doi.org/10.1177/2053951715589417> (Publisher: SAGE Publications Ltd) doi: 10.1177/2053951715589417
- Scharkow, M., Mangold, F., Stier, S., & Breuer, J. (2020). How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1918279117
- Sears, D., & Freedman, J. (1967). Selective Exposure to Information : A Critical Review. *The Public Opinion Quarterly*, 31(2), 194–213.
- Settle, J. E. (2018). *Frenemies: How social media polarizes america*. Cambridge University Press.
- Soliman, A., Hafer, J., & Lemmerich, F. (2019). A characterization of political communities on reddit. In *HT 2019 - Proceedings of the 30th ACM Conference on Hypertext and Social Media* (pp. 259–263). Association for Computing Machinery, Inc. doi: 10.1145/3342220.3343662
- Sunstein, C. R. (2017). *#Republic : divided democracy in the age of social media*. Princeton University Press.
- Tewksbury, D. (2005). The Seeds of Audience Fragmentation: Specialization in the Use of Online News Sites. *Journal of Broadcasting & Electronic Media* 2.
- Tsakalidis, A., Papadopoulos, S., Cristea, A. I., & Kompatsiaris, Y. (2015). Predicting elections for multiple countries using twitter and polls. *IEEE Intelligent Systems*, 30(2), 10–17.
- Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., ... Nyhan, B. (2018, March). *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature* (SSRN Scholarly Paper No. ID 3144139). Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3144139
- Usher, N., & Ng, Y. M. M. (2020, April). Sharing Knowledge and “Microbubbles”: Epistemic

- Communities and Insularity in US Political Journalism. *Social Media + Society*, 6(2).
doi: 10.1177/2056305120926639
- Weber, I., Garimella, V. R. K., & Batayneh, A. (2013). Secular vs. Islamist polarization in Egypt on Twitter. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining - ASONAM '13* (pp. 290–297). New York, New York, USA: ACM Press. Retrieved from
<http://dl.acm.org/citation.cfm?doid=2492517.2492557> doi:
10.1145/2492517.2492557
- Whiting, A., & Williams, D. (2013, January). Why people use social media: a uses and gratifications approach. *Qualitative Market Research: An International Journal*, 16(4), 362–369. Retrieved 2021-08-11, from <https://doi.org/10.1108/QMR-06-2013-0041>
(Publisher: Emerald Group Publishing Limited) doi: 10.1108/QMR-06-2013-0041
- Wojcik, S., & Hughs. (2019, April). *How Twitter Users Compare to the General Public*. Retrieved 2021-05-04, from
<https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/>
- Wong, F. M. F., Tan, C. W., Sen, S., & Chiang, M. (2016). Quantifying Political Leaning from Tweets, Retweets, and Retweeters. *IEEE Transactions on Knowledge and Data Engineering*, 28(8), 2158–2172. Retrieved from
<http://ieeexplore.ieee.org/document/7454756/> doi: 10.1109/TKDE.2016.2553667
- Yardi, S., & Boyd, D. (2010). Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter. *Bulletin of Science, Technology & Society*, 30(5), 316–327. Retrieved from <http://journals.sagepub.com/doi/10.1177/0270467610380011> doi:
10.1177/0270467610380011
- Zamal, F. A., Liu, W., & Ruths, D. (2012). Homophily and Latent Attribute Inference : Inferring Latent Attributes of Twitter Users from Neighbors. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*. doi:
10.9770/jesi.2015.3.1(2)T
- Zillman, D., & Bryant. (1985). *Selective Exposure To Communication*. Routledge.