# The Effect of Social Media on Elections: Evidence from the United States*

Thomas Fujiwara† Karsten Müller‡ Carlo Schwarz§

March 5, 2021

**Abstract**

We study how social media affects election outcomes using variation in the number of Twitter users across U.S. counties induced by participants of the 2007 South by Southwest (SXSW) festival, a key event in Twitter's rise to popularity. We show that this variation, which remains predictive of Twitter use a decade later, is unrelated to electoral outcomes before the platform's mass adoption. Our results suggest that exposure to Twitter lowered the Republican vote share in the 2016 and 2020 presidential elections, but had limited effects on turnout and vote shares in House and Senate races, as well as previous presidential elections. Evidence from two sources of survey data indicates that the effects are driven by independent and moderate voters. Our results are consistent with the idea that Twitter's relatively liberal content can persuade voters to alter their views.

*We are grateful to seminar participants at Princeton University, Imperial College, Warwick University, Bocconi University, University of California San Diego and the 2021 ASSA Conference for their helpful comments.

†Princeton University, Department of Economics, and NBER, fujiwara@princeton.edu.

‡Princeton University, Department of Economics, karstenm@princeton.edu.

§Università Bocconi, Department of Economics, and IGIER, carlo.schwarz@unibocconi.it.

# 1 Introduction

Does social media affect election outcomes? A popular narrative holds that Twitter played a decisive role in both the 2016 presidential election in the United States and the United Kingdom's "Brexit" referendum. Many see this as part of social media's broader influence on political polarization and the recent re-emergence of right-wing populist politicians in many countries. U.S. Federal Election Commissioner Ellen Weintraub, for example, has argued that the social media company Facebook "has no idea how seriously it is hurting democracy" (NPR, 2020a).[1]

An alternative view holds that social media platforms are biased against conservative voices (e.g., NPR, 2020b; Wall Street Journal, 2020) and that the younger, relatively left-leaning user base of platforms such as Twitter is unlikely to tilt elections towards right-wing politicians (e.g., Boxell et al., 2017, 2018). However, there is limited evidence that can be used to evaluate these contrasting (causal) claims. Given how widespread social media use has become in recent years, whether or not it has persuasion effects on voters is viewed as a key priority by policymakers.

This paper focuses on the effects of Twitter, a platform used by around a quarter of American adults. Specifically, we estimate how a county's number of Twitter users affects election results by exploiting the persistent network effect sparked by early Twitter adoption, building on Müller and Schwarz (2019). Although it was launched in March 2006, Twitter's popularity increased rapidly after the platform's advertising campaign at the South by Southwest festival (SXSW) in March 2007. The SXSW festival was also key for Twitter's geographical diffusion: counties with more SXSW followers who joined during the 2007 festival—and thus likely attended the event—saw disproportionately higher growth of Twitter adoption compared to counties with SXSW followers who joined just before the festival. This difference in Twitter use across counties persisted for over a decade, consistent with path dependence in technology adoption with network effects.

Our identification strategy leverages the 2007 SXSW festival as a shock to early Twitter adoption that is uncorrelated with pre-existing election results. Conditional on geographic and socioeconomic controls, and previous interest in the SXSW Twitter account, a county's number of SXSW followers who joined in March 2007 is essentially uncorrelated with a host of county characteristics. It is also unrelated to both levels

[1] See, for example, The New Yorker (2016); New York Times (2017); Allcott and Gentzkow (2017); The Guardian (2018); UK Parliament (2019).

and trends in election outcomes before Twitter's launch (between 2000 and 2004) and during the period it had fewer users (between 2006 and 2012). However, the number of SXSW followers who joined in March 2007 is highly correlated with Twitter usage in 2016, and has predictive power the 2016 and 2020 presidential election results.

We estimate that a 10% increase in a county's number of Twitter users lowered the vote share of Republican presidential candidate Donald Trump by 0.2 percentage points (p.p.) in the 2016 and 2020 presidential elections. The implied persuasion rate is 8.6%, which is smaller than the estimated pro-Republican effect of Fox News (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017), the pro-Democrat effect of the Washington Post (Gerber et al., 2009), or the effect of get-out-the-vote canvassing on voter turnout (Gerber and Green, 2000).

For previous presidential elections in 2008 and 2012, we find effects that are small and statistically indistinguishable from zero. The same holds true for House and Senate races, including the 2016 and 2020 elections. Twitter adoption thus lowered Trump's vote share but did not do so for Republican candidates in congressional races in the same election. Together with other "placebo tests," this pattern bolsters confidence that our estimates are capturing the effect of Twitter, which contains more content on presidential than congressional candidates.[2]

To shed light on the mechanisms behind these results, we estimate Twitter's effect on reported votes in the 2016 Cooperative Congressional Election Survey (CCES) and primary presidential candidates' approval in the Gallup Daily Tracker. Further, we explore data on the partisanship of political content on Twitter.

These exercises yield three findings. First, results based on the CCES indicate that Twitter's effect is driven by independents and moderates switching their votes towards the Democratic candidate (Hillary Clinton). This is consistent with Bayesian updating, since moderates presumably have less extreme priors and are thus more likely to be persuaded. Second, we estimate the effect on Gallup candidate approval rates during the primaries. We find that Twitter decreased Trump's approval ratings and increased Clinton's with only small effects on relatively more moderate Republican candidates.

[2]A violation of our identification condition requires a set of omitted variables that are correlated with a county's number of SXSW followers that joined Twitter in March 2007 but are at the same uncorrelated with: i) the number of SXSW that joined months earlier, ii) levels and trends in election results before Twitter's launch and rise to popularity, iii) a host of observable variables, iv) election results in congressional elections. Our argument is that such omitted variable is implausible to an extent that allows us to interpret our results as causal effects. See Section 5 for further details.

Third, we document that political content on Twitter has a clear pro-Democrat slant using a machine learning approach similar to Gentzkow and Shapiro (2010). We apply this method to the text of over 460 million tweets about the presidential candidates in the 2012, 2016 and 2020 elections. We find a striking shift in the partisan slant of these political tweets toward the left, particularly for tweets about the Republican candidates. We estimate that around 35% of tweets about Mitt Romney sent during the 2012 presidential election had a pro-Democratic slant but almost 80% of tweets about Trump in 2016. Moreover, tweets about Trump are much more numerous than those about Clinton and receive substantially more attention (measured by the number of "likes").[3]

Overall, our results are consistent with Twitter and its primarily pro-Democratic content persuading voters with moderate views that Clinton was preferable to Trump without inducing a more general negative effect on Republicans.

From the outset, we want to stress what our findings do *not* imply. First, they are silent on the effect of social media platforms other than Twitter, such as Facebook. Because we use an identification strategy based on variation in a "shock" that is specific to early Twitter adoption, we do not have a credible research design to estimate the effects of other platforms. While most types of social media share similarities with Twitter, such as being popular among younger and more educated people in urban areas (Pew Research Center, 2019c), other platforms can in principle have different effects on political outcomes. Second, our research design cannot separate the effect of particular types of social media content on Twitter, such as the potential role of foreign governments or misinformation ("fake news"). Third, given our research design, we estimate a "partial equilibrium" effect of adding Twitter users to a county while keeping constant other counties' Twitter use. We thus cannot address whether Twitter had a national-level effect on the election (e.g., Trump's tweets driving traditional media content).

Our work contributes to the literature on how media impacts political outcomes. Expansions of traditional media channels such as newspapers, radio, broadcast television, and cable news have been associated with changes in voter turnout, polarization, and electoral outcomes.[4] There is also evidence that broadband internet (Falck et al., 2014;

[3]Following Gentzkow and Shapiro (2010), we measure the slant of a tweet by the "similarity" of its language to that of a Republican or Democrat Member of Congress.

[4]See, for example, Gentzkow (2006); Huber and Arceneaux (2007); DellaVigna and Kaplan (2007); Gentzkow et al. (2011); Enikolopov et al. (2011); DellaVigna et al. (2014); Larcinese and Miner (2017); Martin and Yurukoglu (2017); Spenkuch and Toniatti (2018).

Gavazza et al., 2019; Campante et al., 2017; Lelkes et al., 2017) and mobile internet (Manacorda and Tesei, 2016; Guriev et al., 2019) exert political effects.[5] The electoral effects of social media have received less attention, although a nascent literature studies the political effects of social media more broadly, for example on protest participation (Howard et al., 2011; Enikolopov et al., 2020; Acemoglu et al., 2017) and xenophobia (Müller and Schwarz, 2020; Müller and Schwarz, 2019; Bursztyn et al., 2019). See Zhuravskaya et al. (2020) for a more extensive review.

Additionally, a burgeoning field of experimental research focuses on social media. Bond et al. (2012) and Jones et al. (2017) provide evidence that online messages on social networks affect voter turnout. Allcott et al. (2020) and Mosquera et al. (2020) find that individuals who deactivate Facebook react along many dimensions, including some measures of political polarization. Levy (2019) studies the effect of randomly assigning Facebook users subscriptions to conservative or liberal media outlets. Bail et al. (2018) estimate the effect of paying politically interested Twitter users to follow a Twitter bot with messages from the opposing political ideology. Chen and Yang (2019) trace out the medium-term impact of exposing Chinese citizens to uncensored internet. Perhaps most immediately related to our paper is recent work by Rotesi (2019), who finds social media negatively affected the Democratic vote share in 2012 by using variation in Twitter adoption resulting from player transfers by popular sports teams that have Twitter accounts.

Taken together, existing research provides an incomplete picture. On one hand, social media has been painted as a key force in the recent rise of right-wing populists in many countries. Experimental studies suggest that social media affects individuals' political beliefs, and observational studies suggest that internet access more broadly can affect political outcomes. On the other hand, whether social media platforms can indeed persuade voters and affect elections results on a larger scale remains largely unclear. Because of social media's widespread use, this lack of clarity raises the question of whether these platforms' potential persuasion effects work similarly to those established for traditional media. Our paper sheds light on this question by focusing on how Twitter affects federal elections in the United States.

[5]For reviews, see DellaVigna and Gentzkow (2010), Napoli (2014), Strömberg (2015), Enikolopov and Petrova (2015), and DellaVigna and Ferrara (2015). Experimental studies include, among others, Gerber and Green (2000), Gerber et al. (2009), and Gerber et al. (2011).

# 2 Background: Social Media and Politics

Most Americans use social media platforms or messaging applications. Data from the Pew Research Center suggest that the most popular services are YouTube (used by 73% of adults in the U.S.), followed by Facebook (69%), and Instagram (37%) (Pew Research Center, 2019c). 22% of adults in the U.S. use Twitter, a rate similar to that of Snapchat (24%) and Whatsapp (20%) users. On average, adult users spend more than an hour a day using social networks (eMarketer, 2019).[6]

One popular perspective is that online networks, and social media in particular, may give rise to so-called "filter bubbles" (Pariser, 2011) or "echo chambers" (Sunstein, 2017). The idea is that social media—unlike traditional mass media outlets—may facilitate the provision and consumption of one-sided information, either through the use of algorithms or by allowing individuals to self-select into preferred content. While there is considerable empirical evidence supporting this idea (e.g. Conover et al., 2011; Weber et al., 2013; Bessi et al., 2015; Del Vicario et al., 2016; Halberstam and Knight, 2016; Schmidt et al., 2017; Levy, 2019), other studies have found that individuals are exposed to a wide range of political opinions on social media (Barberá, 2014; Bakshy et al., 2015; Nelson and Webster, 2017; Beam et al., 2018), perhaps even more so than via traditional media outlets or personal interactions (Gentzkow and Shapiro, 2011). Some work also challenges the notion that increased polarization due to online channels is quantitatively important (Flaxman et al., 2016; Guess, 2018; Boxell et al., 2019).

Much of the recent public discussion about the role of social media platforms has been shaped by controversies, including the consulting firm Cambridge Analytica's involvement in multiple political campaigns (e.g., The Guardian, 2018); the Russian Internet Research Agency's efforts to support Trump's election campaign (e.g., New York Times, 2017); and the role of widely shared false information ("fake news") in both the 2016 U.S. elections (e.g., Allcott and Gentzkow, 2017) and the Brexit referendum in the United Kingdom (e.g., UK Parliament, 2019). Both Clinton and Trump have argued that these factors were instrumental in the 2016 election outcome, as has former president Barack Obama (The New Yorker, 2016). As Brad Parscale, Trump's digital media director in 2016, put it: "Facebook and Twitter were the reason we won this thing. Twitter for Mr. Trump. And Facebook for fundraising" (Wired, 2016). In

[6]Pew bases its usage measures on the share of respondents who state they have ever used one of the online platforms. Twitter reported around 69 million monthly active users in 2019 (see Statista, 2019), which yields a slightly higher share of around a third of the 210 million adults in the U.S.

Online Appendix Figure A.3, we document that discussions of social media have become increasingly frequent in major American news outlets. Mentions of Twitter in particular spiked with the 2016 presidential election when compared to 2012 levels.

Despite its prominence in the political discourse, the empirical relevance of social media for electoral outcomes is largely unknown, and some have suggested that concerns may be overblown. As one example, in the 2016 presidential election, Trump received fewer votes from demographic groups with higher propensities to use social media or the internet more broadly (The Hill, 2016; Boxell et al., 2017, 2018). Indeed, Trump's broadest support came from older white voters without college education in rural areas, who are among the least likely people to use social media actively (Hargittai, 2015; Pew Research Center, 2015, 2018). These patterns seem difficult to square with the idea that online channels were an important driver of the 2016 presidential election result, although such observations also do not rule this out.

Further, the content on social media platforms—particularly on Twitter—is disproportionately left-leaning. While there appears to be a cluster of right-wing networks, Pew Research Center (2019d) estimates that, in 2018, 60% of Twitter users identified as Democrat, and only 35% as Republican. Among Democrats, those on Twitter are considerably more liberal and focus less on finding common ground with Republicans (Pew Research Center, 2020). 26% of American Twitter users follow Obama and 19% follow Trump (Pew Research Center, 2019a). Survey evidence suggests that 80% of Twitter content is produced by people who strongly disapprove of Trump (Pew Research Center, 2019b). "Liberals" are also more likely to get political news on Twitter or Facebook and follow more media and political accounts compared to "conservatives" (Pew Research Center, 2014; Eady et al., 2019). Twitter and Reddit, which are often said to be pro-Trump factors, were considerably more popular among Clinton supporters before the 2016 election (Hargittai, 2015). Although social media allows users to partially select which content they see, Twitter content thus disproportionately leans toward the Democratic party.

We provide additional evidence for the composition of political content on Twitter by analyzing the Twitter reach of Democratic and Republican politicians. We collected data on the Twitter accounts of all Senators and House Representatives from the 110th to the 115th Congresses (2007-2019). In Figure 1, we plot the average number of tweets and followers that members of each party have on Twitter, as well as the average number of retweets and "likes" their tweets receive. The patterns here again clearly indicate that Democratic politicians are more active on Twitter and have larger follower bases

than their Republican counterparts. Tweets by Democrats also receive, on average, three times the number of "likes."[7]

[Figure 1 about here.]

# 3 Data

The main analysis is based on a county-level dataset on election outcomes, political opinions, and Twitter use. It covers 3,065 counties in 48 states (we exclude Alaska and Hawaii) and the District of Columbia (except in congressional elections). County-level election results are from Dave Leip's Atlas of U.S. Presidential Elections and the MIT Election Lab. We complement our analysis with individual-level survey data on approval ratings from the Gallup Daily Tracker and voting data from the Cooperative Congressional Election Study (CCES). Our Twitter use measures is derived from an archive of 475 million geo-located tweets compiled by Kinder-Kurlanda et al. (2017). We combine this with newly collected data on Twitter's early adopters at the 2007 SXSW festival; data on the Twitter activity of U.S. Congress members; and a large corpus of tweets related to the 2012, 2016, and 2020 presidential elections. Additional county characteristics were obtained from the U.S. Census, the U.S. Religious Census, the American Community Survey (ACS), and the Bureau of Labor Statistics (BLS). We describe the individual data sources in more detail below. Online Appendix Table A.2 provides additional details and summary statistics.

**Election Outcomes.** We use county-level data on presidential election outcomes between 1996 and 2020 from Dave Leip's Atlas of U.S. Presidential Elections. We complement this with county-level results on Senate and House elections from the MIT Election Lab for the 1996-2020 period. In all cases, we focus on two-party vote shares. Figure 2a visualizes the Republican party's vote share in the 2016 presidential elections.[8]

[Figure 2 about here.]

[7] In Online Appendix Figure A.2, we confirm that these patterns are not driven by a small group of Members of Congress by showing that they also hold when we compare the median Twitter reach of Democrats and Republicans.

[8] While senatorial and presidential elections are decided at the state level and House elections at the congressional district level, counties are usually smaller geographic units and far more numerous. Additionally, unlike congressional districts, county boundaries are fixed over our sample period, allowing us to observe changes across years.

**Individual-Level Voting Decisions.** The Cooperative Congressional Election Study (CCES) is a nationwide survey that collects information on voter behavior in two waves (before and after the election). We focus on votes for Trump and Clinton in the pivotal 2016 election.[9] The CCES contains a rich set of individual characteristics, including political affiliation, family income (in 12 bins), gender, race, education (in 6 bins), marital status, age, and interest in the news. Table A.3 provides summary statistics (weighted by sample weights). The CCES also uses administrative data on turnout records to verify if its respondents have voted.

**Presidential Candidate Approval.** The Gallup Daily Tracker provides individual-level survey data for a sample of 1,000 individuals per day since 2009.[10] During the 2016 presidential campaign, it fielded survey items regarding approval of Republican and Democratic presidential candidates. This allows us to investigate Trump's pre-election approval relative to other candidates (e.g. Clinton or Ted Cruz). The data also include a rich set of individual characteristics, including political affiliation, county of residence, income (in 10 bins), gender, race, marital status, age, and education (in 6 bins). Table A.4 in the Online Appendix provides summary statistics.[11]

**Twitter Usage.** We construct a measure of county-level Twitter usage based on a sample of 475 million geo-coded tweets collected by Kinder-Kurlanda et al. (2017). The tweets were collected between 2014 and 2015 using the Twitter Streaming API, which continuously returns a 1% sample of all Tweets as long as it is called. The individual tweets from this dataset are already assigned to counties. Additionally, we collected the underlying user profiles for each tweet in the database. This allows us to construct a user-based measure by assigning users to the county from which they tweet most frequently.[12]

The resulting measure, which we use throughout the paper, is a proxy for the number of Twitter users per county, based on 3.7 million individual users (around 7% of the Twitter population). Figure 2b plots the number of Twitter users per capita across

[9] At the time of writing, the CCES for the 2020 election was not yet available.

[10] The Gallup Daily Tracker is only published with a long delay. Data for the 2020 election are not available. As there was no Republican primary in 2020, the data from the Daily Tracker would be less relevant for our analysis.

[11] For some auxiliary estimations in the Online Appendix, we also collapse responses about presidential approval of Trump on the county-level using weighted averages based on the number of survey respondents in each county.

[12] Data is available in the Gesis Datatorium at https://datorium.gesis.org/xmlui/handle/10.7802/1166.

counties. Each user profile further provides us with a short biography and the date that each user joined Twitter.[13] We use the join dates to construct a time-varying proxy of Twitter usage based on how many of the Twitter users had already opened an account at each point in time. To validate this measure, Figure A.1 in the Online Appendix shows that our Twitter usage measure's evolution at each time point closely tracks the number of daily Twitter users from Statista (2019).

**Twitter Data for the South by Southwest Festival.** We collected data for our instrument for Twitter usage, based on early adoption during the SXSW festival, through the Twitter API. More specifically, we scraped the account data for 658,240 users who followed the Twitter account of SXSW Conference & Festivals (@SXSW) at the time of collection (January 2019). We assign these users to counties based on the location people report in their user profile.[14]

A user profile contains the month and year that they joined Twitter, which allows us to determine the number of SXSW followers in each county that joined Twitter in a particular month. The two key variables in our analysis are: i) the number of SXSW followers that joined Twitter in the month of March 2007 and ii) the number of SXSW followers that joined Twitter during 2006 (the year the platform was launched). We refer to (ii) as the number of SXSW that joined before the March 2007 festival. We also scraped the follower lists of SXSW followers who joined in March 2007, which allows us to investigate the connections of Twitter users to the SXSW festival. Further, we additionally collected tweets mentioning the festival, based on the term "SXSW," as well as a proxy for overall Twitter activity based on the 100 common English words.[15] We use these measures to document the SXSW festival's impact on local Twitter adoption.Election security... - I'm writing to CSP to update the even...

**Data on Political Twitter Activity.** We scraped the tweets and user profiles of the 901 Senators and House Representatives from the 110th to 115th (2007-2019) Congress

[13]Note that Kinder-Kurlanda et al. (2017) restrict the data to the sample of tweets that are geo-located. In principle, users can opt out of allowing the platform to determine their location. However, at the time the tweets were collected, the default option was to allow geo-location (i.e., users had to take an action to disable location tagging). As a result, the number of geo-located tweets at this point in time was far larger than today.

[14]Of the 44,625 SXSW followers who joined between 2006 and 2008, we are able to geo-code 25,830 (58%).

[15]We report the full list of words in Table A.7.

who have Twitter accounts. This includes 424 Democrats and 465 Republicans.[16] In total, the data contain 1,999,893 tweets, which we use to analyze the Twitter reach of Democratic and Republican congress members. In Online Appendix Table A.1, we list the 20 Members of Congress with the most Twitter followers.

We complement this dataset with election-related tweets that shed light on the overall partisan slant of Twitter activity during the 2012, 2016, and 2020 elections. For the 2016 election, we use the archive from Littman et al. (2016), which contains 280 million tweets, collected between July 13, 2016, and November 10, 2016, from the Twitter API. For the 2012 election, we use data collected by Diaz et al. (2016), comprising 24 million tweets containing either "Obama" or "Romney" for the period from July 1, 2012 through November 7, 2012. Finally, the 2020 election tweets are based on the archive from Chen et al. (2020), which covers the period from March 2020 to November 2020. To make these datasets comparable, we restrict the 2016 election sample to tweets mentioning either "Clinton" or "Trump" (112 million tweets). Similarly, we restrict the 2020 data set to the time period from July 1, 2020 through November 3, 2020 and tweets mentioning either "Biden" or "Trump" (339 million tweets).

To determine the likely political affiliation of Twitter users, we create measures of political slant based on how "similar" a tweet's content is to tweets sent by Republican or Democratic Members of Congress. More precisely, we train L2 regularized logistic regression models based on 901 congresspeople's Twitter accounts to classify whether a tweet contains language frequently used by either Republican or Democratic politicians. We clean the text of the tweets by removing common words (stopwords) and by reducing the words in each tweets to their morphological roots (lemmatizing). The input is based on unigrams, bigrams, and trigrams from these tweets. We choose the optimal normalization strength using 10-fold cross-validation. The resulting classifier achieves an out-of-sample F1-score of 0.9. We then use this classifier to predict a partisan score between 0 and 1 for each of our election-related tweets. These scores can be interpreted as the probability of a tweet with the same content being sent by a Republican. A score of 1 indicates a very "Republican-sounding" tweet, while 0 means that a tweet exhibits a strong Democratic slant. As such, our approach is similar to how Gentzkow and Shapiro (2010) measure newspaper slant. We provide additional details regarding our machine learning approach in Online Appendix Appendix A.1.1, which also visualizes the most predictive terms identified by the classifier.

[16]The remaining 12 politicians are either Independents or switched their party affiliation at some point.

**Additional County Characteristics** We collect county-level demographic control variables from the U.S. Census and the American Community Survey. In particular, we use information on population, population share by age group and ethnicity, poverty rates, and education levels. We also obtained industry-level employment shares and unemployment rates from the Bureau of Labor Statistics. Additional controls on county media usage patterns are from Simply Analytics. We also construct geographical controls such as the distance from Austin, TX, where SXSW takes place every year; population density; and county size (in square miles). For one set of results we also use donation data from OpenSecrets. We provide a complete description of the variables in Table A.6.

# 4 The 2007 South by Southwest Festival and Early Twitter Adoption

The empirical strategy behind our main results exploits a shock to early-stage Twitter adoption connected to the 2007 SXSW festival, as in Müller and Schwarz (2019). This section discusses the key role of the festival in boosting the platform's popularity and document how it created a persistent effect on its spatial diffusion.[17]

Founded in March 2006, Twitter was largely unknown before SXSW 2007. Twitter's popularity increased dramatically after the festival, where Twitter strategically placed screens in the conference hallways and allowed users to sign-up by simply sending a text message (SMS) to a predefined number. As a result, speakers and bloggers in attendance broadcasted the platform to the outside world, and Twitter went on to win the South by Southwest Interactive Web Award Prize.

SXSW's immediate impact on Twitter's popularity in early 2007 can be seen in Figure 3a, which plots our proxy for the daily number of tweets as well as the number of tweets explicitly mentioning SXSW. The figure shows that Twitter's growth rate accelerated during the festival, visible as the spike in SXSW-related tweets. The month-to-month growth rate of Twitter quadrupled with the start of the SXSW festival. After SXSW 2007, Twitter experienced further rapid growth (Venture Beat, 2008). The

[17]SXSW is an annual conglomeration of parallel film, interactive media, and music festivals and conferences organized jointly that take place in March in Austin, TX.

platform went from an average of 5,000 tweets a day in 2007 to 300,000 in 2008, and 2.5 million in 2009 (Twitter, 2010). In 2019, users sent roughly 500 million tweets a day.[18]

[Figure 3 about here.]

We exploit that the SXSW festival had persistent effects on Twitter's spatial diffusion. This is likely the result of network effects that are key to the social media experience, as a larger number of users makes it more interesting for potential new users to join. Such a mechanism also applies at the local level. For example, a boost in the number of neighbors, personal connections, and/or people who play a prominent role in a locality should also boost the value of joining the platform for those in that locality. Anecdotal evidence suggests that users who joined during SXSW 2007 spread Twitter to their counties of residence.

We provide further support for this hypothesis by investigating whether the inflow of early-stage adopters put these counties on a differential growth path of Twitter usage. Figure 3b plots the estimates of $\beta_\tau$ from the following panel event study regression at the county ($c$) and week ($t$) level:

$$tweets_{ct} = \sum_\tau \beta_\tau SXSW_c^{March2007} \times 1(t=\tau) + \sum_\tau \delta_\tau SXSW_c^{Pre} \times 1(t=\tau) + \theta_c + \gamma_t + \varepsilon_{ct}.$$

where $tweets_{ct}$ is the log of (one plus) the number of tweets in county $c$ on week $t$, $SXSW_c^{March2007}$ is the logarithm of (one plus) the number of SXSW followers in county $c$ that joined Twitter on March 2007 and $SXSW_c^{Pre}$ is a similarly defined variable for followers that joined Twitter before March 2007. $\beta_\tau$ thus illustrates, conditional on county and week fixed effects, the difference in the number of tweets sent from counties with relatively larger numbers of SXSW followers that joined on March 2007. The variables are standardized to have a mean of zero and standard deviation of one. The whiskers represent 95% confidence intervals based on standard errors clustered at the state level. The sample includes the period between the third and fourteenth week of 2007.

Figure 3b illustrates that home counties of SXSW followers who joined during the festival in March 2007 saw a rapid, disproportionate increase in Twitter usage around the time of SXSW. We find no evidence for pre-existing trends prior to March 2007.

[18] Our proxy for Twitter usage is created by scraping tweets that contain any of the 100 most common English words listed in Table A.7. Our data contain any tweet that contains at least one of these words. We should therefore obtain a large fraction of the English speaking tweets at that point in time.

This is consistent with the anecdotal and descriptive evidence that SXSW was a catalyst for the spread of Twitter in the United States.

Online Appendix Figure A.5a presents additional evidence on the long-term adoption effect of the 2007 SXSW festival. It plots estimates from a similar regression as the one in Figure 3b but in a county-quarter panel covering the period from Twitter's launch in 2006 to 2016. The dependent variable is substituted by the logarithm of (one plus) the number of Twitter users in a county based on our baseline measure. The resulting S-shaped pattern in the figure is consistent with models of technology adoption in the presence of network effects. More importantly, we find that the amount of early adopters in a county still matters for the amount of Twitter usage today.

Additionally, Figure A.5b shows just how dominant Twitter users connected to the SXSW festival were among early adopters. In 2007, we estimate that around 60% of Twitter users either followed the SXSW festival or followed someone who followed SXSW and joined in March 2007. As the number of Twitter users increased over time, the importance of SXSW followers in the platform declined. But as Figure A.5a shows, the festival created persistent differences at the county level. The next section outlines how we use the SXSW festival in our 2SLS estimates.

# 5 Empirical Framework

Our identification strategy leverages the 2007 SXSW festival as a shock to early Twitter adoption. We show that, conditional on a set of controls (described in further detail below), a county's number of SXSW followers that joined Twitter in March 2007 is uncorrelated with levels and trends in election outcomes before Twitter's launch and during its early years. It is also uncorrelated with a host of observable county characteristics. This feature of the data can be interpreted as idiosyncratic factors (e.g., who attended the 2007 SXSW, who decided to join Twitter at the time), giving us a "natural experiment" or "exogenous shock" in Twitter adoption that allows to estimate its effect on election outcomes. This interpretation is, of course, not self-evident, and we provide several pieces of evidence in its support.

An important concern is that counties whose population are more interested in the SXSW festival (and its Twitter account) may be systematically different from other counties. To address this issue, our empirical strategy exploits variation in the exact timing of when Twitter users interested in SXSW joined the platform across counties.

In particular, our regressions control for the number of SXSW followers who joined in the months *before* the festival. Intuitively, our empirical strategy compares a "treatment" group of counties with SXSW followers that joined in March 2007 (during the festival) against a "control" of counties with followers that joined before. While both groups of followers were interested in SXSW, we show that only the number of followers that joined on March 2007 are predictive of later Twitter adoption, consistent with the evidence that users that joined during the festival were key in the platform's diffusion. In contrast, the counties in which people joined *before* the festival did not receive additional future Twitter users.[19]

The "treatment" and "control" counties are similar along several characteristics. Table A.5 compares the average characteristics of three types of counties relevant for our identification strategy: 1) the 47 counties with SXSW followers that joined Twitter both in March 2007 and the "pre-period;" 2) the 108 counties with SXSW followers that joined in March 2007 (but none in the "pre-period"); and 3) the 20 counties with SXSW that joined in the "pre-period" (but none in March 2007). Differences in vote shares in the 1996 presidential election, demographics (e.g., race, age, education), and media consumption (e.g., share that watches Fox News) are quantitatively small or zero. This is particularly true for groups (2) and (3) — which are key to the identification — with $t$-tests indicating that differences between the two groups are not statistically different from zero.[20] The geographical variation in the three groups of counties is shown on Online Appendix Figure A.6. As the results in Table A.5 suggest, the counties do not differ systematically in size and how distant they are from major American cities.

Moreover, observable individual characteristics of SXSW followers who joined Twitter in March 2007 and the "pre-period" are also similar. We validate this using data on Twitter user profiles we obtained from the platform. Table A.8 shows that followers who joined in March 2007 have similar first names and profile descriptions compared to those that joined before: users in both groups tend to have common names (e.g., "Michael" or "Chris") and use words such as "founder" or "tech" to describe themselves in their profiles. The correlations of the frequency of first names and terms used in their bios between the two groups are 0.69 and 0.92, respectively.

[19] An alternative approach would be to compare the counties of users who signed up for Twitter during SXSW 2007 with those of users who signed up during *other* festivals. The results from such an exercise are similar and available upon request.

[20] Given the large number of county characteristics, we report Šidàk-corrected $t$-statistics, which are smaller than those generated by applying the Bonferroni correction.

**Specification.** Motivated by the evidence above, our main results are based on estimating the following two equations:

$$Twitter\ users_c = \alpha + \beta \cdot SXSW_c^{March2007} + \gamma \cdot SXSW_c^{Pre} + \mathbf{X_c}\delta + \xi_c \tag{1}$$

$$y_c = \alpha' + \beta' \cdot SXSW_c^{March2007} + \gamma' \cdot SXSW_c^{Pre} + \mathbf{X_c}\delta' + \zeta_c, \tag{2}$$

where $c$ indexes counties, $SXSW_c^{March2007}$ is the logarithm of (one plus) the number of SXSW followers in county $c$ that joined Twitter on March 2007, and $SXSW_c^{Pre}$ is a similarly defined variable for followers who joined Twitter before March 2007. $\mathbf{X_c}$ is a vector of control variables. Note that the right-hand side of both equations is similar. $twitter\ users_c$ is the logarithm of the number of Twitter users in the county (during 2014-2015). $y_c$ are election outcomes (e.g., vote shares), which we estimate in both levels and changes (e.g., $y_c$ can be the vote share in 2016 or the change in vote shares between 2000 and 2016).

In a 2SLS framework, equations (1) and (2) are the first-stage and reduced form, while the second stage is

$$y_c = \phi + \theta \cdot \widehat{Twitter\ users}_c + \pi \cdot SXSW_c^{Pre} + \mathbf{X}'\rho + \varepsilon_c, \tag{3}$$

where $\widehat{Twitter\ users}_c$ is predicted from the first stage regression in equation (1). We weigh observations by turnout (total number of votes cast) in the 2000 presidential election and cluster standard errors at the state level.[21]

**Identification.** Formally, the identification condition for the effect of Twitter users ($\theta$) is that $E(SXSW_c^{March2007} \cdot \varepsilon_c) = 0$ holds. Intuitively, this states that, conditional on the $SXSW_c^{Pre}$ and other controls ($\mathbf{X_c}$), the number of SXSW followers who joined in March 2007 is uncorrelated with other determinants of political outcomes $y_c$, implying that it only affects political outcomes via Twitter usage (the "exclusion restriction").

We provide four pieces of evidence supporting that this condition holds. First, as discussed above, including the $SXSW_c^{Pre}$ control implies that the identifying variation comes from comparing counties with similar observable characteristics.

Second, the coefficient of $SXSW_c^{Pre}$ is small and statistically insignificant in our first stage regressions. This provides us with a "placebo" test based on checking if

[21] We consider spatial standard errors using the methods described in Colella et al. (2019) for robustness.

it is also unrelated to political outcomes in the reduced form and 2SLS regressions. Intuitively, we have two variables that are correlated with interest in the SXSW festival among early Twitter adopters, but only one predicts Twitter users in later years, allowing us to disentangle interest in the festival from its effect via more platform users.

Third, estimating equations (2) and (3) for different time periods shows that $SXSW_c^{March2007}$ does not correlate with both levels and trends in election outcomes before Twitter's launch in 2006 and in its early years, when the platform had few users and was unlikely to affect election outcomes. Intuitively, outcomes in "treatment" and "control" counties behaved similarly before Twitter could plausibly affect elections.

Fourth, we find an effect of $SXSW_c^{March2007}$ on Trump's vote share in 2016 and 2020 but not on House and Senate elections (in neither 2016 nor 2020 and other periods between 2000 and 2018). This pattern is consistent with an effect of Twitter since there is more content on presidential candidates than on congressional elections in the platform.

Stated differently, a violation of the identification condition would require an omitted variable that correlates with $SXSW_c^{March2007}$ and $y_c$ but is uncorrelated with: i) $SXSW_c^{Pre}$, ii) levels and trends in election results before Twitter's launch and rise to popularity, iii) the observable variables presented in Table A.5, and iv) election results in congressional elections both during the Trump elections and before. Our argument is that the existence of such an omitted variable is implausible to an extent that allows us to interpret $\theta$ as the effect of Twitter users on election outcomes.[22]

Moreover, results based on survey data suggest the effects are concentrated among moderate or independent voters, which is also the expected pattern from Twitter having a causal effect due to voter persuasion.

Lastly, another possible concern is that the SXSW adoption shock led to differences in the *composition* of Twitter users when compared to other U.S. counties. We address this concern by comparing the profiles of Twitter users in SXSW home counties with the profiles of Twitter users in the rest of the country. The results are presented in Online Appendix Table A.9. We find that the users profiles in SXSW counties are highly similar to the general Twitter population.

[22]Note also that, while $Twitter\ users_c$ may be measured with error, since it is based on a sub-sample of the universe of tweets, the instrument $SXSW_c^{March2007}$ is based on the universe of SXSW Twitter followers. It is thus unlikely that measurement error concerns affect our 2SLS estimates (although they may affect OLS estimation of equation (3).

# 6 Results

## 6.1 Main results

**First-stage.** Table 1 reports results from estimating equation (1) with different sets of control variables (described in Online Appendix Table A.6). The results indicate that counties with more SXSW followers who joined Twitter in March 2007 have higher numbers of Twitter users during 2014-2015. Since the variables are in logs, the coefficients can be interpreted as elasticities. A one standard deviation increase in SXSW followers in March 2007 (0.32) is associated with around 16% more Twitter users. The results do not seem to be sensitive to the set of included covariates. In contrast, the coefficients for SXSW followers before the 2007 festival are statistically insignificant and small in size: Twitter usage in 2014-2015 is not higher in areas with more SXSW followers who joined Twitter before March 2007.

Figure 4 presents the graphical representation of the estimates in column (5) of Table 1. Specifically, we show a binned scatter plot of $Twitter\ users_c$ against $SXSW_c^{March2007}$ after both variables are "residualized" by partialling out the control variables. The figure is constructed by dividing the x-axis variable into 40 equal-sized bins and plotting the average values of both variables in each bin.[23]

[Table 1 about here.]

[Figure 4 about here.]

**Reduced Form and 2SLS Estimates.** Table 2 shows the reduced form estimates (equation 2) and both OLS and 2SLS estimates of equation (3), focusing on the Republican vote share in the 2016 presidential election. The specifications across columns match those in Table 1. Panel B indicates that the number of SXSW followers who joined Twitter in March 2007 is associated with a lower Republican vote share. Panel C presents the 2SLS effects of Twitter usage on vote shares. The 2SLS estimate in column (5) indicates that a 10% increase in the number of Twitter users in a county lowers

[23]The fitted line is based on the unbinned data. Observations are weighted by turnout in the 2000 presidential election. This procedure guarantees the slope of the fitted line matches the estimate on column (5) of Table 1.

Trump's vote share by 0.21 p.p. (e.g., a reduction from a 46.1% vote share to 45.8%).[24] The results for the 2020 presidential election shown in Table 2 are nearly identical.

Although the estimated effects are always negative and significant at the 1% level, comparing columns shows that the effect sizes vary somewhat with the inclusion of controls, especially demographic and socioeconomic ones. This sensitivity to controls is perhaps expected given their explanatory power over vote shares. The results in other papers that explore effects on vote shares in similar frameworks, such as DellaVigna and Kaplan (2007), Martin and Yurukoglu (2017), and Autor et al. (2020), show a similar sensitivity to controls.

Figure 5a shows the graphical representation of the estimates in column (5) of Panel B of Table 2. Similarly, Figure 5c visualizes the relationship underlying the estimates in column (5) of Panel B of Table 2. Both figures are constructed similarly to Figure 4 but using the 2016 Republican vote share as the y-axis variable. The OLS (non-instrumented) relationship between Twitter users and the Republican vote share from Panel A of Table 2 and Table 2 are shown in Figure 5b and Figure 5d. Both are also negative but slightly smaller in magnitude than the 2SLS estimates.

**Magnitudes and Persuasion Rates.** To interpret the magnitudes of our estimates, we calculate a persuasion rate following DellaVigna and Gentzkow (2010). It can be approximated as $\theta \cdot \frac{t}{e(1-y)}$, where $\theta$ is the 2SLS estimate, $y$ is average Republican vote share, $e$ is the average exposure of American adults to Twitter, and $t$ is the share of adults that turn out to vote. Using the estimate for $\theta$ from column (5) of Table 2, the persuasion rate is 8.6%. It implies that one out of every twelve active Twitter users that voted for Clinton would not have done so if they had not been exposed to the platform.[25]

[24]The F-statistic of our estimated first-stage range from 70 to 120. This suggests that estimation and inference concerns related to weak instruments are unlikely to apply in our case.

[25]The persuasion rate re-scales effect sizes by how many individuals are exposed to the platform and how many are not already persuaded. For marginal changes in exposure, the formula is $f = \frac{dy}{de} \cdot \frac{t}{1-y}$ (DellaVigna and Gentzkow, 2010). Since our estimate $\theta$ is the semi-elasticity $\frac{dy}{de} \cdot e$, we obtain $f = \theta \cdot \frac{t}{e(1-y)}$. In our context, $y = 0.46$, $t = 0.55$, and we assume $e = 0.25$, based on around a quarter of American adults being active Twitter users (Pew Research Center, 2019c). This implicitly assumes that Twitter usage among voters is the same as the overall population. If voters are over-represented among Twitter users, the persuasion rate would be smaller. On one hand, Twitter users are younger (which is associated with lower turnout) but more educated (which is associated with higher turnout) than the general population. Note also that we estimate county-level effects, which may also capture local spillovers and social interaction effects.

This rate is smaller than the estimated pro-Republican persuasion rate of Fox News, which Martin and Yurukoglu (2017) estimate to range between 27% and 58% (depending on the year).[26] It is also smaller than the 19.5% pro-Democrat persuasion rate that Gerber et al. (2009) estimate for the Washington Post. As a further comparison, Gentzkow et al. (2011) estimate a persuasion rate of 12.9% for the effect of reading a local newspaper in the 1869-1928 period on voter turnout. DellaVigna and Gentzkow (2010) report that randomized get-out-the-vote canvassing experiments have estimated persuasion rates within the 10%-15% range.

**Further Robustness and Additional Tests.** The Online Appendix presents a number of additional sensitivity checks. In Table A.10, we consider changes to the baseline regression specification. In particular, we allow for unweighted regressions; alternative functional forms of the pre-SXSW user variable; restrict the sample to the sub-sample of counties where we observe either SXSW followers who joined in March 2007 or the pre-period; and allow for flexible spatial correlation in the standard errors. Further, Table A.12 replaces our baseline time-invariant measure of Twitter usage (making use of all available user data) with a time-varying measure based on how many users were using the platform in a given year. None of these adjustments make a substantial difference in the magnitudes or statistical significance of the estimates.

Figure A.8 provides additional robustness checks for our baseline results for the 2016 and 2020 presidential vote shares. Here, we plot the estimated $\beta$ of our baseline 2SLS equation 3 while flexibly allowing the included control variables to vary. The resulting "specification curves" suggest that our results are highly robust to how our regressions are specified. The estimated coefficients are always negative, almost always statistically significant at the 5% level, and in the overwhelming number of specifications considerably more negative than our baseline estimates.

Online Appendix Table A.14 reports results for additional outcome variables, all of which support the idea that Twitter exerts a pro-Democrat effect. In column (1), we use a probit IV model to investigate the likelihood of a county switching from Obama in 2008 to Trump in 2016. The coefficient suggests that a one standard deviation increase in Twitter usage is associated with a 24% lower probability of a county electing Obama in 2008 and Trump in 2016. Columns 2 and 3 look at changes in campaign donations to Democrats and Republicans between 2000 and 2016. We find a positive and statistically

[26] DellaVigna and Kaplan (2007) estimate a smaller persuasion rate of 11.6% for Fox News.

significant effect for donations to Democrats, and no effect for Republicans. Lastly, columns 4 and 5 look at approval ratings for President Trump in 2017 based on data from the Gallup Daily Poll. We find that exposure to social media is associated with a decrease in Trump's popularity, and more so among Republicans.

[Figure 5 about here.]

[Table 2 about here.]

**Placebo Test.** The coefficients on $SXSW_C^{Pre}$ in Table 2 are statistically insignificant and substantially smaller than those on $SXSW_c^{March2007}$. As discussed in Section 5, this provides support for our identification condition (exclusion restriction). Suppose that our instrument merely captured that counties with an interest in SXSW's Twitter account during the platform's early years also differ in (unobservable) characteristics that predict the 2016 election outcome. If this was the case, the coefficients on $SXSW_C^{Pre}$ should be similar in size to those on $SXSW_c^{March2007}$. Intuitively, we have two variables that are correlated with interest in the SXSW festival, but only one predicts Twitter users in later years, allowing us to disentangle interest in the festival (and its correlates) from its effect via more platform users.

**The 2007 SXSW Shock and Previous Election Outcomes** Table 3 presents estimates repeats the exercise from column (5) of Table 2 using Republican vote share in the previous presidential elections of 2004, 2008, and 2012 as the dependent variable. All estimates for those years are substantially smaller than the one for 2016 and statistically insignificant. For 2004 and 2008, this can be interpreted as a placebo or pre-trends test: conditional on the covariates, our instrument is uncorrelated with outcomes before Twitter's launch (2004) and when the platform had few users (2008). We return to the comparison to 2012, during which Twitter use was more widespread, in Section 6.3. As discussed in Section 5, this result lends additional support for our exclusion restriction. If our instrument merely captured uncontrolled differences across counties, these should also correlate with vote shares in previous elections.[27]

[27]Twitter was officially launched in 2006. Online Appendix Figure A.1 shows the evolution of Twitter users over time.

**Effects on Vote Share Changes.** We also consider specifications of equations (2) and (3) using vote share *changes* instead of *levels* as the dependent variable. All our estimates based on changes take differences relative to the base year 2000 (akin to the approach in Autor et al. (2020)) and use the full set of controls (as in column 5 of Table 2). Figure 6a plots the reduced form estimates for changes in the Republican vote share in presidential elections.

The results corroborate the previously presented evidence based on specification in levels. Our instrument is essentially uncorrelated not only with *levels* but also with *changes* (or *trends*) in election outcomes during the 2000-2012 period. Given our arguments above, this also lends support for our identification strategy. The reduced form effects for 2016 and 2020 are similar to the one estimated using levels instead of changes. The estimated effect ($\theta$) for 2016 using *changes* is -0.017, similar to the one estimated in levels (-0.021).[28]

**Effects on Turnout and Congressional Elections.** Figure 6b, Figure 7a, and Figure 7b replicate the exercise in Figure 6a using voter turnout and vote shares in House and Senate elections as the outcomes. We do not find a statistically significant association between our instrument and election turnout except for 2020. Before 2020, the estimated point effects are usually small. For example, the upper bound on the 95% confidence interval of the 2SLS estimate for the effect of turnout in the 2016 election implies that a 10% increase in Twitter users raises turnout by 0.036 p.p. (Online Appendix Table A.15).

In the highly polarized 2020 elections, which saw the highest turnout rate in more than a century and the highest number of votes ever cast in a U.S. election (NPR, 2020c), we find that Twitter is associated with a larger fraction of voting age population coming out to cast their ballot. However, given the zero effect on turnout in 2016, turnout alone is unlikely to explain the effect on the Republican vote share.

The coefficients for House and Senate races are somewhat more noisily estimated, particularly for the smaller sample of Senate races (where only a third of seats is renewed every two years). Overall, there is little evidence suggesting an effect of Twitter on congressional elections, including in 2016, the 2018 midterm election, and 2020. Finding an effect on presidential vote shares but not in these "down-ballot" races is perhaps

[28]Online Appendix Table A.11 in the Online Appendix also plots the OLS and 2SLS estimates.

expected since content on presidential candidates (and in particular on Trump in 2016 and 2020) is more common on Twitter than content on congressional races.[29]

**Discussion of Identification Condition (Exclusion Restriction).** As discussed in more detail in Section 5, there are four pieces of evidence supporting our identification condition: i) our empirical strategy compares relatively similar counties (Section 5); ii) the placebo test based on the coefficient on $SXSW_c^{Pre}$; iii) the instrument being uncorrelated with levels and changes in election outcomes in the 2000-2012 period; and iv) evidence suggesting Twitter has not affected House and Senate races.

Given this, a violation of the identification condition would require an omitted variable that correlates with the instrument and Trump's vote share in 2016 but is uncorrelated with: i) $SXSW_c^{Pre}$, ii) levels and trends in election results before Twitter's launch and rise to popularity, iii) the observable variables presented in Table A.5, and iv) election results in congressional elections. Our argument that is the existence of such omitted variable is unlikely.

[Table 3 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

**Towards an Average Treatment Effect** As with any instrument, a challenge for interpreting the empirical estimates is that we only identify the local average treatment effect (LATE) instead of the average treatment effect (ATE) of Twitter exposure. In our setting, the "compliers" are counties with higher Twitter usage as a result of the inflow of SXSW attendees. While the negative treatment effect of Twitter usage for these counties is in itself an interesting finding, the ATE for the US overall—and therefore the overall impact of Twitter on elections—is not identified by our estimates, because we only observe variation in SXSW users for a subset of the more than 3,000 U.S. counties.

[29]Online Appendix Table A.16 presents the 2SLS estimates for House and Senate races. To accommodate the Senate's six-year terms, we take changes relative to 2000, 1998, and 1996, instead of always using 2000 (as we do for other outcomes). While we only report in changes, results based on specifications in levels for turnout, House, and Senate elections are similar.

However, two pieces of evidence suggest that our estimates, despite this concern, allow us to infer information about the ATE.

A first indication that our estimates may identify relevant effects comes from comparing the OLS and IV results in Table 2. The estimates are very similar in column 2, and differ by a maximum of around 50% in column 5. This suggests that the IV estimates are in line with what one would expect exploiting the variation in Twitter usage across the entire United States.

Secondly, we build on the approach suggested by Andrews and Oster (2019) and more formally investigate the external validity bias of our estimates. For experimental settings, Andrews and Oster (2019) suggest using the observable heterogeneity in estimated treatment effects within the experimental sample to learn about the ATE in the overall population. Similarly, we can use the heterogeneity of treatment effects within our "treated" counties to approximate the ATE for the United States as a whole.[30] Using all available control variables for the prediction of heterogeneity of the treatment effect, the Andrews and Oster (2019) approach suggests that the ATE should, if anything, be *larger* than the LATE we estimate in our baseline results. More specifically, we find that, for a baseline LATE of $-0.085$, variation in observable characteristics would predict an ATE of $-0.218$. This seems plausible as the more urban counties for which we have instrument variation tend to be Democratic strongholds, and thus likely have fewer independents and moderate Republicans, for which we find the largest persuasion effects. While these findings should only be regarded as suggestive, they provide further support for the idea that Twitter might have increased the vote share of Democrats in the United States.

## 6.2 Effects Are Concentrated Among Moderate Voters

If social media indeed matters for election outcomes, we would expect there to be heterogeneous effects across groups of voters. In particular, Bayesian updating suggest that voters who do not hold strong priors about a particular party should be more likely persuaded. We test this prediction using individuals' voting decision from the 2016

[30] As our setting differs slightly from the one discussed in Andrews and Oster (2019), we need to make some adjustments to our baseline estimation. We estimate the treatment effect exclusively in the subset of counties for which either $SXSW_c^{March\ 2007}$ or $SXSW_c^{Pre}$ are not equal to zero. Then, we define a treatment indicator variable equal to 1 for counties with SXSW followers who joined in March 2007.

CCES. In particular, we estimate the following instrumental variable Probit regression:

$$y_{ic} = \phi + \theta \cdot Twitter\ users_c + \pi \cdot SXSW_c^{Pre} + \mathbf{X}_{ic}\rho + \varepsilon_{ic}, \tag{4}$$

where $y_{ic}$ is an indicator variable equal to 1 if an individual $i$ living in county $c$ voted for Trump in the 2016 election and 0 for Clinton.[31] The definition of the county-level variables $Twitter\ usage_c$ and $SXSW_c^{Pre}$ remains unchanged. $\mathbf{X}_{ic}$ is now a vector of individual-level control variables including age, gender, race, family income, and education. We again instrument for county-level Twitter usage based on the SXSW followers who joined in March 2007 and estimate equation (4) using a maximum likelihood estimator. As such, the coefficient $\theta$ captures whether individuals in counties with more Twitter users were more likely to vote for Trump.

Table 4 presents results from estimating equation (4). In Panel A, Column 1 suggests county-level Twitter usage has a statistically significant negative effect on the likelihood to vote for Trump. The reported marginal effect implies that a 10% increase in the number of Twitter users in a county would lower Trump's vote share by 0.47 p.p..[32]

Columns (2)-(6) report results estimated separately by voters' reported party affiliation. The effect is strongest for voters who identify as independents, and thus likely to not hold strong priors. We also find a negative coefficient for moderate Republicans, which is statistically significant at the 10% level. Overall, the results suggest weaker or zero effects for those with stronger political views, whether Republican or Democrat.

These results on party affiliation further vary with age in a way that may be suggestive of a role for social media. In particular, Panels B and C repeat the exercise of splitting CCES respondents by their party affiliation, but further divide voters into those below and above 40. Because Twitter users tend to be younger than the general population (see Section 2), we may also expect a somewhat larger effect of Twitter on vote outcomes among voters below 40.

Indeed, we find somewhat larger marginal effects for $Log(Twitter\ users)$ among younger voters. For independents, the estimate for young voters is 20% larger than for older voters ($-0.071$ compared to $-0.059$). Among moderate Republicans and

[31] In unreported regressions, we do not find that votes for Jill Stein varied with Twitter usage.

[32] This effect size is within the range of the county-level estimates presented in Table 2. Online Appendix Table A.13 shows this baseline result is robust to using only individuals with validated turnout and/or who stated that they originally intended to vote for Trump in the pre-election wave of the CCES.

Democrats, the estimated coefficients are close to zero for voters 40+ but sizeable for younger voters (although they are not statistically significant at conventional levels). Because young voters are considerably less likely to vote for Trump, this implies much larger effects of Twitter on vote outcomes for those below 40 relative to the baseline probabilities.

We present further support for Twitter having persuasion effects on moderates using county-level data in Online Appendix Table A.17. We estimate our county-level specification (equation 3) for the 2016 elections, but split counties based on how consistently they voted for either the Republican or the Democratic party. Specifically, we define "swing counties" as counties that were not consistently won by one party in the presidential elections from 2000 to 2012. Twitter usage, we find, only negatively impacts the Republican vote share in "swing" counties. We find no evidence for Republican or Democratic strongholds, where people likely have the strongest priors.

[Table 4 about here.]

## 6.3 Social Media and the 2016/2020 Elections

The findings above suggest that Twitter had an effect in 2016 and 2020, but not during previous presidential elections. We address three potential explanations for this pattern: lack of familiarity with social media (a *learning channel*), changes in social media's "slant" (a *content channel*), and Trump's role as an outsider candidate (a *political shock channel*).

The first factor could be the reach of and familiarity with social media content. In 2008, social media was a relatively new type of technology. Only a quarter of American adults used *any* social media platform and only 10% of internet users posted political commentary on social media (Pew Research Center, 2009, 2011). Figure A.1 shows that Twitter, which was founded in 2006, only had around one million users during the 2008 elections, compared to 40 million in 2012, 67 million in 2016 and 69 million in 2020 (Statista, 2019, 2020). Twitter's limited reach and novelty might have initially restricted its impact on voters.

The second possible explanation is that social media's content changed between 2008 and 2016. It is conceivable that, in line with changes in the slant of cable news (e.g., Martin and Yurukoglu, 2017), the typical content to which Twitter users are exposed has become more left-leaning over time.

A third reason is that Trump's political rise constituted a considerable shock to the U.S. political system. In this view, Twitter may not have partisan effects *per se*. Instead, the platform may have served as a conduit for spreading sentiments or information about Trump.

Two pieces of evidence we already presented above are particularly consistent with the political shock channel. First, if Twitter had little effect before 2016 because it was not widely used, its effect on vote shares should systematically increase over time. We do not find evidence supporting this idea in effects for presidential, House, and Senate elections (Figures 6 and 7). Instead, we find a discontinuous negative effect in the 2016 presidential election that persists until 2020. Second, we do not find substantial effects for the 2016 and 2020 House and Senate elections. This implies that voters in counties with higher Twitter usage chose *on the same day* to vote against Trump but not against other members of the Republican party.

**Results from Gallup Approval Ratings.** A complementary pattern emerges when we use data from the Gallup Daily Tracker, which contains approval ratings for the other major Republican presidential candidates who ran alongside Trump during the Republican primaries: Ted Cruz, Marco Rubio, and John Kasich. This analysis allows us to test whether higher Twitter usage was associated with lower approval of all Republican candidates versus Trump in particular. Table 5 shows the results of running individual-level IV probit regressions as in equation (4), where the dependent variable is an indicator variable equal to one if the respondents approved of a specific candidate. As in Table 4, we differentiate between respondents' political affiliation.[33]

Results in equation (4) suggest Twitter usage is not associated with a uniform shift away from Republican candidates. Instead, they indicate a shift towards higher approval of Clinton and lower approval of Trump, especially among independents (who are presumably more likely to be persuaded by social media content). Among Republicans and Democrats, we also find lower approval of Cruz, who is substantially more right-wing than other presidential primary candidates in recent years (FiveThirtyEight, 2015). We find no link between Twitter use and the more moderate Republican candidates, Rubio and Kasich, nor for Bernie Sanders.

Taken together, these results are consistent with Twitter turning voters against voting for Trump in particular and not against the Republican party more generally.

[33]Note that we pool people who identify as leaning Republicans or Democrats with independents, because—in contrast to the CCES data—only a few individuals in the survey are classified as "leaners."

Our results may also explain the absence of an effect in the 2008 and 2012 elections: Obama's opponents John McCain and Mitt Romney were widely considered to be moderate Republicans (e.g., more similar to Kasich than Trump or Cruz).[34]

[Table 5 about here.]

We provide further support for the hypothesis that Trump's 2016 campaign and presidency triggered opposition on Twitter by looking at the content of more than 460 million tweets sent during the 2012, 2016, and 2020 presidential campaigns. Specifically, we analyze the partisan slant of political content on Twitter by asking how "similar" tweets about the presidential candidates (Obama, Romney, Clinton, Trump, Biden) "sound" relative to tweets sent by Democratic or Republican congresspeople. We follow an approach similar to Gentzkow and Shapiro (2010) and train a L2 regularized logistic regression classifier to predict whether a tweet is more likely to approve or disapprove of the candidate, depending on its content's similarity to tweets sent by Members of Congress. As our baseline measure, we assign tweets to the most likely class. If a tweet's content has higher similarity with those of Democratic Members of Congress, it is classified as pro-Democratic, and as pro-Republican otherwise.[35]

Figure 8 plots the amount of Twitter attention directed at the Republican and Democratic presidential candidates in the 2012, 2016, and 2020 elections, as well as the tweets' estimated slant. To account for the amount of attention and popularity tweets receive, we base the graphs on the number of "likes" the tweets received. In Online Appendix Figure A.7, we confirm that the results are similar using the number of tweets.[36]

The graph on the left shows a comparable number of "likes" for tweets about both presidential candidates in the 2012 election, with slightly more "likes" for tweets mentioning Obama. In 2016, the overall volume of political Twitter content increased

[34]As previously discussed, the number of Twitter users in 2008 was relatively small, but by 2012, it was relatively close to its 2016 level (Figure A.1).

[35]See the data section and Appendix A.1.1 in the Online Appendix for more details. In unreported robustness checks, we confirm that these findings are robust using different slant cutoffs for classifying Republican and Democratic tweets, such as only classifying tweets for which the class probabilities are above 75% or 90%.

[36]Twitter users can choose to "like" each tweet they see in their "timeline." A user can only "like" a particular tweet once. "Likes" thus provide an useful metric since they capture the popularity and attention that the content received. For example, if an account sent millions of Republican-slanted tweets about Clinton, but such account had few followers and thus few users who can "like" the message, it would not meaningfully affect measures based on "likes," but could potentially do so for measures based on number of tweets.

10-fold, with four times more tweets about Trump than Clinton. In 2020, the number of election related tweets increased by more than 200% and the number of tweets mentioning Trump is nearly three times larger than tweets about Biden. Note that a "like" for a tweet about a candidate can occur for tweets that are positive or negative about the candidate, so the left panel is not informative about slant or sentiments of Twitter content.

The graph on the right shows the slant of tweets sent during the three election campaigns, specifically the share of tweets with a Republican slant. It indicates a shift in the slant of election-related Twitter content. While 65% of the tweets about Romney in the 2012 election period exhibited a Democratic slant, around 80% of the tweets about Trump appear to come from the Democratic side of the political spectrum. In contrast, the Twitter content about Clinton in 2016 shows less of a Republican slant than the content about Obama in the 2012 election. Taken together with the far larger amount of Twitter content mentioning Trump, these findings suggest Twitter content had a clearly pro-Democratic slant around the 2016 presidential election.

The picture for the 2020 election is similar. While the share of tweets about Trump with a Republican slant increased slightly (by around 30%), we also observe that the share of tweets about Biden with a Democratic slant increased compared to Hillary Clinton in 2016. This finding is in line with the idea that Biden was less "vilified" by the other side of the political aisle and hence also attracted fewer tweets with a Republican slant. The data again indicate that the overwhelming majority of tweets about Trump and Biden exhibit a Democratic slant and could therefore have a "pro-democratic" persuasion effect.

Overall, a potential interpretation of these results is as follows. Twitter users, and users of other social media platforms, are more likely to be young, well-educated, live in dense urban areas, and support the Democratic party (see discussion in Section 2). Perhaps as a result, Democratic politicians are more popular on Twitter than Republicans (Figure 1). This pro-Democrat slant did not manifest itself in a pronounced fashion before the 2016 presidential election because the Republican candidates were relatively moderate. In 2016, however, Twitter became a vehicle for spreading criticism of Trump, which partially continued into 2020. The sheer volume of this slanted content may, in turn, have persuaded voters with weaker priors—independents and perhaps more moderate Republicans—to vote against Trump in the presidential election.

[Figure 8 about here.]

# 7 Conclusion

Election officials around the globe are concerned about social media's increasing influence on voting decisions (e.g. NPR, 2020a). At the time of writing, there is a heated debate about whether platform providers should "moderate" election-related content in the U.S. (e.g. Politico, 2020). Exploiting variation based on a shock to Twitter's initial rise to popularity, our paper provides empirical evidence that social media can affect election outcomes. We find that Twitter likely increased the Democratic party's vote share in the 2016 and 2020 presidential elections. While this finding runs counter to a popular narrative that places social media at the heart of Trump's election win, it is consistent with a growing body of evidence showing that social media users are less, not more likely to vote for Trump in 2016 or hold polarized views (Boxell et al., 2017, 2018). Twitter user demographics make the platform's content partisan: people who use Twitter are 25 percentage points more likely to identify as Democrats rather than Republican and Democratic politicians are more popular on Twitter than Republican ones. Our work suggests that this environment not only reflects selection of like-minded individuals, but also affects voting decisions, particularly for people with more moderate views.

# References


Acemoglu, D., T. A. Hassan, and A. Tahoun (2017, 08). The Power of the Street: Evidence from Egypt's Arab Spring. *The Review of Financial Studies 31*(1), 1–42.

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020, March). The Welfare Effects of Social Media. *American Economic Review 110*(3), 629–76.

Allcott, H. and M. Gentzkow (2017, May). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives 31*(2), 211–36.

Andrews, I. and E. Oster (2019). A simple approximation for evaluating external validity bias. *Economics Letters 178*(C), 58–62.

Autor, D., D. Dorn, G. Hanson, K. Majlesi, et al. (2020). Importing political polarization? the electoral consequences of rising trade exposure. *American Economic Review 110*(10), 3139–83.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121–2168.

Bail, C. A., L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky (2018). Exposure to Opposing Views on Social Media Can Increase Political Polarization. *Proceedings of the National Academy of Sciences 115*(37), 9216–9221.

Bakshy, E., S. Messing, and L. A. Adamic (2015). Exposure to Ideologically Diverse News and Opinion on Facebook. *Science 348*(6239), 1130–1132.

Barberá, P. (2014). How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the US.

Beam, M. A., M. J. Hutchens, and J. D. Hmielowski (2018). Facebook News and (De)polarization: Reinforcing Spirals in the 2016 US Election. *Information, Communication & Society 21*(7), 940–958.

Bessi, A., F. Zollo, M. D. Vicario, A. Scala, G. Caldarelli, and W. Quattrociocchi (2015, 08). Trend of Narratives in the Age of Misinformation. *PLOS ONE 10*(8), 1–16.

Bond, R. M., C. J. Fariss, J. J. Jones, A. D. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler (2012). A 61-Million-Person Experiment in Social Influence and Political Mobilization. *Nature 489*(7415), 295.

Boxell, L., M. Gentzkow, and J. M. Shapiro (2017). Greater Internet Use Is Not Associated with Faster Growth in Political Polarization Among US Demographic Groups. *Proceedings of the National Academy of Sciences of the United States of America*, 201706588.

Boxell, L., M. Gentzkow, and J. M. Shapiro (2018, 07). A Note on Internet Use and the 2016 U.S. Presidential Election Outcome. *PLOS ONE 13*(7), 1–7.

Boxell, L., M. Gentzkow, and J. M. Shapiro (2019). Cross-Country Trends in Affective Polarization. *Working Paper*.

Bursztyn, L., G. Egorov, R. Enikolopov, and M. Petrova (2019, December). Social Media and Xenophobia: Evidence from Russia. Working Paper 26567, National Bureau of Economic Research.

Campante, F., R. Durante, and F. Sobbrio (2017, 12). Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation. *Journal of the European Economic Association 16*(4), 1094–1136.

Chen, E., A. Deb, and E. Ferrara (2020). # election2020: The first public twitter dataset on the 2020 us presidential election. *arXiv preprint arXiv:2010.00600*.

Chen, Y. and D. Y. Yang (2019, June). The Impact of Media Censorship: 1984 or Brave New World? *American Economic Review 109*(6), 2294–2332.

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019). Inference with Arbitrary Clustering. IZA Discussion Papers 12584, Institute of Labor Economics (IZA).

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

Del Vicario, M., A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi (2016). The Spreading of Misinformation Online. *Proceedings of the National Academy of Sciences 113*(3), 554–559.

DellaVigna, S., R. Enikolopov, V. Mironova, M. Petrova, and E. Zhuravskaya (2014, July). Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia. *American Economic Journal: Applied Economics 6*(3), 103–32.

DellaVigna, S. and E. L. Ferrara (2015). Chapter 19 - Economic and Social Impacts of the Media. In S. P. Anderson, J. Waldfogel, and D. Strömberg (Eds.), *Handbook of Media Economics*, Volume 1 of *Handbook of Media Economics*, pp. 723 – 768. North-Holland.

DellaVigna, S. and M. Gentzkow (2010). Persuasion: Empirical Evidence. *Annual Review of Economics 2*(1), 643–669.

DellaVigna, S. and E. Kaplan (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics 122*(3), 1187–1234.

Diaz, F., M. Gamon, J. M. Hofman, E. Kıcıman, and D. Rothschild (2016). Online and Social Media Data as an Imperfect Continuous Panel Survey. *PloS one 11*(1), e0145406.

Eady, G., J. Nagler, A. Guess, J. Zilinsky, and J. A. Tucker (2019). How Many People Live in Political Bubbles on Social Media? Evidence From Linked Survey and Twitter Data. *SAGE Open 9*(1), 2158244019832705.

eMarketer (2019). US Time Spent with Social Media 2019, By Debra Aho Williamson.

Enikolopov, R., A. Makarin, and M. Petrova (2020). Social Media and Protest Participation: Evidence from Russia. *Econometrica 88*(4), 1479–1514.

Enikolopov, R. and M. Petrova (2015). Chapter 17 - Media Capture: Empirical Evidence. In S. P. Anderson, J. Waldfogel, and D. Strömberg (Eds.), *Handbook of Media Economics*, Volume 1 of *Handbook of Media Economics*, pp. 687 – 700. North-Holland.

Enikolopov, R., M. Petrova, and E. Zhuravskaya (2011). Media and Political Persuasion: Evidence from Russia. *The American Economic Review 101*(7), 3253–3285.

Falck, O., R. Gold, and S. Heblich (2014, July). E-lections: Voting Behavior and the Internet. *American Economic Review 104*(7), 2238–65.

FiveThirtyEight (2015). Let's Be Serious About Ted Cruz From The Start: He's Too Extreme And Too Disliked To Win.

Flaxman, S., S. Goel, and J. M. Rao (2016, 03). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly 80*(S1), 298–320.

Gavazza, A., M. Nardotto, and T. Valletti (2019). Internet and Politics: Evidence From UK Local Elections and Local Government Policies. *The Review of Economic Studies 86*(5), 2092–2135.

Gentzkow, M. (2006). Television and Voter Turnout. *The Quarterly Journal of Economics 121*(3), 931–972.

Gentzkow, M. and J. M. Shapiro (2010). What Drives Media Slant? Evidence from US Daily Newspapers. *Econometrica 78*(1), 35–71.

Gentzkow, M. and J. M. Shapiro (2011). Ideological Segregation Online and Offline. *The Quarterly Journal of Economics 126*(4), 1799–1839.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2011, December). The Effect of Newspaper Entry and Exit on Electoral Politics. *American Economic Review 101*(7), 2980–3018.

Gerber, A. S., J. G. Gimpel, D. P. Green, and D. R. Shaw (2011). How Large and Long-lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment. *The American Political Science Review 105*(1), 135–150.

Gerber, A. S. and D. P. Green (2000). The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment. *The American Political Science Review 94*(3), 653–663.

Gerber, A. S., D. Karlan, and D. Bergan (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics 1*(2), 35–52.

Guess, A. (2018). (Almost) Everything in Moderation: New Evidence in Americans' Online Media Diets.

Guriev, S., N. Melnikov, and E. Zhuravskaya (2019, September). 3G Internet and Confidence in Government. CEPR Discussion Papers 14022, C.E.P.R. Discussion Papers.

Halberstam, Y. and B. Knight (2016). Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter. *Journal of Public Economics 143*, 73 – 88.

Hargittai, E. (2015). Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites. *The Annals of the American Academy of Political and Social Science 659*, 63–76.

Howard, P. N., A. Duffy, D. Freelon, M. Hussain, W. Mari, and M. Maziad (2011). Opening Closed Regimes: What Was the Role of Social Media During the Arab Spring? *Working Paper*.

Huber, G. A. and K. Arceneaux (2007). Identifying the Persuasive Effects of Presidential Advertising. *American Journal of Political Science 51*(4), 957–977.

Jones, J. J., R. M. Bond, E. Bakshy, D. Eckles, and J. H. Fowler (2017, 04). Social Influence and Political Mobilization: Further Evidence From a Randomized Experiment in the 2012 U.S. Presidential Election. *PLOS ONE 12*(4), 1–9.

Kinder-Kurlanda, K., K. Weller, W. Zenk-Möltgen, J. Pfeffer, and F. Morstatter (2017). Archiving Information from Geotagged Tweets to Promote Reproducibility and Comparability in Social Media Research. *Big Data & Society 4*(2), 2053951717736336.

Larcinese, V. and L. Miner (2017, June). The Political Impact of the Internet on US Presidential Elections. STICERD - Economic Organisation and Public Policy Discussion Papers Series 63, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Lelkes, Y., G. Sood, and S. Iyengar (2017). The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect. *American Journal of Political Science 61*(1), 5–20.

Levy, R. (2019). Social Media, News Consumption, and Polarization: Evidence from a Field Experiment. *Working Paper*.

Littman, J., L. Wrubel, and D. Kerchner (2016). 2016 United States Presidential Election Tweet Ids.

Manacorda, M. and A. Tesei (2016, May). Liberation Technology: Mobile Phones and Political Mobilization in Africa. CEPR Discussion Papers 11278, C.E.P.R. Discussion Papers.

Martin, G. J. and A. Yurukoglu (2017, September). Bias in Cable News: Persuasion and Polarization. *American Economic Review 107*(9), 2565–2599.

Mosquera, R., M. Odunowo, T. McNamara, X. Guo, and R. Petrie (2020). The Economic Effects of Facebook. *Experimental Economics 23*(2), 575–602.

Müller, K. and C. Schwarz (2019). From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment. *Working Paper*.

Müller, K. and C. Schwarz (2020). Fanning the Flames of Hate: Social Media and Hate Crime. *Journal of the European Economic Association*.

Napoli, P. M. (2014). Measuring Media Impact. *The Norman Lear Center. http://www.learcenter. org/pdf/measuringmedia.pdf*.

Nelson, J. L. and J. G. Webster (2017). The Myth of Partisan Selective Exposure: A Portrait of the Online Political News Audience. *Social Media + Society 3*(3), 2056305117729314.

New York Times (2017). The Fake Americans Russia Created to Influence the Election, By Scott Shane.

NPR (2020a). Facebook Clamps Down On Posts, Ads That Could Undermine U.S. Presidential Election.

NPR (2020b). Facebook Keeps Data Secret, Letting Conservative Bias Claims Persist.

NPR (2020c). President-Elect Joe Biden Hits 80 Million Votes In Year Of Record Turnout.

Pariser, E. (2011). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin Books Limited.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research 12*, 2825–2830.

Pew Research Center (2009). The Internet's Role in Campaign 2008. Technical report.

Pew Research Center (2011). Social Networking Sites and Our Lives. Technical report.

Pew Research Center (2014). Political Polarization & Media Habits. Technical report.

Pew Research Center (2015). Americans' Internet Access: 2000-2015. Technical report.

Pew Research Center (2018). For Most Trump Voters, 'Very Warm' Feelings for Him Endured. Technical report.

Pew Research Center (2019a). About One-In-Five Adult Twitter Users in the U.S. Follow Trump. Technical report.

Pew Research Center (2019b). National Politics on Twitter: Small Share of U.S. Adults Produce Majority of Tweets. Technical report.

Pew Research Center (2019c). Share of U.S. Adults Using Social Media, Including Facebook, Is Mostly Unchanged Since 2018. Technical report.

Pew Research Center (2019d). Sizing Up Twitter Users. Technical report.

Pew Research Center (2020). Democrats On Twitter More Liberal, Less Focused On Compromise Than Those Not On The Platform. Technical report.

Politico (2020). Biden Campaign Lashes Out at New York Post.

Rotesi, T. (2019). The Impact of Twitter on Political Participation. Technical report.

Schmidt, A. L., F. Zollo, M. Del Vicario, A. Bessi, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi (2017). Anatomy of News Consumption on Facebook. *Proceedings of the National Academy of Sciences 114*(12), 3035–3039.

Spenkuch, J. L. and D. Toniatti (2018, 05). Political Advertising and Election Results. *The Quarterly Journal of Economics 133*(4), 1981–2036.

Statista (2019). Twitter: Number of Monthly Active U.S. Users 2010-2019.

Statista (2020). Leading countries based on number of Twitter users as of October 2020.

Strömberg, D. (2015). Media and Politics. *Annual Review of Economics* 7(1), 173–205.

Sunstein, C. R. (2017). *# Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

The Guardian (2018). Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach, by Carole Cadwalladr and Emma Graham-Harrison.

The Hill (2016). Stop Blaming Facebook for Trump's Election Win, By Keith N. Hampton and Eszter Hargittai.

The New Yorker (2016). Obama Reckons with a Trump Presidency, By David Remnick.

Twitter (2010). Measuring Tweets, by Kevin Weil.

UK Parliament (2019). Disinformation and 'fake news': Final Report, By Digital, Culture, Media and Sport Committee.

Venture Beat (2008). Hitwise: Twitter Traffic Is, in Fact, Going up but Still Not Big, by Eric Eldon.

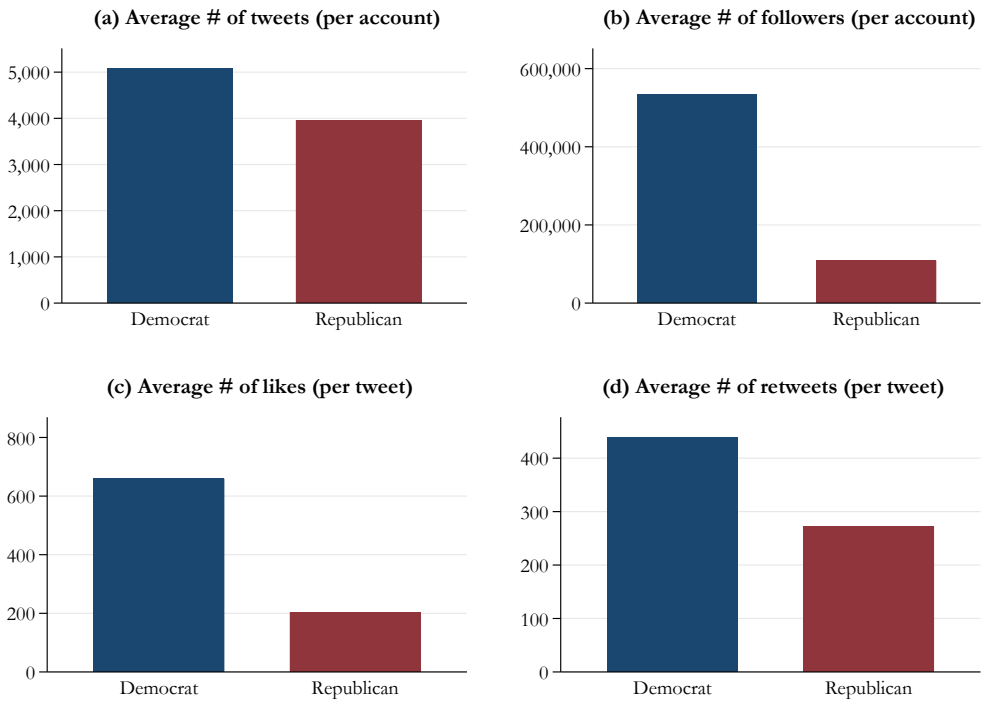
Wall Street Journal (2020). Twitter's Partisan Censors.

Weber, I., V. R. K. Garimella, and A. Batayneh (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, New York, NY, USA, pp. 290–297. Association for Computing Machinery.

Wired (2016). Here's How Facebook Actually Won Trump the Presidency, By Issie Lapowsky.

Zhuravskaya, E., M. Petrova, and R. Enikolopov (2020). Political Effects of the Internet and Social Media. Technical report.

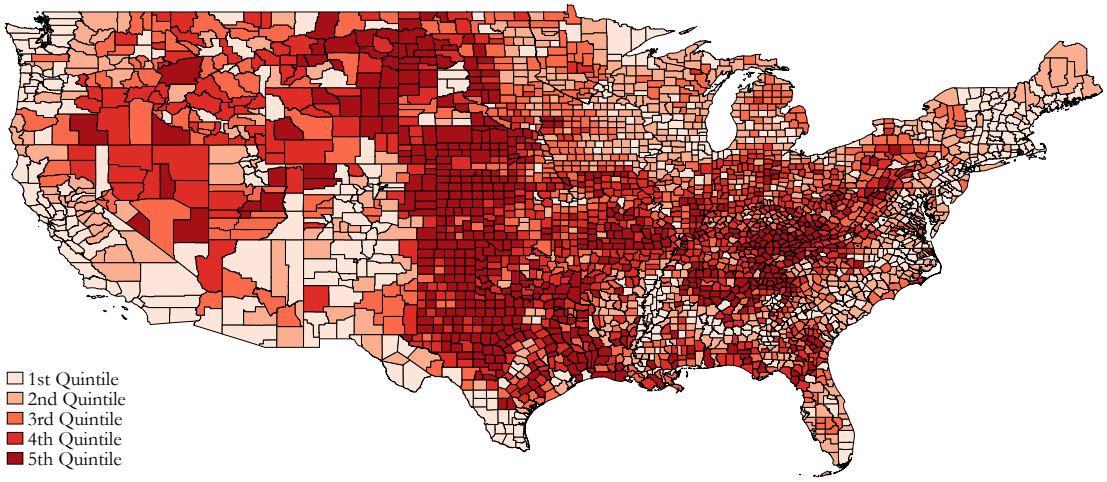
## Figure 1: Twitter Reach by Party



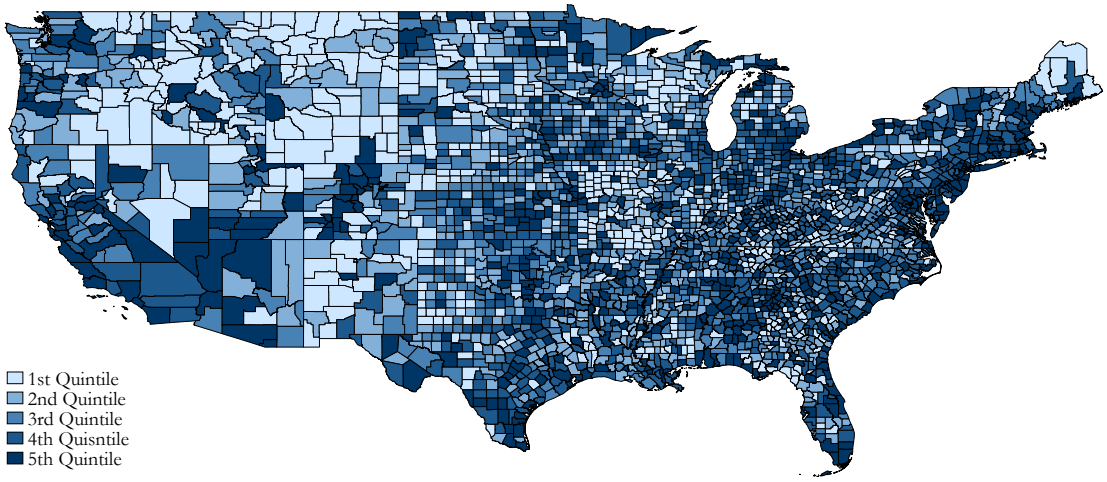

*Notes:* This figure plots data on the Twitter reach of Members of Congress. The sample includes all 901 senators and House representatives who were in office between 2007 and 2019 for whom we could identify a Twitter account. For each account, we present the average number of tweets and followers, and the average number of "likes" and retweets of their tweets. Online Appendix Figure A.2 replicates the figure using medians instead of averages.

**Figure 2: Republican Vote Share and Twitter Usage Across U.S. Counties**

**(a) Republican Vote Share 2016 Presidential Election**



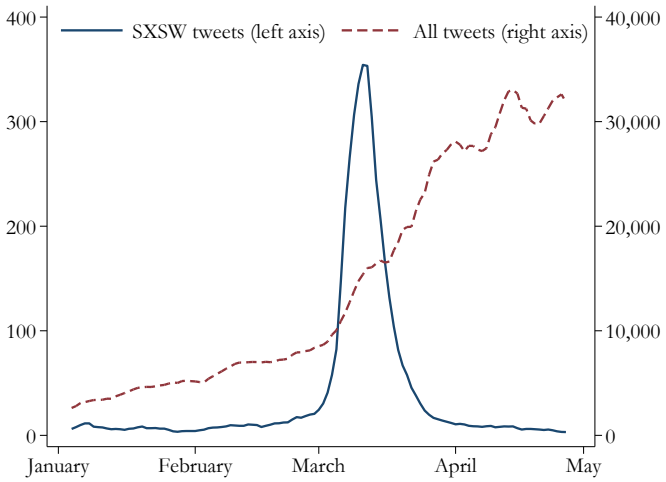

**(b) Twitter Usage per Capita**



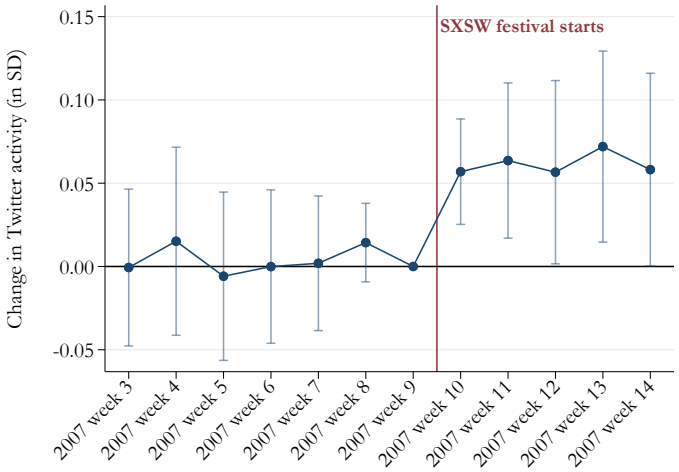

*Notes:* These maps plot the geographical distribution of the Republican vote share in the 2016 presidential election and a proxy for social media usage based on data from Twitter. Panel (a) plots quintiles of the Republican party's vote share in the 2016 presidential election. Panel (b) plots quintiles of the number of Twitter users per capita.

**Figure 3: South by Southwest (SXSW) 2007 and the Spread of Twitter**

**(a) Twitter Activity Around SXSW 2007**

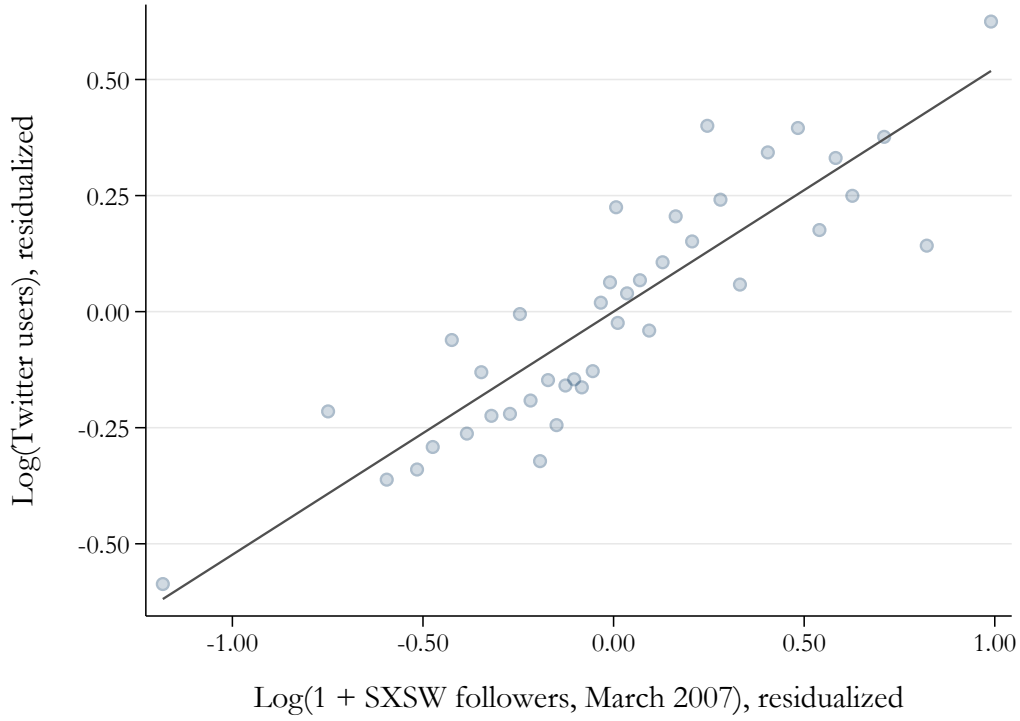



**(b) SXSW and Local Twitter Adoption**



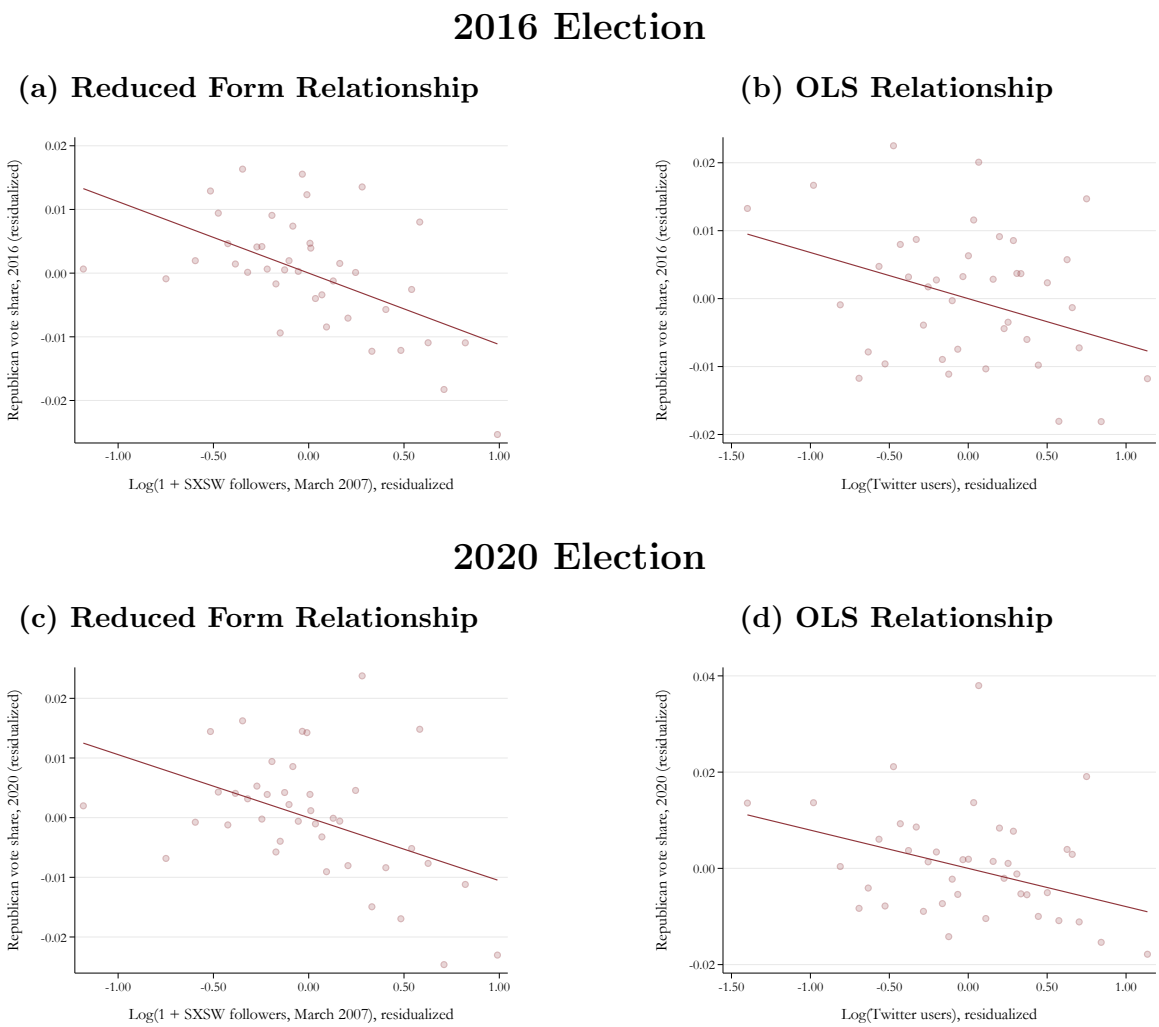

*Notes:* Panel (a) plots the total number of tweets and the number of tweets containing the term "SXSW" over time, smoothed using a 7-day moving average. Panel (b) plots the estimates of $\beta_\tau$ from the panel event study regression $tweets_{ct} = \sum_\tau \beta_\tau SXSW_c^{March2007} \times 1(t = \tau) + \sum_\tau \delta_\tau SXSW_c^{Pre} \times 1(t = \tau) + \theta_c + \gamma_t + \varepsilon_{ct}$ where $tweets_{ct}$ is the log of (one plus) the number of tweets in county $c$ on week $t$, $SXSW_c^{March2007}$ is the logarithm of (one plus) the number of SXSW followers in county $c$ that joined Twitter on March 2007 and $SXSW_c^{Pre}$ is a similarly defined variable for followers that joined Twitter before March 2007. We standardize the variables to have a mean of zero and standard deviation of one. The whiskers represent 95% confidence intervals based on standard errors clustered by state.

**Figure 4: First Stage – South by Southwest (SXSW) and Twitter Usage**




*Notes:* This figure presents a binned scatter plot of the relationship between Twitter users in 2014-2015 and the number of SXSW followers who joined Twitter in March 2007. Variables are residualized by partialling out SXSW followers who joined before March 2007, population deciles, Census region fixed effects, as well as geographical, demographic, socioeconomic, China shock, and 1996 election control variables (see Online Appendix for control variable definitions). The figure is constructed by dividing the x-axis variable into 40 equal-sized bins and plotting the average values of both variables in each bin. The fitted line is estimated using the unbinned data.

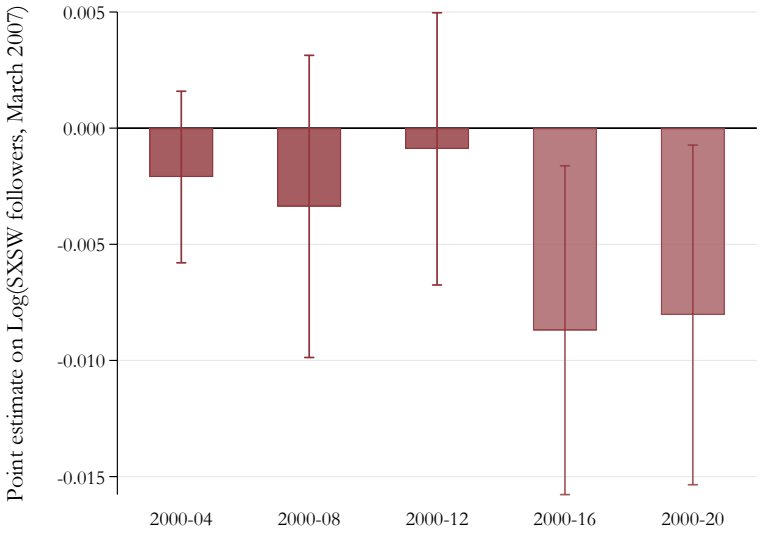
**Figure 5: South by Southwest (SXSW), Twitter, and the 2016/2020 Republican Vote Share**




*Notes:* Panel (a) presents a binned scatter plot of the relationship between the Republican vote share in the 2016 presidential election and the number of SXSW followers who joined Twitter in March 2007. Variables are residualized with respect to SXSW followers who joined before March 2007, population deciles, Census region fixed effects, as well as geographical, demographic, socioeconomic, China shock, and 1996 election control variables. The figure is constructed by dividing the x-axis variable into 40 equal-sized bins and plotting the average values of both variables in each bin. The fitted line is estimated using the unbinned data. Panel (b) replicates the exercise using Twitter users in 2014-2015 as the x-axis variable. Panel (c) and (d) repeats the analysis for the 2020 election

## Figure 6: Twitter and Presidential Elections – Reduced Form

### (a) Changes in Republican Vote Share



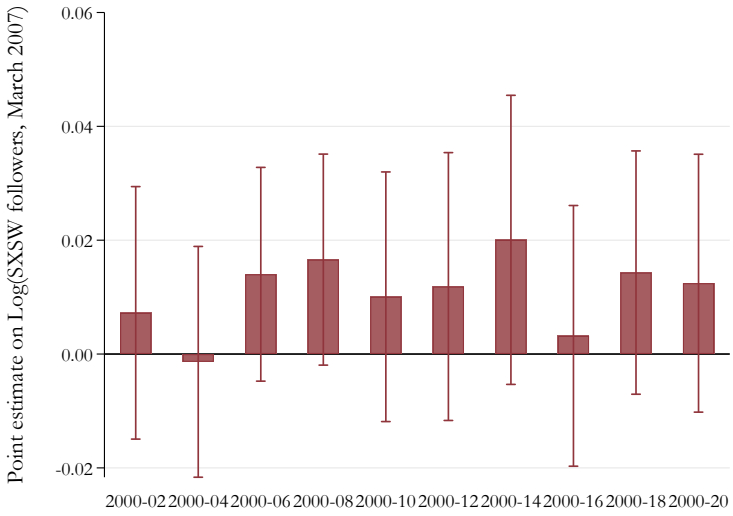

### (b) Change in Voter Turnout



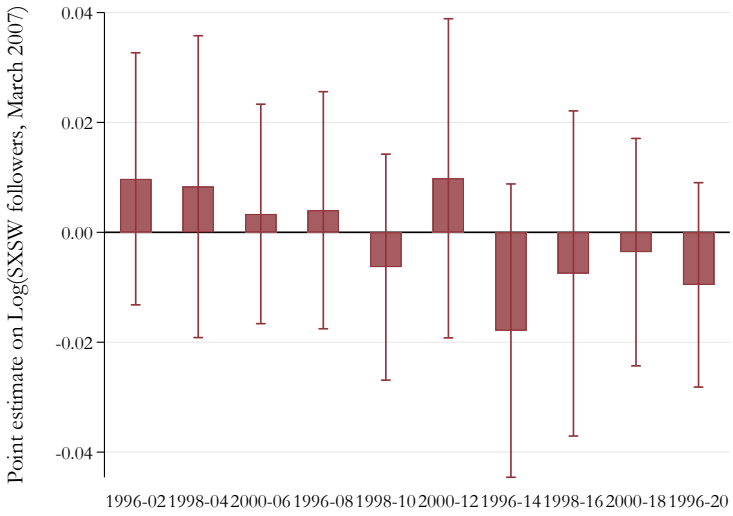

*Notes:* These figures plot reduced form estimates $\hat{\beta}'$ from county-level regressions as in equation (2). They measure the effect of *Log(1 + SXSW followers, March 2007)*, while controlling for *Log(1 + SXSW followers, Pre)*, on changes in the Republican vote share in presidential elections relative to the year 2000 in Panel (a), and changes in the ratio of voter turnout to voting-age population relative to 2000 in Panel (b). All regressions control for population deciles and Census region fixed effects, and the full set of controls (as in column 5 of Table 2). Regressions are weighted by turnout in the 2000 presidential election. Whiskers represent 95% confidence intervals based on standard errors clustered by state.

## Figure 7: Twitter and Congressional Election Results – Reduced Form

### (a) House Elections

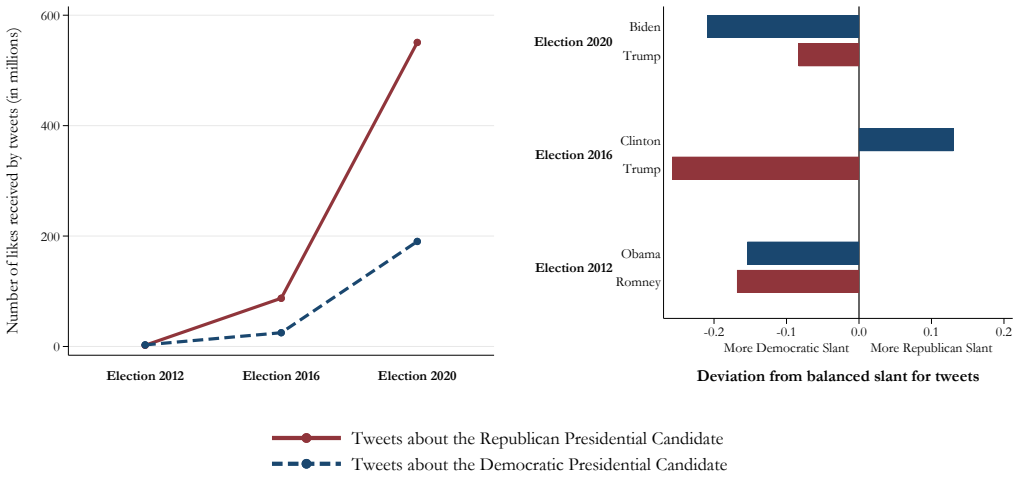



### (b) Senate Elections




*Notes:* These figures plot reduced form estimates $\hat{\beta}'$ from county-level regressions as in equation (2). They measure the reduced form effect of *Log(1 + SXSW followers, March 2007)*, while controlling for *Log(1 + SXSW followers, Pre)*, on the Republican vote share in House and Senate elections since 2000. For House elections in Panel (a), the dependent variable is the change in the Republican vote share since 2000. For Senate elections in Panel (b), the dependent variable is the change in the Republican vote share from six, twelve, or eighteen years ago (to accommodate senators' 6-year terms). All regressions control for population deciles and Census region fixed effects and the full set of controls (as in column 5 of Table 2). Regressions are weighted by turnout in the 2000 presidential election. The whiskers represent 95% confidence intervals based on standard errors clustered by state.

**Figure 8: Twitter's Partisan Slant in the 2012, 2016 and 2020 Presidential Elections**




*Notes:* These figures present the number of "likes" received by tweets that contain the last name of the candidates in the 2012, 2016 and 2020 presidential elections, depending on whether the tweet was classified as having a Republican (instead of Democratic) slant. We classify the slant of a tweet based on similarity in the language to that of a congressional Republican or Democrat, using a L2 regularized logistic regression classifier based on tweets sent by Members of Congress, where optimal normalization strength is chosen using 10-fold cross-validation. See Online Appendix Appendix A.1.1 for details.

## Table 1: South by Southwest 2007 and Twitter Usage

| | *Dep. var.: Log(Twitter users)* | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Log(SXSW followers, March 2007) | 0.726*** | 0.683*** | 0.563*** | 0.524*** | 0.523*** |
| | (0.087) | (0.079) | (0.055) | (0.048) | (0.048) |
| Log(SXSW followers, Pre) | 0.104 | 0.110 | 0.059 | 0.059 | 0.058 |
| | (0.101) | (0.076) | (0.098) | (0.082) | (0.082) |
| Population deciles | Yes | Yes | Yes | Yes | Yes |
| Census region FE | Yes | Yes | Yes | Yes | Yes |
| Geographical controls | | Yes | Yes | Yes | Yes |
| Demographic controls | | | Yes | Yes | Yes |
| Socioeconomic controls | | | Yes | Yes | Yes |
| China shock controls | | | | Yes | Yes |
| 1996 election control | | | | | Yes |
| Observations | 3,065 | 3,065 | 3,064 | 3,064 | 3,064 |
| $R^2$ | 0.92 | 0.93 | 0.95 | 0.95 | 0.95 |
| Mean of DV | 8.22 | 8.22 | 8.22 | 8.22 | 8.22 |
| p-value: March 2007 = Pre | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

*Notes:* This table presents county-level regressions where the dependent variable is the number of Twitter users (in natural logarithm). *Log(SXSW followers, March 2007)* is the number of Twitter users (in logs, with 1 added inside) who joined in March 2007 and follow South by Southwest (SXSW). *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006, defined similarly. The bottom row reports $p$-values from F-tests for the equality of these coefficients. Regressions include the indicated control variables (see the Online Appendix for their descriptions). Observations are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table 2: Twitter and the 2016/2020 Republican Vote Share

| | *Dep. var.: Republican vote share in 2016* | | | | | *Dep. var.: Republican vote share in 2020* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| **Panel A: OLS** | | | | | | | | | | |
| Log(Twitter users) | -0.065*** | -0.067*** | -0.013*** | -0.011*** | -0.007** | -0.058*** | -0.064*** | -0.012*** | -0.011*** | -0.008*** |
| | (0.009) | (0.008) | (0.004) | (0.003) | (0.003) | (0.009) | (0.009) | (0.004) | (0.003) | (0.003) |
| **Panel B: Reduced form** | | | | | | | | | | |
| Log(SXSW followers, March 2007) | -0.053*** | -0.058*** | -0.019*** | -0.014*** | -0.011*** | -0.046*** | -0.055*** | -0.017*** | -0.013*** | -0.011** |
| | (0.011) | (0.012) | (0.005) | (0.004) | (0.004) | (0.009) | (0.010) | (0.006) | (0.005) | (0.005) |
| Log(SXSW followers, Pre) | -0.021 | -0.003 | -0.000 | -0.002 | 0.001 | -0.022 | -0.005 | -0.002 | -0.004 | -0.001 |
| | (0.016) | (0.013) | (0.006) | (0.006) | (0.004) | (0.016) | (0.013) | (0.007) | (0.007) | (0.005) |
| **Panel C: 2SLS** | | | | | | | | | | |
| Log(Twitter users) | -0.072*** | -0.085*** | -0.034*** | -0.027*** | -0.021*** | -0.064*** | -0.080*** | -0.031** | -0.025*** | -0.020** |
| | (0.016) | (0.018) | (0.010) | (0.008) | (0.008) | (0.015) | (0.017) | (0.011) | (0.009) | (0.009) |
| Log(SXSW followers, Pre) | -0.014 | 0.007 | 0.002 | -0.001 | 0.002 | -0.015 | 0.004 | -0.000 | -0.002 | 0.000 |
| | (0.020) | (0.016) | (0.007) | (0.006) | (0.005) | (0.020) | (0.015) | (0.008) | (0.007) | (0.006) |
| Population deciles | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Census region FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographical controls | | Yes | Yes | Yes | Yes | | Yes | Yes | Yes | Yes |
| Demographic controls | | | Yes | Yes | Yes | | | Yes | Yes | Yes |
| Socioeconomic controls | | | Yes | Yes | Yes | | | Yes | Yes | Yes |
| China shock controls | | | | Yes | Yes | | | | Yes | Yes |
| 1996 election control | | | | | Yes | | | | | Yes |
| Observations | 3,065 | 3,065 | 3,064 | 3,064 | 3,064 | 3,065 | 3,065 | 3,064 | 3,064 | 3,064 |
| Mean of DV | 0.46 | 0.46 | 0.46 | 0.46 | 0.46 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 |
| Robust F-stat. | 69.57 | 74.65 | 106.08 | 118.21 | 121.18 | 69.57 | 74.65 | 106.08 | 118.21 | 121.18 |

*Notes:* This table presents county-level regressions where the dependent variable is the Republican vote share in the 2016 or 2020 presidential election. *Log(SXSW followers, March 2007)* is the number of Twitter users (in logs, with 1 added inside) who joined in March 2007 and follow South by Southwest (SXSW). *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006, defined similarly. *Twitter users* are the number of users in 2014-2015. Regressions include the indicated control variables (see the Online Appendix for their descriptions). The first-stage regressions for 2SLS results (Panel C) are presented in Table 1, with the F-stat for the excluded instrument in the bottom row. Observations are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table 3: Twitter and the Republican Vote Share, 2004-2020

| | *Dep. var.: Republican vote share in...* | | | | |
|---|---|---|---|---|---|
| | 2004 | 2008 | 2012 | 2016 | 2020 |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: OLS** | | | | | |
| Log(Twitter users) | -0.003 | 0.000 | 0.001 | -0.007** | -0.008*** |
| | (0.002) | (0.003) | (0.003) | (0.003) | (0.003) |
| **Panel B: Reduced form** | | | | | |
| Log(SXSW followers, March 2007) | -0.005 | -0.006 | -0.003 | -0.011*** | -0.011** |
| | (0.003) | (0.004) | (0.004) | (0.004) | (0.005) |
| Log(SXSW followers, Pre) | 0.001 | -0.000 | -0.002 | 0.001 | -0.001 |
| | (0.003) | (0.006) | (0.005) | (0.004) | (0.005) |
| **Panel B: 2SLS** | | | | | |
| Log(Twitter users) | -0.009 | -0.011 | -0.007 | -0.021*** | -0.020** |
| | (0.006) | (0.009) | (0.008) | (0.008) | (0.009) |
| Log(SXSW followers, Pre) | 0.001 | 0.000 | -0.001 | 0.002 | 0.000 |
| | (0.004) | (0.006) | (0.005) | (0.005) | (0.006) |
| Observations | 3,064 | 3,064 | 3,064 | 3,064 | 3,064 |
| Mean of DV | 0.51 | 0.46 | 0.47 | 0.46 | 0.47 |
| Robust F-stat. | 121.18 | 121.18 | 121.18 | 121.18 | 121.18 |

*Notes:* This table presents county-level regressions where the dependent variable is the Republican vote share in presidential elections. *Log(SXSW followers, March 2007)* is the number of Twitter users (in logs, with 1 added inside) who joined in March 2007 and follow South by Southwest (SXSW). *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006, defined similarly. *Twitter users* are the number of users in 2014-2015. All regressions control for population deciles, Census region fixed effects, and the full set of controls (as in column 5 of Table 2). The first-stage regression for 2SLS results (Panel C) are presented in column (5) of Table 1, with the F-stat for the excluded instrument in the bottom row. Observations are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table 4: Twitter and Individuals' Vote Decisions

| | *Dep. var.: Voted for Trump* | | | | | |
|---|---|---|---|---|---|---|
| | Full Sample | Strong Dem. | Mod. Dem. | Indep. | Mod. Rep. | Strong Rep. |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Full Sample** | | | | | | |
| Log(Twitter users) | -0.129*** | 0.034 | -0.062 | -0.186*** | -0.073* | 0.029 |
| | (0.048) | (0.061) | (0.081) | (0.069) | (0.044) | (0.065) |
| Observations | 94,523 | 27,572 | 20,447 | 9,142 | 18,863 | 17,304 |
| Mean of DV | 0.491 | 0.027 | 0.114 | 0.627 | 0.918 | 0.981 |
| *Marginal effect* | [-0.047] | [0.002] | [-0.011] | [-0.064] | [-0.010] | [0.001] |
| **Panel B: Age Group 18-39** | | | | | | |
| Log(Twitter users) | -0.146*** | 0.080 | -0.171 | -0.198** | -0.118 | 0.115 |
| | (0.047) | (0.103) | (0.119) | (0.090) | (0.076) | (0.111) |
| Observations | 25,177 | 8,420 | 7,317 | 2,050 | 3,732 | 3,042 |
| Mean of DV | 0.376 | 0.043 | 0.089 | 0.466 | 0.859 | 0.956 |
| *Marginal effect* | [-0.051] | [0.006] | [-0.026] | [-0.071] | [-0.024] | [0.009] |
| **Panel C: Age Group 40+** | | | | | | |
| Log(Twitter users) | -0.122** | -0.028 | 0.013 | -0.176** | -0.055 | -0.035 |
| | (0.052) | (0.069) | (0.079) | (0.074) | (0.050) | (0.068) |
| Observations | 69,346 | 19,102 | 13,130 | 7,091 | 15,108 | 14,211 |
| Mean of DV | 0.535 | 0.020 | 0.128 | 0.679 | 0.934 | 0.987 |
| *Marginal effect* | [-0.045] | [-0.001] | [0.003] | [-0.059] | [-0.007] | [-0.001] |

*Notes:* This table presents results estimated using IV probit models, as in equation (4). The dependent variable is a dummy for individuals in the CCES who voted for Trump in 2016. *Log(Twitter users)* is instrumented using the number of SXSW followers that joined Twitter in March 2007. All regressions control for family income, gender, education levels, marital status, news interest, and age, as well as county-level population deciles and Census region fixed effects. Regressions are weighted by survey weights. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

**Table 5: Twitter and Candidate Approval during the 2016 Primaries**

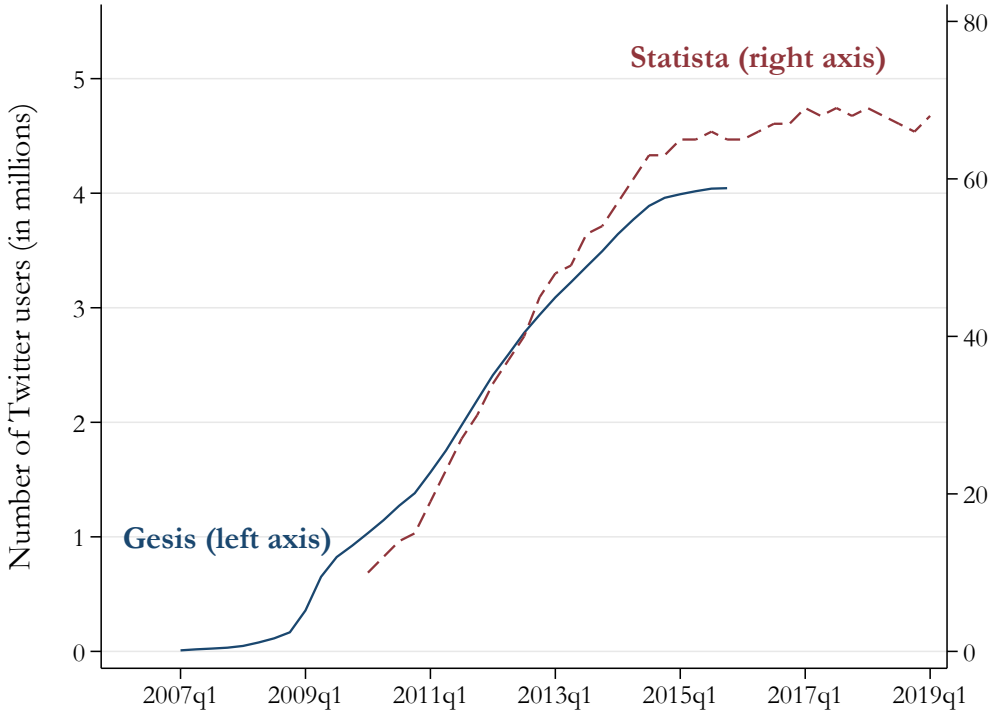
| | *Dep. var.: Approved of candidate during primaries* | | | | | |
|---|---|---|---|---|---|---|
| | Trump | Cruz | Rubio | Kasich | Sanders | Clinton |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Republicans** | | | | | | |
| Log(Twitter users) | -0.108*** | -0.086** | -0.051 | 0.018 | 0.031 | 0.148*** |
| | (0.030) | (0.035) | (0.060) | (0.050) | (0.039) | (0.041) |
| *Marginal effect* | [-0.038] | [-0.029] | [-0.014] | [0.006] | [0.009] | [0.022] |
| Observations | 19,974 | 11,959 | 8,344 | 8,995 | 16,099 | 20,983 |
| Mean of DV | 0.647 | 0.698 | 0.779 | 0.665 | 0.238 | 0.092 |
| **Panel B: Independents and Leaners** | | | | | | |
| Log(Twitter users) | -0.065** | -0.006 | -0.015 | 0.050 | 0.059 | 0.154*** |
| | (0.028) | (0.035) | (0.043) | (0.042) | (0.043) | (0.036) |
| *Marginal effect* | [-0.021] | [-0.002] | [-0.006] | [0.019] | [0.021] | [0.054] |
| Observations | 22,852 | 12,135 | 8,080 | 8,280 | 17,356 | 23,813 |
| Mean of DV | 0.329 | 0.392 | 0.516 | 0.581 | 0.595 | 0.380 |
| **Panel C: Democrats** | | | | | | |
| Log(Twitter users) | -0.052 | -0.116** | -0.036 | 0.076 | 0.004 | 0.081** |
| | (0.051) | (0.054) | (0.056) | (0.051) | (0.050) | (0.038) |
| *Marginal effect* | [-0.009] | [-0.030] | [-0.012] | [0.029] | [0.001] | [0.021] |
| Observations | 20,866 | 11,098 | 7,460 | 7,547 | 16,059 | 21,454 |
| Mean of DV | 0.107 | 0.195 | 0.271 | 0.502 | 0.808 | 0.807 |

This table presents results estimated using IV probit models, as in equation (4). The dependent variable is a dummy for individuals who approved the respective presidential candidate during the presidential primaries in 2015 and 2016. We restrict the sample to the period before Trump became the presumptive nominee in June 2016. *Log(Twitter users)* is instrumented using the number of SXSW followers that joined Twitter in March 2007. All regressions control for income, gender, education, and marital status, as well as county-level population deciles and Census region fixed effects. Regressions are weighted by survey weights. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

# A Online Appendix

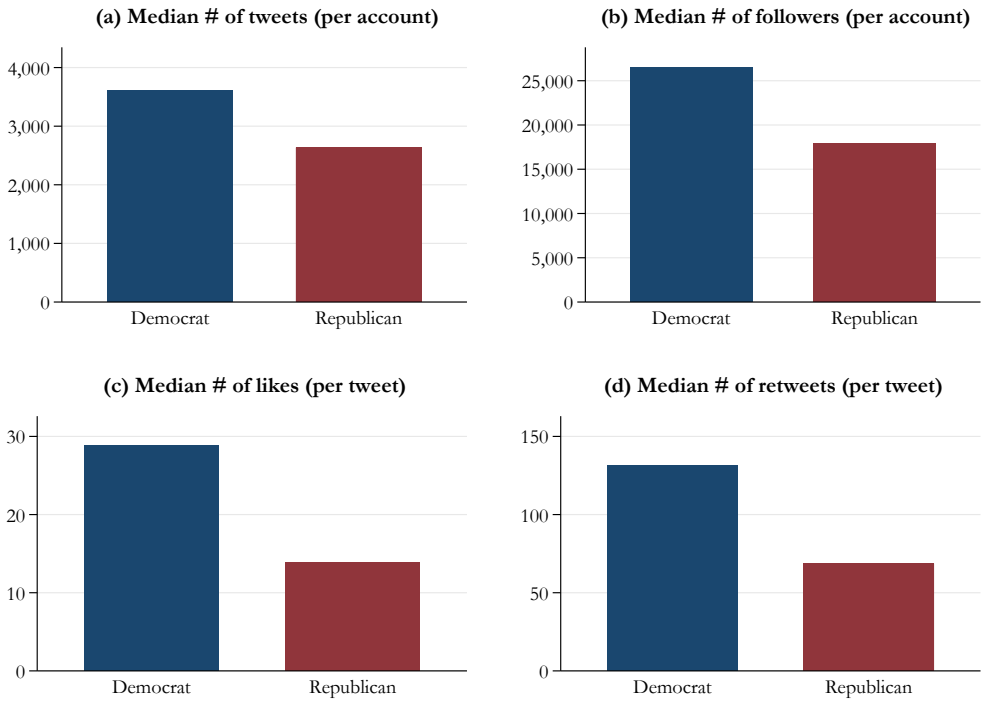
## A.1. Appendix 1: Additional Details on Data

**Figure A.1: Twitter Usage over Time**



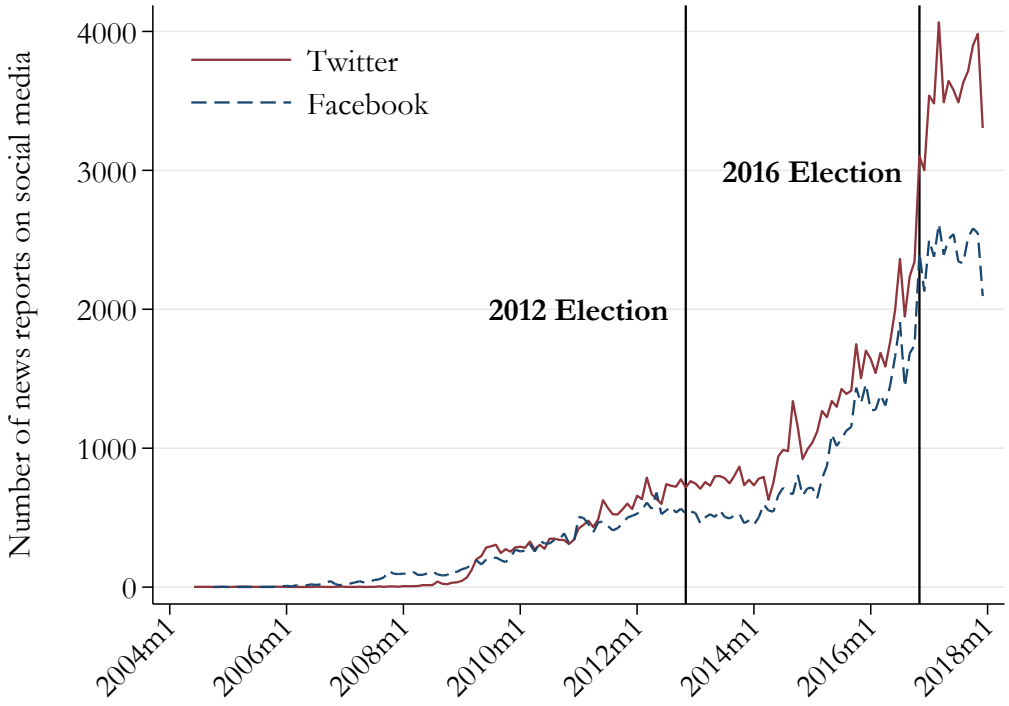

*Notes:* This graph plots the number of Twitter users in the Gesis data (Kinder-Kurlanda et al., 2017) and data on active monthly users reported from Statista based on Twitter's reports.

## Figure A.2: Twitter Reach by Party (Medians)




*Notes:* This figure plots data on the Twitter reach of Members of Congress. The sample includes all 901 senators and House representatives who were in office between 2007 and 2019 for whom we could identify a Twitter account. For each account, we present the median number of tweets and followers, and the median number of "likes" and retweets of their tweets.

## Figure A.3: News Reports About Social Media




*Notes:* This graph plots the number of times the terms "Twitter" and "Facebook" are mentioned in USA Today, The Washington Post, The New York Post, and The New York Times based on data from Nexis.

## Table A.1: List of Congresspeople with the Most Twitter Followers

| Rank | Name | Twitter Handle | Party | Tweets | Followers | Likes | Retweets |
|---|---|---|---|---|---|---|---|
| 1. | Barack Obama | barackobama | D | 15,754 | 114,369,395 | 38,496 | 8,475 |
| 2. | Hillary Rodham Clinton | hillaryclinton | D | 11,491 | 27,026,749 | 22,413 | 6,780 |
| 3. | Bernard Sanders | BernieSanders | D | 17,667 | 11,436,732 | 19,324 | 3,625 |
| 4. | Alexandria Ocasio-Cortez | AOC | D | 10,087 | 6,596,922 | 30,668 | 8,634 |
| 5. | Elizabeth Warren | SenWarren | D | 5,345 | 5,743,064 | 5,845 | 1,863 |
| 6. | Joseph R. Biden Jr. | joebiden | D | 4,228 | 4,507,744 | 11,789 | 2,472 |
| 7. | Cory A. Booker | CoryBooker | D | 65,242 | 4,468,595 | 2,802 | 798 |
| 8. | Mike Pence | mike_pence | R | 8,069 | 4,397,219 | 2,768 | 4,176 |
| 9. | Nancy Pelosi | SpeakerPelosi | D | 10,154 | 4,149,922 | 9,022 | 2,693 |
| 10. | Marco Rubio | MarcoRubio | R | 12,789 | 4,098,788 | 2,642 | 1,608 |
| 11. | Paul D. Ryan | SpeakerRyan | R | 14,755 | 3,646,397 | 781 | 260 |
| 12. | Kamala D. Harris | KamalaHarris | D | 13,119 | 3,476,952 | 10,366 | 2,229 |
| 13. | John F. Kerry | johnkerry | D | 2,608 | 3,360,092 | 773 | 328 |
| 14. | John McCain | SenJohnMcCain | R | 14,409 | 3,075,281 | 2,551 | 721 |
| 15. | Rand Paul | randpaul | R | 13,642 | 2,700,813 | 3,926 | 1,491 |
| 16. | Charles E. Schumer | SenSchumer | D | 17,004 | 2,145,147 | 6,457 | 2,171 |
| 17. | Adam B. Schiff | RepAdamSchiff | D | 5,623 | 2,087,003 | 19,009 | 6,849 |
| 18. | Ilhan Omar | IlhanMN | D | 14,678 | 1,947,411 | 6,644 | 2,649 |
| 19. | Lindsey Graham | LindseyGrahamSC | R | 10,872 | 1,535,623 | 5,698 | 1,616 |
| 20. | Mike Pompeo | SecPompeo | R | 2,148 | 1,503,151 | 4,936 | 1,711 |

*Notes:* This table lists the 20 Members of Congress (who were in office during 2007-2019) with the most Twitter followers, at time of collection. "likes" and Retweets refer to the average number of interactions the congresspeople receive for their average tweet.

## Table A.2: Summary Statistics (County-Level)

| | Mean | Std. Dev. | Min. | Median | Max. | N |
|---|---|---|---|---|---|---|
| **Vote outcomes and Twitter data** | | | | | | |
| Republican two-party vote share (2016) | 0.46 | 0.17 | 0.08 | 0.45 | 0.95 | 3,065 |
| Change in Republican two-party vote share, 2000-2016 | -0.02 | 0.10 | -0.33 | -0.03 | 0.45 | 3,065 |
| Log(Twitter users) | 8.22 | 1.99 | 0.00 | 8.45 | 12.35 | 3,065 |
| Log(SXSW followers, March 2007) | 0.69 | 1.13 | 0.00 | 0.00 | 4.98 | 3,065 |
| Log(SXSW followers, Pre) | 0.33 | 0.73 | 0.00 | 0.00 | 3.61 | 3,065 |
| **Geographical controls** | | | | | | |
| Population density | 1925.15 | 6342.94 | 0.10 | 508.30 | 69468.40 | 3,065 |
| Log(County area) | 6.72 | 0.92 | 3.26 | 6.64 | 9.91 | 3,065 |
| Distance from Austin, TX (in miles) | 1731.21 | 653.61 | 5.04 | 1750.86 | 3098.88 | 3,065 |
| Distance from Chicago (in miles) | 1246.18 | 813.51 | 7.16 | 1103.75 | 3040.38 | 3,065 |
| Distance from NYC (in miles) | 1600.47 | 1255.14 | 6.48 | 1285.98 | 4191.67 | 3,065 |
| Distance from San Francisco (in miles) | 2841.16 | 1231.98 | 41.11 | 3157.16 | 4565.01 | 3,065 |
| Distance from Washington, DC (in miles) | 1448.51 | 1175.55 | 7.88 | 1047.13 | 3983.08 | 3,065 |
| **Demographic controls** | | | | | | |
| % aged 20-24 | 0.07 | 0.02 | 0.01 | 0.06 | 0.27 | 3,065 |
| % aged 25-29 | 0.07 | 0.01 | 0.03 | 0.07 | 0.15 | 3,065 |
| % aged 30-34 | 0.07 | 0.01 | 0.03 | 0.06 | 0.12 | 3,065 |
| % aged 35-39 | 0.06 | 0.01 | 0.03 | 0.06 | 0.10 | 3,065 |
| % aged 40-44 | 0.06 | 0.01 | 0.02 | 0.06 | 0.10 | 3,065 |
| % aged 45-49 | 0.06 | 0.01 | 0.02 | 0.06 | 0.09 | 3,065 |
| % aged 50+ | 0.36 | 0.06 | 0.11 | 0.35 | 0.75 | 3,065 |
| Population growth, 2000-2016 | 0.14 | 0.19 | -0.43 | 0.10 | 1.32 | 3,065 |
| % white | 0.65 | 0.21 | 0.03 | 0.68 | 0.98 | 3,065 |
| % black | 0.12 | 0.12 | 0.00 | 0.08 | 0.85 | 3,065 |
| % native American | 0.01 | 0.03 | 0.00 | 0.00 | 0.90 | 3,065 |
| % Asian | 0.05 | 0.06 | 0.00 | 0.03 | 0.37 | 3,065 |
| % Hispanic | 0.15 | 0.15 | 0.01 | 0.09 | 0.96 | 3,065 |
| % unemployed | 5.31 | 1.42 | 1.80 | 5.10 | 24.10 | 3,065 |
| **Socioeconomic controls** | | | | | | |
| % below poverty level | 15.11 | 5.34 | 1.40 | 15.10 | 53.30 | 3,065 |
| % employed in IT | 0.02 | 0.02 | 0.00 | 0.02 | 0.21 | 3,065 |
| % employed in construction/real estate | 0.07 | 0.03 | 0.00 | 0.07 | 1.00 | 3,065 |
| sh_emp_manufac | 0.11 | 0.08 | 0.00 | 0.08 | 0.72 | 3,065 |
| % adults with high school degree | 28.13 | 7.41 | 8.30 | 27.50 | 54.80 | 3,065 |
| % adults with college degree | 20.98 | 3.66 | 8.70 | 20.90 | 35.60 | 3,065 |
| % watching Fox News | 0.26 | 0.01 | 0.23 | 0.26 | 0.30 | 3,064 |
| % watching prime time TV | 0.43 | 0.01 | 0.40 | 0.43 | 0.47 | 3,064 |
| **China shock controls** | | | | | | |
| Exposure to Chinese import competition | 2.63 | 2.02 | -0.63 | 2.10 | 43.08 | 3,065 |
| Share of routine occupations | 31.87 | 2.36 | 22.23 | 32.14 | 36.66 | 3,065 |
| Average offshorability index | -0.02 | 0.50 | -1.64 | 0.09 | 1.24 | 3,065 |
| **1996 election control** | | | | | | |
| Republican two-party vote share (1996) | 0.41 | 0.11 | 0.10 | 0.41 | 0.79 | 3,065 |

*Notes:* This table presents descriptive statistics for the main estimation sample, weighted by turnout in the 2000 presidential elections.

## Table A.3: Summary Statistics (CCES Individual-Level)

| | Mean | Std. Dev. | Min. | Median | Max. | N |
|---|---|---|---|---|---|---|
| **Vote outcome** | | | | | | |
| Voted for Trump | 0.49 | 0.50 | 0.00 | 0.00 | 1.00 | 94,523 |
| **Twitter data** | | | | | | |
| Log(Twitter users) | 8.31 | 1.92 | 0.69 | 8.45 | 12.35 | 94,523 |
| Log(SXSW followers, March 2007) | 0.69 | 1.13 | 0.00 | 0.00 | 4.98 | 94,523 |
| Log(SXSW followers, Pre) | 0.32 | 0.71 | 0.00 | 0.00 | 3.61 | 94,523 |
| **Individual control variables** | | | | | | |
| Log(Age) | 3.88 | 0.37 | 2.89 | 3.99 | 4.55 | 94,523 |
| Family income dummies | 7.08 | 3.62 | 1.00 | 7.00 | 13.00 | 94,523 |
| Female dummy | 1.53 | 0.50 | 1.00 | 2.00 | 2.00 | 94,523 |
| Education dummies | 3.50 | 1.54 | 1.00 | 3.00 | 6.00 | 94,523 |
| Marital status dummies | 2.34 | 1.71 | 1.00 | 1.00 | 5.00 | 94,523 |
| Interest in news dummies | 1.59 | 0.92 | 1.00 | 1.00 | 7.00 | 94,523 |

*Notes:* This table presents descriptive statistics for the CCES estimation sample, weighted by survey weights.

## Table A.4: Summary Statistics (Gallup Individual-Level)

| | Mean | Std. Dev. | Min. | Median | Max. | N |
|---|---|---|---|---|---|---|
| **Candidate approval outcomes** | | | | | | |
| Approve of Trump? | 0.34 | 0.48 | 0.00 | 0.00 | 1.00 | 64,764 |
| Approve of Kasich? | 0.60 | 0.49 | 0.00 | 1.00 | 1.00 | 8,735 |
| Approve of Rubio? | 0.50 | 0.50 | 0.00 | 1.00 | 1.00 | 6,201 |
| Approve of Cruz? | 0.41 | 0.49 | 0.00 | 0.00 | 1.00 | 11,504 |
| Approve of Sanders? | 0.57 | 0.50 | 0.00 | 1.00 | 1.00 | 27,137 |
| Approve of Clinton? | 0.43 | 0.50 | 0.00 | 0.00 | 1.00 | 36,367 |
| **Twitter data** | | | | | | |
| Log(Twitter users) | 8.29 | 1.97 | 0.00 | 8.48 | 12.35 | 64,764 |
| Log(SXSW followers, March 2007) | 0.72 | 1.15 | 0.00 | 0.00 | 4.98 | 64,764 |
| Log(SXSW followers, Pre) | 0.34 | 0.73 | 0.00 | 0.00 | 3.61 | 64,764 |
| **Individual control variables** | | | | | | |
| Income dummies | 6.99 | 2.38 | 1.00 | 7.00 | 10.00 | 64,764 |
| Female dummy | 1.50 | 0.50 | 1.00 | 2.00 | 2.00 | 64,764 |
| Education dummies | 3.58 | 1.60 | 1.00 | 4.00 | 6.00 | 64,764 |
| Marital status dummies | 1.98 | 0.94 | 1.00 | 2.00 | 5.00 | 64,764 |
| Age deciles | 4.45 | 2.68 | 1.00 | 4.00 | 10.00 | 64,764 |

*Notes:* This table presents descriptive statistics for the Gallup estimation sample, weighted by survey weights.

## Table A.5: Instrument Balancedness

| | March 2007 *and* Pre (1) | March 2007 *only* (2) | Pre *only* (3) | Difference in means (2) - (3) | p-value | Šidàk p-value |
|---|---|---|---|---|---|---|
| Population density | 5192.27 | 1021.39 | 1998.35 | -976.96 | 0.07* | 0.91 |
| Log(County area) | 6.30 | 6.63 | 6.54 | 0.09 | 0.73 | 1.00 |
| Distance from Austin, TX (in miles) | 1775.99 | 1749.38 | 1626.64 | 122.74 | 0.48 | 1.00 |
| Distance from Chicago (in miles) | 1439.45 | 1329.47 | 1214.42 | 115.05 | 0.53 | 1.00 |
| Distance from NYC (in miles) | 1685.31 | 1594.99 | 1510.05 | 84.94 | 0.78 | 1.00 |
| Distance from San Francisco (in miles) | 2751.83 | 2900.11 | 2833.01 | 67.10 | 0.83 | 1.00 |
| Distance from Washington, DC (in miles) | 1558.55 | 1450.23 | 1397.05 | 53.18 | 0.85 | 1.00 |
| % aged 20-24 | 0.07 | 0.08 | 0.08 | 0.00 | 0.92 | 1.00 |
| % aged 25-29 | 0.09 | 0.07 | 0.07 | -0.00 | 0.51 | 1.00 |
| % aged 30-34 | 0.08 | 0.07 | 0.07 | -0.00 | 0.58 | 1.00 |
| % aged 35-39 | 0.07 | 0.06 | 0.06 | -0.00 | 0.82 | 1.00 |
| % aged 40-44 | 0.06 | 0.06 | 0.06 | 0.00 | 0.82 | 1.00 |
| % aged 45-49 | 0.07 | 0.06 | 0.06 | 0.00 | 0.89 | 1.00 |
| % aged 50+ | 0.32 | 0.35 | 0.35 | -0.00 | 0.97 | 1.00 |
| Population growth, 2000-2016 | 0.18 | 0.18 | 0.15 | 0.03 | 0.56 | 1.00 |
| % white | 0.50 | 0.65 | 0.67 | -0.02 | 0.62 | 1.00 |
| % black | 0.18 | 0.12 | 0.08 | 0.04 | 0.20 | 1.00 |
| % native American | 0.01 | 0.01 | 0.02 | -0.02 | 0.02** | 0.45 |
| % Asian | 0.10 | 0.05 | 0.05 | -0.01 | 0.55 | 1.00 |
| % Hispanic | 0.20 | 0.16 | 0.15 | 0.01 | 0.80 | 1.00 |
| % unemployed | 4.86 | 5.05 | 4.51 | 0.54 | 0.07* | 0.91 |
| % below poverty level | 15.71 | 15.82 | 13.69 | 2.14 | 0.17 | 1.00 |
| % employed in IT | 0.04 | 0.02 | 0.02 | -0.00 | 0.98 | 1.00 |
| % employed in construction/real estate | 0.06 | 0.07 | 0.07 | 0.01 | 0.39 | 1.00 |
| sh_emp_manufac | 0.07 | 0.09 | 0.07 | 0.02 | 0.16 | 0.99 |
| % adults with high school degree | 21.76 | 25.99 | 25.77 | 0.22 | 0.88 | 1.00 |
| % adults with college degree | 18.89 | 21.16 | 21.20 | -0.04 | 0.97 | 1.00 |
| % watching Fox News | 0.25 | 0.26 | 0.26 | -0.00 | 0.91 | 1.00 |
| % watching prime time TV | 0.42 | 0.43 | 0.43 | 0.00 | 0.91 | 1.00 |
| Exposure to Chinese import competition | 2.55 | 2.46 | 2.79 | -0.32 | 0.54 | 1.00 |
| Share of routine occupations | 32.47 | 31.38 | 31.25 | 0.13 | 0.82 | 1.00 |
| Average offshorability index | 0.37 | -0.07 | -0.05 | -0.02 | 0.84 | 1.00 |
| Republican two-party vote share (1996) | 0.36 | 0.42 | 0.42 | -0.00 | 0.90 | 1.00 |

*Notes:* This table presents the averages for the main control variables separately for the three types of counties relevant for our identification strategy: 1) the 47 counties with SXSW followers that joined Twitter both in March 2007 and the "pre-period"; 2) the 108 counties with SXSW followers that joined in March 2007 (but none in the "pre-period"); and 3) the 20 counties with SXSW that joined in the "pre-period" (but none in March 2007). We report p-values from a two-sided $t$-test for the equality of means between the counties with the key identifying variation, as well as Šidàk-corrected values to account for multiple hypothesis testing. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table A.6: Description of Main Variables

| Variable | Description | Source |
|---|---|---|
| **Vote outcomes** | | |
| Republican vote share (2016) | The vote share of the Republican party in the 2016 presidential election. | Dave Leip Election Atlas |
| ΔRepublican vote share, 2000-2016 | Change in the vote share of the Republican party between the 2000 and 2016 presidential elections. | Dave Leip Election Atlas |
| **Twitter data** | | |
| Twitter users | The number of Twitter users per county (in natural logarithm), based on tweets collected using the Twitter streaming API in a 12 month period from June to November 2014 and June to November 2015. | Gesis Datatorium |
| SXSW followers, March 2007 | The number of Twitter users following the @SXSW account in each county that joined Twitter in March 2007. | Twitter Search API |
| SXSW followers, Pre | The total number of Twitter users following the @SXSW account in each county that joined Twitter at any point in 2006. | Twitter Search API |
| **County-level control variables** | | |
| Geographical controls | Include the distance (in miles) from the country centroid from Austin (TX), Chicago, NYC, San Francisco, and Washington, DC, population density, and the logarithm of the land area for each county. | U.S. Census Tigerline File |
| Demographic controls | Include the share of people in the age buckets 20-24, 25-29, 30-34, 40-44, 45-49 and 50+, as well as the percentage change in county population between 2000 and 2016. | |
| Socioeconomic controls | Include a county's poverty rate; unemployment rate; the share of the population employed in construction/real estate, manufacturing, and information technology; the share of adults over 25 with at least a high school degree or some college education; as well as prime time TV viewership and the share of Fox News viewership. | U.S. Census/BLS/SimplyAnalytics |
| China shock controls | Include the measure of import competition from China, routine tasks, and offshorability from Autor et al. (2013). | |
| Election control | Contains the vote share of the Republican party in the 1996 presidential election. | Dave Leip Election Atlas |

**Table A.7: Search Terms Used to Create a Proxy for Total Tweets**

| | | | | | |
|---|---|---|---|---|---|
| 0 | but | his | one | these | would |
| 1 | by | how | only | they | year |
| 2 | can | if | or | think | you |
| 3 | come | in | other | this | your |
| 4 | could | into | our | time | |
| 5 | day | it | out | two | |
| 6 | do | its | over | up | |
| 7 | even | just | people | us | |
| 8 | first | know | say | use | |
| 9 | for | like | see | want | |
| I | from | look | she | way | |
| about | get | make | so | we | |
| after | give | me | some | well | |
| all | go | most | take | what | |
| also | good | my | than | when | |
| any | have | new | that | which | |
| as | he | no | their | who | |
| at | he | not | them | with | |
| back | her | now | then | with | |
| because | him | on | there | work | |

*Notes:* This table lists the 100 most common English words that were used as search terms to generate a proxy of "total tweets" used in Figure 3b.

### A.1.1 Additional Details on the Logistic Regression Classifier

We train a machine learning classifier to determine whether tweets are more likely to be sent by Democratic-leaning or Republican-leaning users using the Python sci-kit package (Pedregosa et al., 2011). The classification process starts with the preparation of the underlying Twitter data. The inputs are the text of each of the 1,999,893 tweets from U.S. congress members. The target variable $y$ for the classifier is an indicator variable equal to one for tweets sent by Republicans and zero otherwise. The feature matrix $X$ for the classifier are created by count-vectorizing the texts of the tweets. In other words, we transform the text of the tweets into $n \times v$ matrix, where $n$ is the number of tweets and $v$ is the number of unique 1,2,3-grams that occur in the tweets. In preparation for this step, we removed common words (stopwords), links, and special characters from the tweets. Additionally, we reduced the words in the tweets to their morphological roots using a lemmatizer, which improves the performance of the classifier. As an example, the lemmatizer changes words like "walking" and "walked" to "walk". Lastly, we reweight the n-grams $v$ of tweet $i$ using term frequency–inverse document frequency (tf-idf):

$$tfidf(f_{i,v}) = (1 + \ln(f_{i,v}) \cdot (\ln(\frac{1+T}{1+d_v}) + 1) \tag{A.1}$$

where $d_v$ is the number of tweets n-gram $v$ appears in. This reweighting reduces the importance of words that appear frequently in many tweets, which help little to discriminate between tweets. The tf-idf vectorizer also normalizes the feature matrix by its L2-norm.

The vector $y$ and the matrix $X$ then serve as the input for a $L2$ regularized logistic regression classifier. The optimization of the classifier involves the minimization of the following cost function[37]:

$$\min_{w,c} \; C \sum_{i=1}^{n} \log(\exp(-y_i(X_i^T w + c) + 1) + \frac{1}{2} w^T w \tag{A.2}$$

where $w$ are the weights (coefficients) of the logistic regression, $c$ is a constant (intercept), and $C$ is the inverse of the regularization strength. Larger values of $C$ imply weaker regularization. For $C \longrightarrow \infty$, the classifier converges towards a normal logistic regression. As is standard in most machine learning applications, we choose the optimal

[37]Note that this formulation of the cost function assumes that $y_i$ takes values $-1; 1$. We use this formulation in line with the sci-kit documentation

regularization strength $C$ using 10-fold cross-validation. This involves randomly splitting the training data into ten equal slices. Nine of the ten slices are then used to train the classifier, while the out-of-sample performance is evaluated against the remaining slice using F1-scores.

The final classifier achieves an out-of-sample F1-score of 0.9 and therefore very accurately predicts the party affiliation of Congress members. We then take this classifier and apply it to the universe of tweets sent during the 2012 and 2016 presidential election. For each tweet in the election data, the classifier provides us with a predicted class (either Democrat or Republican) and a probability for this class label. In spirit, this approach is similar to the work of Gentzkow and Shapiro (2010). While they identify expressions that are more frequently used by Democrats and Republicans by hand, we use a machine learning classifier to identify n-grams in the tweets of Congress members that help us to differentiate between the two parties.

We visualize the most predictive n-grams identified by the classifier in Figure A.4. Overall, the classifier identifies words, hashtags, and Twitter handles that are intuitively associated with a Republican slant. The most predictive term are the hashtags "tcot" (Top Conservatives on Twitter) and "ms01" (Mississippi's 1st District). Further, retweeting the "housegop", "senategop", or "realdonaltrump" hashtags predict Twitter content with a more Republican slant as do the terms "buildthewall", "illegal immigration", or "taxcutsandjobsact".

**Figure A.4: Most Predictive Terms Of "Republican" Tweets**




*Notes:* This word cloud plots the n-grams most predictive of tweets sounding like those of Republican Members of Congress, as identified by the logistic regression classifier. The size of the word represents the magnitude of the coefficients.

## A.2. Appendix 2: Additional Details on the SXSW Festival

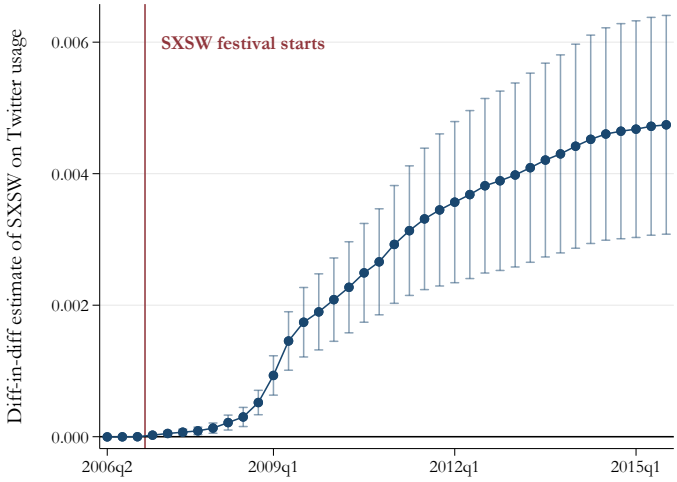
**Table A.8: Balancedness of SXSW Counties' User Characteristics**

| **First names (Corr. = 0.69)** | | **Terms used in bio (Corr. = 0.92)** | |
|---|---|---|---|
| Pre-period | March 2007 | Pre-period | March 2007 |
| michael | michael | http | http |
| mike | john | founder | com |
| paul | chris | com | digital |
| chris | jeff | co | founder |
| ryan | matt | tech | medium |
| eric | brian | design | director |
| david | david | director | tech |
| matthew | alex | product | music |
| john | jason | digital | social |
| jeff | kevin | designer | marketing |
| robert | paul | medium | design |
| mark | mike | music | co |
| andrew | dan | social | writer |
| daniel | andrew | love | love |
| james | peter | marketing | lover |
| kevin | jim | web | dad |
| jay | tom | geek | creative |
| jonathan | jennifer | writer | tweet |
| rob | steve | technology | author |
| rachel | todd | dad | designer |

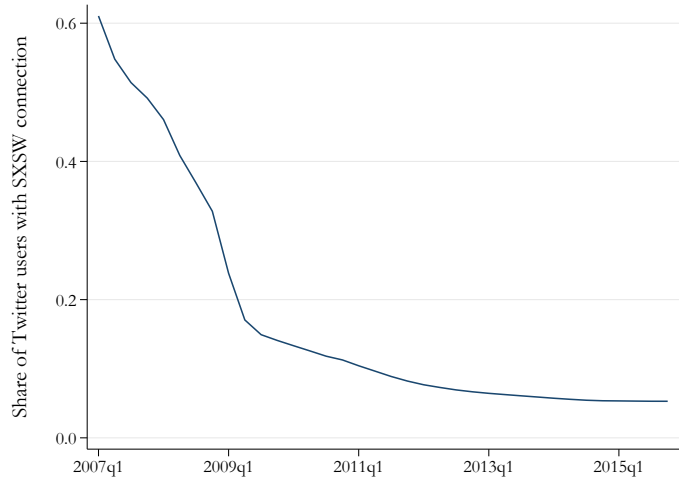
*Notes:* This table presents the ranking of the most common first names and terms used in a Twitter user's "bio" among users who follow "South by Southwest" on Twitter, depending on whether they joined during March 2007 or in the pre-period.

**Figure A.5: Additional Evidence for the Impact of the SXSW Festival**

**(a) Long-term Effects of the 2007 SXSW on Twitter Adoption**




**(b) Connections to the SXSW festival**



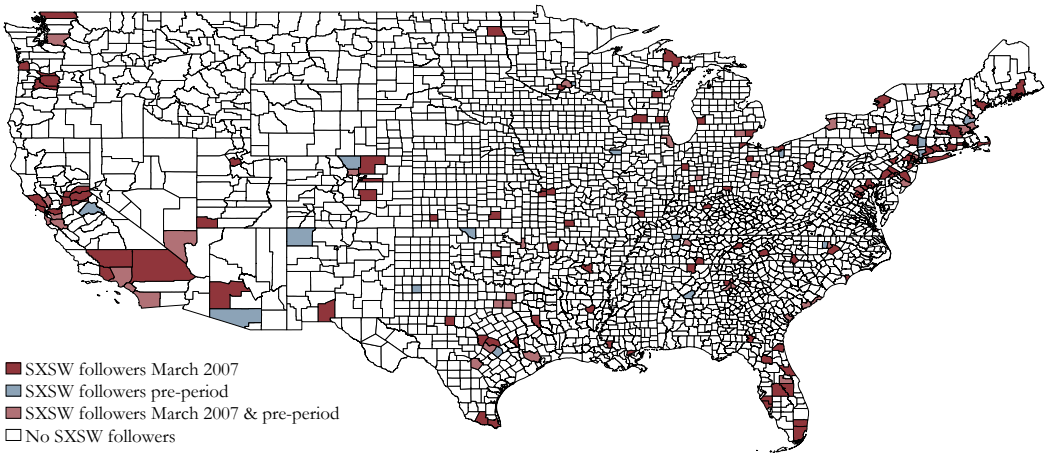

*Notes:* The figures provide evidence on the long-term impact of the SXSW festival on Twitter usage across the United States. Panel (a) plots the $\beta_\tau$ from the panel event study regression $users_{ct} = \sum_\tau \beta_\tau SXSW_c^{March2007} \times 1(t=\tau) + \sum_\tau \delta_\tau SXSW_c^{Pre} \times 1(t=\tau) + \theta_c + \gamma_t + \varepsilon_{ct}$ where $users_{ct}$ is the number of Twitter users per capita in county $c$ on quarter $t$, $SXSW_c^{March2007}$ is the logarithm of (one plus) the number of SXSW followers in county $c$ who joined Twitter in March 2007 and $SXSW_c^{Pre}$ is a similarly defined variable for followers who joined Twitter before March 2007. We standardize the variables to have a mean of zero and standard deviation of one. The whiskers represent 95% confidence intervals based on standard errors clustered by state, where 2006q4 serves as excluded period. While the confidence intervals for 2006q2 and 2006q3 cannot be seen, they include zero. Panel (b) plots the share of Twitter who either follow SXSW or follow an user that follows SXSW.

**Table A.9: Are Twitter Users in Counties With SXSW Followers Different?**

| **User first names (Corr. = 0.97)** | | **Terms used in user bio (Corr. = 0.94)** | |
|---|---|---|---|
| Other counties | SXSW counties | Other counties | SXSW counties |
| michael | michael | love | co |
| chris | david | life | love |
| john | chris | co | life |
| david | john | http | http |
| sarah | alex | http co | http co |
| mike | mike | god | music |
| emily | matt | ig | lover |
| ryan | sarah | music | ig |
| matt | ryan | university | de |
| alex | andrew | like | like |
| taylor | emily | fan | fan |
| ashley | brian | live | world |
| nick | jessica | lover | instagram |
| jessica | james | mom | thing |
| tyler | kevin | husband | la |
| hannah | daniel | time | live |
| katie | ashley | follow | time |
| amanda | jason | one | com |
| lauren | lauren | wife | artist |
| brian | mark | thing | one |

*Notes:* This table compares the individual characteristics of Twitter users from counties with "South by Southwest" followers who joined in March 2007 ("SXSW counties") to Twitter users from all other U.S. counties ("Other counties"). We plot the ranking of the most common first names and terms used in a Twitter user's "bio".

## Figure A.6: Identifying Variation




*Notes:* This map plots three types of counties relevant for our identification strategy: 1) the 47 counties with SXSW followers that joined Twitter both in March 2007 and the "pre-period" (light red) ; 2) the 108 counties with SXSW followers that joined in March 2007, but none in the "pre-period" (dark red); and 3) the 20 counties with SXSW that joined in the "pre-period," but none in March 2007 (blue).

## A.3. Appendix 3: Additional Robustness Checks

**Table A.10: Twitter and the Republican Vote Share 2016 – Robustness**

| | *Dep. var.: Republican vote share in 2016* | | | | |
|---|---|---|---|---|---|
| | No regression weights | Pre-period control polynomial | Pre-period control deciles | No zero SXSW user counties | Spatial standard errors |
| | (1) | (2) | (3) | (4) | (5) |
| Log(Twitter users) | -0.099*** | -0.079*** | -0.076*** | -0.080*** | -0.099*** |
| | (0.015) | (0.016) | (0.013) | (0.029) | (0.019) |
| Observations | 3,065 | 3,065 | 3,065 | 165 | 3,065 |
| Mean of DV | 0.64 | 0.46 | 0.46 | 0.34 | 0.64 |
| Robust F-stat. | 84.87 | 90.78 | 88.40 | 21.09 | 70.68 |

*Notes:* This table presents 2SLS results estimated using equation (3). The dependent variable is the vote share of the Republican party in the 2016 presidential election. *Log(Twitter users)* is instrumented using the number of users who started following SXSW in March 2007 (in logs with 1 added inside). *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006. All regressions include the controls from column 2 in Table 2. All regressions except columns 1 and 5 are weighted by turnout in the 2000 presidential election. Columns 2 and 3 control for a fifth-order polynomial or deciles of SXSW followers who joined Twitter before the SXSW 2007 event. Column 4 drops all counties that had no SXSW followers joining Twitter in March 2007 or in the period before. Column 5 uses spatial standard errors based on the method proposed in Colella et al. (2019), implemented in Stata as *acreg*, using a 200 miles cutoff. In columns 1 to 4, standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

**Table A.11: Twitter and Changes in the Republican Vote Share, 2004-2020**

| | *Dep. var.: ΔRepublican vote share between...* | | | | |
|---|---|---|---|---|---|
| | 2000-04 | 2000-08 | 2000-12 | 2000-16 | 2000-20 |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: OLS** | | | | | |
| Log(Twitter users) | 0.004** | 0.007*** | 0.008*** | -0.000 | -0.001 |
| | (0.001) | (0.002) | (0.002) | (0.003) | (0.002) |
| **Panel B: Reduced form** | | | | | |
| Log(SXSW followers, March 2007) | -0.002 | -0.003 | -0.001 | -0.009** | -0.008** |
| | (0.002) | (0.003) | (0.003) | (0.004) | (0.004) |
| Log(SXSW followers, Pre) | -0.000 | -0.001 | -0.002 | 0.000 | -0.002 |
| | (0.003) | (0.005) | (0.003) | (0.004) | (0.005) |
| **Panel C: 2SLS** | | | | | |
| Log(Twitter users) | -0.004 | -0.006 | -0.002 | -0.017** | -0.015** |
| | (0.004) | (0.006) | (0.006) | (0.007) | (0.007) |
| Log(SXSW followers, Pre) | 0.000 | -0.001 | -0.002 | 0.001 | -0.001 |
| | (0.003) | (0.005) | (0.003) | (0.005) | (0.005) |
| Observations | 3,064 | 3,064 | 3,064 | 3,064 | 3,064 |
| Mean of DV | 0.03 | -0.02 | -0.01 | -0.02 | -0.01 |
| Robust F-stat. | 121.18 | 121.18 | 121.18 | 121.18 | 121.18 |

*Notes:* This table presents county-level regressions where the dependent variable is the change in the vote share of the Republican party between 2000 and the indicated year. *Log(SXSW followers, March 2007)* is the number of Twitter users (in logs, with 1 added inside) who joined in March 2007 and follow South by Southwest (SXSW). *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006, defined similarly. *Twitter users* are the number of users in 2014-2015. All regressions control for population deciles, Census region fixed effects, and the full set of controls (as in column 5 of Table 2). The first-stage regressions for 2SLS results (Panel C) are presented in Table 1, with the F-stat for the excluded instrument in the bottom row. Observations are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

**Table A.12: Time-Varying Twitter Usage and Changes in Vote Shares**

| | *Dep. var.: ΔRepublican vote share between...* | | | | |
|---|---|---|---|---|---|
| | 2000-04 | 2000-08 | 2000-12 | 2000-16 | 2000-20 |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: OLS** | | | | | |
| Log(Twitter users) | 0.000 | -0.001 | 0.007*** | -0.000 | -0.001 |
| | (0.001) | (0.002) | (0.002) | (0.003) | (0.002) |
| **Panel B: First stage** | | | | | |
| Log(SXSW followers, March 2007) | 0.644*** | 0.644*** | 0.546*** | 0.523*** | 0.523*** |
| | (0.055) | (0.055) | (0.047) | (0.048) | (0.048) |
| **Panel C: Reduced form** | | | | | |
| Log(SXSW followers, March 2007) | -0.002 | -0.003 | -0.001 | -0.009** | -0.008** |
| | (0.002) | (0.003) | (0.003) | (0.004) | (0.004) |
| Log(SXSW followers, Pre) | -0.000 | -0.001 | -0.002 | 0.000 | -0.002 |
| | (0.003) | (0.005) | (0.003) | (0.004) | (0.005) |
| **Panel D: 2SLS** | | | | | |
| Log(Twitter users) | -0.003 | -0.005 | -0.002 | -0.017** | -0.015** |
| | (0.003) | (0.005) | (0.005) | (0.007) | (0.007) |
| Log(SXSW followers, Pre) | 0.000 | -0.001 | -0.002 | 0.001 | -0.001 |
| | (0.003) | (0.005) | (0.003) | (0.005) | (0.005) |
| Mean of DV | 0.026 | -0.024 | -0.008 | -0.018 | -0.009 |
| *Twitter usage measured in* | *2008* | *2008* | *2012* | *2016* | *2016* |

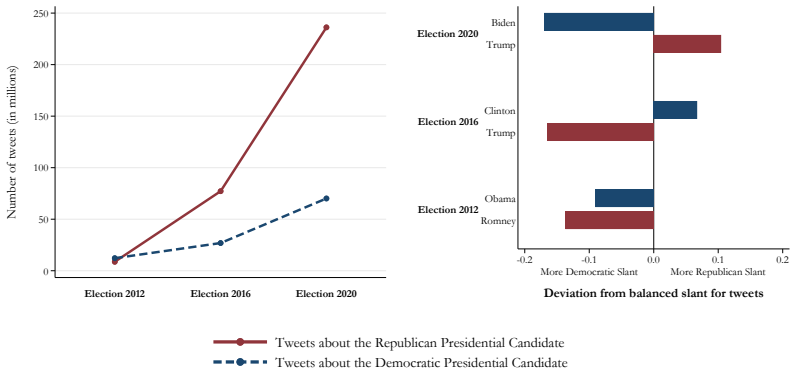
*Notes:* This table presents county-level regressions where the dependent variable is the change in the vote share of the Republican party between 2000 and the indicated year (except for Panel B, where the dependent variable is *Twitter users*. *Log(SXSW followers, March 2007)* is the number of Twitter users (in logs, with 1 added inside) who joined in March 2007 and follow South by Southwest (SXSW). *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006, defined similarly. Differently from other tables, *Twitter users* varies over time, as opposed to being fixed to 2014-2015. Regressions include the indicated control variables (see the Online Appendix for their descriptions). Observations are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

**Table A.13: Twitter and Individual's Vote Decisions – Robustness**

| | *Dep. var.: Voted for Trump* | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | Baseline | Verified vote | Vote intention | Intended Trump vote | Intended other vote |
| Log(Twitter users) | -0.129*** | -0.140** | -0.133** | -0.231*** | -0.064* |
| | (0.048) | (0.056) | (0.052) | (0.082) | (0.034) |
| Observations | 94,523 | 28,413 | 46,418 | 14,723 | 24,354 |
| Mean of DV | 0.491 | 0.495 | 0.455 | 0.991 | 0.137 |
| *Marginal effect* | [-0.047] | [-0.051] | [-0.048] | [-0.005] | [-0.013] |

*Notes:* This table presents results estimated using IV probit models, as in equation (4). The dependent variable is a dummy for individuals in the CCES who voted for Trump in 2016. *Log(Twitter users)* is instrumented using the number of SXSW followers that joined Twitter in March 2007. All regressions control for family income, gender, education levels, marital status, news interest, and age, as well as county-level population deciles and Census region fixed effects. Regressions are weighted by survey weights. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

**Figure A.7: Twitter's Partisan Slant in the 2012 and 2016 Presidential Elections**




*Notes:* These figures present the number of tweets that contain the last name of the candidates in the 2012 and 2016 presidential elections, depending on whether the tweet was classified as having a Republican (instead of Democratic) slant. We classify the slant of a tweet based on similarity in the language to that of a congressional Republican or Democrat, using a L2 regularized logistic regression classifier based on tweets sent by Members of Congress, where optimal normalization strength is chosen using 10-fold cross-validation. See Online Appendix Appendix A.1.1 for details.

# A.4. Appendix 4: Further Results

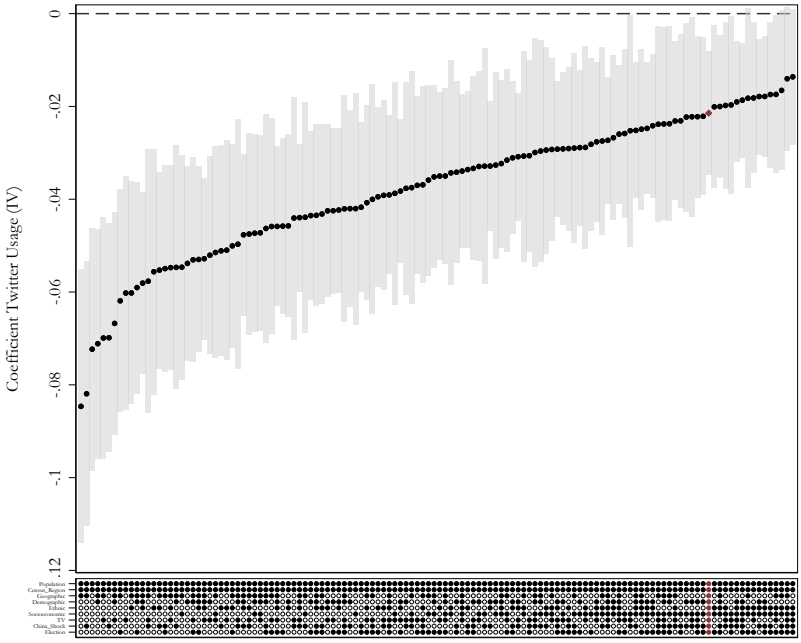
**Table A.14: Additional Outcomes**

| | *Switching prob.* | *ΔCampaign don., 2000-16* | | *Trump approval, 2017* | |
|---|---|---|---|---|---|
| | Obama to Trump | Democrats | Republicans | Democrats | Republicans |
| | (1) | (2) | (3) | (4) | (5) |
| Log(Twitter users) | -0.138*** | 0.866*** | 0.168 | -0.011** | -0.037*** |
| | (0.048) | (0.185) | (0.235) | (0.005) | (0.009) |
| Observations | 3,065 | 2,250 | 2,446 | 2,727 | 2,920 |
| Mean of DV | 0.105 | 1.943 | 1.096 | 0.066 | 0.850 |

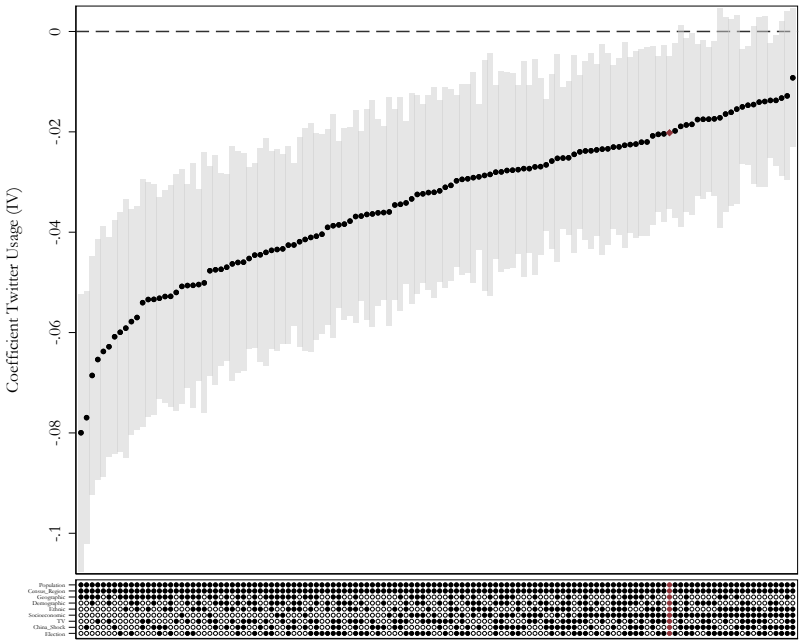
*Notes:* This table presents results from county-level regressions of equation (3). Column 1 shows results from an IV probit regression where the dependent variable is a dummy equal to 1 for the 217 counties for which both Obama and Trump gained the majority of votes in 2008 and 2016, respectively. In columns 2 and 3, the dependent variable is the difference in the natural logarithm of campaign distribution to the Democratic and Republican party, respectively, between 2000 and 2016. In columns 4 and 5, the dependent variable is the share of respondents in the Gallup Daily Poll approving of Trump in 2017. *Log(Twitter users)* is instrumented using the number of users who started following SXSW in March 2007. All regressions control for population deciles and Census region fixed effects and geographical controls See the Online Appendix for variable descriptions. Regressions are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Figure A.8: Specification Curves

### (a) 2SLS Estimates 2016




### (b) 2SLS Estimates 2020




*Notes:* These figures plot the 2SLS estimates and 95% confidence intervals from a regression of the Republican vote share in 2016 (panel a) and 2020 (panel b) on $Log(Twitter\ users)$, instrumented with $SXSW^{March\ 2007}$. All regressions include population deciles, census region fixed effects, and $SXSW^{Pre}$. The combination of the other included control variables is shown at the bottom; filled circles mean a set of controls was included.

**Table A.15: Twitter and Changes in Voter Turnout, 2004-2020**

| | $\Delta$ *Votes cast/voting age pop.* | | | | |
|---|---|---|---|---|---|
| | 2000-04 | 2000-08 | 2000-12 | 2000-16 | 2000-20 |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: OLS** | | | | | |
| Log(Twitter users) | -0.001 | -0.001 | -0.003 | -0.001 | 0.005 |
| | (0.003) | (0.003) | (0.003) | (0.004) | (0.003) |
| **Panel B: Reduced form** | | | | | |
| Log(SXSW followers, March 2007) | -0.000 | -0.000 | -0.001 | 0.001 | 0.007** |
| | (0.003) | (0.003) | (0.004) | (0.004) | (0.003) |
| Log(SXSW followers, Pre) | -0.000 | 0.000 | -0.002 | -0.001 | -0.005 |
| | (0.006) | (0.004) | (0.005) | (0.005) | (0.005) |
| **Panel C: 2SLS** | | | | | |
| Log(Twitter users) | -0.000 | -0.000 | -0.001 | 0.002 | 0.014** |
| | (0.006) | (0.006) | (0.008) | (0.008) | (0.006) |
| Log(SXSW followers, Pre) | -0.000 | 0.000 | -0.002 | -0.001 | -0.006 |
| | (0.006) | (0.005) | (0.005) | (0.005) | (0.005) |
| Observations | 3063 | 3063 | 3063 | 3063 | 3063 |
| Mean of DV | 0.088 | 0.079 | 0.053 | 0.057 | 0.126 |
| Robust F-stat. | 121.23 | 121.23 | 121.23 | 121.23 | 121.23 |

*Notes:* This table presents county-level regressions where the dependent variable is the change in the voter turnout (as a share of voting age population) between 2000 and the indicated year. *Log(SXSW followers, March 2007)* is the number of Twitter users (in logs, with 1 added inside) who joined in March 2007 and follow South by Southwest (SXSW). *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006, defined similarly. *Twitter users* are the number of users in 2014-2015. All regressions control for population deciles, Census region fixed effects, and the full set of controls (as in column 5 of Table 2). The first-stage regressions for 2SLS results (Panel C) are presented in Table 1, with the F-stat for the excluded instrument in the bottom row. Observations are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table A.16: Twitter and Congressional Elections – 2SLS Estimates

**Panel A: House elections**

| | *ΔRepublican vote share in House election between...* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2000-02 | 2000-04 | 2000-06 | 2000-08 | 2000-10 | 2000-12 | 2000-14 | 2000-16 | 2000-18 | 2000-20 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Log(Twitter users) | 0.014 | -0.003 | 0.027 | 0.032 | 0.019 | 0.023 | 0.039 | 0.006 | 0.028 | 0.023 |
| | (0.018) | (0.019) | (0.020) | (0.022) | (0.021) | (0.023) | (0.024) | (0.024) | (0.023) | (0.023) |
| Observations | 2,982 | 2,955 | 2,971 | 2,957 | 2,969 | 2,983 | 2,978 | 2,980 | 2,983 | 2,982 |
| Mean of DV | 0.02 | 0.01 | -0.04 | -0.06 | 0.03 | -0.01 | 0.02 | 0.00 | -0.04 | -0.01 |
| Robust F-stat. | 109.63 | 110.42 | 109.20 | 109.49 | 109.83 | 109.68 | 110.07 | 109.87 | 109.68 | 109.53 |

**Panel B: Senate elections**

| | *ΔRepublican vote share in Senate election between...* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1996-02 | 1998-04 | 2000-06 | 1996-08 | 1998-10 | 2000-12 | 1996-14 | 1998-16 | 2000-18 | 2000-20 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Log(Twitter users) | 0.019 | 0.015 | 0.007 | 0.008 | -0.011 | 0.021 | -0.035 | -0.013 | -0.008 | -0.019 |
| | (0.023) | (0.024) | (0.021) | (0.021) | (0.018) | (0.031) | (0.026) | (0.026) | (0.022) | (0.018) |
| Observations | 2,183 | 1,985 | 1,832 | 2,183 | 1,985 | 1,832 | 2,183 | 1,985 | 1,832 | 2,183 |
| Mean of DV | 0.01 | -0.03 | -0.06 | -0.05 | 0.01 | -0.06 | 0.03 | -0.07 | -0.10 | -0.00 |
| Robust F-stat. | 79.24 | 85.91 | 72.85 | 79.24 | 85.91 | 72.85 | 79.24 | 85.91 | 72.85 | 79.24 |

*Notes:* This table presents results estimated using 2SLS, as in equation (3). For House elections in Panel A, the dependent variable is the change in the Republican vote share since 2000. For Senate elections in Panel B, the dependent variable is the change in the Republican vote share from six, twelve, or eighteen years ago (to accommodate senators' 6-year terms). *Log(Twitter users)* is instrumented using the number of users who started following SXSW in March 2007 (in logs with 1 added inside). All regressions control for population deciles and Census region fixed effects and the full set of controls (as in column 5 of Table 2). See the Online Appendix for variable descriptions. Regressions are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

**Table A.17: Social Media Persuasion in Swing and Safe Counties**

| | $\Delta$*Republican vote share in 2016* | | | |
|---|---|---|---|---|
| | Swing counties (1) | Republican counties (2) | Democratic counties (3) | Safe counties (4) |
| Log(Twitter users) | -0.073*** | -0.006 | -0.008 | -0.005 |
| | (0.024) | (0.008) | (0.008) | (0.006) |
| Log(SXSW followers, Pre) | 0.013 | -0.001 | 0.000 | -0.002 |
| | (0.015) | (0.014) | (0.008) | (0.005) |
| Observations | 716 | 1,990 | 358 | 2,348 |
| Mean of DV | -0.033 | 0.021 | -0.040 | -0.012 |
| Robust F-stat. | 14.70 | 11.97 | 105.57 | 99.87 |

*Notes:* This table presents results estimated using 2SLS, as in equation (3). The dependent variable is the change in the vote share of the Republican party between the 2000 and 2016 presidential elections. *Swing counties* are those that were not consistently won by either Republicans or Democrats between 2000 and 2012; *Republican* and *Democratic* counties are those who voted consistently. *Log(Twitter users)* is instrumented using the number of users who started following SXSW in March 2007. *SXSW followers, Pre* is the number of SXSW followers who registered at some point in 2006. All regressions control for population deciles and Census region fixed effects and the full set of controls (as in column 5 of Table 2). See the Online Appendix for variable descriptions. Regressions are weighted by turnout in the 2000 presidential election. Standard errors in parentheses are clustered by state. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.