



From Tweets to the Streets: Twitter and Extremist Protests in the United States

Gisli Gylfason

► To cite this version:

Gisli Gylfason. From Tweets to the Streets: Twitter and Extremist Protests in the United States. 2023. halshs-04188189

HAL Id: halshs-04188189

<https://shs.hal.science/halshs-04188189>

Preprint submitted on 25 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



WORKING PAPER N° 2023 – 29

**From Tweets to the Streets: Twitter and Extremist
Protests in the United States**

Gisli Gylfason

JEL Codes: D70, L86, L82, P16

**Keywords: twitter; collective action; protests; social media; information
technology.**



From Tweets to the Streets: Twitter and Extremist Protests in the United States*

Gísli Gylfason[†]

August 25, 2023

Abstract

How does social media affect the composition of political protests in the United States? Using early adoption of Twitter at the 2007 South by Southwest (SXSW) festival as a plausibly exogenous source of variation in county-level Twitter penetration (Müller & Schwarz, 2023), and comprehensive data on protest events, this paper finds that Twitter penetration increases the frequency of protests overall, but also radicalizes them. Twitter disproportionately fuels protests with participation of “extreme” groups—groups that are particularly militant, radical, or hateful. These effects do not depend on the topic of the protest nor political leaning. I also present survey evidence suggesting that coordination is not the only mechanism driving these results: An increase in county-level Twitter penetration implies an increase in respondent’s willingness to justify violence against other people, normalizing the participation in extreme groups and extreme protests.

Keywords: Twitter, Collective action, Protests, Social media, Information Technology

JEL codes: D70, L86, L82, P16.

*I am grateful to Ekaterina Zhuravskaya for supervision and guidance on this project. I thank Oliver Vanden Eynde, Hillel Rapoport, Sandra Ponchet, Karsten Müller, Carlo Schwarz, Nikita Melnikov, participant at the PSE GPET dissertation seminar, the PSE Applied Economics Lunch seminar, and the 2nd Academic Workshop for Icelandic Economists Abroad at the University of Iceland, for helpful feedback and comments.

[†]PhD Candidate, Paris School of Economics & École des Hautes Études en Sciences Sociales (EHESS). Email: gisli.gylfason@psemail.eu

1 Introduction

Since the late 2000s, the world has witnessed increasing waves of protest movements such as the Arab Spring or the Black Lives Matter Movement (Ortiz et al., 2013). These developments have coincided with the rapid expansion of Information and Communication Technologies (ICTs) in general and social media in particular, creating the widespread perception of their role in facilitating protest coordination (Economist, 2019; Ovide, 2020). More recently in western democracies, particularly the United States where far-right demonstration activity has been surging (ACLED, 2022), reports document not only rising protest activity but increased hostility and aggression during protest events (ACLED, 2020c,a,b). Journalists and politicians have suggested that social media may also play an important role in these trends (Brewster, 2021; Hadero, 2023). While increased violence during demonstrations is worrying, understanding the influences of social media on the composition of political protests is important as protests can influence attitudes, voting behaviour and policy (Madestam et al., 2013; Skoy, 2021; Klein Teeselink & Melios, 2021; Caesmann et al., 2021; Casanueva, 2022).

The literature has explored the mechanisms through which ICTs can facilitate protests and confirmed them empirically (e.g., Little, 2016; Enikolopov, Makarin, & Petrova, 2020). However, the best available evidence is limited to either rigorous studies focusing on protests related to one particular social movement (e.g., Enikolopov, Makarin, & Petrova, 2020) or studies taking a more global approach, treating the nature of different movements as a black box (e.g., Manacorda & Tesei, 2020). Does social media disproportionately fuel protest mobilization among political fringes? Does social media influence aggression and violence during protest events? This paper contributes to the understanding of these issues by studying the effect of Twitter penetration on various protest movements, depending on the characteristics the movements and of the demonstration events themselves in the United States. It finds that Twitter penetration causes a disproportionate increase in protest events that involve “extreme” groups—groups that are particularly militant, radical, or hateful.

Like traditional media, social media can provide citizens with politically relevant information increasing their willingness to demonstrate for a given cause. In addition, social media can facilitate coordination through horizontal interaction between users, alleviating collective-action problems (Olson, 1965). These coordination effects can stem from reduced costs of acquiring and exchanging logistical and tactical information on demonstration events (Little, 2016; Enikolopov, Makarin, & Petrova, 2020), from updating beliefs about how many others are willing to demonstrate (Edmond, 2013;

Barbera et al., 2020; González, 2020; Passarelli & Tabellini, 2017) or by altering how one can project their social image via protest participation (Enikolopov, Makarin, Petrova, & Polishchuk, 2020; Cantoni et al., 2019).

Social media may disproportionately fuel extremists mobilization through various channels. Prominent social commentators and scientists have long warned that through fierce competition for users attention, social media platforms create informational environments that trigger fear, anger and stoke public outrage (CBS, 2022; Rose-Stockwell, 2023). These warnings contend that algorithmic content recommendations designed to maximize engagement may disproportionately suggest content which triggers emotional responses, or more radical versions of whatever content the user was already consuming (Tufekci, 2018; Rose-Stockwell, 2017). Social psychologists have also argued that the way social media turns much communication into public performance, evaluated by the public in forms of likes and other metrics, may tilt incentives towards expressing more anger, aggression and radical positions than one would in other forms of communication (Haidt & Rose-Stockwell, 2019).¹ This constant push towards more radical content and exposure to aggressive online communication may drive partisan hostility, persuade individuals towards taking more radical positions and thus disproportionately fuel extreme protest movements.²

Even if no radicalization process occurs, social media may be a more powerful coordination tool among those already radicalized. First, as extreme voices diffuse more widely (Barberá et al., 2015; Brady et al., 2017), social media could be especially conducive in connecting extremists to like-minded individuals, thus facilitating tactical coordination among extremists disproportionately. Second, the “outrage machine” may

¹Empirical investigations have shown that exposure to hate speech online is prevalent and Twitter users who post about politics tend to have more extreme ideological preferences than the general population (Oksanen et al., 2014; Barberá et al., 2015). Evidence suggests that tweets using more moral or emotional language—plausibly correlated with the intensity or extremity of opinion—diffuse more intensely on Twitter (Brady et al., 2017). Finally, Beknazar-Yuzbashev et al. (2023) show in a randomized experiment that hiding “toxic” text content from users on Facebook and Twitter reduces content consumption, implying that the toxic content indeed boosts engagement.

²It is worth noting that the empirical literature on the implications for social media and political polarization and partisan hostility remains inconclusive. For example, Boxell et al. (2017) find that recent growth in political polarization in the United States is driven by demographic groups that are the least likely to use the Internet and social media and Barberá (2015) finds that in Germany, Spain and the United States, most social media users have ideologically diverse networks online and that exposure to this diversity increases political moderation. On the other hand, Melnikov (2023) finds that, access to 3G internet contributed to increased political polarization between Democratic and Republican voters, Lelkes et al. (2017) find that broadband internet availability increases partisan hostility and segregation in the consumption of partisan media, and Allcott et al. (2020) find that deactivation of Facebook accounts reduced political polarization of the account owner. See Zhuravskaya et al. (2020) for a review of the literature.

alter people’s beliefs about the popularity of extreme views. If individuals’ protest participation are strategic compliments as often is theorized in the literature ([Edmond, 2013](#); [Passarelli & Tabellini, 2017](#); [Barbera et al., 2020](#)), this may increase extremists’ willingness to act upon their opinions in the form of public demonstration.³

Following [Müller & Schwarz \(2023\)](#) and [Fujiwara et al. \(2021\)](#), this paper relies on a plausibly exogenous shock to Twitter’s early adoption for identification: The South By Southwest festival (SXSW) in Austin, Texas in March 2007. Promotion of Twitter at the festival sparked its popularity among attendees and, through network effects, in their home counties. The promotion was so successful that county-level festival attendance, proxied by the number of current Twitter users who created their Twitter account in March 2007 and follow the official account of SXSW, is still to this day predictive of contemporary county-level Twitter penetration. Further, counties with SXSW Twitter followers who created their accounts in 2006 and follow SXSW, and counties with SXSW Twitter followers who created their accounts in March 2007 are statistically similar in terms of observable characteristics, while only the latter Twitter users are predictive of contemporary Twitter penetration. This paper uses the variation in Twitter penetration predicted by SXSW followers who created their accounts in March 2007, while controlling for pre-2007 interest in the festival proxied by the SXSW followers who created their accounts in 2006, in an instrumental variable framework.

Using the Armed Conflict Location and Event data, aggregated at the county-level, this paper classifies demonstration events depending on whether “extreme” groups—groups that are particularly militant, radical, or hateful—are involved, depending on the topic and political alignment of the protest, and depending on aggressive behaviour. In accordance to the literature (e.g., [Enikolopov, Makarin, & Petrova, 2020](#)), it finds that Twitter penetration can substantially fuel protest mobilization. A one percent increase in Twitter penetration implies around a one percent increase in the frequency of demonstration events overall between January 2020 and 12th of November 2021. In addition, Twitter penetration disproportionately fuels protest events where “extreme”

³The idea of strategic complementary in protest participation assumes that the cost of participation is lower when the protest is larger (e.g., the likelihood of being arrested may be lower in a larger crowd) or that the likelihood of “success” (e.g., the extent to which demonstrators can influence policy or the attitudes of the general population) increases with size. Evidence of strategic complementarity in protest participation is found in some settings [Manacorda & Tesei \(2020\)](#) but not always. [Cantoni et al. \(2019\)](#) find experimental evidence of strategic substitutability in the context of Hong Kong’s ongoing anti-authoritarian movement. They provide suggestive evidence of 3 sources of strategic substitutability: Public Good characteristics of protest participation, greater perceived likelihood of government crackdown in larger protests, and that subjects perceive participating in smaller protests as a stronger signal of their ideological “type.”

groups are involved, both for extreme left and extreme right groups. In counties with at least one demonstration event over the period, moving from the median to the third quartile of the log number of Twitter users (residualized by population) implies an increase in the share of demonstrations with extreme group involvement by about 1.3 percentage points, from an average of 2.4%. Twitter penetration also implies an increase in the relative frequency of demonstration events where demonstrators exhibit aggressive or violent behaviour, although this effect are somewhat imprecisely estimated. The effect of Twitter penetration on protest frequency does not vary substantially by the topic for protest nor by the political alignment of the protest events, implying that the disproportionate effects for extreme groups are not purely topic driven.

Analysing individual level data from the 2020 Cooperative Election Study this paper finds some evidence supporting the idea that county-level Twitter penetration increases the intensity of opinion concerning statements on racial and gender issues among county residents. However, these effects are only found for a minority of statements on these issues. With data from the 2017 wave of the World Value Survey, it finds that county-level Twitter penetration implies an increased willingness to justify violence against others, but these effects are not found for willingness to justify “political violence.” While the literature has highlighted how social media fuels protest activity by reducing the cost of coordination (e.g., [Enikolopov, Makarin, & Petrova, 2020](#); [Fergusson & Molina, 2019](#)), this individual-level evidence suggest that Twitter may not only serve as a more powerful coordination tool in the hands of extreme groups, but that Twitter penetration could increase the appeal of such groups—many of which are defined here as “extreme” due to explicit willingness to use violence. However, the effects are not consistently found across outcomes. Analysing the effects of regional level treatments on individual level outcomes also limits interpretability as we do not learn about the effects of individuals’ social media use, but about the effects of Twitter usage in respondents’ county. Therefore, these results should be interpreted only as suggestive of this mechanism.

This paper contributes to the literature on the effects of media on political outcomes ([Enikolopov et al., 2011](#); [Gentzkow et al., 2011](#); [Gentzkow, 2006](#); [DellaVigna & Gentzkow, 2010](#)), and especially the role of the Internet and social media ([Falck et al., 2014](#); [Guriev et al., 2021](#); [Zhuravskaya et al., 2020](#)). By examining specially demonstrations concerning particularly salient topics on which the US population holds polarized views on and those where extreme groups are involved, it contributes to understanding on the links between the internet and social media, and political polarization

(Gentzkow & Shapiro, 2011; Halberstam & Knight, 2016; Boxell et al., 2017; Barberá, 2015; Yanagizawa-Drott et al., 2020; Melnikov, 2023; Lelkes et al., 2017). Further, by examining if Twitter penetration increases the relative frequency of involvement of extreme groups (in particular, the involvement of white supremacy groups on the far-right) during demonstration events, and the relative frequency of aggressive and violent behaviour, it contributes to the literature on media's, in particular social media's, role in propagating violence and hateful sentiments (Müller & Schwarz, 2021, 2023; Bursztyn et al., 2019; Yanagizawa-Drott, 2014; Adena et al., 2015; DellaVigna et al., 2014).

This paper is directly related to the growing literature on the information and communication technologies', and social media's, ability to facilitate protest activity. Amorim et al. (2022) find that broadband Internet availability increased the probability of an Occupy Movement protest event. Enikolopov, Makarin, & Petrova (2020) show that VKontakte (VK), the most popular social media application in Russia, facilitated protests during a protest wave in 2011 triggered by electoral fraud benefiting the incumbent party, United Russia. Casanueva et al. (2021) show that Twitter penetration played an important role in mobilizing new protesters and broadening coalitions for the Black Lives Matter movement in the United States. However, all of these papers study a single protest movement and thus cannot explore differential effects depending on the characteristics of movements. Other studies take a more aggregate approach. Mobile phone coverage has been shown to increase protest activity across Africa (Manacorda & Tesei, 2020) and globally (Christensen & Garfias, 2018). Fergusson & Molina (2019) find that Facebook's release in a new language implies increasing protest activity in countries where the language is commonly spoken. These studies also find that ICTs can limit violent conflict (Fergusson & Molina, 2019) and government repression during protest events (Christensen & Garfias, 2018), consistent with the conjecture that ICTs can increase visibility of and thus deter violent action (Durante & Zhuravskaya, 2018).

The main contribution to the literature is to examine directly the effects of social media on protest activity, depending on the characteristics of protests movements themselves. This paper classifies protests events depending on the involvement of extreme groups, on the topic of demonstration, and the political alignment of demonstration, and studies differential effects of social media penetration depending on these characteristics. Further, compared to Christensen & Garfias (2018) and Fergusson & Molina (2019) who examine government repression and violent conflict, it distinguishes between non-peaceful behaviour of demonstrators themselves, versus outside parties, where heterogeneities are found.

2 Data and Background

2.1 Protests in the United States

Contemporary United States provide a good setting to test whether social media influences the composition of demonstrations with respect to the groups demonstrating, the topics for protest, and with respect to aggressive or violent behaviour during demonstration events. The United States has been a vibrant protest environment throughout the years, and recently demonstration activity has been surging. As an example, [ACLED \(2020c\)](#) recorded more demonstrations in the United States than any other country, except for India, covered in their almost-global data set during the initial 3 months of data collection in the United States. US citizens do not only protest a lot, they also engage in protests around a diverse set of issues. In 2020, three issues were a particularly common cause for protests: a) Demonstrations associated with the Black Lives Matter (BLM) movement, protesting police brutality against Black people in the United States, following the police murder of George Floyd. b) Demonstrations related to the Presidential elections in 2020. These include rallies in support of or against one of the presidential candidates, as well as protests against the election results, alleged Democratic voter fraud and “illegal” ballot counting, or protests demanding the results to be respected. c) Demonstration linked to the Coronavirus pandemic ([ACLED, 2020a](#)). In addition to these three highly visible and oftentimes dividing topics, the ACLED data records numerous other causes for protests in the United States, ranging from demands concerning laborer’s rights or compensation, local or federal environmental policy, human or animal rights topics, or immigration issues. Further, there are increasing signs of political instability and aggression during demonstration events. Comparing 2020 to 2019, the start of their data collection in the United States, ACLED reports that counter-protests became more frequent, government authorities increasingly engaged with and exhibited force against protesters, and armed non-state actors were increasingly visible in demonstration events, either participating in protests, engaging in “peacekeeping,” or intimidating demonstrators ([ACLED, 2020b](#)).

2.1.1 Protest data and main variables

Data on protest come from the Armed Conflict Location and Event Data (ACLED) project, introduced in the descriptive analysis in the section above ([Raleigh et al., 2010](#)). ACLED is an event-level dataset that documents the location, date, and characteristics of each event. ACLED researchers manually compile events from regional, national,

and international media outlets deemed reliable and are supplemented by reports by international institutions or non-governmental organizations (NGOs). Unfortunately, limited information on the number of individual participants is available in the data set. The data used is from January 2020 to 12th of November 2021, the latest available data at the time of analysis.

ACLED documents various political violent events and conflict, defined as use of force by a group with a political purpose or motivation. It also documents events that are “potential pre-cursors or critical junctures of a violent conflict,” such as protests and riots, these events will be the focus of this study. Demonstration events are defined as all physical congregations of three or more people when they are directed against a political entity, government institution, policy, group or individual, tradition or event, businesses, or other private institutions.

Three variables are particularly helpful for protest categorization. First, for each event the associated groups, organizations, or (where relevant) the salient racial/ethnic identity of demonstrators are documented as “associated actors.” There can be many such associated actors associated with one event. There are 574 unique associated actors coded in the data, where each event is associated to 0-18 of these actors. Second, each event is accompanied by a “note” with qualitative information about the event.⁴

The main focus is to identify demonstrations events where “extreme” groups are involved, versus more mainstream demonstration events. A political group is defined here as “extreme” if one of the following conditions hold: a) The group is aggressive and willing to use violence or militant tactics to achieve its’ goals. For example, the Black Lives Matter movement is defined as mainstream, while “Detroit Will Breathe”, a political group also protesting against police brutality towards black lives but explicitly self-defining as militant, is classified as extreme. b) The group is “revolutionary”, in the sense that it poses radical opposition to current political systems or institutions (e.g., electoral institutions, property rights, communist ideologies). c) The groups subscribes to nationalists ideologies or is a hate group. The full list of political groups and organizations defined as extreme is found in Table A1 in the Appendix.

Next, events including a set of keywords in their “note” and events with associated actors corresponding to a particular topic are coded as such. First, events including keywords such as “anti-abortion,” “pro-life,” or “reproductive rights” are coded as demonstrations concerning abortion rights. Second, all events including a “tag” in

⁴An example of a note accompanying an event is the following: “On 12 November 2021, a group of protesters gathered outside of Ascension Via Christi hospital in Wichita (Kansas) to protest against coronavirus vaccine mandates for health workers. Unvaccinated employees joined the protest after finishing their final shift before the mandate took effect that day.”

the note representing that the protest was associated with the 2020 post-presidential election unrest (“stop the steal” or “count every vote”) are coded as part of a 2020 post-presidential election unrest topic. Third, events including keywords such as “vaccine,” “mask mandate,” or “coronavirus prevention measures” are coded as part of a corona virus pandemic topic. The goal is to disentangle demonstrations strictly related to public health measures and those related to economic relief of specific groups *due to* the pandemic. General public health measures were a deeply dividing and politicized topic in the United States, while demands of economic relief, better working conditions, higher wages or banning of evictions for specific groups due to the pandemic might be seen as simply advocating for specific interest groups and not as polarizing a topic. Fourth, events related to the Black Lives Matter (BLM) movement (and police brutality against minorities) are coded using associated actors. Associated actors related to the movement were manually coded and include for example “Black Lives Matter,” “Antifa,” “Blue Lives Matter,” or “Back the Blue.” Fifth, a broader racial issues topic-category is defined which includes in addition to the BLM associated actors such as “Stop Asian Hate,” “The National Association for the Advancement of Colored People,” “The White Defence Force,” or “The Aryan Cowboys Brotherhood.” Note that in general, the categories described are not mutually exclusive. One protest event can be included in e.g., both the “Covid related” category if it includes the covid-related keywords, and in the “BLM related” category if organizations and movement usually demonstrating against police brutality against black people are involved. Further, a protest event can be defined as extreme and e.g., “racial issues related”, if the extreme group concerns racial issues.

For the extreme protests and all topics except for the “Covid related” category, events can also separate along broad left-right political alignment. E.g. Pro-Choice abortion related protests, or those where “Black Lives Matter” is an associated actor would be defined as left-leaning, while Pro-Life or “Blue Lives Matter” protests would be defined as right-leaning. A full list of keywords, phrases, and associated actors used to code each of the topics is provided in Tables [A2](#) to [A6](#) in the Appendix.

Finally, events are described by the *event type* and *sub-event type*. Demonstration event types are Protests or Riots. Protests are events where demonstrators do not engage in violence or disruptive acts such as property destruction, although violence may be used against them. Riots are events where demonstrators engage in disruptive acts such as property destruction or violence without the *use* of lethal weapons. Protest are further categorized into three sub-types: *Peaceful protests* are events where demonstrators are not engaging in violence or any other form of rioting behaviour, and are not

faced with any sort of force or engagement. *Protests with intervention* are events where demonstrators are peaceful, but there is an attempt to disperse or suppress the protest without serious/lethal injuries being reported. *Protests with excessive force against protesters* are events where peaceful protesters are targeted with violence leading to (or if it could lead to) serious/lethal injuries. Riots on the other hand are categorized into two sub-types: *Violent demonstrations* are events where demonstrators engage in disruptive or violent behaviour, examples of disruptive behaviour would include vandalism, road-blocking using barricades, or burning tires. *Mob violence* are events when rioters violently interact with other rioters, another armed groups or civilians, outside of demonstrations and without the *use* of lethal weapons. This paper examines heterogeneous effects for a) *Peaceful protests*, b) Demonstrations where *peaceful protesters met with intervention or violence*, which comprise of the sub-events *Protests with intervention* and *Protests with excessive force against protesters*, and c) Demonstrations with *violent protesters*, which comprise of all Riot event types.

Figure 1 shows the spacial distribution of the main outcome variables. The upper map shows the number of overall demonstration events (per capita) between January 2020 and November 2021. The lower map shows the share of demonstrations where extreme groups are involved, in counties where at least 1 such demonstration occurred. In total, 57% of counties had demonstration events over the period of analysis, and 12% had demonstrations with extreme group involvement. Figures A1 to A3 in the Appendix show histograms of the distribution, and additional maps of the spacial distribution, of these main outcome variables.

[Figure 1 around here]

2.2 Twitter and the South by Southwest festival

In 2021, surveys indicate that a large majority of Americans self-report ever using any social media site, and that their share has remained relatively stable over half a decade. These data suggests that that the most popular platforms are YouTube, Facebook and Instagram, used by 81%, 69% and 40% of those who respond to survey questions. 23% of U.S. adults used Twitter, a similar number of users as for Snapchat (25%) and WhatsApp (23%). Like with most other social media platforms, Twitter is most popular among young, educated and urban adults ([Pew Research Center, 2021](#)).

Although Twitter was founded in March 2006, it remained largely unknown before the South by Southwest festival (SXSW hereafter) in March 2007. The importance of

the festival for early adoption has been explicitly stated by Evan Williams, a co-founder of the platform. At the festival, Twitter had screens in hallways and created an event-specific feature where one could easily create an account and broadcast posts on these screens. Those who created accounts through this feature were automatically made to follow “Twitter ambassadors,” festival attendees already on the platform (Quora, 2011). Fujiwara et al. (2021) and Müller & Schwarz (2023) document quite thoroughly the immediate growth in Twitter penetration after the festival in 2007, and show that this growth occurs primarily in the counties of festival goers. But of course the platform did not reach its later popularity immediately. Activity went from an average of 5,000 tweets per day in 2007, to 300,000 in 2008, numbers dwarfing in comparison to the 500 million tweets sent each day on average in 2019.

The SXSW festival itself is an ambitious project combining film, interactive media and music festivals with professional conferences with a focus on creative industries and technology. These bundles of events take place annually in mid-March in Austin, Texas. The festival advertises itself as an essential destination for global professionals where, in addition to film screenings, music and comedy shows, there are ample opportunities for professional development and networking (see the official homepage of the festival: <https://www.sxsw.com/about/>). The festival is wildly popular, in 2016 it attracted around 300,000 official attendees (Theis, 2016). Already by 2007, the festival attracted speakers, performers and attendees from all around the United States and beyond. Broadly speaking, the festival’s content can be described as combining a wide range of western pop-culture entertainment with professional themes of technological innovation.⁵

2.2.1 Twitter and South by Southwest festival attendance data

I am indebted to and thank Karsten Müller and Carlo Schwarz for generously sharing their data for both proxies of Twitter penetration and for SXSW festival attendance.

Twitter penetration at the county level is proxied utilizing a sample of 475 million geo-coded tweets collected by Kinder-Kurlanda et al. (2017). These tweets are all from within the United States, and were collected from 1st of June to the end of

⁵For example, in 2007 the festival included performances from musicians from all around the United States and international acts such as Peter, Bjorn & John from Sweden, Amy Winehouse from the UK, and YB from South Korea. The festival premiered Judd Apatow’s popular comedy film “Knocked Up” to the world, and held panel sessions on topics such as “Virtual Worlds and Virtual Humans” with more than 450 speakers from the media and information technology sector. Examining the program for 2006 is suggestive of similar scope, size and themes, and can be found here: <https://www.sxsw.com/about/history/>.

November, in 2014 and 2015 through the Twitter Streaming API, which returns a 1% sample of daily tweets as long as it is called. Only tweets that can be geo-located are included in the sample. Twitter allows users to “geo-tag” their tweets, adding the latitude/longitude information from one’s device’s GPS sensor to the post. This information therefore refers to the location from which each tweet was posted. The tweets in the dataset are already assigned to counties. To get a measure of users per county, the user profiles for each tweet in the database is collected and the users are then assigned to the county from which they tweet most frequently.

It is worth noting that, as geo-tagging tweets is optional, the observed sample is highly self-selected. In fact, only about 1% of users chose to geo-locate their tweets. Further selection issues arise as during the data-collection period, Twitter introduced a new way for users to share their location by tagging tweets with the name of their location from a dropdown list, as opposed to the latitude/longitude pair. The data collection did not catch tweets sharing this location information, but this new option substantially reduced the popularity of “geo-tagging” tweets (Kinder-Kurlanda et al., 2017). However, Fujiwara et al. (2021) show that measures using these tweets closely resemble other measures of Twitter penetration, such as estimates of the number of Twitter users over time from Statista or survey estimates of county-level number of Twitter users, alleviating some concerns about self-selection biases.

[Figure 2 around here]

South by Southwest festival attendance is proxied using data collected through the Twitter API. User data on each account following the official Twitter account of the SXSW festival (@SXSW) is webscraped, and assigned to counties based on the location users report in their user profile. The two crucial variables constructed with this data are: a) The county-level number of SXSW followers who created their Twitter account in March 2007, 2007 SXSW followers hereafter, and b) the number of SXSW followers that created their Twitter account in 2006 (the founding year of Twitter), 2006 SXSW followers hereafter. The number of SXSW followers who created accounts in March 2007 are taken as a proxy for county-level attendance at the festival, the excluded instrument, while those creating accounts in 2006 are used to control for general interest in the festival. A total of 855 SXSW followers who created their accounts in March 2007 could be located within the United States, into a total of 155 counties. A total of 209 followers created their accounts in 2006 were located, into a total of 67 counties. Figure 2 shows the spacial distribution of Twitter users per capita and of the SXSW followers across continental United States, while Figures A4 to A6 in

the Appendix show histograms of the distributions and additional maps of the spacial distribution of the Twitter users.

2.3 Additional data sources

Data to measure indications of radicalization come from two surveys: The 2020 wave of the Cooperative Election Study (CES) and the World Value Survey (WVS) wave 7, administered in 2017 in the United States (Schaffner et al., 2023; Haerpfer, 2022). In the CES, respondents answered both a pre-election wave conducted between September 29 to November 2, and a post-election wave conducted between November 8 to December 14. It includes 61,000 respondents and was sampled to represent the general adult population from each state. The WVS used stratified sampling, made to represent the population’s age, race/ethnicity, education, and gender distributions in the United States (46 strata in total). 5,375 eligible US panelists were invited to participate in the survey and offered the cash equivalent of \$15 for completing the survey. 2,596 respondents completed the interview. Both surveys include geographic information on respondents.

The CES does not include questions that directly capture radicalization, or would be clearly indicative of involvement in the “extreme” political groups defined above. However, respondents were asked about their attitudes towards various statements on the police, and racial or gender issues, and to rate themselves and political actors on a ideology scale from very liberal to very conservative. Intensity of opinions is measured using dummy variables equal to one when individuals answer the options with the highest intensity in either end (e.g., “strongly agree” to a statement rather than “somewhat agree”). While the WVS includes a much smaller sample, it includes questions that could more directly be linked to radicalization. In particular, respondents are asked to rate whether a) Violence against other people, b) Political violence, c) Terrorism as a political, ideological or religious mean, can be justified on a 10-point scale, from “Never justifiable” to “Always justifiable.”

Additionally, various county-level demographic, geographical, socio-economic, and voting result variables are included as controls. For the most part, these variables are from 2009 and 2010, which can be considered a “pre-treatment” period as Twitter did not gain immense popularity until later.⁶ Data on population size, age, race, and ethnic composition from the 2010 U.S. Census and also population size data from the 2020 U.S.

⁶In 2010, worldwide Twitter users were around 30-50 million, about 10-15% of the 300 million users estimated in 2015. Since 2015, the growth of Twitter users has remained relatively modest, with between 300 and 350 million users until 2019 (Statista, 2019)

Census to calculate population growth from 2010-2016. Next geographical variables such as the square kilometer area of each county, the linear distance between Austin, Texas (the location of the SXSW festival) and the centroid of each county, as well as the population density in each county, are calculated using the U.S. Census tigerline 2020 shapefile. Estimates of poverty rates and median household income are from the 2009 US Census Bureau’s SAIPE program. Estimates of educational attainment (in particular, share of adult population with at least a high school diploma and with at least a Bachelor degree) in 2010 come from the American Community Survey, and Unemployment and Employment-to-Population rates in 2009 from the Bureau of Labor Statistics. Finally, data on county-level presidential election results in 2000 (George W. Bush), 2004 (George W. Bush), and 2008 (John McCain) are from the MIT Election Lab. Table A7 in the Appendix describes the full set of controls.

3 Empirical Strategy

If Twitter were an important facilitator of protests, we would expect to see greater demonstration activity in areas where Twitter penetration is higher. The relationship of interest is:

$$Y_c = \beta_0 + \beta_1 \text{Log}(\text{TwitterUsers}_c) + \beta_2 \text{Log}(1 + \text{SXSW}_c^{2006}) + \mathbf{X}'_c \boldsymbol{\beta} + \epsilon_c \quad (1)$$

Where Y_c represents a demonstration variable in county c (e.g., the share of demonstrations involving extreme groups), $\text{Twitter penetration}_c$ represents a measure of Twitter penetration, here the natural logarithm of unique Twitter users in the [Kinder-Kurlanda et al. \(2017\)](#) data set, in county c , \mathbf{X}'_c is a vector of covariates, β ’s are coefficients to be estimated and ϵ_c is the error term. When looking at overall demonstration activity, we would expect β_1 to be positive, representing the positive effect of Twitter penetration on demonstration activity. However, Twitter penetration may well be endogenous to demonstration activity for various reasons, giving alternative interpretations to β_1 . First, unobserved county characteristics might drive both Twitter adoption and demonstration activity. As an example, as Twitter allows users to follow politicians and express their political opinions, differences in inherent interest in political activism might drive county-level differences in Twitter adoption—leading to a spurious correlation and overestimation of the impact of Twitter penetration on demonstration activity. Second, as social medias such as Twitter facilitate protest event coordination, one might expect counties high regional protest activity to drive

individuals to join Twitter, to ease coordination for demonstrations they would have attended even without the existence of Twitter.⁷

To circumvent these concerns, identification relies the plausibly exogenous shock to early Twitter adoption in the United States connected to the 2007 SXSW festival, inspired by [Fujiwara et al. \(2021\)](#) and [Müller & Schwarz \(2023\)](#). As discussed above, the 2007 conferences provided a tipping point in the rise of Twitter’s popularity. Consistent with the literature on path dependence in technology adoption ([Arrow, 2000](#); [Liebowitz & Margolis, 1995](#)) and following [Fujiwara et al. \(2021\)](#) and [Müller & Schwarz \(2023\)](#), a persistent effect of festival attendance on the spacial distribution of Twitter penetration is found. This setting lends itself well to an instrumental variable strategy where county-level Twitter penetration is instrumented for by the county-level attendance at SXSW in 2007 and subsequent adoption, measured by SXSW Twitter account followers who created their account around the time of the festival.

It is not immediately obvious that county-level festival attendance and subsequent early adoption of Twitter is exogenous to protest participation. One crucial concern is that counties with festival goers or general interest in the festival might be systematically different from other counties, in ways that correlate with demonstration activity. Information on the exact timing of Twitter adoption of the SXSW followers helps alleviate these concerns. Consistent with prior evidence, SXSW followers that created their accounts in March 2007, presumably after attending the festival, are robustly and persistently predictive of the spacial distribution of Twitter penetration, while the county-level number of SXSW followers who created their accounts in 2006 are not predictive of Twitter penetration. Further, counties with SXSW March 2007 followers are statistically similar to counties with SXSW 2006 followers as shown in [Table A8](#). This allows to disentangle the effect of general interest in the SXSW festival and early adoption of Twitter from the effect of current Twitter penetration. By including the number of SXSW followers who created their account prior to the festival (in 2006), general interest in SXSW, and early interest in Twitter, are effectively controlled for. Moreover, this provides an placebo test by testing whether inherent interest in the festival has an effect on protest activity. It turns out that throughout the analysis, SXSW followers who created their accounts in 2006 rarely correlate with demonstration related outcomes—in the few cases where they correlate to demonstration outcomes, the correlation is of the opposite sign compared to 2007 followers, i.e., negative. Formally,

⁷Note that this reverse causality problem can only occur in the data if demonstration activity pre-2014 drives Twitter adoption, and pre-2014 demonstration activity is predictive of demonstrations in 2020 and 2021.

I estimate the following equations:

$$\begin{aligned} \text{Log}(\text{TwitterUsers}_c) = & \alpha_0 + \alpha_1 \text{Log}(1 + \text{SXSW}_c^{\text{March2007}}) + \alpha_2 \text{Log}(1 + \text{SXSW}_c^{2006}) \\ & + \mathbf{X}'_c \boldsymbol{\alpha} + v_c \end{aligned} \quad (2)$$

$$\begin{aligned} Y_c = & \gamma_0 + \gamma_1 \text{Log}(1 + \text{SXSW}_c^{\text{March2007}}) + \gamma_2 \text{Log}(1 + \text{SXSW}_c^{2006}) \\ & + \mathbf{X}'_c \boldsymbol{\gamma} + \mu_c \end{aligned} \quad (3)$$

Where equation (2) represents the first stage and equation (3) the reduced-form estimation.⁸ $\text{SXSW}^{\text{March2007}}$ represents the number of 2007 SXSW followers and SXSW^{2006} represents the number of 2006 SXSW followers in county c . α 's and γ 's are coefficients to be estimated, and u_c and μ_c are error terms. Other notation remains the same as for equation (1). The accompanying second stage equation can be written as:

$$Y_c = \lambda_0 + \lambda_1 \text{Log}(\hat{\text{TwitterUsers}}_c) + \lambda_2 \text{Log}(1 + \text{SXSW}_c^{2006}) + \mathbf{X}'_c \boldsymbol{\lambda} + \phi_c \quad (4)$$

Where $\text{Log}(\hat{\text{TwitterUsers}}_c)$ refers to the predicted value of the county-level measure of Twitter penetration from equation (2), λ 's are coefficients to be estimated, ϕ is the error term, and other notation is the same as before. Here, λ_1 represents the Local Average Treatment Effect (LATE), the effect of increased Twitter penetration for counties where Twitter adoption would be strongly affected by being the home of an SXSW festival goer in 2007. Intuitively, early adoption of Twitter among festival attendees is likely to increase Twitter penetration in attendees home counties if the SXSW attendees have large offline social networks in their home counties and their networks are receptive to early adoption of a new social media platform. This is likely the case in counties with a non-negligible population of “tech savvy” individuals interested in adopting a new social media platform, similar to SXSW enthusiasts in terms of socio-demographics or cultural interest. However, as discussed above, the festival offers a wide variety of popular culture attractions and should attract the interest of diverse groups of people. The LATE should thus not be excessively particular, although one should not overgeneralize the results.

For individual level survey data, the empirical strategy follows the county-level

⁸In the Appendix I repeat the main analyses using the inverse hyperbolic sine of SXSW followers instead of the logarithm of one plus SXSW followers in the first and second stage. Main results are robust to this alternative specification.

analysis closely. I estimate the following equations:

$$\begin{aligned} \text{Log}(\text{TwitterUsers}_{ic}) &= \theta_0 + \theta_1 \text{Log}(1 + \text{SXSW}_c^{\text{March2007}}) + \theta_2 \text{Log}(1 + \text{SXSW}_c^{2006}) \\ &\quad + \mathbf{X}'_{ic} \boldsymbol{\theta} + \eta_{ic} \end{aligned} \quad (5)$$

$$\begin{aligned} \text{Attitude}_{ic} &= \pi_0 + \pi_1 \text{Log}(\hat{\text{TwitterUsers}}_c) + \pi_2 \text{Log}(1 + \text{SXSW}_c^{2006}) \\ &\quad + \mathbf{X}'_{ic} \boldsymbol{\pi} + \psi_{ic} \end{aligned} \quad (6)$$

Where equation (5) represents the first stage and equation (6) the second stage. $\text{Log}(\text{TwitterUsers}_{ic})$ are the natural logarithm of unique Twitter users in individual i 's county c , $\text{SXSW}_c^{\text{March2007}}$ and SXSW_c^{2006} are the number of SXSW followers in March 2007 and 2006 respectively, in individual i 's county c , and \mathbf{X}'_{ic} is a vector of county-level and individual level covariates. $\text{Log}(\hat{\text{TwitterUsers}}_c)$ is the predicted value of $\text{Log}(\text{TwitterUsers}_{ic})$ from the first stage and Attitude_{ic} are the answers to survey questions for individual i residing in county c . θ 's and π 's are coefficients to be estimated and η_{ic} and ψ_{ic} are error terms.

Note that the individual level analysis studies the effect of county-level Twitter penetration on individual-level attitudes. Thus, they cannot tell us directly how individual's usage of Twitter influences their attitudes, but rather whether Twitter popularity in one's local environment influences attitudes. In addition, the surveys used have either only indirect questions on radicalization, or quite small sample size. It is therefore natural to interpret results from this analysis not as an authoritative answer to the question on Twitter's contribution to radicalization, but only as indicative evidence on the question.

4 Baseline results

4.1 The first stage

Table 1 shows regression results for the first-stage relationship between the county-level number of SXSW 2007 followers and Twitter penetration, measured as the natural logarithm of the number of Twitter users from the sample of tweets from 2014-2015. It shows that, as discussed above, the SXSW 2007 followers correlate quite robustly to the logarithm of Twitter users, while the SXSW 2006 followers do not. The relationship between the instrument and Twitter users becomes more precisely estimated once main controls are included. Quantitatively, a standard deviation increase in the natural logarithm of SXSW 2007 followers plus one (0.319) implies a 14% increase in the

number of Twitter users, given the estimate of 0.417 in column (5).

The robust F-statistic is well above the traditional rule-of-thumb threshold of 10 and takes values around 35-40 in the most stringent specifications. As the model is just-identified (i.e., does not include more excluded instruments than instrumented variables of interest), the robust F-statistic (sometimes also called the Kleibergen Paap) corresponds to the effective F-statistic developed by [Olea & Pflueger \(2013\)](#). The F-statistics are thus above the threshold for a potential bias of 10% and a significance level of 5%, which is 23. Further, [Andrews et al. \(2019\)](#) show that for higher-than-10 F-statistics the weak-instrument problem is unlikely to affect the validity of conventional t-statistics in the case of clustered standard errors. This alleviates concerns that the results presented while controlling properly for county characteristics suffer from a weak first stage. Panel A in Table [A9](#) in the Appendix reports the first stage using the inverse hyperbolic sine of SXSW 2007 and 2006 followers instead of the logarithm of one plus SXSW followers, where results remain robust: SXSW 2007 followers correlate robustly with the logarithm of Twitter users while the SXSW 2006 followers do not, and F-statistics are around 40 in the most stringent specifications.

[Table [1](#) around here]

This relationship is unlikely to be driven by selection on unobservables. Applying the approach from [Oster \(2019\)](#) and comparing a specification with only a control for $\text{Log}(1 + \text{SXSW}_i^{2006})$ to the one with all controls (column 5) and, assuming unobserved factors could explain all variation in demonstration frequency, gives an “Oster- δ ” of 2.23. This suggests, that for the true effect of SXSW 2007 followers on the logarithm of Twitter users to be zero, unobservable variables would have to be about 2.2 times as important as the controls added in column (5) in terms of selection into “treatment”.⁹ Given the inclusion of numerous controls, the scale of selection on unobservables needed to explain away the relationship is implausibly high, which is reassuring for causal interpretation in what follows.

Figure [3](#) visualizes the first stage of the main specification. The upper panel presents a binned scatter plot of the relationship between the natural logarithm of Twitter users and the number of SXSW 2007 followers, residualized by partialling out the full set of controls. The yellow dots represent bins, the blue line represents a linear

⁹Note that assuming unobservables can fully explain variation in Twitter penetration is very conservative. This might not be the case in the presence of outcome measurement error. Assuming for example, that 5% of variation in Twitter penetration is left unexplained by unobservables and doing the same comparison gives an “Oster- δ ” of 18.2.

regression relationship and the shaded area shows 95% confidence intervals, calculated using heteroscedasticity robust standard errors clustered at the state-level. The lower panel plots the same relationship between the natural logarithm of Twitter users and the SXSW 2007 followers. We see a statistically significant relationship between SXSW 2007 followers and Twitter penetration, but no relationship exists between the 2006 SXSW followers and the natural logarithm of Twitter users. Figure A7 in the Appendix shows the same plots with additional grey dots representing individual observations.

[Figure 3 around here]

4.2 IV results

Table 2 presents estimations from the Two Stage Least Square (2SLS) model using the same five specifications as before, the table provides three sets of results: Panel A shows county-level OLS correlation between Twitter users and the number of demonstrations, Panel B shows the reduced-form, and Panel C reports the second stage results, the local average treatment effect of Twitter penetration on demonstration frequency. As the relationship between Twitter penetration and the frequency of demonstration events is likely proportional, and a substantial proportion of counties have zero events, the table presents results using the inverse hyperbolic sine of the number of demonstrations as the outcome variable. Panel B in Table A9 in the Appendix reports the second stage results using the inverse hyperbolic sine of SXSW followers instead of one plus the logarithm of SXSW followers, and results remain robust. Further, as the inverse hyperbolic sine transformation has come under scrutiny recently, in particular as estimates may be sensitive to arbitrary choices regarding the unit of measurement (see e.g., [Aihounton & Henningsen \(2021\)](#)), results from Table 2 are replicated using alternative functional forms for the outcome variable in Table A10 in the Appendix, where results remain comparable.

In Panel A we see a positive, significant, but quantitatively somewhat unimportant correlation between Twitter users and the number of demonstrations. After adding the main controls, a percentage increase in Twitter users is associated with approximately a 0.3% increase in the frequency of demonstration events.

Panel B presents reduced-form relationship between the county-level number of SXSW 2007 followers and the total number of demonstrations. Across all reported specifications, there is a positive relationship between the number of SXSW 2007 followers and the number of demonstrations significant at the 1% level. After the inclusion

of state fixed-effects and population controls, coefficient estimates remain largely stable across specifications. On the contrary, the coefficient estimate of the effect of SXSW 2006 followers is small, imprecisely estimated and statistically indistinguishable from zero. Figure 4 visualizes the regression from column (5) using binned scatter plots. The upper panel presents a binned scatter plot where the x-axis represents the natural logarithm of (one plus) SXSW 2007 followers, while in the lower panel the natural logarithm of (one plus) SXSW 2006 followers are on the x-axis. Both variables on the y-axis and the x-axis are residualized by partialling out the full set of controls. Figure A8 in the Appendix shows the same plots with additional grey dots representing individual observations.

[Figure 4 around here]

Panel C presents the second stage specification and the Local Average Treatment effect of Twitter penetration for compliers. The effect is statistically significant at the 1% significance level and quite stable across specifications. The effect is very large: A percentage increase in Twitter penetration implies around a percentage increase in the frequency of demonstrations between January 2020 and November 2021. SXSW 2006 followers do not seem to correlate with demonstration frequency. Coefficient estimates are much lower in the OLS estimates in Panel A, where endogeneity is not dealt with, compared to 2SLS estimates in Panel C. The comparison of OLS and 2SLS results from Enikolopov, Makarin, & Petrova (2020), who use a similar identification strategy—a shock to early adoption as a plausibly exogenous source of variation in regional VK penetration—to estimate the effect of social media on protest participation in Russia gave similar results. These differences could be due to large negative selection of counties into Twitter penetration, for example that people who are in general less interested in societal issues or political activism and thus less likely to participate in demonstrations are more likely to join Twitter. Alternatively, the effects estimated in Panel C could reflect particularities in the local average treatment (LATE), i.e., that the effect of Twitter on demonstration activity is larger in counties where the effect of SXSW followers on Twitter penetration was also stronger.

[Table 2 around here]

Taken together, these results can be taken as evidence that Twitter may facilitate demonstration activity. These correlations are unlikely to be driven by selection of SXSW enthusiasts and attendees into areas with high demonstration activity as the

number of SXSW followers who created their accounts in 2006 are controlled for, and they also do not correlate with demonstration activity.

As pointed out by Müller & Schwarz (2023), some uncertainty remains on the extent to which the causal effects estimated can be exclusively attributed to Twitter, as opposed to other social media platforms. While the SXSW festival in 2007 likely only influenced Twitter adoption initially, it is possible that spillovers in the adoption of other social media platforms existed in the medium-run. Therefore, the estimated effect might not be pure “Twitter effect.” Such spillovers should however not change the core argument of the paper, i.e., that these estimates and the ones presented below refer to the causal effect of social media diffusion on demonstration frequency and the composition of political protests.

5 The effect on the composition of political protests

5.1 Salient topics and extreme groups

After establishing a causal relationship between Twitter penetration and the frequency of demonstrations overall, we can explore how this effect depends on the nature of the protest movement. Table 3 estimates the effect of Twitter penetration on the frequency of demonstrations that fall under the topic categories explained in section 2.1. Panel A shows the estimates for the same right-hand side specifications as is column (5) in Tables 1 and 2, i.e., with the full set of controls, and outcome variables that are the inverse hyperbolic sine of the number of demonstrations that fall in a given demonstration category. Panel A is replicated in Table A11 in the Appendix using alternatives to the inverse hyperbolic sine as the outcome variable, where results remain comparable. Panel B restricts the sample to only counties where at least 1 demonstration event occurred over the period, and runs the same right-hand side specification on the share of demonstrations, among all demonstration in the county, that fit the indicated demonstration category. Column (1) shows the corresponding first stage and columns (2) to (7) show the second stages for a given protest category.

Panel A shows that Twitter penetration has an positive effect on demonstrations concerning all topics. Focusing on columns (2) to (6) we see that the effects are precisely estimated for all different topics and do not exhibit much heterogeneity. Point estimates are never far from the average effect on demonstrations overall, i.e., close to a unit elasticity, and are all less than or around one standard error away from the average effect presented in Panel C, column (5) of Table 2. The main exception is that in column

(7), demonstrations where extreme groups are involved seem to be disproportionately affected. A percentage increase in Twitter penetration implies around 1.6% increase in the frequency of protests where extreme groups are involved. However, the 95% confidence interval of this effect intersects the 95% confidence interval of the effect on protests overall, so homogeneous effects cannot be ruled out.

[Table 3 around here]

Panel B, on the other hand, estimates the effect of Twitter penetration on the share of all demonstrations falling into a given demonstration category, for counties where at least 1 demonstration occurs over the sample period. First, we cannot reject a null effect, for the share of “Stop the Steal”, abortion rights, BLM, or racial issues related demonstrations, at any reasonable significance threshold. That is, we cannot rule out that demonstrations of these categories increase on average proportionately to demonstrations overall. Twitter penetration implies a decrease in the share of protests which are coronavirus related. To give some sense of magnitude, residualizing the log number of Twitter users using population deciles (as population is highly predictive of the number of Twitter users) and moving from the first to the third quartile implies an increase of roughly 0.6 in the log number of Twitter Users,¹⁰ which implies a decrease in the share of covid related protests by about 3 percentage points from an average of 9%. Finally and most importantly, compared the the range of salient protest topics, we see in column (7) that an increase in Twitter penetration implies an increase in the share of demonstration events where extreme groups are involved, statistically significant at the 1% level. This is a sizable effect, moving from the first to third quartile of the log number of Twitter users (residualized by population) implies around a 2.5 percentage points increase in the share of these protests, from an average of 2.5%. Further, moving from the median to the third quartile implies around a 50% increase in the share of demonstrations where extreme groups are involved. Results on the effect of Twitter users on the number and share of demonstrations where extreme groups are involved remain robust to instrumenting Twitter users via the inverse hyperbolic sine of SXSW followers, as reported in Table A12 in the Appendix.

Figure 5 visualizes the reduced form corresponding to the regression in column (7) panel B, the main result. It presents binned scatter plots where the y-axis shows the share of demonstrations involving extreme groups and the x-axis represents the natural logarithm of (one plus) SXSW 2007 followers in the upper panel and the natural

¹⁰See distribution of Twitter users in figures A4 and A5

logarithm of (one plus) SXSW 2006 followers in the lower panel. The yellow dots represent bins, the blue line represents a linear regression line and the shaded area shows 95% confidence intervals, calculated using heteroscedasticity robust standard errors clustered at the state-level. Figure A9 in the Appendix shows the same plots with additional grey dots representing individual observations. We see a positive relationship between the 2007 SXSW followers, but a smaller and negative one for the 2006 SXSW followers.

[Figure 5 around here]

Taken together, these results indicate that Twitter penetration disproportionately fuels demonstrations where extreme groups are involved. Overall, the effects of Twitter penetration do not vary substantially for various salient protest topics in the United States during the period, making it unlikely that the large effect for extreme demonstrations is purely topic driven.

5.2 Political leaning

Table 4 studies whether the effect of Twitter on protests depends on the political leaning of the protest event. That is, depending on whether the topic of protest, or the groups involved in the protest align with the political left or the right in the United States. Again, column (1) shows the respective first stage and columns (2) to (5) show second stages where the outcome variables are demonstrations in the column-category. Panel A shows effects on the inverse hyperbolic sine of the number of protests in a given category while Panel B shows regressions on a sub-sample of counties where at least 1 protest occurred and the outcome variable is the share of all protests in the county which fall into a given demonstration category. In column (2) we look at “Overall left leaning” demonstrations: These are all left leaning protests that fall in any topic category from Table 3 (including those associated with extreme groups, if they are left leaning), in addition to if the protest is associated with the Democratic party and a small set of additional left-leaning protest movements. Column (3) in contrast, shows effects on protests where an extreme left group was involved. Columns (4) and (5) show the same comparison for “Overall right leaning” demonstrations versus demonstrations where extreme right groups are involved. A full description of protest categorization is found in Tables A1 to A6. Table A12 in the Appendix reports estimates of the effects on the number and share of demonstrations with extreme left or extreme right group involvement when Twitter users are instrumented for using

the inverse hyperbolic sine of SXSW followers, where results remain comparable. As before, the same specifications as in Panel A are replicated using alternatives to the inverse hyperbolic sine transformation in Table A13 in the Appendix. Those results are consistent with what is presented here.

In Panel A we see that the point estimates for all protest categories, except for those where extreme left groups are involved, are quite similar and within one standard error distance from the average effect on all protests, an approximate unit elasticity, found in column (5) of Table 2. The effects are very similar and statistically indistinguishable for overall left leaning versus overall right leaning protests. For protests where extreme left groups are involved, the point estimate is around 50% larger than for others, a percentage increase in Twitter penetration implies a 1.6% increase in protests where extreme left groups are involved. However, as in Table 3, the 95% confidence interval of this effect intersects the 95% confidence interval of the effect on protests overall, so homogeneous effects cannot be ruled out with this evidence.

Panel B on the other hand, studies counties where at least 1 demonstration occurs. For overall left and overall right leaning demonstrations, null effects on shares—i.e., that the effect of Twitter penetration on these demonstration categories is proportional to the average effect—cannot be rejected. Further, Tables A14 to A17, take separately the topic categories “Stop the Steal”, Abortion related, BLM related, and Racism related,¹¹ split them up by political leaning and compare effects on e.g., left leaning “Stop the Steal” related protests (that is, the “Count every vote” movement) versus right leaning “Stop the Steal” related protests. Heterogeneities depending on political leaning, within any of the salient topics studied, are not found. However, increased Twitter penetration implies an increase in the share of demonstrations where extreme left groups are involved, and also the share of demonstration where extreme right groups are involved, statistically significant at, at least, the 10% level. Again, to give sense of magnitudes, moving from the first to third quartile of the log number of Twitter users (residualized by population) implies around a 0.8 percentage points increase in the share of protests where extreme left groups are involved, from an average of 1%, while the same exercise implies about a half percentage point increase in the share of protests where extreme right groups are involved, close to a 50% increase. The 95% confidence intervals for effect sizes for extreme right and extreme left demonstrations intersect, implying that we cannot rule out homogeneous effects depending on political

¹¹The way Coronavirus related protests are identified via keywords does not allow to distinguish demonstrations calling for *stricter* public health measures versus those calling for more lax public health measures, sadly. Thus, it is not possible to compare left-leaning versus right-leaning coronavirus related demonstrations

leaning of extreme groups.

[Table 4 around here]

While the conceptual definition of extreme groups remains the same irrespective on political alignment, the following pattern emerges: On the left, extreme groups can be classified into two broad categories. First, there are political groups who protest the same societal conditions as the broad Black Lives Matter movement (i.e., police brutality and discrimination against the black population), but are more aggressive and willing to use militant tactics than the mainstream movement, or envision revolutionary remedies to these problems. Second, there are various revolutionary political groups which contest in one way or another current political systems or property rights institutions. These groups include, for example, communist and socialist parties, and radical environmentalist groups. Table A18 shows that the effects found for extreme left groups hold both for the “Extreme BLM” groups, and for the “Other extreme left” groups. On the right, extreme groups are unified by white supremacy ideologies. One can thus conceptualize the effects on these groups as social media’s effects on individuals’ willingness to publicly display their hateful ideologies. Prior research has found that social media can act as a propagating mechanism for hateful sentiments and hate crime (Bursztyn et al., 2019; Müller & Schwarz, 2021). In particular in the United States, this effect may be driven by hateful messages from influential opinion makers to the general public (Müller & Schwarz, 2023), thus influencing norms about what is socially acceptable (Bursztyn et al., 2020). As social image concerns likely influence protest participation decisions (Enikolopov, Makarin, Petrova, & Polishchuk, 2020; Cantoni et al., 2019), this potential shift of social norms may be a driving factor for the disproportionate effects found for white supremacy groups.

Taken together, these results confirm and strengthen the main finding: Twitter penetration facilitates protest mobilization, it does so irrespective of the topic of protest and the effect do not depend on political leaning. However, Twitter penetration disproportionately fuels demonstrations where extreme groups are involved, both extreme groups to the left and to the right.

5.3 Violent v.s. peaceful protests

Table 5 analyses whether Twitter plays a role in reported increases of hostility, violence, and repression during demonstration events in the United States. It examines whether there are heterogeneous effects on protests that are peaceful versus those

where outside groups or individuals interact non-peacefully with demonstrators, versus demonstrations where the demonstrators themselves are violent or exhibiting disruptive behaviour. Finally, it studies specifically events where demonstrators appear with firearms. Again, column (1) shows the respective first stage and columns (2) to (5) show second stages where the outcome variables are demonstrations in the column-category. Panel A shows effects on the inverse hyperbolic sine of the number of protests in a given category while Panel B shows regressions on a subsample of counties where at least 1 protest occurred and the outcome variable is the share of all protests in the county which fall into a given demonstration category. Column (2) studies entirely peaceful protests, column (3) protests where peaceful protesters are met by outside groups or individuals interact non-peacefully with demonstrators, column (4) studies demonstrations where demonstrators themselves are violent and column (5) studies events where demonstrators appear with firearms. Table A19 in the Appendix repeats the analysis in Panel A with alternatives to the inverse hyperbolic sine transformation, and results remain consistent.

Panel A shows that Twitter penetration increases both peaceful and non-peaceful protests significantly. The point estimates are somewhat larger for events where peaceful protesters are met with intervention or violence, and where demonstrators are violent, than the point estimate for peaceful protests. However, these are not statistically significantly different from one another. In Panel B, we see some indication that Twitter penetration increases the relative frequency of non-peaceful or aggressive behaviour during demonstration events. Point estimates are rather imprecisely estimated, but imply that the shares of demonstrations where peaceful protesters are met with interference or violence from outside groups, where protesters themselves are violent, and where they carry firearms, are all positive. For the behaviour of protesters themselves, Twitter penetration increases the relative frequency of non-peaceful behaviour and firearm carrying, significantly distinguishable from a null effect at the 10% significance level.

Summing up the evidence in Table 5, there are some indications that Twitter penetration does not only disproportionately fuel protests associated with extreme groups, but that it may also imply more aggressive behaviour among demonstrators.

[Table 5 around here]

6 Mechanisms

As discussed in the introduction, social media may fuel extremists' protests disproportionately through two mechanisms: It may be a source of radicalization or alternatively, social media may be a more powerful coordination tool in the hands of extremists. To gauge at the likelihood of the former mechanism, this section studies the effect of county-level Twitter penetration on opinions and attitudes of individuals in those counties.

Table 6 studies the effect of county-level Twitter penetration on strong opinions concerning statements on racial issues. Column (1) shows the first stage in these individual-level regressions, and columns (2) to (6) show the effects of Twitter penetration on the likelihood of selecting an answer option that indicates a strong opinion on the statement. Panel A shows the effects on the likelihood of holding a strong opinion that aligns with conservative political views in the United States, and Panel B shows effects on the likelihood of holding a strong liberal opinion. For most statements, estimated effects of Twitter penetration are small and statistically indistinguishable from zero. The exceptions are that Twitter penetration implies an increase in the likelihood of strongly agreeing with the statement "Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favours," significantly different from 0 at the 10% level, and strongly agreeing to "Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class," significant at the 1% level. A percentage increase in Twitter penetration implies an increase in the likelihood of strongly agreeing to the former statement by about 0.02 percentage points, and an increase of 0.03 percentage points for the latter statement, where 6 and 30 percent of respondent strongly agreed with the statements respectively. Therefore, while there are not consistent effects across the range of statements, there is some evidence of increased intensity of agreement with both anti-racial minority statements, and with pro-racial minority statements.

Tables A20, A21 in the Appendix present the effects of Twitter penetration on strong opinions concerning statements on gender issues and on rating oneself or the Democratic and Republican parties as very liberal or conservative (as opposed to somewhat liberal, conservative, or "middle of the road"). These results paint a similar picture: There is some, but inconsistent across statements, evidence of increased intense agreement with conservative statements on gender issues, while Twitter penetration implies an increase in defining oneself as "very liberal," with no effects on ideologi-

cal ratings of the two parties nor affecting the likelihood of defining oneself as “very conservative,”

Overall, these results provide some support of the idea that social media penetration in one’s local environment increases one’s intensity of opinion, but this evidence is quite weak.

[Table 6 around here]

Next, Table 7 presents the effects of county-level Twitter penetration on the extent to which respondents in the World Value Survey consider political violence, terrorism, or violence against other people more generally justifiable, on a scale from 1-10 where 1 indicates “never” and 10 indicates “always.” Column (1) shows the first stage and columns (2) to (4) show second stages. In Panel A, the outcome is a dummy variable equal to 1 if respondents rate the corresponding action with a number higher than 1 (i.e., the action is not “never justifiable”), in Panel B the outcome is a dummy equal to 1 if respondents rate the action with a number higher than 5, and in Panel C the outcome is a dummy equal to 1 if respondents rate the action with 10 (“always justifiable”). In all panels, the point estimates for justifiability of political violence are very small and statistically indistinguishable from zero. For the justifiability of terrorism, effects are imprecisely estimated, but point estimates are positive especially for higher justifiability thresholds in Panels B and C. Finally, in column 3 we find positive effects of county-level Twitter penetration on the justifiability of violence against other people, especially for the higher thresholds in Panels B and C. A percentage increase in Twitter penetration implies around a 0.06 percentage point increase in the likelihood of rating the justifiability of violence higher than 5, while around 7 percent of respondents rate the justifiability of violence in such a way. A percentage increase in Twitter penetration implies around a 0.02 percentage point increase in the likelihood of rating violence as “always justifiable” while 0.9 percent of respondents did so. Tables A22-A24 replicate this exercise with outcome dummies for other thresholds. The picture painted remains similar when examining the other thresholds: There seem to be no effects on the justifiability of political violence, positive but imprecisely estimated coefficients for higher thresholds of justifiability of terrorism, and positive and precisely estimated effects on the justifiability of violence for all thresholds higher than 5.

The difference in how attitudes towards “political violence” and “violence against other people” respond to Twitter penetration in one’s local environment is complex to interpret. However, this analysis presents somewhat consistent evidence that Twitter penetration in one’s local environment seems to increase individuals’ willingness to jus-

tify violence. Therefore, this section finds evidence, albeit weak, of Twitter’s influence on the intensity of opinion as well as on willingness to justify violence. As the bulk of the “extreme political groups” examined in the main analysis are explicitly willing to use violence or militant tactics to achieve their goals, this individual level analysis suggests that Twitter may not only a more powerful coordination tool in the hands of extremists, but that it may also *radicalize* the population to some extent.

[Table 7 around here]

7 Conclusion

This paper investigates whether social media can effect the composition of political protests. Following Müller & Schwarz (2023) and Fujiwara et al. (2021), it leverages a plausibly exogenous shock in early adoption of Twitter stemming from county-level attendance at the South by Southwest festival in 2007 as an instrument for contemporary county-level Twitter penetration. Consistent with prior research (e.g., Enikolopov, Makarin, & Petrova, 2020; Manacorda & Tesei, 2020), it finds that Twitter penetration increased demonstration activity considerably: A one percent increase in Twitter penetration led to around a one percent increase in demonstration frequency between January 2020 and 12th of November 2021. Twitter penetration disproportionately fuels demonstrations where extreme groups are involved. This is the case both for aggressive and militant black power movements, radical communist, socialist or environmentalist movements on the far-left, and for white supremacists on the far-right. Moving from the median to the third quartile of Twitter penetration (residualized by population) increases the share of demonstrations where extreme groups are involved by around 1.3 percentage points, from an average of 2.4%. Further, this paper finds indications that Twitter penetration increases the relative frequency of violent or aggressive behaviour on behalf of protesters during protest events.

While social medias’ ability to facilitate the coordination of collective action has been highlighted in the literature (e.g., Enikolopov, Makarin, & Petrova, 2020; Ferguson & Molina, 2019), I find evidence from survey data suggesting that Twitter may not only be a more powerful coordination tool in the hands of political fringes, but that county-level Twitter penetration may contribute to radicalization. County-level Twitter penetration increases the likelihood of holding strong opinions on statements related to racial or gender issues among county residents. It also increases county residents’ willingness to justify violence against other people. However, effects are only

found for a minority of statements related to racial and gender issues, while increased willingness to justify violence only holds for a general category of “violence against other people,” but not for “political violence.” This evidence should thus be treated with caution and only interpreted as suggestive.

These results speak to current debates on social media as an “outrage machine.” They show that social media can increase the visibility of political fringes, not only online but also off-line through demonstrations.

References

- ACLED. (2020a). *Cdt spotlight: Covid-19 us protest patterns*. Retrieved from <https://acleddata.com/2020/11/30/cdt-spotlight-covid-19-us-protest-patterns/> (Last accessed 21 May 2022)
- ACLED. (2020b). *Demonstration trends in the united states*. Retrieved from <https://acleddata.com/2020/09/23/demonstration-trends-in-the-united-states/> (Last accessed 21 May 2022)
- ACLED. (2020c). *Mid-year update: 10 conflicts to worry about in 2020*. Retrieved from <https://acleddata.com/2020/08/18/mid-year-update-10-conflicts-to-worry-about-in-2020/#1597759650330-f3890636-9943> (Last accessed 21 May 2022)
- ACLED. (2022). *From the capitol riot to the midterms: Shifts in american far-right mobilization between 2021 and 2022*. Retrieved from `@misc{ACLED2020UStrends,author={ACLED},title={DEMONSTRATIONTRENDSINTHEUNITEDSTATES},year={2020},note={Lastaccessed21May2022},url={https://acleddata.com/2020/09/23/demonstration-trends-in-the-united-states/}}` (Last accessed 11 July 2023)
- Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., & Zhuravskaya, E. (2015). Radio and the rise of the nazis in prewar germany. *The Quarterly Journal of Economics*, 130(4), 1885–1939.
- Aihounton, G. B., & Henningsen, A. (2021). Units of measurement and the inverse hyperbolic sine transformation. *The Econometrics Journal*, 24(2), 334–351.
- Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, 110(3), 629–76.
- Amorim, G., Lima, R. C., & Sampaio, B. (2022). Broadband internet and protests: Evidence from the occupy movement. *Information Economics and Policy*, 60, 100982.
- Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11, 727–753.
- Arrow, K. J. (2000). Increasing returns: historiographic issues and path dependence. *The European Journal of the History of Economic Thought*, 7(2), 171–180.

- Barberá, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J., & González-Bailón, S. (2015). The critical periphery in the growth of social protests. *PloS one*, 10(11), e0143611.
- Barbera, S., Jackson, M. O., et al. (2020). A model of protests, revolution, and information. *Quarterly Journal of Political Science*, 15(3), 297–335.
- Barberá, P. (2015). How social media reduces mass political polarization. evidence from germany, spain, and the us. *Job Market Paper, New York University*, 46.
- Beknazar-Yuzbashev, G., Jiménez Durán, R., McCrosky, J., & Stalinski, M. (2023). Toxic content and user engagement on social media: Evidence from a field experiment. *Available at SSRN*.
- Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). Greater internet use is not associated with faster growth in political polarization among us demographic groups. *Proceedings of the National Academy of Sciences*, 114(40), 10612–10617.
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318.
- Brewster, T. (2021). Sheryl sandberg downplayed facebook’s role in the capitol hill siege—justice department files tell a very different story. *Retrieved from <https://www.forbes.com/sites/thomasbrewster/2021/02/07/sheryl-sandberg-downplayed-facebooks-role-in-the-capitol-hill-siege-justice-department-files-tell-a-very-different-story/> (Accessed: 2020- 03-19)*.
- Bursztyn, L., Egorov, G., Enikolopov, R., & Petrova, M. (2019). Social media and xenophobia: evidence from russia. *NBER Working paper no. 26567*.
- Bursztyn, L., Egorov, G., & Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American Economic Review*, 110(11), 3522–3548.
- Caesmann, M., Caprettini, B., Voth, H.-J., & Yanagizawa-Drott, D. (2021). Going viral: propaganda, persuasion and polarization in 1932 hamburg.
- Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2019). Protests as strategic games: experimental evidence from hong kong’s antiauthoritarian movement. *The Quarterly Journal of Economics*, 134(2), 1021–1077.

- Casanueva, A. (2022). Can chants in the street change politics' tune? evidence from the 15m movement in Spain. *Available at SSRN: <https://ssrn.com/abstract=3831874> or <http://dx.doi.org/10.2139/ssrn.3831874>.*
- Casanueva, A., Avetian, V., Sardoschau, S., & Saxena, K. (2021). Social media and the broadening of social movements: Evidence from black lives matter. *Available at SSRN 3831819.*
- CBS, N. (2022). Social media's role in America's polarized political climate. *Retrieved from <https://www.cbsnews.com/news/social-media-political-polarization-60-minutes-2022-11-06/> (Accessed: 2023-07-15).*
- Christensen, D., & Garfias, F. (2018). Can you hear me now? how communication technology affects protest and repression. *Quarterly journal of political science*, 13(1), 89.
- DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., & Zhuravskaya, E. (2014). Cross-border media and nationalism: Evidence from Serbian radio in Croatia. *American Economic Journal: Applied Economics*, 6(3), 103–32.
- DellaVigna, S., & Gentzkow, M. (2010). Persuasion: empirical evidence. *Annu. Rev. Econ.*, 2(1), 643–669.
- Durante, R., & Zhuravskaya, E. (2018). Attack when the world is not watching? us news and the Israeli-Palestinian conflict. *Journal of Political Economy*, 126(3), 1085–1133.
- Economist, T. (2019). Something in the air: Why are so many countries witnessing mass protests? *Retrieved from <https://www.economist.com/international/2019/11/04/why-are-so-many-countries-witnessing-mass-protests> (Accessed: 2023-07-11).*
- Edmond, C. (2013). Information manipulation, coordination, and regime change. *Review of Economic Studies*, 80(4), 1422–1458.
- Enikolopov, R., Makarin, A., & Petrova, M. (2020). Social media and protest participation: Evidence from Russia. *Econometrica*, 88(4), 1479–1514.
- Enikolopov, R., Makarin, A., Petrova, M., & Polishchuk, L. (2020). Social image, networks, and protest participation. *Available at SSRN: <https://ssrn.com/abstract=2940171> or <http://dx.doi.org/10.2139/ssrn.2940171>.*

- Enikolopov, R., Petrova, M., & Zhuravskaya, E. (2011). Media and political persuasion: Evidence from russia. *American Economic Review*, 101(7), 3253–85.
- Falck, O., Gold, R., & Heblich, S. (2014). E-lections: Voting behavior and the internet. *American Economic Review*, 104(7), 2238–65.
- Fergusson, L., & Molina, C. (2019). Facebook causes protests. *Documento CEDE*(41).
- Fujiwara, T., Müller, K., & Schwarz, C. (2021). The effect of social media on elections: Evidence from the united states. *NBER Working Paper no. 28849*.
- Gentzkow, M. (2006). Television and voter turnout. *The Quarterly Journal of Economics*, 121(3), 931–972.
- Gentzkow, M., & Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4), 1799–1839.
- Gentzkow, M., Shapiro, J. M., & Sinkinson, M. (2011). The effect of newspaper entry and exit on electoral politics. *American Economic Review*, 101(7), 2980–3018.
- González, F. (2020). Collective action in networks: Evidence from the chilean student movement. *Journal of Public Economics*, 188, 104220.
- Guriev, S., Melnikov, N., & Zhuravskaya, E. (2021). 3g internet and confidence in government. *The Quarterly Journal of Economics*, 136(4), 2533–2613.
- Hadero, H. (2023). Why social media is being blamed for fueling the riots in france. Retrieved from <https://apnews.com/article/france-shooting-riots-social-media-tiktok-snapchat-c6328bdfbaf8be851c84416993e09cbf> (Accessed: 2023-07-11).
- Haerpfer, I. R. M. A. W. C. K. K. D.-M. J. M. L. P. N. E. P. . B. P., C. (2022). *World Values Survey: Round Seven – Country-Pooled Datafile Version 5.0.0*. Madrid, Spain, & Vienna, Austria: JD Systems Institute & WVSA Secretariat. Retrieved from <https://doi.org/10.14281/18241.20> doi: doi:10.14281/18241.20
- Haidt, J., & Rose-Stockwell, T. (2019). The dark psychology of social networks. *The Atlantic*, 6–60.
- Halberstam, Y., & Knight, B. (2016). Homophily, group size, and the diffusion of political information in social networks: Evidence from twitter. *Journal of Public Economics*, 143, 73–88.

- Kinder-Kurlanda, K., Weller, K., Zenk-Möltgen, W., Pfeffer, J., & Morstatter, F. (2017). Archiving information from geotagged tweets to promote reproducibility and comparability in social media research. *Big Data & Society*, 4(2), 2053951717736336.
- Klein Teeselink, B., & Melios, G. (2021). Weather to protest: The effect of black lives matter protests on the 2020 presidential election. *Available at SSRN 3809877*.
- Lelkes, Y., Sood, G., & Iyengar, S. (2017). The hostile audience: The effect of access to broadband internet on partisan affect. *American Journal of Political Science*, 61(1), 5–20.
- Liebowitz, S. J., & Margolis, S. E. (1995). Path dependence, lock-in, and history. *Journal of Law, Economics, & Organization*, 205–226.
- Little, A. T. (2016). Communication technology and protest. *The Journal of Politics*, 78(1), 152–166.
- Madestam, A., Shoag, D., Veuger, S., & Yanagizawa-Drott, D. (2013). Do political protests matter? evidence from the tea party movement. *The Quarterly Journal of Economics*, 128(4), 1633–1685.
- Manacorda, M., & Tesei, A. (2020). Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2), 533–567.
- Melnikov, N. (2023). Mobile internet and political polarization. *Available at SSRN 3937760*.
- Müller, K., & Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4), 2131–2167.
- Müller, K., & Schwarz, C. (2023). From hashtag to hate crime: Twitter and antiminority sentiment. *American Economic Journal: Applied Economics*, 15(3), 270–312.
- Oksanen, A., Hawdon, J., Holkeri, E., Näsi, M., & Räsänen, P. (2014). Exposure to online hate among young social media users. In *Soul of society: a focus on the lives of children & youth*. Emerald Group Publishing Limited.
- Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3), 358–369.
- Olson, M. (1965). *The logic of collective action*. Harvard University Press, MA.

- Ortiz, I., Burke, S., Berrada, M., & Cortés, H. (2013). World protests 2006-2013. *Initiative for Policy Dialogue and Friedrich-Ebert-Stiftung New York Working Paper*(2013).
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2), 187–204.
- Ovide, S. (2020). How social media has changed civil rights protests. Retrieved from <https://www.nytimes.com/2020/06/18/technology/social-media-protests.html> (Accessed: 2023-07-11).
- Passarelli, F., & Tabellini, G. (2017). Emotions and political unrest. *Journal of Political Economy*, 125(3), 903–946.
- Pew Research Center. (2021). *Social media use in 2021*. Technical report.
- Quora. (2011). *What is the process involved in launching a start-up at sxsw*. Retrieved from <https://www.quora.com/What-is-the-process-involved-in-launching-a-start-up-at-SXSW> (Last accessed 21 May 2022)
- Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). Introducing acled: an armed conflict location and event dataset: special data feature. *Journal of Peace Research*, 47(5), 651–660.
- Rose-Stockwell, T. (2017). This is how your fear and outrage are being sold for profit. Retrieved from <https://tobiasrose.medium.com/the-enemy-in-our-feeds-e86511488de> (Accessed: 2023-07-12).
- Rose-Stockwell, T. (2023). *Outrage machine: How tech amplifies discontent, disrupts democracy and what we can do about it*. Grand Central Publishing. Retrieved from <https://books.google.fr/books?id=3E5jEAAAQBAJ>
- Schaffner, B., Ansolabehere, S., & Shih, M. (2023). *Cooperative Election Study Common Content, 2022*. Harvard Dataverse. Retrieved from <https://doi.org/10.7910/DVN/PR4L8P> doi: 10.7910/DVN/PR4L8P
- Skoy, E. (2021). Black lives matter protests, fatal police interactions, and crime. *Contemporary Economic Policy*, 39(2), 280–291.
- Statista. (2019). *Number of monthly active twitter users worldwide from 1st quarter 2010 to 1st quarter 2019*. Retrieved from <https://www.statista.com/statistics/>

[282087/number-of-monthly-active-twitter-users/](#) (Last accessed 21 May 2022)

Theis, M. (2016). *Sxsw economic impact up slightly in 2016; hotel rates hit new high*. Austin Business Journal. Retrieved from <https://www.bizjournals.com/austin/news/2016/09/07/sxsw-economic-impact-up-slightly-in-2016-hotel.html> (Last accessed 21 May 2022)

Tufekci, Z. (2018). How social media took us from tahrir square to donald trump. *MIT Technology Review*, 14, 18.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the rwandan genocide. *The Quarterly Journal of Economics*, 129(4), 1947–1994.

Yanagizawa-Drott, D., Rao, A., Petrova, M., Enikolopov, R., & L., B. (2020). Echo chambers: Does online network structure affect political polarization? *Working Paper University of Zurich*.

Zhuravskaya, E., Petrova, M., & Enikolopov, R. (2020). Political effects of the internet and social media. *Annual Review of Economics*, 12, 415–438.

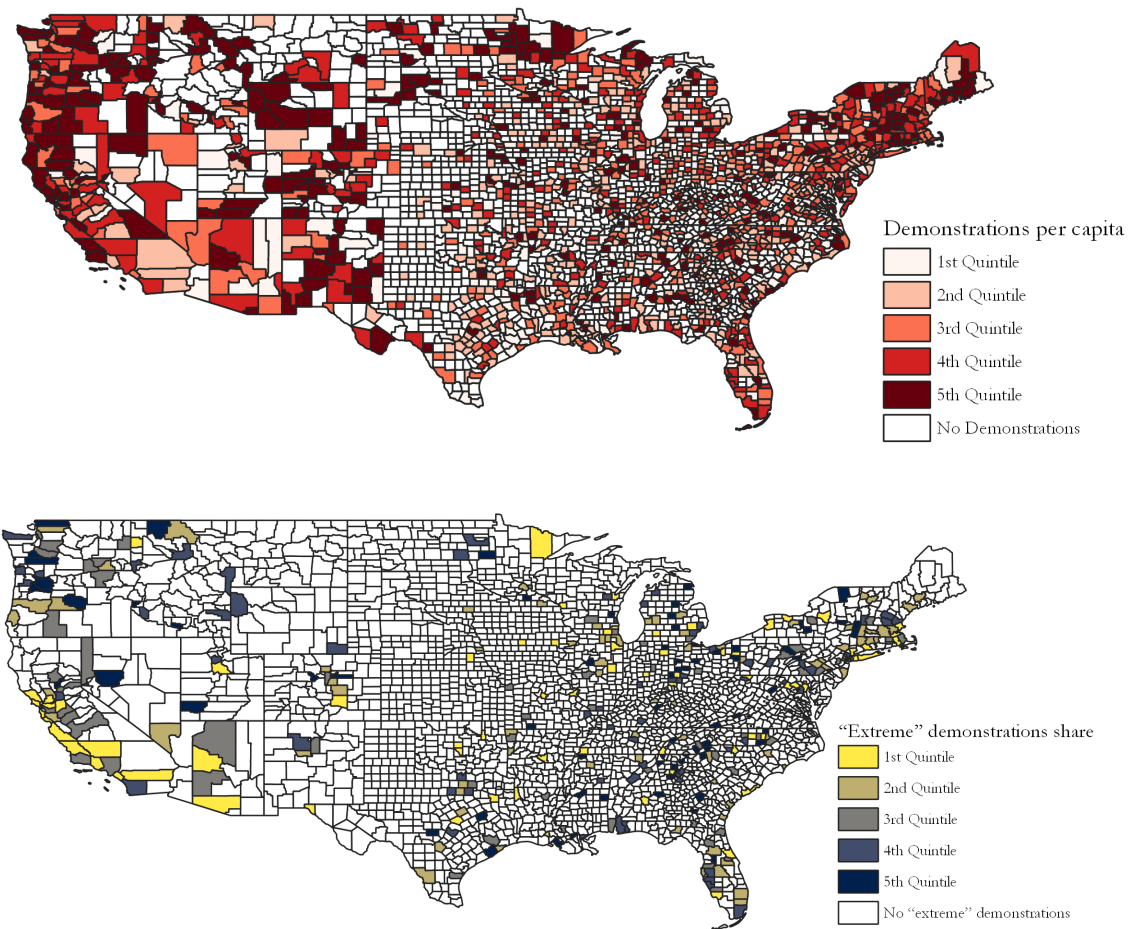


Figure 1: Spacial distribution of demonstration events

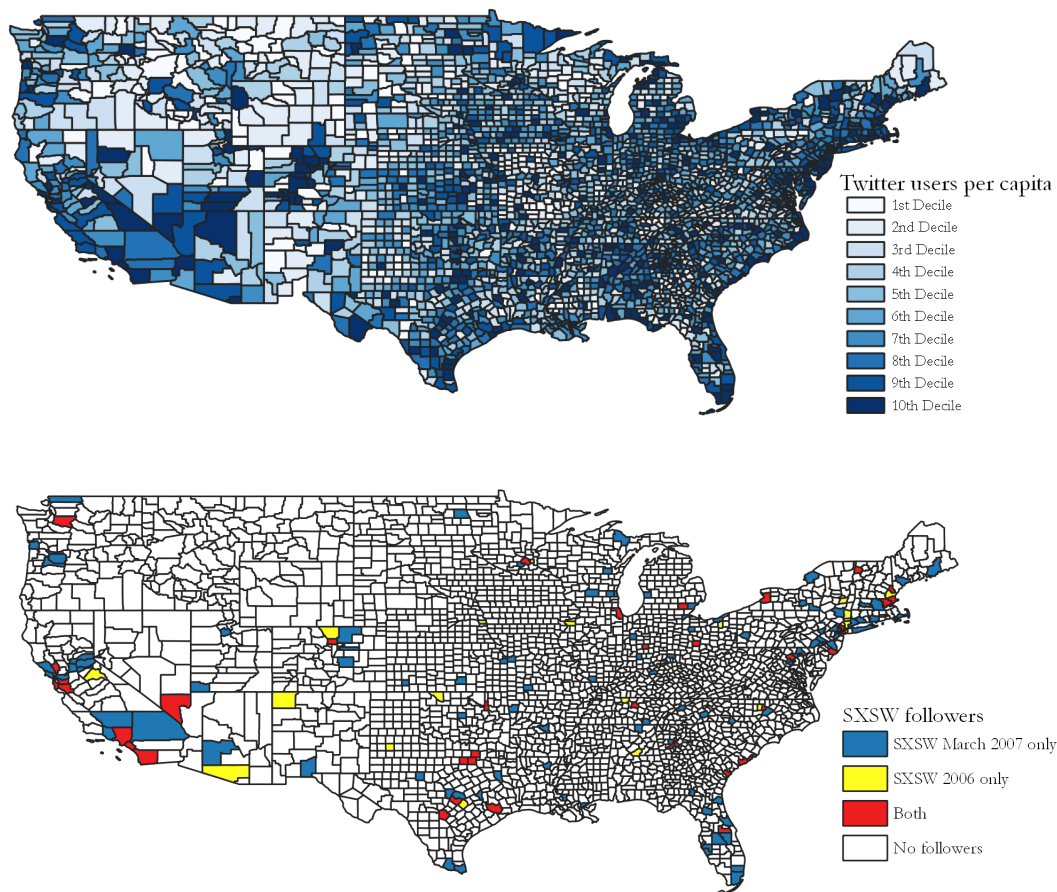


Figure 2: Spatial distribution of Twitter users per capita and SXSW followers

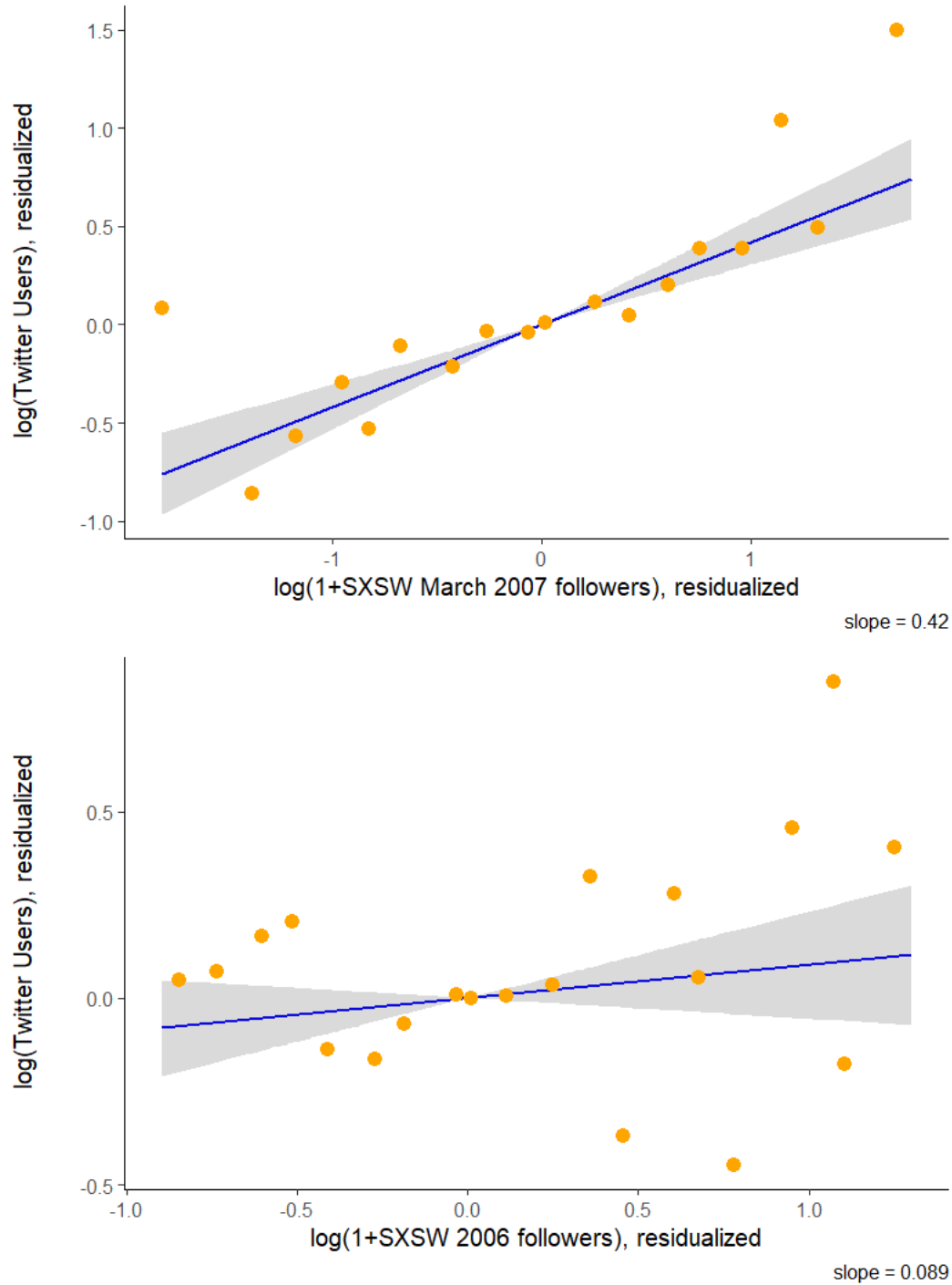


Figure 3: First stage visualization

Note: This figure presents binned scatter plots of the relationship between county-level Twitter users in 2014-2015 and the number of SXSW followers. In the upper panel, the x-axis represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in March 2007 and in the lower panel it represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in 2006. Variables are residualized by partialling out population deciles, demographic, geographic, socio-economic, and pre-2010 election controls (see description of control variables in Table A7 in the Appendix) and the SXSW follower variable not represented in the x-axis of each panel. Yellow dots represent average values of both variables within a bin. The blue line represents a line of best fit using the unbinned data, and the shaded blue area are 95% confidence intervals calculated using robust standard errors clustered at the state-level.

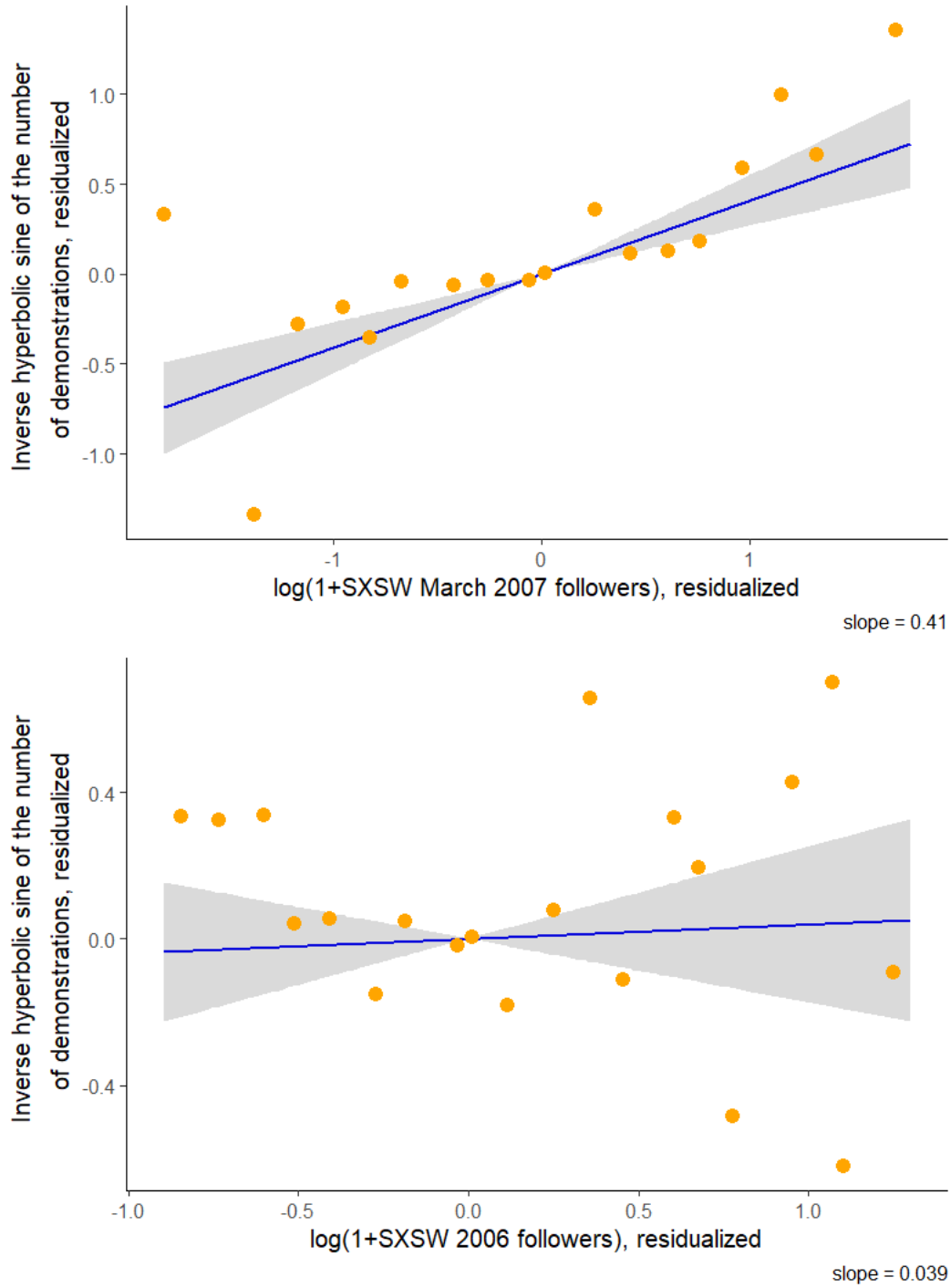


Figure 4: Reduced form visualization, number of demonstrations

Note: This figure presents binned scatter plots of the relationship between county-level number of demonstrations between Jan 2020 and Nov 2021, and the number of SXSW followers. In the upper panel, the x-axis represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in March 2007 and in the lower panel it represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in 2006. Variables are residualized by partialling out population deciles, demographic, geographic, socio-economic, and pre-2010 election controls (see description of control variables in Table A7 in the Appendix) and the SXSW follower variable not represented in the x-axis of each panel. Yellow dots represent average values of both variables within a bin. The blue line represents a line of best fit using the unbinned data, and the shaded blue area are 95% confidence intervals calculated using robust standard errors clustered at the state-level.

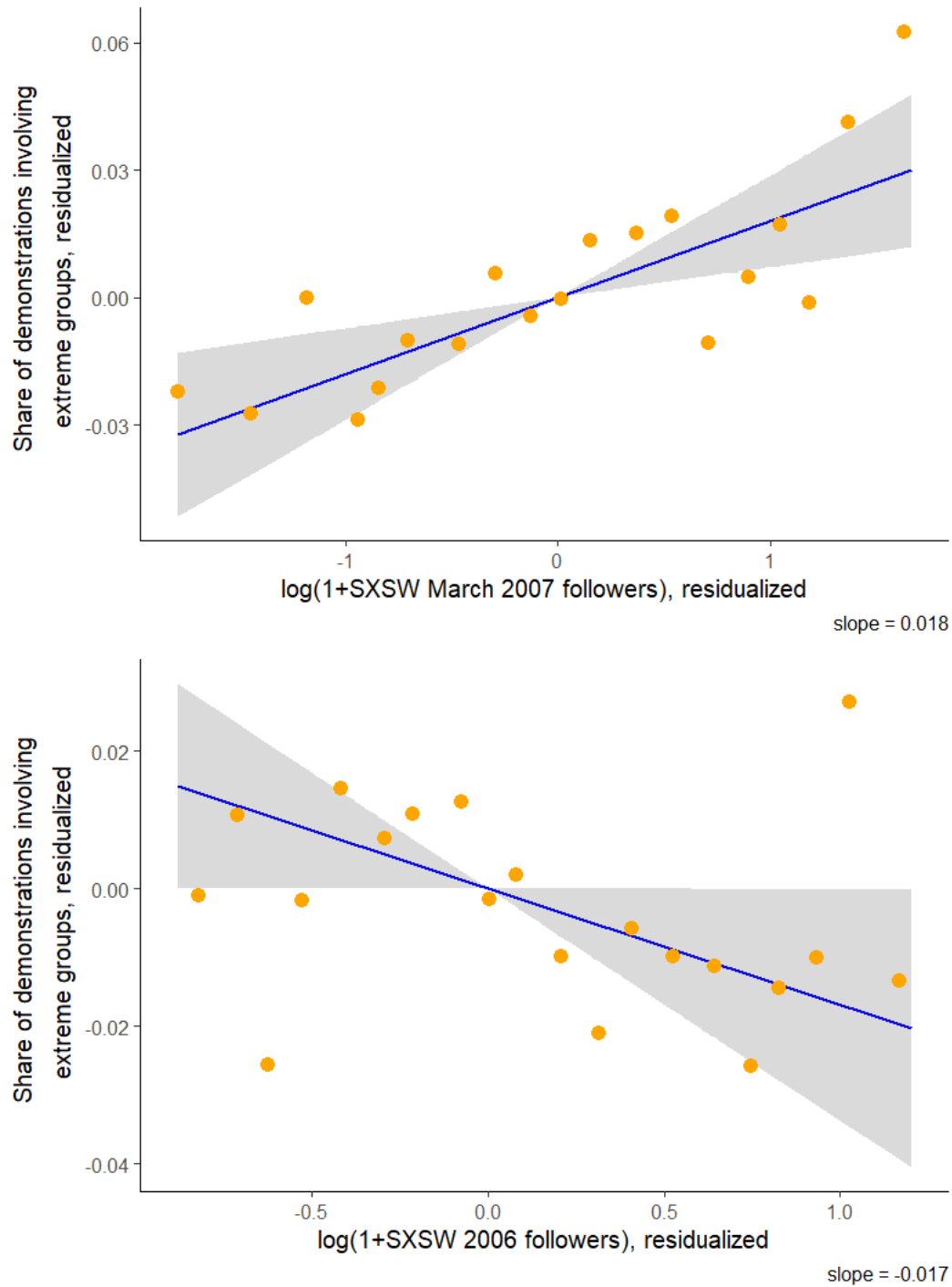


Figure 5: Reduced form visualization, share of demonstrations involving extreme groups

Note: This figure presents binned scatter plots of the relationship between county-level share of demonstrations involving extreme groups and the number of SXSW followers, in counties where at least 1 demonstration occurred during the period ($N=1,776$). In the upper panel, the x-axis represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in March 2007 and in the lower panel it represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in 2006. Variables are residualized by partialling out population deciles, demographic, geographic, socio-economic, and pre-2010 election controls (see description of control variables in Table A7 in the Appendix) and the SXSW follower variable not represented in the x-axis of each panel. Yellow dots represent average values of both variables within a bin. The blue line represents a line of best fit using the unbinned data, and the shaded blue area are 95% confidence intervals calculated using robust standard errors clustered at the state-level.

Table 1: First stage analysis

	(1)	(2)	(3)	(4)	(5)
	<i>Dependent variable: Log(Twitter Users)</i>				
Log(1+SXSW March2007)	0.619*** (0.063)	0.481*** (0.058)	0.455*** (0.060)	0.426*** (0.059)	0.417*** (0.059)
Log(1+SXSW 2006)	0.206*** (0.073)	0.121 (0.074)	0.090 (0.076)	0.096 (0.073)	0.087 (0.074)
R ²	0.926	0.938	0.939	0.944	0.945
Cluster robust F-statistic	84.77	53.83	48.09	38.22	36.30
State FE	Yes	Yes	Yes	Yes	Yes
Population deciles	Yes	Yes	Yes	Yes	Yes
Demographic controls		Yes	Yes	Yes	Yes
Geographic controls			Yes	Yes	Yes
Socio-economic controls				Yes	Yes
Pre 2010 election controls					Yes
Observations	3,105	3,105	3,105	3,105	3,104

Note: This table presents county-level regressions where the natural logarithm of our Twitter users from 2014-2015 is the dependent variable. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Regressions include the indicated control variables (see Table A7 in the Appendix for their descriptions). Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table 2: Effect of Twitter penetration on demonstrations

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	<i>Inverse hyperbolic sine of number of demonstrations</i>				
Panel A: OLS					
Log(Twitter Users)	0.464*** (0.051)	0.341*** (0.038)	0.345*** (0.040)	0.292*** (0.039)	0.286*** (0.039)
R ²	0.804	0.818	0.818	0.824	0.825
Panel B: Reduced-form					
Log(1+SXSW March2007)	0.614*** (0.095)	0.463*** (0.067)	0.479*** (0.066)	0.424*** (0.072)	0.408*** (0.073)
Log(1+SXSW 2006)	0.083 (0.112)	−0.002 (0.112)	0.008 (0.112)	0.046 (0.108)	0.036 (0.110)
R ²	0.789	0.812	0.812	0.820	0.821
Panel C: 2SLS Second Stage					
Log(Twitter Users)	0.992*** (0.127)	0.961*** (0.125)	1.054*** (0.142)	0.994*** (0.142)	0.978*** (0.147)
Log(1+SXSW2006)	−0.121 (0.113)	−0.118 (0.104)	−0.087 (0.098)	−0.049 (0.094)	−0.049 (0.092)
Cluster robust F-statistic	84.77	53.83	48.09	38.22	36.30
Observations	3,105	3,105	3,105	3,105	3,104
Mean number of demonstrations	11.49	11.49	11.49	11.49	11.49
State FE	Yes	Yes	Yes	Yes	Yes
Population deciles	Yes	Yes	Yes	Yes	Yes
Demographic controls		Yes	Yes	Yes	Yes
Geographic controls			Yes	Yes	Yes
Socio-economic controls				Yes	Yes
Pre 2010 election controls					Yes

Note: This table presents county-level regressions for an OLS and a 2SLS model of the effect of Twitter penetration on demonstration frequency. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Regressions include the indicated control variables (see Table A7 in the Appendix for their descriptions). Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table 3: 2SLS demonstrations by topic and extremity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Topic of protest:		Coronavirus	“Stop the Steal”	Abortion rights	BLM related	Racism related	“Extreme”
Specification:	First Stage	Second Stage					
Panel A		Inverse hyperbolic sine of number of demonstrations					
Dependent variable:	Log(Twitter Users)						
Log(1+SXSW March 2007)	0.417*** (0.059)						
Log(Twitter Users)		0.814*** (0.166)	0.902*** (0.215)	0.796*** (0.147)	1.033*** (0.152)	1.041*** (0.156)	1.628*** (0.286)
Log(1+SXSW 2006)	0.087 (0.074)	-0.013 (0.125)	0.029 (0.104)	-0.024 (0.123)	-0.113 (0.103)	-0.115 (0.101)	-0.081 (0.144)
R ²	0.945						
Cluster robust F-statistic	36.30						
Observations	3,104	3,104	3,104	3,104	3,104	3,104	3,104
Mean number of demonstrations		1.096	0.157	0.193	4.027	4.341	0.593
Panel B		Share among all demonstrations					
Dependent variables:	Log(Twitter Users)						
Log(1+SXSW March 2007)	0.433*** (0.056)						
Log(Twitter Users)		-0.054*** (0.020)	0.007 (0.012)	-0.002 (0.007)	-0.030 (0.037)	-0.023 (0.037)	0.041*** (0.014)
Log(1+SXSW 2006)	0.075 (0.075)	0.012 (0.015)	-0.008 (0.005)	-0.001 (0.006)	0.007 (0.024)	-0.001 (0.023)	-0.020*** (0.009)
R ²	0.939						
Cluster robust F-statistic	43.34						
Mean dependent variable		0.092	0.010	0.017	0.464	0.480	0.024
Observations	1,776	1,776	1,776	1,776	1,776	1,776	1,776

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March 2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(7) present second stages. Coronavirus refers to demonstrations related to public health measures due to the Covid pandemic. "Stop the Steal" refers to demonstrations related to the 2020 post-presidential election unrest (including those calling for results to be respected). Abortion rights refers to demonstrations related to reproductive rights. BLM related refers to demonstrations concerning police brutality against minorities, and counter movements. Racial issues refers to demonstrations concerning racial tensions while Extreme Left or Right refers to demonstrations involving a political group defined as extreme. A detailed description of topic categorization and the set of extreme groups is presented in Tables A2 to A6 in the Appendix. Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column 5 in Table 1), see Table A7 in the Appendix for their descriptions. Robust standard errors in parentheses are clustered by state. * p<0.1; ** p<0.05; *** p<0.01

Table 4: 2SLS demonstrations by political leaning

	(1)	(2)	(3)	(4)	(5)
Political leaning:		Left		Right	
		Overall	Extreme	Overall	Extreme
Specification:	First Stage	Second Stage			
Panel A					
Dependent variable:	Log(Twitter Users)	<i>Inverse hyperbolic sine of number of demonstrations</i>			
Log(1+SXSW March 2007)	0.417*** (0.059)				
Log(Twitter Users)		1.132*** (0.152)	1.573*** (0.292)	1.043*** (0.197)	1.041*** (0.210)
Log(1+SXSW 2006)	0.087 (0.074)	-0.107 (0.098)	-0.073 (0.161)	-0.072 (0.115)	-0.017 (0.122)
R ²	0.945				
Cluster robust F-statistic	36.30				
Observations	3,104	3,104	3,104	3,104	3,104
Mean number demonstrations		4.751	0.424	0.845	0.190
Panel B					
Dependent variable:	Log(Twitter Users)	<i>Share among all demonstrations</i>			
Log(1+SXSW March 2007)	0.433*** (0.056)				
Log(Twitter Users)		-0.009 (0.034)	0.028*** (0.010)	0.017 (0.020)	0.015** (0.007)
Log(1+SXSW 2006)	0.075 (0.075)	0.009 (0.020)	-0.012 (0.007)	-0.017 (0.011)	-0.009** (0.004)
R ²	0.939				
Cluster robust F-statistic	43.34				
Observations	1,776	1,776	1,776	1,776	1,776
Mean dependent variable		0.497	0.011	0.081	0.012

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(5) present second stages. Mainstream left refers to demonstrations which are left-leaning, mainstream right refers to right-leaning demonstrations, Extreme left and right refers to demonstrations where extreme left or extreme right groups are involved, respectively. A detailed description of protest categorization can be found in Tables A2 to A6 in the Appendix. Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions.

Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table 5: 2SLS demonstrations by aggression

Type of demonstration:	(1)	(2)	(3)	(4)	(5)
Specification:	First Stage	Peaceful	Peaceful Protesters met with intervention or violence	Violent Protesters	Armed demonstrators
Panel A				Second Stage	
Dependent variable:	Log(Twitter Users)	<i>Inverse hyper sine of number of demonstrations</i>			
Log(1+SXSW March 2007)	0.417*** (0.059)				
Log(Twitter Users)		0.945*** (0.149)	1.125*** (0.200)	1.200*** (0.236)	1.088*** (0.228)
Log(1+SXSW 2006)	0.087 (0.074)	-0.035 (0.092)	0.082 (0.128)	-0.018 (0.169)	-0.287* (0.153)
R ²	0.945				
Cluster robust F-statistic	36.30				
Observations	3,104	3,104	3,104	3,104	3,104
Mean number of demonstrations		10.796	0.322	0.301	0.194
Panel B				<i>Shares among all demonstrations</i>	
Dependent variable:	Log(Twitter Users)				
Log(1+SXSW March 2007)	0.433*** (0.056)				
Log(Twitter Users)			0.031 (0.019)	0.022* (0.012)	0.012* (0.007)
Log(1+SXSW 2006)	0.075 (0.075)		-0.012* (0.006)	-0.007 (0.008)	-0.008* (0.005)
R ²	0.939				
Cluster robust F-statistic	43.34				
Observations	1,776	1,776	1,776	1,776	1,776
Mean dependent variable			0.018	0.009	0.017

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(5) present second stages. Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decade dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions. Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table 6: Effects on strong opinions on statements related to racial issues, Cooperative Election Study 2020

	(1)	(2)	(3)	(4)	(5)	(6)
Statement:		Police safe	No white privilege	Racial problems rare	Black people don't deserve favors	Lingering effects of slavery
Specification:	First Stage	Second Stage				
Panel A						
Dependent variable:	Log(Twitter Users)	<i>D=1 if strongly conservative opinion</i>				
Log(1+SXSW March 2007)	0.492*** (0.048)					
Log(Twitter Users)		-0.006 (0.012)	0.005 (0.011)	0.012 (0.009)	0.018* (0.011)	0.007 (0.012)
Log(1+SXSW 2006)	0.018 (0.075)	0.003 (0.008)	-0.008 (0.008)	-0.009 (0.006)	-0.009 (0.008)	-0.006 (0.008)
Mean dependent variable		0.45	0.18	0.09	0.06	0.23
R ²	0.956					
Cluster robust F-statistic	104.89					
Panel B						
Dependent variable:		<i>D=1 if strongly liberal opinion</i>				
Log(Twitter Users)		0.001 (0.008)	0.020 (0.013)	-0.001 (0.014)	0.013 (0.009)	0.033*** (0.012)
Log(1+SXSW 2006)		0.009* (0.005)	-0.006 (0.008)	-0.007 (0.008)	-0.003 (0.007)	-0.009 (0.008)
Mean dependent variable		0.08	0.35	0.39	0.36	0.30
Observations	54,317	54,285	46,425	45,308	46,215	46,420

Note: This table presents individual-level regressions. Log(Twitter Users) is the natural logarithm of unique Twitter users from a large sample of Tweets from 2014-2015 SXSW March2007 and SXSW 2006 are the number of Twitter users who follow South by Southwest (@SXSW) and joined Twitter in March 2007 and in 2006 respectively. Column (1) presents first stage regressions, while columns (2)-(6) present second stages. All regressions include the following individual level controls: Age, gender, education, race, ethnicity, employment status, home-ownership, family income, number of children and internet connection, and the full set of county-characteristics (as in column (5) in Table 1, see Table A7 in the Appendix for their descriptions). The statements are the following: "Police safe": "Do the police make you feel...?" where "Mostly safe" is considered the strongly conservative option and "Mostly unsafe" is the strongly liberal option. "No white privilege": "White people in the U.S. have certain advantages because of the color of their skin," where "strongly disagree" is considered the strongly conservative opinion and "strongly agree" is the strongly liberal option. "Racial problems rare": "Racial problems in the U.S. are rare, isolated situations," "Black people don't deserve favors": "Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors," "Lingering effects of slavery": "Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class," where "strongly agree" is the strongly liberal option and "strongly disagree" is the strongly conservative option. Observations are weighted to represent adult Americans and robust standard errors are clustered by county. *p<0.1; **p<0.05; ***p<0.01

Table 7: Effects on attitudes towards violence and terrorism, World Value Survey 2017

	(1)	(2)	(3)	(4)
		Political violence can be justified	Terrorism can be justified	Violence can be justified
Specification:	First Stage	Second Stage		
Panel A				
Dependent variable:	Log(Twitter Users)	<i>D=1 if justifiable>1 (1=never, out of 10)</i>		
Log(1+SXSW March 2007)	0.493*** (0.062)			
Log(Twitter Users)		0.002 (0.044)	0.015 (0.037)	0.019 (0.040)
Log(1+SXSW 2006)	0.065 (0.084)	0.021 (0.032)	-0.044* (0.026)	-0.007 (0.028)
Mean dependent variable		0.59	0.82	0.53
R ²	0.732			
Cluster robust F-statistic	58.08			
Panel B				
Dependent variable:		<i>D=1 if justifiable>5 (out of 10)</i>		
Log(Twitter Users)		-0.007 (0.027)	0.125 (0.085)	0.056** (0.025)
Log(1+SXSW 2006)		-0.007 (0.018)	-0.107** (0.054)	-0.033* (0.019)
Mean dependent variable		0.06	0.04	0.07
Panel C				
Dependent variable:	Log(Twitter Users)	<i>D=1 if always justifiable (10 out of 10)</i>		
Log(Twitter Users)		0.003 (0.007)	0.226 (0.154)	0.019*** (0.007)
Log(1+SXSW 2006)		0.005 (0.005)	-0.193** (0.097)	-0.006 (0.005)
Mean dependent variable		0.009	0.009	0.009
Observations	2,245	2,219	2,224	2,220

Note: This table presents individual-level regressions. Log(Twitter Users) is the natural logarithm of unique Twitter users from a large sample of Tweets from 2014-2015 in the respondents county, SXSW March2007 and SXSW 2006 are the number of Twitter users who follow South by Southwest (@SXSW) and joined Twitter in March 2007 and in 2006 respectively, in the respondents county. Column (1) presents first stage regressions, while columns (2)-(4) present second stages. All regressions include the following individual level controls: Rural/Urban dummy, settlement population, language of interview, interview privacy, sex, age, whether individual was born in the United States, household size, language at home, relationship status, dummy for having children, education, occupation, self-reported "class" and income decile, religiosity, ethnicity, and the full set of county-characteristics (as in column (5) in Table 1, see Table A7 in the Appendix for their descriptions). The outcome variable is constructed from the question: "Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card." Where respondents rate the following on a scale from 1-10: "Political violence" (Column 1), "Terrorism as a political, ideological or religious mean" (Column 2), "Violence against other people" (Column 3). Robust standard errors are clustered by county. *p<0.1; **p<0.05; ***p<0.01

A Appendix

Distribution of number of demonstrations, alternative visualization

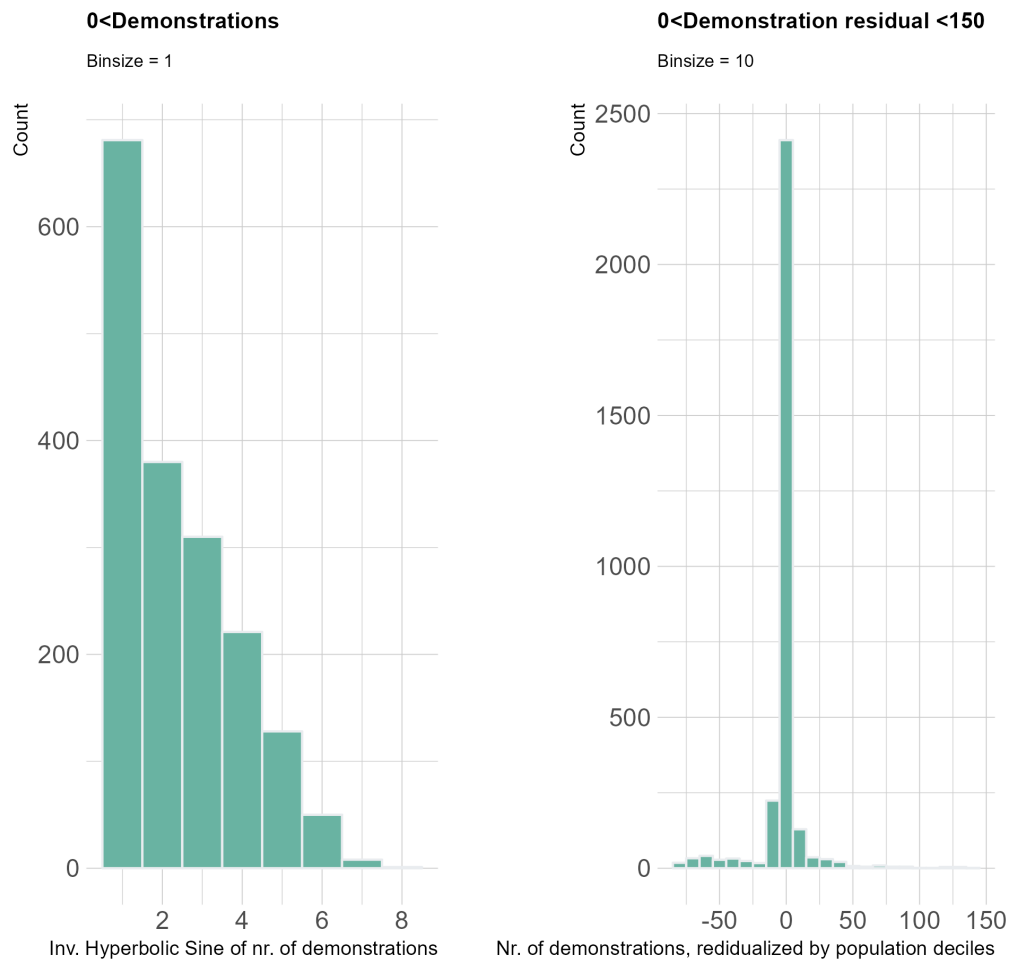


Figure A1: Distribution of the number of demonstrations

Note: The figures show the distribution of the number of demonstrations in the dataset.

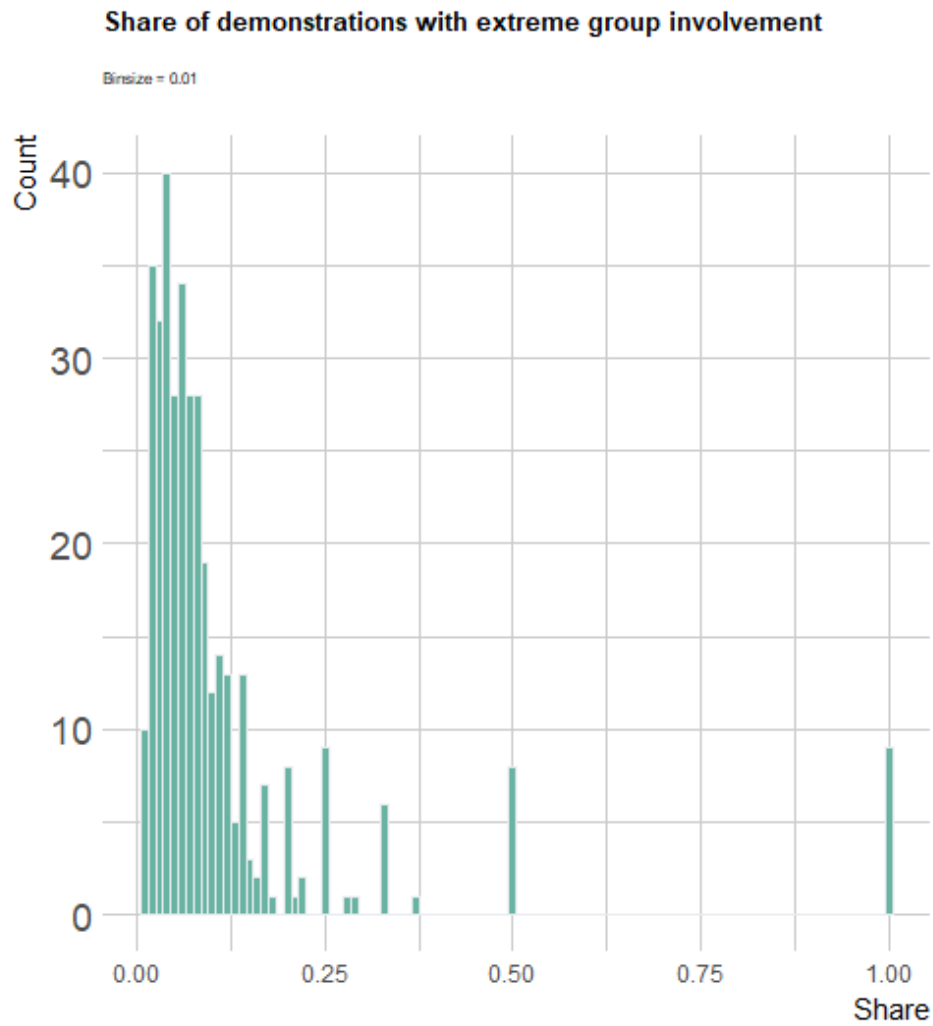


Figure A2: Distribution of the share of "extreme" demonstrations

Note: The figures show the distribution of the share of demonstrations with extreme group involvement, among the 370 counties with at least 1 such demonstration in the dataset.

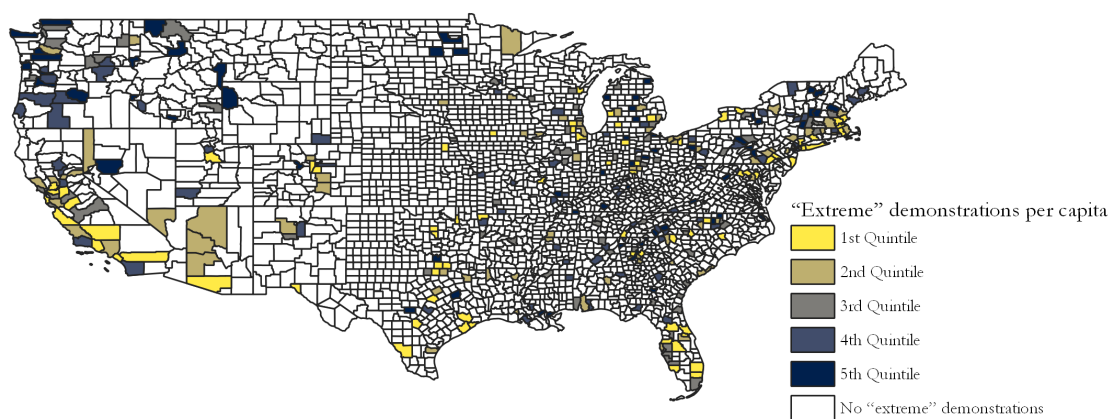
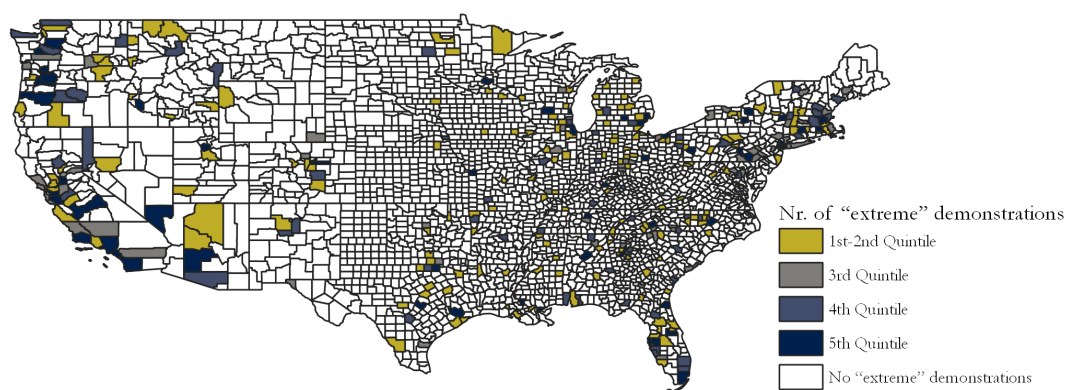
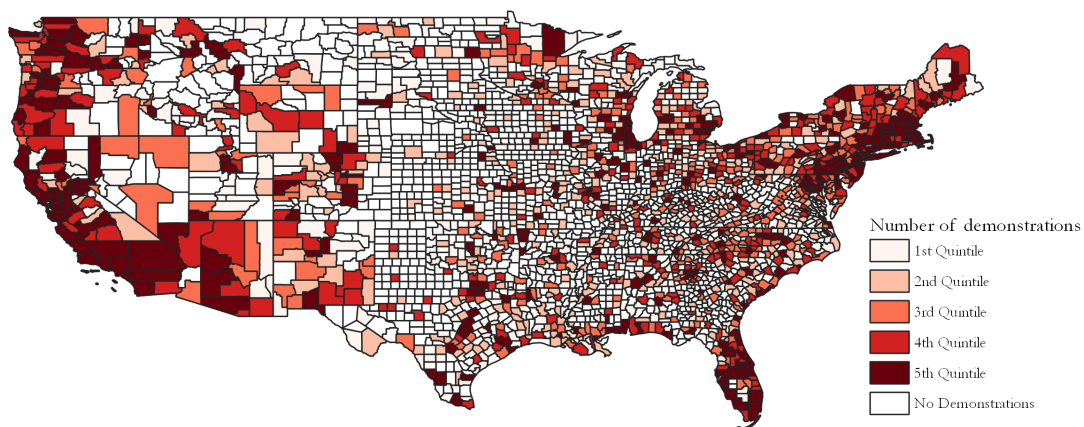


Figure A3: Spacial distribution of demonstration events and extreme demonstration events

Histograms of log(Twitter Users)

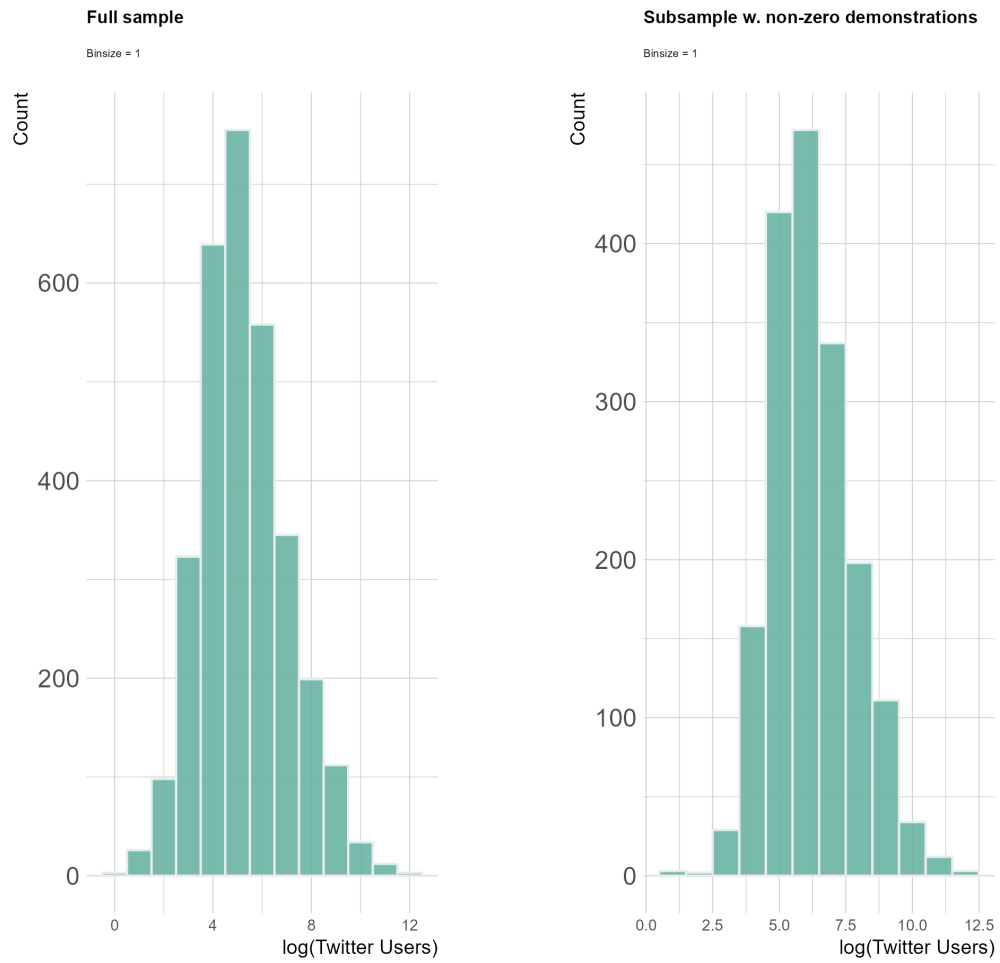


Figure A4: Distribution of tweets

Note: The figures show the distribution of Tweets in the dataset.

Histograms of log(Twitter Users), residualized

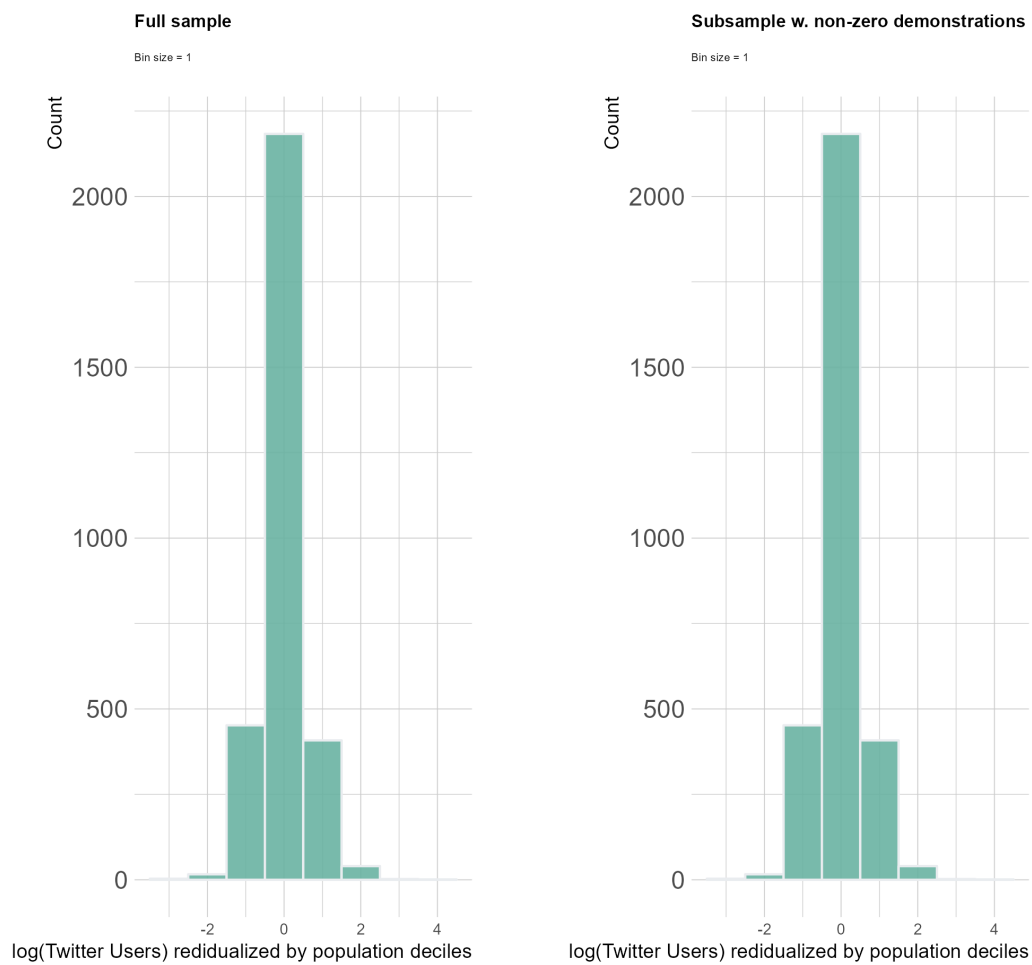


Figure A5: Distribution of tweets

Note: The figures show the distribution of Tweets, residualized by population deciles, in the dataset.

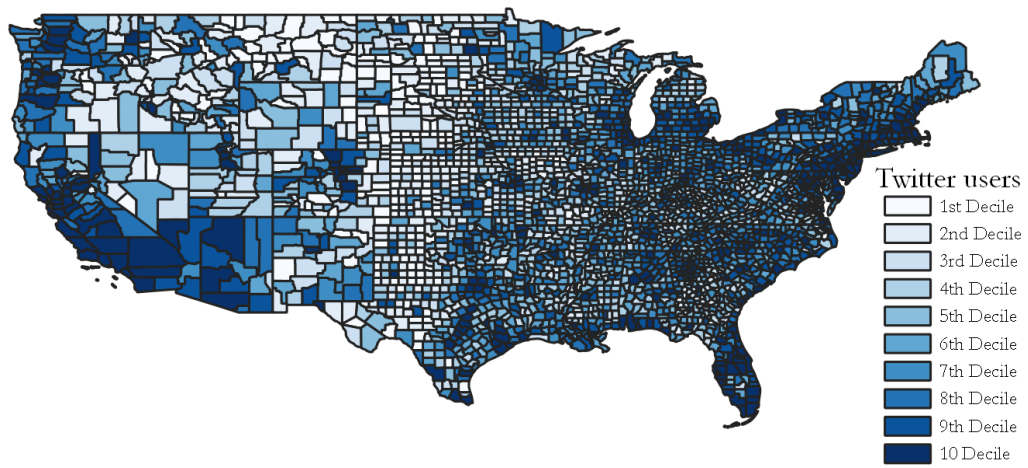


Figure A6: Spacial distribution of Twitter users

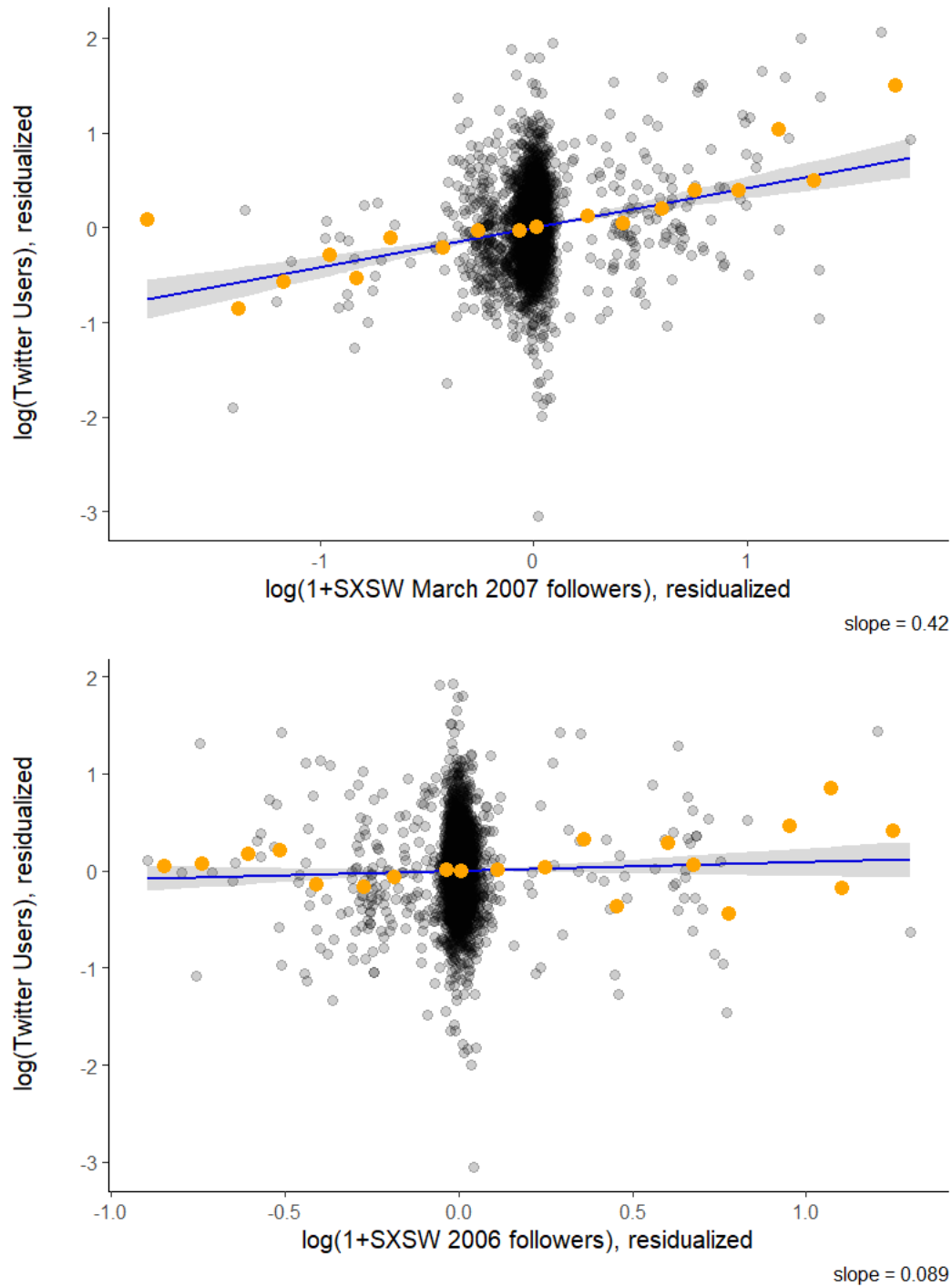


Figure A7: First stage visualization

Note: This figure presents binned scatter plots of the relationship between county-level Twitter users in 2014-2015 and the number of SXSW followers. In the upper panel, the x-axis represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in March 2007 and in the lower panel it represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in 2006. Variables are residualized by partialling out population deciles, demographic, geographic, socio-economic, and pre-2010 election controls (see description of control variables in Table A7 in the Appendix) and the SXSW follower variable not represented in the x-axis of each panel. Grey dot's represent individual observations and yellow dots represent average values of both variables within a bin. The blue line represents a line of best fit using the unbinned data, and the shaded blue area are 95% confidence intervals calculated using robust standard errors clustered at the state-level.

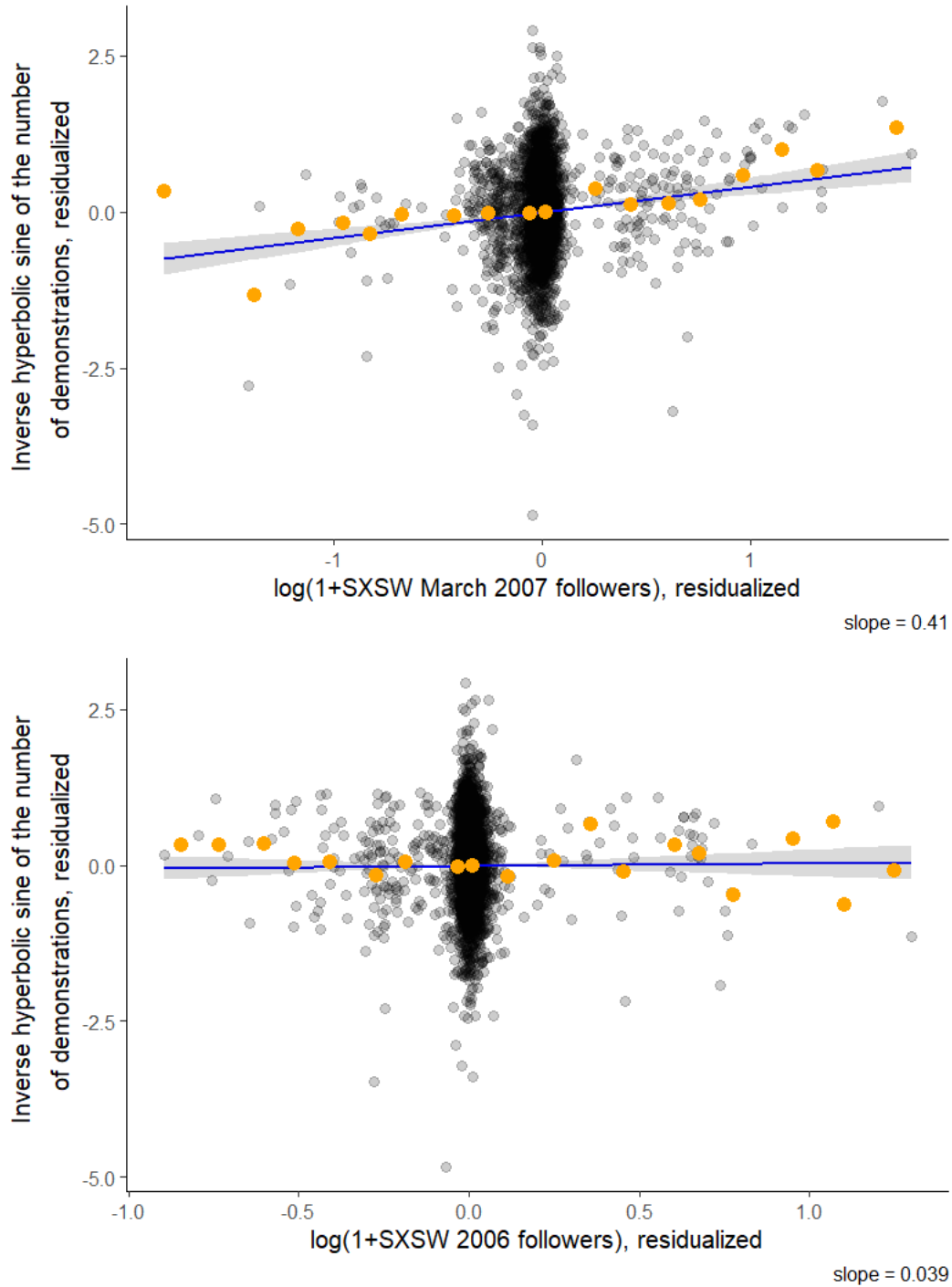


Figure A8: Reduced form visualization, number of protests

Note: This figure presents binned scatter plots of the relationship between county-level number of demonstrations between Jan 2020 and Nov 2021, and the number of SXSW followers. In the upper panel, the x-axis represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in March 2007 and in the lower panel it represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in 2006. Variables are residualized by partialling out population deciles, demographic, geographic, socio-economic, and pre-2010 election controls (see description of control variables in Table A7 in the Appendix) and the SXSW follower variable not represented in the x-axis of each panel. Grey dots represent individual observations and yellow dots represent average values of both variables within a bin. The blue line represents a line of best fit using the unbinned data, and the shaded blue area are 95% confidence intervals calculated using robust standard errors clustered at the state-level.

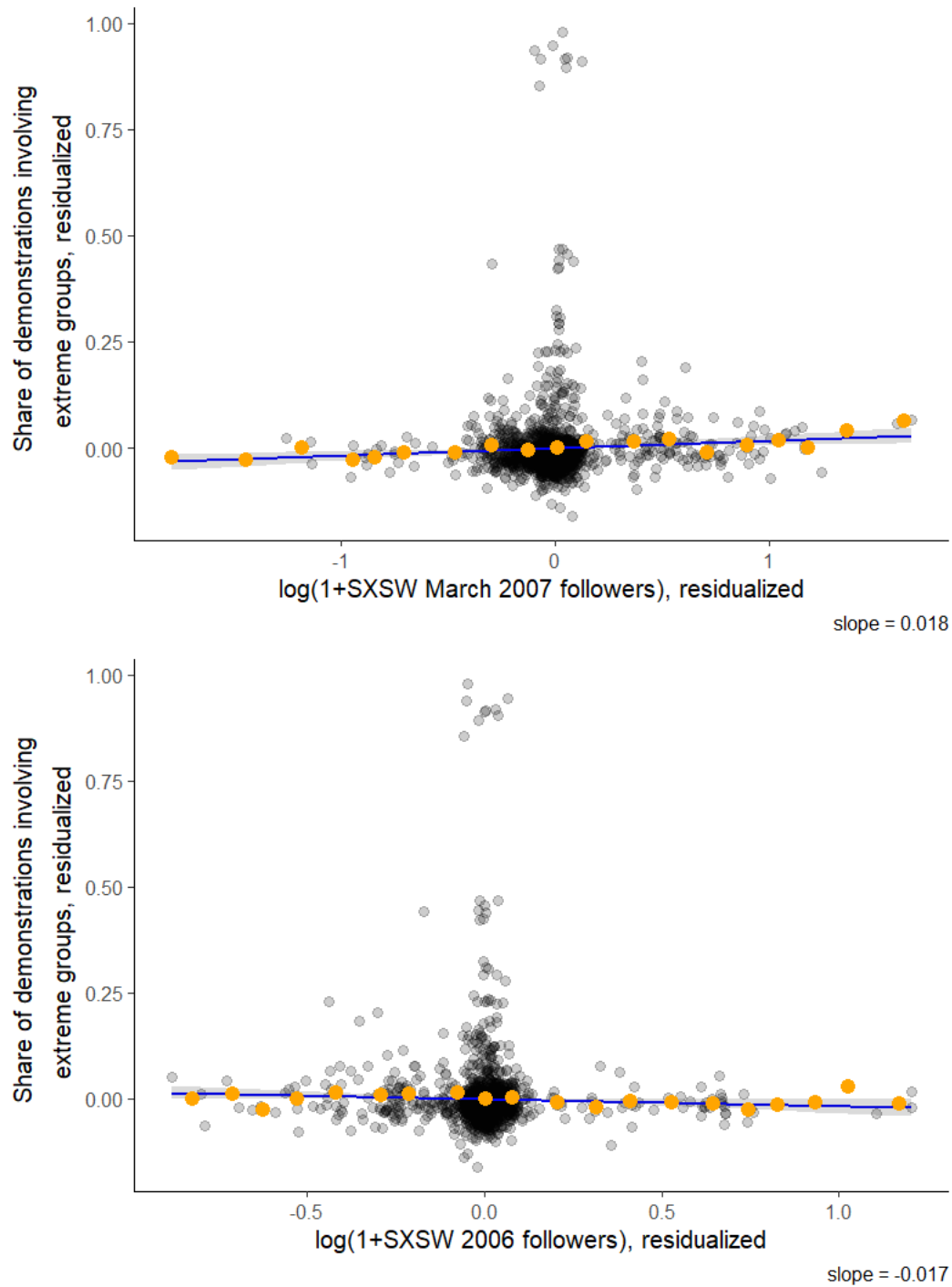


Figure A9: Reduced form visualization, share of demonstrations involving extreme groups

Note: This figure presents binned scatter plots of the relationship between county-level share of demonstrations involving extreme groups and the number of SXSW followers. In the upper panel, the x-axis represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in March 2007 and in the lower panel it represents the logarithm of (one plus) the number of SXSW followers who joined Twitter in 2006. Variables are residualized by partialling out population deciles, demographic, geographic, socio-economic, and pre-2010 election controls (see description of control variables in Table A7 in the Appendix) and the SXSW follower variable not represented in the x-axis of each panel. Grey dot's represent individual observations and yellow dots represent average values of both variables within a bin. The blue line represents a line of best fit using the unbinned data, and the shaded blue area are 95% confidence intervals calculated using robust standard errors clustered at the state-level.

Table A1: Description of filtering processes for topic categorization: Extreme groups

Category	Filter by
Extreme left	
a) "Extreme" BLM:	"Antifa" <i>or</i> "RBPP: Revolutionary Black Panther Party" <i>or</i> "NFAC: Not Fucking Around Coalition" <i>or</i> "TPR: The People's Revolution" <i>or</i> "Take 'Em Down" <i>or</i> "NBPP: New Black Panther Party" <i>or</i> "Black Panthers" <i>or</i> "Black Man Army" <i>or</i> "NABPP: New Afrikan Black Panther Party" <i>or</i> "Detroit Will Breathe" <i>or</i> "No Sleep Till Justice" <i>or</i> "ADOS: American Descendants of Slavery" <i>or</i> "NAIP: New Afrikan Independence Party"
b) Other:	"Cancel the Rents Movement" <i>or</i> "PSL: Party for Socialism and Liberation" <i>or</i> "Refuse Fascism" <i>or</i> "DSA: Democratic Socialists of America" <i>or</i> "Extinction Rebellion" <i>or</i> "APSP: African People's Socialist Party" <i>or</i> "SAP: Socialist Alternative Party" <i>or</i> "CPUSA: Communist Party of the United States of America" <i>or</i> "RCP: Revolutionary Communist Party" <i>or</i> "SWP: Socialist Workers Party"
Extreme right	
a) White supremacists:	"White Nationalists" <i>or</i> "Boogaloo boys" <i>or</i> "Proud Boys" <i>or</i> "Groypers" <i>or</i> "White Defense Force" <i>or</i> "Aryan Nations" <i>or</i> "ABC: Aryan Cowboys Brotherhood" <i>or</i> "QAnon" <i>or</i> "Super Happy Fun America" <i>or</i> "GDL: Goyim Defense League" <i>or</i> "NSC: Nationalist Social Club" <i>or</i> "Boogaloo Boys" <i>or</i> "KKK: Ku Klux Klan" <i>or</i> "Minnesota Patriot Alliance" <i>or</i> "NSM: National Socialist Movement" <i>or</i> "Patriot Front" <i>or</i> "UADF: United American Defence Force" <i>or</i> "Oath Keepers" <i>or</i> ("Washington State III%ers" <i>or</i> "Patriot Prayer" <i>and not</i> "coronavirus" <i>or</i> "gun regulations" in notes)

Table A2: Description of filtering processes for topic categorization: Coronavirus and Presidential elections

Topic	Filter by
Coronavirus public health measures	Notes include: "vaccine" <i>or</i> "vaccinated" <i>or</i> "vaccination" <i>or</i> "mask mandate" <i>or</i> "mask-mandate" <i>or</i> "must wear masks" <i>or</i> "coronavirus mandate" <i>or</i> "coronavirus prevention measures" <i>or</i> "COVID-related restrictions" <i>or</i> "COVID restrictions" <i>or</i> "Covid restrictions" <i>or</i> "public health protocols" <i>or</i> "coronavirus protocols" <i>or</i> "health protection" <i>or</i> "slow the spread" <i>or</i> "wearing masks" <i>or</i> "mask wearing" <i>or</i> "anti-mask" <i>or</i> "pro-mask" <i>or</i> "universal masking" <i>or</i> "coronavirus related safety" <i>or</i> "mask policy" <i>or</i> ("mask" <i>and</i> "coronavirus") <i>or</i> "coronavirus pandemic restrictions" <i>or</i> "restrictions related to the coronavirus" <i>or</i> "contact tracing" <i>or</i> "health and safety measures" <i>or</i> "coronavirus protection" <i>or</i> "coronavirus-related policies" <i>or</i> "public health measures" <i>or</i> "coronavirus test" <i>or</i> "coronavirus precautions" <i>or</i> "quarantine" <i>or</i> "Quarantine" <i>or</i> ("safety protocols" <i>and</i> "corona") <i>or</i> "coronavirus-related mandates" <i>or</i> "stricter measures" <i>or</i> (("mandates" <i>or</i> "measures") <i>and</i> ("corona" <i>or</i> "virus" <i>or</i> "covid")) <i>or</i> "coronavirus restrictions" <i>or</i> "policies on coronavirus" <i>or</i> "protections against the spread" <i>or</i> ("restrictions" <i>and</i> ("corona" <i>or</i> "pandemic" <i>or</i> "covid" <i>or</i> "COVID")) <i>or</i> "in-person learning" <i>or</i> "remote learning" <i>or</i> "Reopen Our Cities" <i>or</i> "reopen businesses" <i>or</i> "reopening businesses" <i>or</i> "business closure" <i>or</i> "school closure" <i>or</i> ("lockdown" <i>and</i> "corona") <i>or</i> "reopening schools" <i>or</i> "reopen schools" <i>or</i> "in-person school" <i>or</i> "reopening of schools" <i>or</i> "in-person classes" <i>or</i> "reopen bars" <i>or</i> "reopening of bars" <i>or</i> "closing of bars" <i>or</i> "restaurant closure" <i>or</i> ("stay-at-home order" <i>and</i> "corona")
Count Every Vote (Left leaning)	Notes include: count every vote" tag
Stop the Steal (Right leaning)	Notes include: "stop the steal" tag

Table A3: Description of filtering processes for topic categorization: Abortion related

Category	Filter by
Abortion rights, Pro-Choice (Left-leaning)	Notes include: ("reproductive rights" or "support of abortion") or Associated actors include: ("PP: Planned Parenthood")
Abortion rights, Pro-Life (Right-leaning)	Notes include: ("pro-life" or "right to live" or "anti abortion" or "Anti abortion" or "anti-abortion" or "against abortion" or "protesting abortion" or ("abortion" and "murder")) or Associated actors include: ("RTL: Right To Life" or "Family Policy Alliance") or (associated actors include "Christian Group" and notes include "abortion")

Table A4: Description of filtering processes for topic categorization: BLM related

Category	Filter by
BLM or related (Left-leaning)	Associated actors include: "BLM: Black Lives Matter" or "Antifa" or "BU: Black Unity" or "RBPP: Revolutionary Black Panther Party" or "NFAC: Not Fucking Around Coalition" or "Expect US" or "NOI: Nation of Islam" or "TPR: The People's Revolution"
Pro-Police/Counter to BLM (Right-leaning)	Associated actors include: ("Back the Blue" or "Blue Lives Matter" or Gastonia Watchmen" or "West Virginia 34 Mountain Milita" or "Pro-Police Group") or Notes include: "back the blue"

Table A5: Description of filtering processes for topic categorization: Racial issues

Category	Filter by
Anti Racism (Left leaning)	The same associated actors as "BLM or related," and in addition: "Stop Asian Hate" or "Take 'Em Down" or "BVM: Black Voters Matter" or "NAACP: National Association for the Advancement of Colored People" or "NBPP: New Black Panther Party" or "Black Panthers"
Pro-Police/Counter to BLM/Racist (Right leaning)	The same associated actors as "Pro-Police/Counter to BLM," and in addition: "White Nationalists" or "Boogaloo boys" or "Proud Boys" or "Groypers" or "White Defence Force" or "Aryan Nations" or "ABC: Aryan Cowboys Brotherhood" or "QAnon" or "Super Happy Fun America" or "GDL: Goyim Defense League" or "NSC: Nationalist Social Club" or "KKK: Ku Klux Klan" or "Minnesota Patriot Alliance" or "NSM: National Socialist Movement" or "Patriot Front" or "UADF: United American Defence Force" or ("Washington State III%ers" or "Patriot Prayer" and not "coronavirus" or "gun regulations" in notes)

Table A6: Description of filtering processes for topic categorization

Category	Filter by
Left leaning	<p>Protest is "Count Every Vote" related <i>or</i> Pro-Choice related <i>or</i> BLM or Anti Racism related <i>or</i> "Extreme left."</p> <p>In addition: Associated actors include: "LGBT" <i>or</i> "NLM: Native Lives Matter" <i>or</i> "FFF: Fridays For Future" <i>or</i> "Cancel the Rents Movement" <i>or</i> "Abolish ICE" <i>or</i> "DEM: Democratic Party"</p>
Right leaning	<p>Protest is "Stop The Steal" related <i>or</i> Pro-Life related <i>or</i> Pro-Police/Counter to BLM/Racist <i>or</i> "Extreme right."</p> <p>In addition: Associated actors include: "Occupy Movement" <i>or</i> "Save Our Children" <i>or</i> "GOA: Gun Owners Of America" <i>or</i> "Everytown for Gun Safety" <i>or</i> "Sons of Liberty" <i>or</i> "Iron City Citizen Response Unit" <i>or</i> "GOP: Republican Party" <i>or</i></p> <p>Notes include: "right to bear arms" <i>or</i> "2nd amendment" <i>or</i> "second amendment" <i>or</i> "against gun regulation" <i>or</i> "gun-rights" <i>or</i> "pro-gun"</p>

Table A7: Description of control variables

Variables	Description	Source
Demographic controls	Include total population, the share of people in age buckets 20-24, 25-29, 30-34, 35-39, 40-44, 45-49 and 50+, the share of Women, the share of Black or African American, American Indian and Alaska Native, Asian, Native Hawaiian or other Pacific Islander and the share of Hispanic among the total population.	2010 US Census
Geographic controls	$\text{Log}(\text{Area})$ in square kilometers, $\text{Log}(\text{Distance from Austin, Texas})$, population density 2016	US Census Tigerline File
Socio-economic controls 1	Poverty rates and median household income	2009 US Census Bureau's SAIP program
Socio-economic controls 2	Share of adult population with at least high school diploma, share of adult population with at least Bachelor degree	2010 American Community Survey
Socio-economic controls 3	Unemployment rate and Employment-to-Population rate in 2010	Bureau of Labor Statistics
Pre-2010 election controls	Share of votes towards the Republican candidate in 2000 (George W. Bush), 2004 (George W. Bush), and 2008 (John McCain)	MIT Election Lab

Table A8: IV balance

Variable:	(1) Both	(2) March 2007 <i>only</i>	(3) 2006 <i>only</i>	(4) T-test (2)-(3) P-value
Population in 2016	1,484	560	427	0.370
Share aged 20-24	0.069	0.076	0.075	0.930
Share aged 25-29	0.086	0.071	0.073	0.475
Share aged 30-34	0.083	0.067	0.068	0.772
Share aged 35-39	0.073	0.065	0.065	0.908
Share aged 40-44	0.063	0.059	0.059	0.931
Share aged 45-49	0.063	0.060	0.059	0.603
Share 50 or older	0.033	0.356	0.0356	0.978
Share women	0.51	0.509	0.504	0.124
Share white	0.662	0.783	0.792	0.781
Share black	0.190	0.131	0.094	0.211
Share native american	0.009	0.009	0.032	0.010**
Share Asian	0.105	0.048	0.053	0.624
Share Pacific Islanders	0.002	0.002	0.001	0.340
Share Hispanic	0.202	0.160	0.154	0.869
Population growth 2000-2016	0.181	0.177	0.155	0.640
Log(Area)	21.07	21.40	21.31	0.726
Log(Distance from Austin, Texas)	14.13	14.22	14.09	0.466
Population Density	0.015	0.004	0.006	0.103
Share Poor	12.13	12.53	11.58	0.447
Log median hh income	11.24	11.09	11.13	0.590
Share adults with highschool degree	0.886	0.896	0.902	0.674
Share adults with bachelor degree	0.429	0.356	0.367	0.701
Unemployment rate	0.033	0.035	0.034	0.774
Employment to population ratio	0.515	0.489	0.494	0.690
Republican vote share 2000	0.374	0.479	0.489	0.748
Republican vote share 2004	0.383	0.501	0.511	0.756
Republican vote share 2008	0.327	0.444	0.470	0.420
Observations	47	108	20	

Table A9: Baseline results, inverse hyperbolic sine of SXSX followers

	(1)	(2)	(3)	(4)	(5)
Panel A: The First Stage					
Dependent variable:	<i>Log(Twitter Users)</i>				
Inverse hyperbolic sine of SXSX March2007	0.497*** (0.046)	0.386*** (0.044)	0.365*** (0.046)	0.342*** (0.044)	0.335*** (0.044)
Inverse hyperbolic sine of SXSX 2006	0.191*** (0.054)	0.122** (0.058)	0.094 (0.060)	0.098* (0.059)	0.092 (0.060)
R ²	0.927	0.938	0.939	0.944	0.945
Panel B: 2SLS Second Stage					
Dependent variable:	<i>Inverse hyperbolic sine of number of demonstrations</i>				
Log(Twitter Users)	0.992*** (0.128)	0.962*** (0.126)	1.047*** (0.141)	0.986*** (0.142)	0.969*** (0.147)
Inverse hyperbolic sine of SXSX 2006	−0.096 (0.090)	−0.093 (0.083)	−0.067 (0.079)	−0.037 (0.075)	−0.036 (0.074)
Cluster robust F-statistic	92.24	55.67	49.81	43.68	41.83
Observations	3,105	3,105	3,105	3,105	3,104
State FE	Yes	Yes	Yes	Yes	Yes
Population deciles	Yes	Yes	Yes	Yes	Yes
Demographic controls		Yes	Yes	Yes	Yes
Geographic controls			Yes	Yes	Yes
Socio-economic controls				Yes	Yes
Pre 2010 election controls					Yes

Note: This table presents county-level regressions. In Panel A the natural logarithm of our Twitter users from 2014-2015 is the dependent variable and in Panel B the inverse hyperbolic sine of the number of demonstrations overall is the dependent variable. SXSX March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by South-west (@SXSX). SXSX 2006 is the number of Twitter users who joined in 2006 and follow SXSX. Regressions include the indicated control variables (see Table A7 in the Appendix for their descriptions). Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A10: Baseline results second stage, alternative outcome specification

	(1)	(2)	(3)	(4)
Dependent variable:	<i>Number of demonstrations</i>			
Functional form of dependent variable:	<i>Linear</i>	<i>Log(1+X)</i>	<i>Log(X)</i>	\sqrt{X}
Panel A: OLS				
Log(Twitter Users)	15.145*** (3.885)	1.009*** (0.138)	0.466*** (0.045)	0.825*** (0.141)
R ²	0.706	0.833	0.808	0.829
Panel B: 2SLS				
Log(Twitter Users)	133.744*** (20.711)	1.009*** (0.138)	0.989*** (0.127)	5.489*** (0.605)
Log(1+SXSW 2006)	34.679 (22.075)	-0.048 (0.084)	-0.093 (0.082)	0.588 (0.482)
Observations	3,104	3,104	1,776	3,104
Mean number of demonstrations	11.49	11.49	20.07	11.49

Note: This table presents county-level regressions from an OLS model and the second stage from a 2SLS model of the effect of Twitter penetration on demonstration frequency. Log(Twitter Users) is the natural logarithm of unique Twitter users from a sample of Tweets from 2014-2015. SXSW 2006 is the number of Twitter users who joined in 2006 and follow South by Southwest (@SXSW). All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions. Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A11: 2SLS demonstrations by topic and extremity, alternative outcome specification

Topic of protest: Specification:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First Stage	Coronavirus	"Stop the Steal"	Abortion rights	BLM related	Racism related	"Extreme"
Panel A							
Dependent variable:	Log(Twitter Users)						
Log(1+SXSW March 2007)	0.417*** (0.059)						
Log(Twitter Users)		9.057*** (2.145)	2.584*** (0.727)	1.548*** (0.319)	45.521*** (9.642)	50.203*** (9.857)	13.807*** (3.052)
Log(1+SXSW 2006)	0.087 (0.074)	4.421* (2.459)	0.145 (0.351)	0.122 (0.319)	3.768 (7.354)	5.226 (7.973)	1.038 (2.646)
R ²	0.945						
Cluster robust F-statistic	36.30						
Observations	3,104						
Panel B							
Dependent variable:							
Log(Twitter Users)		0.708*** (0.133)	0.711*** (0.168)	0.615*** (0.113)	1.009*** (0.135)	1.022*** (0.138)	1.371*** (0.240)
Log(1+SXSW 2006)		0.014 (0.102)	0.019 (0.082)	-0.017 (0.096)	-0.103 (0.092)	-0.102 (0.091)	-0.058 (0.127)
Panel C							
Dependent variable:							
Log(Twitter Users)		1.239*** (0.190)	0.922*** (0.220)	0.780*** (0.142)	3.150*** (0.449)	3.336*** (0.446)	2.223*** (0.406)
Log(1+SXSW 2006)		0.219 (0.204)	0.014 (0.108)	-0.032 (0.121)	-0.029 (0.315)	0.011 (0.320)	-0.009 (0.263)
Observations		3,104	3,104	3,104	3,104	3,104	3,104
Mean number of demonstrations		1.096	0.157	0.193	4.027	4.341	0.424

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March 2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(7) present second stages. Coronavirus refers to demonstrations related to public health measures due to the Covid pandemic. "Stop the Steal" refers to demonstrations related to the 2020 post-presidential election unrest (including those calling for results to be respected). Abortion rights refers to demonstrations related to reproductive rights. BLM related refers to demonstrations concerning police brutality against minorities, and counter movements. Racial issues refers to demonstrations concerning racial tensions while Extreme Left or Right refers to demonstrations involving a political group defined as extreme. A detailed description of topic categorization and the set of extreme groups is presented in Tables A2 to A6 in the Appendix. All regressions control for population via decile dummies, and the full set of controls (as in column 5 in Table 1), see Table A7 in the Appendix for their descriptions. Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A12: 2SLS demonstrations by extreme group involvement, inverse hyperbolic sine of SXSW followers

	(1)	(2)	(3)	(4)
Political leaning:		Extreme overall	Extreme left	Extreme right
Specification:	First Stage	Second Stage		
Panel A				
Dependent variable:	Log(Twitter Users)	<i>Inverse hyperbolic sine of number of demonstrations</i>		
Inverse hyperbolic sine of SXSW March2007	0.335*** (0.044)			
Log(Twitter Users)		1.605*** (0.285)	1.551*** (0.292)	1.005*** (0.205)
Inverse hyperbolic sine of SXSW 2006	0.092 (0.092)	-0.059 (0.116)	-0.054 (0.130)	-0.003 (0.097)
R ²	0.945			
Cluster robust F-statistic	41.83			
Observations	3,104	3,104	3,104	3,104
Mean number demonstrations		0.593	0.424	0.190
Panel B				
Dependent variable:	Log(Twitter Users)	<i>Share among all demonstrations</i>		
Inverse hyperbolic sine of SXSW March 2007	0.347*** (0.043)			
Log(Twitter Users)		0.040*** (0.014)	0.028*** (0.010)	0.014** (0.007)
Inverse hyperbolic sine of SXSW 2006	0.083 (0.061)	-0.016** (0.007)	-0.009 (0.006)	-0.007** (0.003)
R ²	0.939			
Cluster robust F-statistic	45.70			
Observations	1,776	1,776	1,776	1,776
Mean dependent variable		0.024	0.011	0.012

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(5) present second stages. Extreme left and right refers to demonstrations where extreme left or extreme right groups are involved, respectively. A list of extreme groups included can be found in Table A1 in the Appendix. Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions. Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A13: 2SLS demonstrations by political leaning, alternative outcome specification

	(1)	(2)	(3)	(4)	(5)
Political leaning:		Left		Right	
		Overall	Extreme	Overall	Extreme
Specification:	First Stage	Second Stage			
Panel A					
Dependent variable:	Log(Twitter Users)	<i>Number of demonstrations</i>			
Log(1+SXSW March 2007)	0.417*** (0.059)				
Log(Twitter Users)		58.204*** (11.038)	10.344*** (2.782)	8.473*** (1.452)	4.182*** (0.887)
Log(1+SXSW 2006)	0.087 (0.074)	8.764 (9.425)	1.275 (2.478)	0.099 (1.024)	-0.300 (0.554)
R ²	0.945				
Cluster robust F-statistic	36.30				
Panel B					
Dependent variable:		<i>Log(1+Number of demonstrations)</i>			
Log(Twitter Users)		1.104*** (0.137)	1.293*** (0.243)	0.888*** (0.163)	0.835*** (0.168)
Log(1+SXSW 2006)		-0.089 (0.088)	-0.044 (0.141)	-0.057 (0.095)	-0.016 (0.099)
Panel C					
Dependent variable:		<i>√Number of demonstrations</i>			
Log(Twitter Users)		3.706*** (0.470)	1.954*** (0.392)	1.431*** (0.247)	1.121*** (0.229)
Log(1+SXSW 2006)		0.151 (0.345)	0.028 (0.275)	-0.071 (0.148)	-0.009 (0.136)
Observations	3,104	3,104	3,104	3,104	3,104
Mean number of demonstrations		4.751	0.424	0.845	0.190

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(5) present second stages. Mainstream left refers to demonstrations which are left-leaning, mainstream right refers to right-leaning demonstrations, Extreme left and right refers to demonstrations where extreme left or extreme right groups are involved, respectively. A detailed description of protest categorization can be found in Tables A2 to A6 in the Appendix. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions.

Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A14: 2SLS Presidential election related demonstrations

	(1)	(2)	(3)
Political leaning:		"Count Every Vote"	"Stop The Steal"
Specification:	First Stage	Second Stage	
Panel A			
Dependent variable:	Log(Twitter Users)	<i>Inverse hyperbolic sine of number of demonstrations</i>	
Log(1+SXSW March 2007)	0.417*** (0.059)		
Log(Twitter Users)		0.420*** (0.136)	0.639*** (0.202)
Log(1+SXSW 2006)	0.087 (0.074)	0.028 (0.079)	0.033 (0.123)
R ²	0.945		
Cluster robust F-statistic	36.30		
Observations	3,104	3,104	3,104
Mean number of demonstrations		0.050	0.157
Panel B			
Dependent variable:	Log(Twitter Users)	<i>Share among all demonstrations</i>	
Log(1+SXSW March 2007)	0.433*** (0.056)		
Log(Twitter Users)		0.0003 (0.003)	0.007 (0.010)
Log(1+SXSW 2006)	0.075 (0.075)	-0.002 (0.002)	-0.006 (0.005)
R ²	0.939		
Cluster robust F-statistic	43.34		
Observations	1,776	1,776	1,776
Mean dependent variable		0.002	0.010

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(3) present second stages. A detailed description of protest categorization can be found in Table A2.

Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions.

Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A15: 2SLS Abortion related demonstrations

	(1)	(2)	(3)
Political leaning:		Pro-Choice	Pro-Life
Specification:	First Stage	Second Stage	
Panel A			
Dependent variable:	Log(Twitter Users)	<i>Inverse hyperbolic sine of number of demonstrations</i>	
Log(1+SXSW March 2007)	0.417*** (0.059)		
Log(Twitter Users)		0.483*** (0.144)	0.612*** (0.117)
Log(1+SXSW 2006)	0.087 (0.074)	0.059 (0.124)	-0.049 (0.100)
R ²	0.945		
Cluster robust F-statistic	36.30		
Observations	3,104	3,104	3,104
Mean number of demonstrations		0.074	0.134
Panel B			
Dependent variable:	Log(Twitter Users)	<i>Share among all demonstrations</i>	
Log(1+SXSW March 2007)	0.433*** (0.056)		
Log(Twitter Users)		-0.003 (0.003)	0.001 (0.007)
Log(1+SXSW 2006)	0.075 (0.075)	0.001 (0.003)	-0.002 (0.005)
R ²	0.939		
Cluster robust F-statistic	43.34		
Observations	1,776	1,776	1,776
Mean dependent variable		0.005	0.011

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(3) present second stages. A detailed description of protest categorization can be found in Table A3.

Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions.

Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.001

Table A16: 2SLS BLM demonstrations

	(1)	(2)	(3)
Political leaning:		BLM or related	Pro-police/counter to BLM
Specification:	First Stage	Second Stage	
Panel A			
Dependent variable:	Log(Twitter Users)	<i>Inverse hyperbolic sine of number of demonstrations</i>	
Log(1+SXSW March 2007)	0.417*** (0.059)		
Log(Twitter Users)		1.073*** (0.149)	0.728*** (0.169)
Log(1+SXSW 2006)	0.087 (0.074)	-0.107 (0.104)	-0.210* (0.124)
R ²	0.945		
Cluster robust F-statistic	36.30		
Observations	3,104	3,104	3,104
Mean number of demonstrations		3.822	0.298
Panel B			
Dependent variable:	Log(Twitter Users)	<i>Share among all demonstrations</i>	
Log(1+SXSW March 2007)	0.433*** (0.056)		
Log(Twitter Users)		-0.037 (0.039)	0.008 (0.015)
Log(1+SXSW 2006)	0.075 (0.075)	0.017 (0.025)	-0.009 (0.009)
R ²	0.939		
Cluster robust F-statistic	43.34		
Observations	1,776	1,776	1,776
Mean dependent variable		0.435	0.038

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(3) present second stages. A detailed description of protest categorization can be found in Table A4.

Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions.

Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A17: 2SLS Racism related demonstrations

	(1)	(2)	(3)
Political leaning:		BLM/Anti Racism	Pro-police/counter to BLM/Racist
Specification:	First Stage	Second Stage	
Panel A			
Dependent variable:	Log(Twitter Users)	<i>Inverse hyperbolic sine of number of demonstrations</i>	
Log(1+SXSW March 2007)	0.417*** (0.059)		
Log(Twitter Users)		1.059*** (0.150)	1.058*** (0.202)
Log(1+SXSW 2006)	0.087 (0.074)	-0.102 (0.102)	-0.196 (0.121)
R ²	0.945		
Cluster robust F-statistic	36.30		
Observations	3,104	3,104	3,104
Mean number of demonstrations		4.045	0.426
Panel B			
Dependent variable:	Log(Twitter Users)	<i>Share among all demonstrations</i>	
Log(1+SXSW March 2007)	0.433*** (0.056)		
Log(Twitter Users)		-0.037 (0.039)	0.016 (0.015)
Log(1+SXSW 2006)	0.075 (0.075)	0.014 (0.024)	-0.016* (0.009)
R ²	0.939		
Cluster robust F-statistic	43.34		
Observations	1,776	1,776	1,776
Mean dependent variable		0.447	0.044

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(3) present second stages. A detailed description of protest categorization can be found in Table A5.

Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions.

Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A18: 2SLS Extreme left groups

	(1)	(2)	(3)
Political leaning:		“Extreme BLM”	Extreme left other
Specification:	First Stage	Second Stage	
Panel A			
Dependent variable:	Log(Twitter Users)	<i>Inverse hyperbolic sine of number of demonstrations</i>	
Log(1+SXSW March 2007)	0.417*** (0.059)		
Log(Twitter Users)		0.784*** (0.211)	1.509*** (0.264)
Log(1+SXSW 2006)	0.087 (0.074)	-0.109 (0.126)	-0.053 (0.161)
R ²	0.945		
Cluster robust F-statistic	36.30		
Observations	3,104	3,104	3,104
Mean number of demonstrations		0.076	0.362
Panel B			
Dependent variable:	Log(Twitter Users)	<i>Share among all demonstrations</i>	
Log(1+SXSW March 2007)	0.433*** (0.056)		
Log(Twitter Users)		0.008** (0.004)	0.022*** (0.008)
Log(1+SXSW 2006)	0.075 (0.075)	-0.002 (0.002)	-0.010 (0.007)
R ²	0.939		
Cluster robust F-statistic	43.34		
Observations	1,776	1,776	1,776
Mean dependent variable		0.002	0.009

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(3) present second stages. A detailed description of protest categorization can be found in Table A1.

Panel A presents 2SLS results where the inverse hyperbolic sine of the number of demonstrations fitting the given column category is the dependent variable. Panel B presents 2SLS results on a subsample of counties where at least 1 demonstration event occurred between Jan 2020 and Nov 2021, where the dependent variable is the share of demonstration that fit the column-category. All regressions control for population via decile dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions.

Robust standard errors in parentheses are clustered by state. *p<0.1; **p<0.05; ***p<0.01

Table A19: 2SLS demonstrations by aggression, Alternative outcome specification

Type of demonstration:	(1)	(2)	(3)	(4)	(5)
Specification:	First Stage	Peaceful	Peaceful Protesters met with intervention or violence	Violent Protesters	Armed demonstrators
Panel A					
Dependent variable:	Log(Twitter Users)	<i>Number of demonstrations</i>			
Log(1+SXSW March 2007)	0.417*** (0.059)				
Log(Twitter Users)		116.323*** (15.998)	5.993*** (1.473)	10.085* (5.633)	3.765*** (1.116)
Log(1+SXSW 2006)	0.087 (0.074)	35.198* (19.345)	0.911 (1.069)	-1.897 (3.615)	-1.254* (0.707)
R ²	0.945				
Cluster robust F-statistic	36.30				
Observations	3,104				
Panel B					
Dependent variable:	<i>Log(1+Number of demonstrations)</i>				
Log(Twitter Users)		0.976*** (0.138)	0.926*** (0.165)	0.979*** (0.202)	0.867*** (0.184)
Log(1+SXSW 2006)		-0.034 (0.083)	0.067 (0.103)	-0.015 (0.142)	-0.232* (0.123)
Panel C					
Dependent variable:	<i>√Number of demonstrations</i>				
Log(Twitter Users)		5.062*** (0.509)	1.341*** (0.247)	1.526*** (0.449)	1.170*** (0.260)
Log(1+SXSW 2006)		0.654 (0.440)	0.118 (0.154)	-0.073 (0.298)	-0.332* (0.172)
Observations		3,104	3,104	3,104	3,104
Mean number of demonstrations		10.796	0.322	0.301	0.194

Note: This table presents county-level regressions. Log(Twitter Users) is the natural logarithm of our Twitter users from 2014-2015. SXSW March2007 is the number of Twitter users who joined Twitter in March 2007 and follow South by Southwest (@SXSW). SXSW 2006 is the number of Twitter users who joined in 2006 and follow SXSW. Column (1) presents first stage regressions, while columns (2)-(5) present second stages. All regressions control for population via decade dummies, and the full set of controls (as in column (5) in Table 1), see Table A7 in the Appendix for their descriptions. Robust standard errors in parentheses are clustered by state.

*p<0.1; **p<0.05; ***p<0.01

Table A20: Effects on strong opinions on statements related to gender issues, Cooperative Election Study

	(1)	(2)	(3)
Statement:		Women too easily offended	Women seek power
Specification:	First Stage	Second Stage	
Dependent variable:	Log(Twitter Users)	<i>D=1 if strongly conservative opinion</i>	
Log(1+SXSW March 2007)	0.492*** (0.048)		
Log(Twitter Users)		0.002 (0.011)	0.015** (0.007)
Log(1+SXSW 2006)	0.018 (0.075)	0.009 (0.007)	-0.007 (0.005)
Mean dependent variable		0.10	0.27
R ²	0.956		
Cluster robust F-statistic	104.89		
Dependent variable:		<i>D=1 if strongly liberal opinion</i>	
Log(Twitter Users)		0.012 (0.010)	0.011 (0.012)
Log(1+SXSW 2006)		-0.015* (0.008)	-0.002 (0.009)
Mean dependent variable		0.28	0.23
Observations	54,317	46,173	46,379

Note: This table presents individual-level regressions. Log(Twitter Users) is the natural logarithm of unique Twitter users from a large sample of Tweets from 2014-2015 SXSW March2007 and SXSW 2006 are the number of Twitter users who follow South by Southwest (@SXSW) and joined Twitter in March 2007 and in 2006 respectively. Column (1) presents first stage regressions, while columns (2)-(6) present second stages. All regressions include the following individual level controls: Age, gender, education, race, ethnicity, employment status, home-ownership, family income, number of children and internet connection, and the full set of county-characteristics (as in column (5) in Table 1, see Table A7 in the Appendix for their descriptions. The statements are the following: “Women too easily offended” which is not shortened and “Women seek power”: “Women seek to gain power by getting control over men” where “strongly agree” is the strongly conservative option and “strongly disagree” is the strongly liberal option. Observations are weighted to represent adult Americans and robust standard errors are clustered by county.

*p<0.1; **p<0.05; ***p<0.01

Table A21: Effects on strong ideology ratings

	(1)	(2)	(3)	(4)
Who to rate:		Oneself	Democratic party	Republican party
Specification:	First Stage	Second Stage		
Dependent variable:	Log(Twitter Users)	<i>D=1 if Very liberal</i>		
Log(1+SXSX March 2007)	0.492*** (0.048)			
Log(Twitter Users)		0.021*** (0.008)	0.002 (0.011)	-0.005 (0.005)
Log(1+SXSX 2006)	0.018 (0.075)	-0.008 (0.006)	0.002 (0.009)	0.006 (0.004)
Mean dependent variable		0.13	0.33	0.02
R ²	0.956			
Cluster robust F-statistic	104.89			
Dependent variable:		<i>D=1 if Very conservative</i>		
Log(Twitter Users)		-0.005 (0.008)	-0.004 (0.004)	0.003 (0.012)
Log(1+SXSX 2006)		0.005 (0.005)	0.003 (0.003)	-0.006 (0.009)
Mean dependent variable		0.11	0.02	0.40
Observations	54,354	54,298	54,298	54,292

Note: This table presents individual-level regressions. Log(Twitter Users) is the natural logarithm of unique Twitter users from a large sample of Tweets from 2014-2015 SXSX March2007 and SXSX 2006 are the number of Twitter users who follow South by Southwest (@SXSX) and joined Twitter in March 2007 and in 2006 respectively. Column (1) presents first stage regressions, while columns (2)-(6) present second stages. All regressions include the following individual level controls: Age, gender, education, race, ethnicity, employment status, home-ownership, family income, number of children and internet connection, and the full set of county-characteristics (as in column (5) in Table 1, see Table A7 in the Appendix for their descriptions. Observations are weighted to represent adult Americans and robust standard errors are clustered by county. *p<0.1; **p<0.05; ***p<0.01

Table A22: Effects on attitudes towards violence and terrorism alternative thresholds, World Value Survey 2017

	(1)	(2)	(3)	(4)
		Political violence can be justified	Terrorism can be justified	Violence can be justified
Specification:	First Stage	Second Stage		
Dependent variable:	Log(Twitter Users)	<i>D=1 if justifiable>1 (1=never, out of 10)</i>		
Log(1+SXSW March 2007)	0.493*** (0.062)			
Log(Twitter Users)		0.002 (0.044)	0.015 (0.037)	0.019 (0.040)
Log(1+SXSW 2006)	0.065 (0.084)	0.021 (0.032)	-0.044* (0.026)	-0.007 (0.028)
Mean dependent variable		0.41	0.18	0.47
R ²	0.732			
Cluster robust F-statistic	58.08			
Dependent variable:		<i>D=1 if justifiable>2 (out of 10)</i>		
Log(Twitter Users)		0.042 (0.041)	0.050 (0.034)	-0.015 (0.037)
Log(1+SXSW 2006)		-0.008 (0.028)	-0.043** (0.022)	-0.008 (0.028)
Mean dependent variable		0.28	0.11	0.33
Dependent variable:		<i>D=1 if justifiable>3 (out of 10)</i>		
Log(Twitter Users)		0.045 (0.034)	0.075 (0.051)	0.023 (0.033)
Log(1+SXSW 2006)		-0.017 (0.024)	-0.064** (0.032)	-0.014 (0.025)
Mean dependent variable		0.20	0.08	0.22
Observations		2,219	2,224	2,220

Note: This table presents individual-level regressions. Log(Twitter Users) is the natural logarithm of unique Twitter users from a large sample of Tweets from 2014-2015 in the respondents county, SXSW March2007 and SXSW 2006 are the number of Twitter users who follow South by Southwest (@SXSW) and joined Twitter in March 2007 and in 2006 respectively, in the respondents county. Column (1) presents first stage regressions, while columns (2)-(4) present second stages. All regressions include the following individual level controls: Rural/Urban dummy, settlement population, language of interview, interview privacy, sex, age, whether individual was born in the United States, household size, language at home, relationship status, dummy for having children, education, occupation, self-reported “class” and income decile, religiosity, ethnicity, and the full set of county-characteristics (as in column (5) in Table 1, see Table A7 in the Appendix for their descriptions. The outcome variable is constructed from the question: “Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card.” Where respondents rate the following on a scale from 1-10: “Political violence” (Column 1), “Terrorism as a political, ideological or religious mean” (Column 2), “Violence against other people” (Column 3). Robust standard errors are clustered by county. *p<0.1; **p<0.05; ***p<0.01

Table A23: Effects on attitudes towards violence and terrorism alternative thresholds, World Value Survey 2017

	(1)	(2)	(3)	(4)
		Political violence can be justified	Terrorism can be justified	Violence can be justified
Specification:	First Stage	Second Stage		
Dependent variable:	Log(Twitter Users)	<i>D=1 if justifiable > 4 (out of 10)</i>		
Log(1+SXSW March 2007)	0.493*** (0.062)			
Log(Twitter Users)		0.040 (0.033)	0.100 (0.068)	0.042 (0.036)
Log(1+SXSW 2006)	0.065 (0.084)	-0.021 (0.024)	-0.086** (0.043)	-0.018 (0.026)
Mean dependent variable		0.15	0.07	0.18
R ²	0.732			
Cluster robust F-statistic	58.08			
Dependent variable:		<i>D=1 if justifiable > 5 (out of 10)</i>		
Log(Twitter Users)		-0.007 (0.027)	0.125 (0.085)	0.056** (0.025)
Log(1+SXSW 2006)		-0.007 (0.018)	-0.107** (0.054)	-0.033* (0.019)
Mean dependent variable		0.06	0.04	0.07
Dependent variable:		<i>D=1 if justifiable > 6 (out of 10)</i>		
Log(Twitter Users)		0.002 (0.018)	0.150 (0.102)	0.035** (0.018)
Log(1+SXSW 2006)		-0.005 (0.014)	-0.129** (0.065)	-0.024* (0.013)
Mean dependent variable		0.04	0.02	0.05
Observations		2,219	2,224	2,220

Note: This table presents individual-level regressions. Log(Twitter Users) is the natural logarithm of unique Twitter users from a large sample of Tweets from 2014-2015 in the respondents county, SXSW March2007 and SXSW 2006 are the number of Twitter users who follow South by Southwest (@SXSW) and joined Twitter in March 2007 and in 2006 respectively, in the respondents county. Column (1) presents first stage regressions, while columns (2)-(4) present second stages. All regressions include the following individual level controls: Rural/Urban dummy, settlement population, language of interview, interview privacy, sex, age, whether individual was born in the United States, household size, language at home, relationship status, dummy for having children, education, occupation, self-reported "class" and income decile, religiosity, ethnicity, and the full set of county-characteristics (as in column (5) in Table 1, see Table A7 in the Appendix for their descriptions). The outcome variable is constructed from the question: "Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card." Where respondents rate the following on a scale from 1-10: "Political violence" (Column 1), "Terrorism as a political, ideological or religious mean" (Column 2), "Violence against other people" (Column 3). Robust standard errors are clustered by county. *p<0.1; **p<0.05; ***p<0.01

Table A24: Effects on attitudes towards violence and terrorism alternative thresholds, World Value Survey 2017

	(1)	(2)	(3)	(4)
		Political violence can be justified	Terrorism can be justified	Violence can be justified
Specification:	First Stage	Second Stage		
Dependent variable:		<i>D=1 if justifiable > 7 (out of 10)</i>		
Log(1+SXSX March 2007)	0.493*** (0.062)			
Log(Twitter Users)		0.001 (0.010)	0.176 (0.119)	0.034*** (0.011)
Log(1+SXSX 2006)	0.065 (0.084)	0.001 (0.008)	-0.150** (0.076)	-0.018** (0.008)
R ²	0.732			
Cluster robust F-statistic	58.08			
Mean dependent variable		0.02	0.01	0.03
Dependent variable:		<i>D=1 if justifiable > 8 (out of 10)</i>		
Log(Twitter Users)		0.009 (0.009)	0.201 (0.136)	0.033*** (0.008)
Log(1+SXSX 2006)		0.006 (0.006)	-0.172** (0.086)	-0.006 (0.006)
Mean dependent variable		0.01	0.01	0.01
Dependent variable:	Log(Twitter Users)	<i>D=1 if always justifiable (10 out of 10)</i>		
Log(Twitter Users)		0.003 (0.007)	0.226 (0.154)	0.019*** (0.007)
Log(1+SXSX 2006)		0.005 (0.005)	-0.193** (0.097)	-0.006 (0.005)
Mean dependent variable		0.009	0.009	0.009
Observations	2,245	2,219	2,224	2,220

Note: This table presents individual-level regressions. Log(Twitter Users) is the natural logarithm of unique Twitter users from a large sample of Tweets from 2014-2015 in the respondents county, SXSX March2007 and SXSX 2006 are the number of Twitter users who follow South by Southwest (@SXSX) and joined Twitter in March 2007 and in 2006 respectively, in the respondents county. Column (1) presents first stage regressions, while columns (2)-(4) present second stages. All regressions include the following individual level controls: Rural/Urban dummy, settlement population, language of interview, interview privacy, sex, age, whether individual was born in the United States, household size, language at home, relationship status, dummy for having children, education, occupation, self-reported "class" and income decile, religiosity, ethnicity, and the full set of county-characteristics (as in column (5) in Table 1, see Table A7 in the Appendix for their descriptions. The outcome variable is constructed from the question: "Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card." Where respondents rate the following on a scale from 1-10: "Political violence" (Column 1), "Terrorism as a political, ideological or religious mean" (Column 2), "Violence against other people" (Column 3). Robust standard errors are clustered by county. *p<0.1; **p<0.05; ***p<0.01