I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15812

Social Media and the Broadening of Social Movements: Evidence from Black Lives Matter

Annalí Casanueva Artís
Vladimir Avetian
Sulin Sardoschau
Kritika Saxena

DECEMBER 2022

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15812

Social Media and the Broadening of Social Movements: Evidence from Black Lives Matter

Annalí Casanueva Artís ®
Paris School of Economics

Vladimir Avetian ®
Paris Dauphine University - PSL

Sulin Sardoschau ®
Humboldt University and IZA

Kritika Saxena
University of Groningen

DECEMBER 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Social Media and the Broadening of Social Movements: Evidence from Black Lives Matter*

How do modern social movements broaden their base? Prompted by the viral video footage of George Floyd's murder, the Black Lives Matter (BLM) movement gained unprecedented scope in the spring of 2020. In this paper, we show that pandemic exposure (COVID-19 related deaths) significantly increased the take-up of social media and subsequently mobilized protesters in whiter, more affluent and suburban counties with low ex-ante probability of protesting. We exploit Super Spreader Events in the early stages of the pandemic as a source of plausibly exogenous variation at the county level and develop a novel index of social media penetration, using information from more than 45 million tweets, google searches and mobility data. We show that a one standard deviation increase in pandemic exposure increased the number of new Twitter accounts by 27% and increased protest propensity by 9 percentage points. Our results suggest that social media can be persuasive and inspire action outside of traditional coalitions.

JEL Classification: P16, D7
Keywords: social media, BLM, protest, COVID-19

Corresponding author:
Sulin Sardoschau
Humboldt University
Unter den Linden 6
10117 Berlin
Germany
E-mail: sulin.sardoschau@hu-berlin.de

* The authors thank Maja Adena, Hunt Allcott, Abhijit Banerjee, François Bourguignon, Eric Brousseau, Davide Cantoni, Ruben Durante, Roberto Galbiati, Sergei Guriev, Pamina Koenig, Jonathan Lehne, Alexey Makarin, Benjamin Marx, Nikita Melnikov, Raúl Sánchez de la Sierra, Lore Vandewalle, Carlos Winograd, Maiting Zhuang, and Ekaterina Zhuravskaya for their useful comments and suggestions. We also thank Marlon Seror for his invaluable inputs. We are thankful to all the participants of the 25th Annual conference of the Society for Institutional and Organizational Economics (SIOE), the annual meeting of the European Economics Association (EEA-ESEM), the Households in Conflict (HiCN) conference, the AFSE conference, the Lisbon Meeting on Economics and Political Science 2022, the brown bag seminar at The Graduate Institute of International and Development Studies, the Science Po internal seminar, the Empirical Economics Seminar at Ludwigs Maximilian University Munich, the Development Economics Network Berlin, the Applied Young Economists Webinar, the Milan Labor Lunch Seminar, the Berlin Applied Micro Seminar, and the SSES Annual Congress of 2022 for their useful feedback. The order of authors has been randomized as indicated by the superscript ®.

1 Introduction

There is a far more representative cross-section of America out on the streets [...] That didn't exist back in the 1960s. That broad coalition.
- Barack Obama, June 3rd 2020

The effectiveness of social movements depends on their ability to mobilize allies, build coalitions and inspire reform through collective action (Olson, 1989; Ostrom, 1990; Della Porta and Diani, 2015, 2020). Traditionally, mobilization was carried out at the local level via face-to-face interactions (Morris, 1986). Today, political activism has moved to the virtual space and modern social movements communicate, mobilize and persuade through social media (McKersie, 2021).

The Black Lives Matter (BLM) movement was born on Twitter in 2013. In May of 2020, #BlackLivesMatter became the most popular hashtag on Twitter, peaking at 8.8 million mentions per day (PEW, 2020). Videos on Twitter about the murder of George Floyd at the hands of police officer Derek Chauvin were watched over 1.4 billion times and the ensuing protest was labeled the "largest" and "broadest" social movement in the history of the United States (NYT, WP, 2020).

In the weeks leading up to the protests, social media platforms announced record growth of new users and online activity, which they attributed to the outbreak of the pandemic. Meta reported a 50% increase in time spent across their platforms Facebook, Instagram and Whatsapp (Meta, 2020). In a letter to their shareholders, Twitter reported an increase in daily active users by 24% that "was driven by an increased engagement due to the COVID-19 pandemic" (Twitter, 2020).

In this paper, we provide evidence that the broadening of the BLM movement during the pandemic was driven by an increase in the use of social media. We approach this question in two parts. First, we show "reduced form" evidence on the link between pandemic exposure and BLM protest participation at the county level. We focus on the set of counties that had never protested for a BLM related cause before and leverage Super Spreader Events (SSE) in the early stages of the pandemic to establish a causal link. We find that a one standard deviation increase in pandemic exposure (23 deaths by 100K inhabitants) led to an increase in protest probability by 9 percentage points.

In the second part of the paper, we investigate whether the uptake in social media can account for the pandemic-induced broadening of the BLM movement and rule out alternative explanations (such as a rise in the salience of racial inequality, lower opportunity costs of protesting or higher overall propensity to protest). Our estimates suggest that a one standard deviation increase in pandemic exposure led to a 27% increase in new Twitter accounts. We also show that BLM protests only arise in places that increase the use of social media in response to the pandemic. Based on this evidence, we conjecture that new segments of the population became active on social media during the pandemic and were subsequently exposed to the viral protest trigger that ultimately motivated them to take to the streets.

Previous work has shown that social media can solve the collective action and coordination problem for individuals already sympathetic to a political cause (Cantoni et al., 2019; Bursztyn et al., 2021; Enikolopov et al., 2020; Manacorda and Tesei, 2020). Most similar to our study, Enikolopov et al. (2020) show that social media helps to solve the collective action problem in a one-shot setting, where the roll-out of a social media platform coincides with a contested election in Russia. In contrast, we focus on the role of social media as a tool that can *broaden* alliances and mobilize new fractions of society.

In addition, previous papers exploit supply side constraints (informal networks or infrastructure) in the early stages of internet or social media roll-out going back to the early 2000s (Guriev et al., 2019; Müller and Schwarz, 2020; Enikolopov et al., 2020; Manacorda and Tesei, 2020). However, initial constraints become less relevant over time and do not account for more recent determinants of social media penetration. To the best of our knowledge, we are the first to show that COVID-19 acted as a demand shock for social media, overcoming important challenges in identifying the effects of social media in saturated markets.

Our empirical strategy is based on a small time window between the end of March and mid April of 2020 during which COVID-19 was prevalent enough but lockdown stringency lax enough to allow for so-called SSEs to occur. These events are characterized by the presence of one highly infectious individual (a super-spreader) and took place mainly at birthday parties, nursing homes or prisons. We exploit cross-sectional variation in the number of SSEs within a 50 kilometer radius from the county border but not within the county 6 weeks prior to the murder of George Floyd to construct our instrument for exposure to COVID-19 at the county level.

We perform several exercises to probe the plausibility of the exclusion restriction of our instrument conditional on state fixed effects and set of political and socio-economic controls. Most importantly, we *i*) show in a placebo test that SSEs do not predict past BLM events or past social media use, *ii*) using LASSO, we weight SSEs by their inverse probability of occurrence and *iii*) generate a control variable that captures the pre-pandemic protest propensity based on characteristics of counties with prior BLM protest.[1] Our results hold for various iterations of our SSE instrument (varying distance, time lag, and cases associated with SSEs), sample composition,

[1]We describe the LASSO selected model in detail in Appendix C.3.

spatial correlation, and definition of the treatment and outcome variables.

In addition, we propose three alternative identification strategies and show that our results replicate. First, using large scale mobile phone mobility data by *SafeGraph*, we instrument pandemic exposure with tourist flows to one of the largest SSEs in the US – Florida spring break in March 2020. Second, we employ a difference in differences approach, for which we scrape information on all similar BLM protest triggers since 2014 to estimate the differential response to a protest trigger before and after the pandemic. Third, we use a LASSO-based matching approach, comparing counties with similar pre-pandemic protest probabilities.

As previewed above, we find robust evidence that pandemic exposure significantly increased the probability of observing BLM protests. In line with the idea of a broadening movement, we show that this is driven by white, suburban, affluent counties that have a low ex-ante probability of protesting. We also show that this activism extends to the virtual space with more tweets about BLM and more followers of the official BLM Twitter account.

In the second part of the paper, we investigate the role of social media in the broadening of the BLM movement. We start by repeating the above analysis, this time using a novel index of social media penetration as our main outcome variable. The index is measured *before* the protest trigger but after the outbreak of the pandemic in the United States. We use the first principle component of multiple variables: *i*) the (log) cumulative number of new Twitter accounts, which we obtain by scraping and geo-coding information on the creation date of new Twitter accounts at the county level from approximately 45 million tweets, *ii*) Google searches for the term "Twitter", hypothesizing that new users will Google the term first to create an account and *iii*) Google mobility data at the county level, assuming that increased residential stays (time spent at home) as well as lower social, work and leisure mobility is associated with more time spent online.[2] We find that the pandemic has a positive and significant effect on our social media penetration index.

In a next step, we focus on the role of Twitter in mobilizing BLM protest. We interact pandemic exposure with three different measures of Twitter usage: i) new Twitter accounts created during the pandemic, ii) baseline Twitter penetration, i.e. the number of Twitter users in December 2019 (scraped from a random sample of tweets, using the 100 most common words in English) and iii) instrumented baseline Twitter penetration, replicating the Müller and Schwarz (2020) instrument. We show

[2]We use a normalized index of search activity for the term 'Twitter' provided by Google Trends. Search activity indices are provided as integers from zero to 100 with an unreported privacy threshold. Each observation is the number of searches of the given term divided by the total searches from the geography and time range, which is then normalized between regions such that the region with the largest measure is set to 100. The Google Trends data is defined on a designated market area (DMA) level.

for all three measures that the effect of COVID-19 on protest participation is driven by counties with higher Twitter usage. In addition, we find a network effect: the demand shock to social media was largest in places where the marginal return to joining a platform was large enough (existing Twitter network) but markets were not fully saturated yet. This led to a disproportionate increase in the number of Twitter users in counties with no prior BLM protest history and consequently mobilized them to join the movement for the first time during the pandemic.

To probe the social media mechanism further, we use individual-level survey data. We find that individuals living in a county with a higher number of COVID-19 related deaths are more likely to receive news about George Floyd through social media than through other channels. Respondents are also more sympathetic towards the BLM movement without changing attitudes towards other progressive issues.[3]

In the last part of our paper, we look at competing mechanisms that could have affected BLM protest participation. First, we verify that BLM protests were not substituted away from some locations to others. Second, we test whether the pandemic increased the overall salience of racial inequality *before* the murder of George Floyd by using information on the COVID-19 death burden on Blacks and BLM-related search terms on Google before the protest trigger.[4] Third, we investigate whether the pandemic has decreased the economic and social opportunity cost of protesting, exploiting information on COVID-19 related unemployment and lockdown stringency. Fourth, we check whether the pandemic led to an overall agitation and higher propensity to protest by looking at other types of protests and counter movements to BLM.

We contribute to the nascent literature on the effects of the internet on political outcomes (Lelkes et al., 2017; Boxell et al., 2017; Campante et al., 2018; Guriev et al., 2019) and the effect of social media on xenophobia, polarization, political preferences, social capital and networks (Acemoglu et al., 2018; Enikolopov et al., 2018; Bursztyn et al., 2019; Müller and Schwarz, 2020; Zhuravskaya et al., 2020; Müller and Schwarz, 2021; Fujiwara et al., 2021; Campante et al., 2021; Melnikov, 2021).

Our analysis also contributes to a large literature that analyzes the determinants of social movements and protests, ranging from macro level drivers, such as local institutions or socio-economic conditions (Lipsky, 1968; Eisinger, 1973; McCarthy and Zald, 1977; Besley and Persson, 2011; Dube and Vargas, 2013; Berman et al., 2017), to micro level drivers, including individual decision making processes (Ellis and Fender,

[3] The data set does not contain information on the location of the respondent but only whether they live in a low, medium or high COVID-19 county. Therefore, we cannot employ our instrument for exposure to COVID-19 and can only present correlations, conditional on a large set of individual controls.

[4] Throughout this paper we follow the AP Stylebook when it comes to the capitalization of race and ethnicity. See https://www.apstylebook.com/race-related-coverage for more details.

2011; Guriev and Treisman, 2015; Sangnier and Zylberberg, 2017; Chenoweth et al., 2022) and different aspects of individual and social psychology (Guriev and Treisman, 2015; Sangnier and Zylberberg, 2017; Passarelli and Tabellini, 2017; González and Prem, 2020; Hager et al., 2020; Kutlaca et al., 2022; García-Jimeno et al., 2022; Cantoni et al., 2019; Bursztyn et al., 2021).

The remainder of the paper is organized as follows. In section 2, we provide a background on the BLM movement, present motivating evidence and describe our main data sources. We explain our empirical strategy in section 3 before moving to our main results in section 4. Section 5 provides various pieces of evidence for the social media mechanism. Section 6 addresses competing mechanisms. Section 7 concludes the paper.

2 Background and Data

2.1 BLM History and Motivating Evidence

The Black Lives Matter (BLM) movement emerged on social media after the acquittal of George Zimmerman in the deadly shooting of a Black teenager named Trayvon Martin. The movement was founded by three Black activists, Alicia Garza, Patrisse Cullors, and Opal Tometi in July of 2013 with the aim to end systemic racism, abolish white supremacy and state-sanctioned violence (Black Lives Matter, 2020), and more generally, to "fundamentally shape whites' attitudes toward Blacks" (Mazumder, 2019).

Over the following months, an ever-increasing but small number of activists coalesced under the hashtag #BlackLivesMatter on Twitter and Facebook. In August of 2014, after a court decision to not indite the responsible police officer in the fatal shooting of Michael Brown in Ferguson, #BLM became one of the most widely used hashtags on Twitter (the hashtag was used 1.7 million times in the three weeks following the court decision, compared to 5000 tweets in all of 2013, see Freelon et al. (2016); Anderson and Hitlin (2016)), confirming its status as a mainstream social media phenomenon. The shooting of Michael Brown was followed by a large and protracted protest in the city of Ferguson. The consequences of this shooting rippled throughout American society, generating counter-movements under the hashtag #AllLivesMatter and #BlueLivesMatter and mobilizing protesters (for and against the cause) far beyond the city's borders.

BLM played a crucial role in transforming localized activism into a coordinated movement across various locations within and outside of the United States. The *Black Lives Matter Global Network Infrastructure* was designed to provide decentralized actors with resources and guidelines to organize protests, receive information about the movement, and coordinate through social media.[5]

In the following years, the BLM movement expanded geographically and demographically, attracting an unprecedented number of participants after the murder of George Floyd in Minneapolis on May 25th 2020. Protesters took to the streets when a video of the murder of George Floyd went viral on social media, showing how police officer Derek Chauvin suffocated George Floyd using a choke-hold. The video spurred unrest in Minneapolis but the protests quickly expanded to other parts of the United States, including communities that had never engaged in BLM protests before. The number of BLM protests quadrupled in May and June of 2020, compared to previous peaks in 2016 (see Figure **??**).

The surge in BLM protests in the spring of 2020 is all the more remarkable as the COVID-19 pandemic was well underway. At the time of George Floyd's murder almost 100,000 COVID-19-related deaths had been recorded in the United States and the country was reeling under the first wave of the pandemic (see Figure A2).[6] Tough lockdown and social distancing measures were imposed in many counties to prevent the spread of the virus. Average lockdown stringency peaked in May (Hale et al., 2020) and the Center for Disease Control and Prevention urged the public to "remain out of congregate settings, avoid mass gatherings, and maintain distance from others when possible" (CDC, 2020). In Appendix Figure A4, we look at the evolution of BLM protest offline and online, splitting the sample into counties with above and below median exposure to COVID-19. In the left panel, we find that pandemic exposure is related to a higher number of BLM protests. The right panel tracks the evolution of tweets that mention the hashtags *#BLM* or *#BlackLivesMatter* in the three weeks following the murder of George Floyd, showing that locations that were more affected by COVID-19 engage more with the BLM hashtag online.

In Figure 1, we plot all counties that observed a BLM protest after the murder of George Floyd, splitting the sample into counties that had and did not have a BLM-related protest before the pandemic.[7] In line with public perception, the BLM movement has indeed broadened it's base in 2020. The geographic distribution of protesting counties does not follow the typical coastal divides but is spread across all of the US. Most importantly, we find that counties with no prior BLM protest make up half of the counties that protest in response to the murder of George Floyd.

[5] The founders state that "[...] when it was time for us to leave, inspired by our friends in Ferguson, organizers from 18 different cities went back home and developed Black Lives Matter chapters in their communities and towns — broadening the political will and movement building reach catalyzed by the #BlackLivesMatter project" (Black Lives Matter, 2020).

[6] We present the cumulative and weekly COVID-19 deaths per sub-sample in Appendix Figure A3.

[7] We use data from *Elephrame* on BLM events between 2014 and 2020 and describe this data set in more detail in the next section and in Appendix A

2.2 Main Data Sources

In this section, we present the main data sources. We also describe the additional data sources in Appendix A and provide an overview of all the data sources in Appendix Table A1.

COVID-19. Data on COVID-19 related deaths and cases in the USA at the county level comes from the New York Times. This data set provides the cumulative count of cases and deaths every day for each county in the USA, starting from January 21, 2020 when the country's first COVID-19 case was reported. A key limitation of COVID-19 cases data is that it depends on the testing facility and availability of the test kits in the region. We therefore mainly rely on COVID-19 related deaths as a measure of exposure to the pandemic. We also obtain data on daily COVID-19 hospitalizations and deaths by race and ethnicity at the state-level from the Center for Disease Control and Prevention.

Super spreader events. We collect data on COVID-19 SSEs from a project started by independent investigators and researchers from London School of Hygiene and Tropical Medicine (Leclerc et al., 2020). Data are put together based on scientific journals and news reports on SSEs, which are defined as "clusters" or "outbreaks" of COVID-19 infections with a minimum of 2 infections outside of the home. For the whole period (January 2020 to January 2021), we identify a total of 1074 SSEs in the USA. Most commonly, events occur in nursing homes, prisons, factories, and retribution (correction facility) or medical centers. Figure A5 shows the distribution of these events by their type and Table A2 provides descriptive statistics about each type of event. We describe the nature of these events in more detail in Section 3 and lay out the limitations of the SSE data set and how we address those in Appendix Section A.

Black Lives Matter. This data comes from the crowd-sourced platform Elephrame. It provides information on the place and date of each BLM protest and estimated number of participants, as well as a link to a news article covering the protest. We extracted and geo-located all protests from June 2014 to September 2020. These protests are decidedly pro Black Lives Matter. We add information on other protests from the US Crisis Monitor, a joint project between ACLED and the Bridging Divides Initiative (BDI) at Princeton University, that collects real-time data on different types of political violence and protests in the US from Spring 2020 to present day.

Twitter. We collect three types of Twitter data at different points in time (before the pandemic, during the pandemic but before the murder of Floyd and in the three weeks after the murder of Floyd). First, from the Twitter API we collect the universe of tweets with BLM related hashtags. This includes the hashtags #BlackLivesMatter, #BlackLifeMatters, #BLM, #AllLivesMatter, and #BlueLivesMatter. We present a selection of tweet examples from our collected sample in Appendix Table A3. Second, we collect data to proxy the baseline Twitter penetration in December 2019 by taking a random sample of tweets that use the most common 100 words in the English language. Third, we scrape information on all followers of the official Black Lives Matter Twitter account (as of March 2022). With the help of a geo-location algorithm, we can assign about 5 to 20% of Twitter users (depending on the sample) to counties. We show in Appendix Table A4 that the characteristics of counties for which we have geo-located tweets are indistinguishable to characteristics of the full sample. Using this data we are able to proxy *i*) online protest for and against BLM with the number of tweets containing the relevant hashtags and the number of followers of the official BLM account *ii*) pandemic Twitter penetration, using the creation date of the Twitter accounts and *iii*) information on baseline Twitter penetration. We describe this in more detail in Appendix A.

2.3 Descriptive Statistics

Summary statistics are presented in Table 1. As outlined above, we use information that is available at different points in time. In Table 1, we present 5 panels that split the variables according to when they are measured: *i*) three weeks after George Floyd's murder, *ii*) the day of the murder, *iii*) before the murder but after the pandemic started in January 2020, *iv*) later outcomes and *v*) baseline county characteristics before the outbreak of the pandemic.

The average likelihood of observing a BLM-related protest at the county level between May 25th and June 14th lies at about 5%. There are on average 0.06 events per county in the three weeks following George Floyd's murder and the average number of participants is approximately 21 with a maximum of 5.5K participants.[8] If an event occurs, the average number of participants per event is about 355. In the three weeks following George Floyd's murder we can identify over 300 tweets per county using BLM-related hashtags and about 1.8 new Twitter users per county (those created after the outbreak of the pandemic but before the murder of Floyd) that tweet about BLM.

The per county average number of cumulative COVID-19 related deaths is 8.4

[8]The average sets the number of participants in places with no BLM protests as zero.

(or 0.1 per 1000 population) by May 25th 2020. Absolute cumulative cases are approximately 164 per county (or 2.4 per 1000). The maximum number of deaths in a county at the time was 1000, compared to 31,000 deaths in March 2022. While COVID-19 cases and deaths were comparatively low, the salience of the pandemic was particularly high. In fact, lockdown stringency in the United States peaked in late April 2020. We also report the Black Death Burden (BDB) and find that Blacks were disproportionately affected by the pandemic. The average BDB index is 1.3 indicating that Blacks died at a rate 30% higher than their share of the population would predict. The average county experienced about three SSEs in its immediate surroundings between January 2020 and April 2020.

In addition, we report detailed summary statistics for the different sub-samples in Table A5. We report the full sample in the left-hand columns and present a breakdown of the summary statistics by sub-sample in the middle and right-hand side of the table. We distinguish between counties with no BLM events before the pandemic and those with prior BLM events. The vast majority of counties where there was no history of protest for a BLM-related cause continue to not protest after the murder of George Floyd (2,635 counties, which is approximately 85% of all counties). However, we observe that among the sample of "no BLM event before" 133 counties start to protest for the first time during the pandemic. We also report summary statistics on the traditional protesters, i.e. counties that have had a prior BLM protest. Among those 339 traditional protesters, 123 counties did not protest after the murder of George Floyd and 176 counties continue to protest. As expected, the average probability of observing a protest in response to the murder of George Floyd is 10 times higher among traditional protesters compared to other counties. Remarkably, however, the first-time protesters make up nearly 50 percent of all counties that protested during the pandemic. Counties that traditionally protest have a higher Black population share; a higher median household income; are more urban and have a higher Democratic leaning than counties with no prior BLM protest.

3 Empirical Strategy

3.1 Baseline Estimating Equation

To study the effect of exposure to COVID-19 on BLM protests in counties with no prior BLM events, we estimate

$$BLM_c = \beta_0 + \beta_1 Covid_{cs} + \mathbf{X}_c \beta_{\mathbf{X}} + \delta_s + \epsilon_{cs} \tag{1}$$

where BLM_c is a dummy variable for the presence of a BLM protest in county c during the three weeks following the murder of George Floyd.[9] We are interested in the coefficient β_1, which captures the effect of one additional COVID-19 related death per 1000 inhabitants in county c of state s at the time of George Floyd's murder on May 25th 2020. In addition to state fixed effects δ_s, the vector $\mathbf{X}_c$ includes an array of county level controls (we describe all these variables in detail in Table 1). Specifically, we include variables that are associated with participation in the BLM movement, such as a dummy for urban counties and Black population share and the poverty rate among Blacks. Most importantly, we also include determinants of BLM protests following the murder of George Floyd, namely the use of deadly force by police (i.e. number of Black people who died during an encounter with the police, excluding suicides, for two time periods: from summer 2014 to 2019 and in 2020 up to May 25th).[10] We also control for underlying political and attitudinal factors and socioeconomic drivers of protest and social media use, such as the vote share for Republicans in the 2012 and 2016 presidential elections, median household income, unemployment rate, community resilience, and two proxies for social capital (number of civil organizations and number of religious organizations). We cluster standard errors at the state level.

3.2 Instrumental Variable Estimation

A key empirical challenge in ascertaining the causal impact of exposure to COVID-19 on BLM protests is that both could be driven by unobserved factors. For instance, tight-knit and socially active communities may both increase the spread of the virus and protest more for a BLM-related cause. Alternatively, counties that are in favor of lax social distancing rules (and thus more aligned with the president's views at the time) are less likely to engage in BLM protests. Additionally, we may be concerned that BLM protests themselves could lead to COVID-19 infections. While we can assuage the latter concern by measuring COVID-19 exposure at baseline (e.g. before the murder of George Floyd and the onset of BLM protests), we address the former concern with an instrumental variable approach.

We exploit plausibly exogenous variation in the occurrence of SSEs to causally identify the effect of COVID-19 on BLM protests at the county level. Specifically, we construct the IV as the sum of all SSEs that occur within 50 km of the county

[9]We restrict the sample for our main outcome of interest to the three weeks after the death of George Floyd, that is the period from May 25th to June 14th for several reasons: we can capture a large share of the protest behavior (66 percent of BLM protests following George Floyd's murder can be observed in this three week window) while limiting the potential for confounding factors to arise.

[10]When analzing the full sample, we always include as a control the number of BLM events between 2014 and 2019 to account for general propensity to protest for a BLM related cause.

border but not within the county until 6 weeks before the murder of George Floyd. The first stage is written as:

$$Covid_{cs} = \zeta_0 + \zeta_1 Z_{cs} + \mathbf{X}_c \zeta_{\mathbf{X}} + \gamma_c + \eta_{cs}, \tag{2}$$

$$Z_{cs} = \sum_{m=1}^{t-6} SSE_{csm}^{neighbor} \tag{3}$$

The key identifying assumption of this instrument is that - given the set of controls and state fixed effects - SSEs only affect BLM protests through an increase in exposure to COVID-19. The state fixed effects capture unobserved time-invariant characteristics at the state level that could be related to both the prevalence of SSEs and the propensity to observe a BLM protest. These include, for instance, testing capacities, lockdown stringency, differences in state laws, or policing strategies, all of which are mandated at the state level. In addition, we include control variables at the county level that account for within state heterogeneity. We control for the political alignment with the presidency and baseline sympathy towards the BLM movement with the Republican vote shares in 2012 and 2016 as well as deaths of Blacks at the hand of police between 2014 and 2020. In addition, we control for demographic factors that determine participation in the BLM movement and may also be related to the likelihood of observing a SSE such as the Black and non-white population share and the degree of urbanization. We also include control variables that capture potential socio-economic determinants of both SSEs and protest participation with the unemployment rate, median household income and importantly an index developed by the US census bureau that captures the "vulnerability of a county in the face of a pandemic". This index captures the age and household structure, health conditions (share of population with certain pre-conditions), as well as health care coverage and health infrastructure.[11]

We also show that including state fixed effects and the full set of control does not significantly change the magnitude or significance of the coefficient of interest. We take this as first suggestive evidence that SSEs are a plausibly exogenous source of variation in exposure to COVID-19. In the following we argue which features of the instrument justify the plausibility of the exclusion restriction and test this more explicitly in Section 3.3.

Event types. Super-spreaders are individuals who are an order of magnitude more contagious than others. This phenomenon, well-known in epidemiology, is instru-

[11]We describe this index in more detail in A.

mental in infectious disease spread (e.g. Galvani and May (2005)) and of particular importance for COVID-19, where 70–80% of transmissions can be traced back to just 10–20% of cases (Adam et al., 2020; Endo et al., 2020; Miller et al., 2020). SSEs are characterized by the presence of highly contagious individuals. More specifically, the data set includes "outbreaks" and "clusters" with "two or more test-confirmed cases of COVID-19 among individuals associated with a specific non-residential setting with illness onset dates within a 14-day period" (see Appendix A for more details). Importantly for our context, these SSEs are not necessarily mass gatherings, alleviating concerns about SSEs as a proxy for a county's propensity to organize large public events, including BLM events. The majority of the approximately 1000 SSEs in our data take place at birthday parties, prisons and in the medical care sector(see Figure A5).[12]

Window of opportunity. Next, we illustrate in Figure 2 that the overwhelming majority of SSEs (solid blue line) occurred between the second week of March and the last week of April. This was an opportune period for SSEs for two main reasons. First, infections were sufficiently high to introduce a significant number of superspreader individuals. Second, lock-down measures were not yet stringent enough (in addition to the lack of public awareness) to restrict group gatherings and encourage mask-wearing. The red dotted line of Figure 2 shows that the increase in the number of new COVID-19 cases coincided with the increase in SSEs. The green dashed line illustrates that state-issued stringency measures (as measured by the stringency index from the Oxford COVID-19 Government Response Tracker) peaked around the time that SSEs leveled off. We argue that during this time window, the occurrence of SSEs was mainly driven by the presence of a highly infectious person, rather than heterogeneity in risk preferences or other underlying factors that could drive both SSEs and BLM protests. We only include SSEs until April 13th 2020 - 6 weeks prior to George Floyd's murder, to account for the fact that SSEs further into the pandemic may be more endogenous. In Figure A6, we illustrate this time period relative to the spread of COVID-19 (measured as the cumulative number of COVID-19 related deaths) and the surge in BLM protests.

Geographic proximity. Lastly, we improve on the plausibility of the exclusion restriction by exploiting SSEs *outside* the county and not within the county. Specifically, we use the number of SSEs within a 50km (or approximately 30 mile) radius from the county border in which we measure exposure to COVID-19 and BLM

[12]We exclude specific event locations in a robustness check and control for testing capacities and the presence of health care facilities with state fixed effects and the community resilience index in all specifications.

protests. We illustrate the construction of our instrument in Figure 3 using the example of Arizona. We argue that SSEs in geographic proximity but not in the county itself are less likely to be correlated with unobserved characteristics of the county in which we observe BLM protests.[13]

In Figure 4 we show the geographical distribution of our instrument across US counties. In the top panel, we map the cumulative number of SSEs 6 weeks prior to Floyd's murder at the county level. In the bottom panel, we illustrate the identifying variation of our instrument, i.e. the number of SSEs in 50 km proximity to the county border up to April 13th. We present the first stage results in Table B1. We show that one additional SSE increases the number of COVID-19 deaths per thousand population by between 0.7 and 0.8 percentage points, depending on the specification. For all specifications the F-statistic is well above the standard threshold.

3.3 Instrument Robustness and Validity

Overall, the features of our instrument (epidemiological feature, small window of opportunity, geographic distance) lend confidence to a causal interpretation of our IV estimation. We provide a detailed description of all robustness checks in Appendix B and preview some of the most important checks here.

In a first set of exercises in Appendix Table B2, we probe the robustness of our instrument to changes in the timing of and distance to SSEs. We always present baseline results in column 1. In columns 2 to 4, we consider the baseline time lag of 6 weeks, i.e. SSEs until April 13th 2020, but vary the distance of SSEs to the border between 25km and 200km. Next, we keep the 50km distance but vary the time lag of SSEs until the protest trigger, reducing it to five weeks and expanding it to seven and eight weeks in columns 6 to 8. Our results hold for all of these iterations.

Next, in Appendix Table B3, we probe the robustness of our instrument with respect to changes in the definition and various weighting schemes. Again, we report our baseline specification in column 1. In column 2, we exclude SSEs in prisons as they may impact the public perception of exposure to the pandemic differently and may also be related to factors that drive BLM protests. Next, in column 3, we also include the number of SSEs in the own county to account for correlation between neighboring and own SSEs. Then we vary the importance of distance to the neighboring SSE. We include both the simple linear distance and squared distance to the SSE in columns 4 and 5. Lastly, in column 6 we weight each observation

[13] In line with this, we show in a robustness check that our results barely change when controlling for SSEs within the county.

(i.e. each county) by their inverse probability of being treated by the instrument, using LASSO.[14] In doing so, we give more weight to counties that had a low a-priori likelihood of having a SSEs in close proximity. Our results remain robust to changes in the definition of the instrument.

We continue with an additional exercise that validates the plausibility of the exclusion restriction in column 2 of Table B4. We show that SSEs in neighboring counties do not predict the likelihood of past BLM protest between 2014 and 2019 for the full sample. If our instrument was related to some unobserved heterogeneity that drives BLM events, we should observe a direct effect of SSEs on past BLM events. Reassuringly, this is not the case. We provide additional robustness checks for our main results in Subsection 4.4.

4 COVID-19 and BLM

4.1 Main Results

We present our main results in Table 2.[15] As previewed above, we estimate the effect of exposure to COVID-19 on the likelihood of observing a BLM protest in the weeks following George Floyd's murder. Panels A, B and C report 2SLS, reduced form and OLS estimates respectively. In column 1, we present the coefficient without fixed effects and controls. We subsequently add state fixed effects and the full set of control variables described in section 3. Our preferred specification is presented in column 3. We focus on the three week period after George Floyd's murder because it captures a substantial proportion of all BLM protests in 2020 (see Figure A6) but confirm in columns 4 and 5 also hold for a 6 and 9 week window.[16]

Overall, we find a consistently strong and positive effect of COVID-19 on BLM protest. In terms of magnitude, we estimate that a one standard deviation increase in COVID-19 exposure (approximately 23 deaths per 100,000 inhabitants) increases the likelihood of observing a protest between 9 and 13 percentage points depending on the specification.Across all columns the coefficients are statistically indistinguishable from each other.

Throughout all of our estimations the IV estimates exhibit larger coefficients compared to the OLS. In the absence of exogenous variation in pandemic exposure, counties with fewer COVID-19 related deaths are more likely to experience a BLM

[14]We describe this approach in more detail in Appendix C.3

[15]Throughout the paper, we will show the corresponding results for the full sample and the sample of counties with prior BLm events in Appendix D. Main results for the sub-samples are presented in Appendix Table D1.

[16]We also focus on the three weeks after George Floyd to remain consistent with subsequent analyses using big data from social media. There, we focus on the three weeks after the murder of George Floyd to reduce the data volume and manage computational capacities since we leverage information from more than 45 million tweets.

protest – given state-fixed effects and the full set of control variables. The OLS would then underestimate the true effect.[17] For instance, individuals in more progressive counties could hold more favorable attitudes towards the BLM movement and at the same time behave more cautiously with regards to the pandemic by adhering to stricter social distancing rules. Using mobile phone mobility data, we find that counties with BLM after the murder of George Floyd also decrease their workplace and leisure mobility while increasing residential stay in the weeks leading up to the protest. This is in line with Dave et al. (2020) who show that BLM protesters adhere more to social distancing measures.[18]

4.2 Heterogeneity

What are the characteristics of counties that protest in response to the murder George Floyd when pandemic exposure is high? In Table 3, we interact exposure to COVID-19 with baseline characteristics for the full sample of counties and report the coefficient of the interacting variable in the bottom row.[19] We instrument both COVID and the interaction term with our SSE variable and report the respective F statistics at the bottom.

In column 1 of Table 3, we show the baseline effect for the full sample for reference. In column 2, we interact exposure to COVID-19 with an index for the baseline propensity to protest for a county. We construct this index using a large battery of county characteristics from various sources and employ a LASSO approach to choose the set of variables that best predict past protest behavior (we describe the design in more detail in Appendix C.3). This yields an continuous indicator for ex-ante protest probability that ranges between zero and one with higher values indicating higher predicted protest propensity before 2020. We find that ex-ante protest probability is negative, large and highly significant, indicating that pandemic exposure was particularly effective in mobilizing protest in counties with a low ex-ante protest probability. We take this as evidence that the pandemic caused a true broadening of the BLM movement into segments of the population with low baseline engagement.

In columns 3-10, we follow a more naive approach by focusing on county characteristics that are typically associated with BLM protest. In columns 3 and 4,

[17]Since the treatment (exposure to COVID) is measured before the protest trigger, reverse causality is not the driver of the difference in magnitude.

[18]In addition, part of the gap between OLS and IV can stem from differences in the local average treatment effect and the average treatment effect. In our IV estimation, we observe the effect of COVID on BLM for compliers, i.e. counties that experience an increase in COVID-19 related deaths due to SSEs in neighboring counties. Counties whose COVID-19 death toll was not primarily determined by surrounding SSEs will be discounted. These counties may be different in terms of baseline propensity to protest, overall salience of the pandemic or along unobserved dimensions that decrease the marginal effect of the pandemic on protest in response to Geroge Floyd's murder.

[19]In in Table D2 and Table D3, we also run the heterogeneity exercise for the sub-sample of counties with and with no prior BLM events We obtain similar result for the sub-sample of counties with no prior BLM events.

we consider heterogeneity by race as recorded in the American Community Survey in 2018.[20] The coefficient of the interacting variable indicates that - as expected - counties with a higher non-Black and non-white population share are less likely to protest overall. This is in line with our prior that those who are most affected by the movement's grievances are typically protesting. However, counties with a higher non-Black population share (including whites, Hispanics, Asians and "others") are more likely to protest in response to pandemic exposure. Interestingly, the effect of counties with higher non-white population shares (this includes other minorities beyond Blacks) exhibits the opposite sign and is insignificant, indicating that the share of whites is driving the result in the previous column.

In column 5, we move to the economic prosperity of the county, as proxied by the median household income - again measured in 2018 from the American Community Survey. Richer counties are more likely to protest overall and these counties protest even more in response to the pandemic. This is in line with two mutually non-exclusive interpretations. First, the literature on protest and conflict highlights that individuals need basic resources to be able to engage in protest in the first place (Bates et al., 2002; Bazzi and Blattman, 2014; Besley and Persson, 2011). In the midst of a pandemic, only affluent household may be able to stem the (opportunity and logistical) costs of protesting. Second, it is possible that - similar to the results on the ex-ante protest probability and racial composition of a county - the BLM movement broadens into more affluent spaces.

As expected, counties with higher vote shares for Donald Trump in the 2016 elections (vote share Republican reported in column 5) are less likely to participate in BLM protests overall. However, the coefficient of the interaction term is negative, not significant and very noisy, indicating that the political leaning is less relevant for the likelihood of a BLM event occurring in response to higher exposure to COVID-19. Conditional on state fixed effects this may not be surprising, as they capture a large share of the variation in political leaning. Interestingly, recent work by Engist and Schafmeister (2022) also does not find any evidence of BLM protests mobilizing voters.

In columns 7 to 10, we consider different classifications for a county's degree of urbanization as defined by the 2013 NCHS Urban-Rural Classification Scheme for Counties. Typically, BLM protests occur in large metropolitan areas, like New York or Los Angeles and less frequently in smaller cities, suburban or rural areas. In column 6, we look at the effect of the pandemic on counties that are not a part of a large city. This encompasses fairly big sub-urban areas like Bergen County, New

[20] Self reported racial identification with the categories: white, Black, Asian, Hispanic and "other".

Jersey (adjacent to Bronx County in New York) to small rural areas like Mariposa County, California. Similarly, we consider only suburban counties in column 7. Both of these county types experience an increase in BLM protests in response to the pandemic. We report results on protest participation in small towns and rural areas in columns 8 and 9 and find no differential response to COVID-19 exposure.

Overall, we find evidence that the pandemic mobilized protesters in counties with lower ex-ante protest probabilities. These include counties with a higher share of whites, more affluent counties and counties in suburbs outside of large cities and small towns.

4.3 Alternative Outcomes

Our main variable of interest, so far, was the likelihood of observing any BLM protest in the three weeks following the murder of George Floyd. In Table 4, we turn to alternative outcomes, including the frequency and scope of BLM protests as well as BLM protest online. We report our baseline result in column 1 for reference.[21] In column 2, we look at the intensive margin, i.e. the number of BLM protest in the three weeks following the murder of George Floyd, and find an increase in the number of BLM protest. Focusing on the scope of the protests in columns 3 and 4, we investigate the effect of COVID-19 on the total number of participants and the average number of participants. We find a negative but non significant and very noisy estimates for the effect of COVID-19 on both measures for the scope of BLM protests. We conclude that pandemic exposure led to an increase in the frequency of BLM protests without significantly diminishing its scope.

Next, we investigate the impact of pandemic exposure on online protest, specifically BLM-related content on Twitter. In column 5, we report as an outcome the total number of geo-localized BLM-related tweets in a county in the three weeks following George Floyd's murder. These are based on the universe of tweets that use the hashtags #BlackLivesMatter #BlackLifeMatters or #BLM. We end up with a total of 2.5 million tweets that we aggregate to the county level. We find a large effect of pandemic exposure on the number of BLM tweets. A one standard deviation increase in pandemic exposure leads to 600 more BLM-related tweets per county, and 1.6 million across all counties without prior BLM protest. These tweets could however be either in favor or against the movement. In order to proxy online support for the BLM movement more explicitly, we scrape information on all followers of the official BLM account and geo-localize each of those Twitter users. We present the results

[21] Results for all counties and those that protested before the murder of George Floyd are reported in Appendix Table D4

using this outcome in column 6.[22] We find that places that were more exposed to the pandemic had a higher number of followers of the BLM account. A one standard deviation increase in pandemic exposure leads to 40 more Twitter users following the official BLM account per county, and 110 000 in total.[23]

An increase in the number of BLM followers has potential implications for the medium-run mobilizing potential and success of the movement. The official Twitter account serves as a primary coordination, communication and mobilization tool for BLM (Black Lives Matter, 2020). Therefore, the expansion of the follower base may help activate these groups, when similar protest triggers arise in the future. In addition, the creation of an online network around the movement can keep users exposed to certain information, discourse, narrative and ideology that could continue to change their preference, awareness, beliefs and ultimately actions, including electoral choices (Casanueva, 2021).

4.4 Summary of Robustness Checks

Robustness of main results. We continue to probe the robustness of our main results in Appendix Tables B4 and B5. We provide a brief description of these exercises here and describe them in more detail in Appendix B. All robustness checks are performed on the full sample (Panel A) and our sample of interest (Panel B), i.e. counties with no prior BLM protest. Baseline results are always presented in column 1 for reference.

In columns 3 to 9 of Appendix Table B4, we consider the robustness of our main result with respect to changes in the sample composition and the definition of the treatment variable. In columns 3 and 4, we exclude counties and whole states on the coasts and our results hold. We do this for two reasons: first, counties and states next to the ocean will mechanically have fewer neighboring counties with SSEs. Second, when thinking about a "broadening" of the BLM coalition, we want to verify that this does not just apply to states with pre-existing progressive leanings. In columns 5 to 7, we shorten the time horizon for BLM protests after the murder of George Floyd from 3 to 2 weeks and expand it to 6 and 8 weeks. In column 8, we change the definition of pandemic exposure from COVID-19 related deaths to cases. Column 9 includes, as an additional control, the number of COVID-19 related deaths in the past seven days. In doing so, we account for heterogeneity in the trajectory of the COVID-19 pandemic when cumulative deaths over the whole period are similar. All

[22] The scraping was conducted on March 2022 meaning: i) that we are not able to capture users that stopped following BLM; ii) that we are capturing users that may have started following well after the murder of George Floyd and iii) that we are capturing users that started following before the pandemic.

[23] A 10% increase in pandemic exposure would similarly lead to 25 more tweets and 1.7 more followers per county, or 70 000 tweets and 4 700 followers in total.

of these checks yield consistent results.

In Table B5, we continue to validate our main findings. In column 2, we run a IV probit estimation instead of an OLS estimation. In column 3, we expand on the idea of comparing counties with similar ex-ante BLM protest probabilities and go beyond the socio-demographic and political controls. Using LASSO, we select the subset of relevant county-level variables that determine past BLM events and create a propensity score for protesting, based on the selection of these variables.[24] We include this variable as an additional control and confirm that our results remain robust. In columns 4 to 6, we include fixed effects for various ranges of the pre-pandemic protest probabilities. We split the fixed effects along the thresholds that produce county groups of different sizes: 3 groups (with 1000 counties each), 30 groups (with 100 counties each) and 300 groups (with 10 counties each). This allows us to compare counties within narrow bands of ex-ante protest probabilities. In columns 7 and 8, we replace the state clustering with spatial clustering, allowing correlation in a 50 km radius for column 7, and between neighbors for column 8. Column 9 omits clustering altogether. Reassuringly, our results are not sensitive to any of these changes.

Alternative identification strategies. In a last step, we complement our preferred identification strategy in three ways. We summarize them here briefly and explain them in more detail in Appendix C. First, we design an alternative instrument, using one large SSE rather than multiple small SSEs. Specifically, we use mobile phone mobility data for 45 million phones to instrument pandemic exposure with touristic flows to Florida spring break in March of 2020. We report the results in Table C1. Second, in Table C2 we exploit the panel dimension of our data set to estimate an instrumented difference-in-differences model. For this we scrape information on "notable deaths" since 2014 and look at the differential effect of BLM protest in response to a protest trigger before and after the pandemic, using county and state-week fixed effects.[25] Third, in Table C3 we perform a LASSO matching approach comparing counties with a similar ex-ante protest probability. All of these approaches confirm the baseline results.

[24] We describe this approach in more detail in Appendix C.3

[25] Names of the victims are listed in the Data Appendix Table A6

5 Social Media and BLM

5.1 COVID-19 and the Use of Social Media

The literature on the effect of social media on protest and other political outcomes typically exploits supply side constraints to the access to social media, leveraging the staggered roll-out of social media going back to the mid 2000s (Enikolopov et al., 2020; Manacorda and Tesei, 2020; Müller and Schwarz, 2020). However, this approach is less suitable to account for the more recent expansion of social media use, including the reported rise in the number of Twitter users during the pandemic.[26] In this section, we argue that the pandemic shifted a substantial proportion of the population to the digital space. Social media - and particularly Twitter - users were then exposed to an unexpected and highly viral protest trigger - the murder of George Floyd - which subsequently mobilized them to take to the streets for the first time.

To test this hypothesis, we create a novel index of social media penetration that comprises the first principle component of three main variables: *i*) the (log) cumulative number of new Twitter accounts, which we obtain by scraping and geo-coding information on the creation date of new Twitter accounts at the county level from approximately 45 million tweets, *ii*) the normalized index of search activity for term "Twitter" provided by Google Trends, hypothesizing that new users will Google the term and then create an account , and *iii*) Google mobility data at the county level, assuming that increased residential stay (time spent at home) as well as lower social, work and leisure mobility is associated with more time spent online.

All of these variables are measured between January 2020 and May 24th 2020, i.e. after the outbreak of the pandemic but before the murder of George Floyd. We limit the observation period such that the BLM events themselves do not impact online activity. We show the features of our index in Table A7, presenting the correlation between the different sub-components in Panel a), the eigenvalues of the principle components in Panel b) and the factor loadings in Panel c). We show that all components of the index are significantly correlated but that coefficients are small (ranging from 0.05 to 0.5), lending confidence to the relevance of all sub-components without capturing the same dimensions of social media use. This is also supported by the fact that factor loadings for each sub-component are rather balanced in the first principle component.

[26] We show in Figure 2 and Appendix Figure A7 that in the weeks leading up to George Floyd's murder, the stringency of social distancing measures increased and workplace and leisure mobility decreased substantially. Moreover, many online services - including Twitter - reported substantial increases in the number of users during the first months of the pandemic.In the Twitter letter to shareholders of April 30th 2020, Twitter states: "Average monetizable DAU [daily active users] grew 24% year over year... The increase in mDAU was driven by ... an increased engagement due to the COVID-19 pandemic."

In column 1 of Table 5, we show that pandemic exposure, i.e. cumulative COVID-19 deaths per 1000 population until May 24th, had a positive and significant impact on our index of social media penetration. Next, we focus on the specific subcomponents of the index and find that pandemic exposure increased the raw number of new Twitter accounts (column 2) and the log number of new Twitter accounts (column 3).[27] The effect is large: the coefficient of column 3 suggests that a one standard deviation increase in pandemic exposure led to a 27% increase in the number of new Twitter accounts created between January and May 2020. To rule out concerns that our instrument could be correlated with unobserved factors that drive both COVID-19 and Twitter use, we show in Appendix Table B6 that SSEs do not predict past Twitter penetration. In column 4, we find that pandemic exposure significantly increased the Google search for Twitter. Lastly, using high frequency mobility data from Google we show that the average time spent at home in the four weeks leading up to George Floyd's murder increases with pandemic exposure.

Taken together, we take this as evidence that the pandemic has increased online activity and particularly the use of Twitter among counties that never protested for a BLM-related cause before. In Appendix Table D5, we also show that this not the case for other counties and explore the reasons for this in the next section. Taken together, we conjecture that the pandemic acted as a demand shock to social media in areas with lower ex-ante BLM salience.

5.2 Twitter and BLM protest

In this section, we focus on the role of Twitter in mobilizing protest. To do so, we interact exposure to COVID with three measures of Twitter penetration. First, new Twitter accounts, i.e. the (log) number of Twitter users that tweet about BLM and have created their account during the pandemic. Second, baseline Twitter penetration, i.e. the log number of Twitter users in a county in December 2019, using a random sample of tweets containing the 100 most commonly used words in English.[28] Third, instrumented baseline Twitter penetration, replicating the Müller and Schwarz (2020) instrument.

We present our results in Table 6. In column 1, we include the number of new

[27]We include both the absolute number of accounts and the log number of accounts (new Twitter accounts) for two reasons. On the one hand, we do not have a prior as to whether the absolute number Twitter users or share of Twitter users is important for the occurrence of a BLM event. It is possible that irrespective of county size or Twitter penetration at the county level, there is a threshold level of individuals that need to be mobilized for a BLM event to occur. The average number of protesters at a BLM event in counties with no prior BLM events is about 350 individuals. On the other hand, in the absence of a good measure for relative importance of Twitter (by population, baseline Twitter usage, overall social media users) we want to give less weight to counties with higher Twitter penetration. Including both in the principle component will allow us to account for distributional features of Twitter penetration. The principle component will only capture the residual correlation between the two variables.

[28]We describe the sampling in more detail in Appendix A.

accounts as an additional control. We find that counties with a higher number of new Twitter users are also more likely to protest in response to the murder of George Floyd. In column 2, we interact COVID-19 with new Twitter accounts and show that the effect of pandemic exposure on the likelihood of observing a BLM protest is driven by counties with a higher take-up of social media during the pandemic. However, these estimates may suffer from bias amplification since – as shown above – new Twitter accounts are determined by pandemic exposure and hence a bad control (Cinelli et al., 2021).

Therefore, we explore heterogeneity by baseline Twitter penetration. In column 3, we add the number of geo-localized Twitter accounts in December 2019 and find a positive and significant correlation between Twitter usage at baseline and propensity to protest. In the next column, we interact pandemic exposure with baseline Twitter penetration and find a positive and significant effect for the interaction term. Similar to our results in the previous columns, we confirm that the effect of COVID-19 on BLM protest is primarily driven by counties with higher baseline Twitter penetration.

This result is in line with two possible interpretations. For one, COVID-19 may shift Twitter usage at the intensive margin. Baseline Twitter users may spend more time on the platform in places that are more exposed to the pandemic and in turn become more likely to join the protest. Another interpretation is that the take-up of Twitter during the pandemic crucially depends on the existing network size. The latter interpretation is in line with the literature on path dependence in technology adoption (Arthur, 1989), which posits that the marginal utility of joining a network increases with the size of the network. We test the latter interpretation in Appendix Table B7. Specifically, in columns 3 and 4, we estimate the effect of pandemic exposure interacted with baseline Twitter penetration on the number of new accounts created during the pandemic. We find that COVID-19 disproportionately increases the take-up of Twitter in places with large existing networks, confirming the hypothesis on path dependency in technology adoption.

In a next step, we leverage plausibly exogenous variation in Twitter penetration at baseline, replicating an instrument used in Müller and Schwarz (2020) that leverages a film, interactive media, and music festival and conference held annually in Austin Texas. The 2007 edition heavily promoted Twitter and crucially determined the spread of Twitter in the early stages of its roll-out. We describe the construction and validity of the instrument in detail in Appendix C and show in Appendix Table C4 that the first stage is sufficiently strong. Our coefficient in column 5 suggest that a doubling of baseline Twitter penetration increases the likelihood of observing a BLM event by approximately 5 percentage points. In column 6, we include the interaction term between instrumented pandemic exposure and instrumented baseline Twitter penetration. With this exercise, we leverage exogenous variation in both COVID-19 and Twitter usage. The results confirm our previous findings: COVID-19 has a more significant effect in predicting a BLM protest in counties with higher (exogenous) baseline Twitter penetration.

Lastly, we try to answer the following question: why did COVID-19 not lead to an increase in the likelihood of observing a BLM protest among traditional protesters? As previewed above, in Appendix Table D1 we show that there is no effect of pandemic exposure on the propensity to protest among counties that have protested before. Similarly, we show in Appendix Table D6 that – for counties with prior BLM protest – there is also no effect of the interaction between pandemic exposure and Twitter on the likelihood of observing a BLM protest.

We hypothesize that counties that protested before also have highly saturated social media markets. Consequently, the pandemic may not have encouraged social media take-up at the margin and therefore did not mobilize new protesters. We provide several pieces of evidence consistent with this hypothesis. First, Appendix Table A5 reveals that the number of Twitter users at baseline is approximately 15 times higher in counties with prior BLM protest than in those without. Second, we show in Appendix Table D5 that counties with prior BLM protest do not experience a higher social media penetration in response to the pandemic. Third, we show in Panel B of Appendix Table B7 that baseline Twitter penetration does not predict new Twitter accounts for counties with prior BLM protests. As expected, we also find that the Müller and Schwarz (2020) instrument is too weak for the sub-sample of counties with prior BLM protest (see Panel B of Table D6). This confirms the idea that instruments that exploit the early stages of social media roll-out under-perform in saturated markets.

Overall, our results suggest that the pandemic was only able to act as a demand shock to social media in places *i*) that are not fully saturated but *ii*) offer a higher marginal utility of joining social media in the form of sufficiently large existing networks. These conditions apply specifically to the sub-set of suburban, affluent and white counties with low ex-ante salience of BLM that join the protest for the first time in 2020.

5.3 News Consumption and Attitudes towards BLM

One important caveat of our analysis is that we only observe protest at the county level and cannot identify individual protesters. In this subsection we examine the social media mechanisms more closely by exploiting individual-level survey data. We investigate whether exposure to COVID-19 at the individual level is correlated with a shift in news consumption away from traditional media and towards social media. We then investigate whether this shift is accompanied by a change in attitudes towards Blacks and the Black Lives Matter movement more generally.

We cannot claim that these results are causal, because we cannot apply identification strategies discussed in previous sections: the location of the respondent is anonymized in the survey. The only available information is the severity of exposure to COVID-19 in respondent's county of residence in June 2020. However, the rich set of individual-level controls and placebo checks assuage concerns about omitted variable bias.

We use survey data from the Pew Research Center to conduct individual-level multivariate regressions on different outcomes, controlling for respondent characteristics: race, whether or not they live in a metropolitan area, gender, age, education, income and whether or not they lean towards the Democratic party. We describe the underlying data in more detail in Appendix Section A. Table 7 shows the results. Columns 1 - 3 show the intensity and form of news consumption in the context of George Floyd's murder. Higher levels of COVID-19 are positively and significantly associated with more news consumption about George Floyd and more social media news consumption about George Floyd. In column 3, we show that individuals in counties with higher COVID-19 exposure also consume relatively more news about George Floyd on social media, confirming a change in the information set – or at least their source.

Then, we analyze whether this change in mode of news consumption is accompanied by a change in attitudes. In column 4, we find that individuals are more likely to report that higher hospitalization rates of Blacks during the pandemic are caused by circumstances beyond their control, rather than personal choices or lifestyle. Respondents are also more likely to agree with the statement that the BLM protests arise because of structural racism and not as an excuse for criminal behavior. To rule out that exposure to COVID-19 in the earlier stages of the pandemic is just a proxy for more progressive leaning counties, we use an additional question that deals with an unrelated progressive issue: legal status for undocumented immigrants. Individuals living in counties with higher exposure to COVID-19 are not more likely to prefer more rights for undocumented immigrants, alleviating some of the concern about unobserved heterogeneity.

6 Competing Mechanisms

We have established a clear link between pandemic exposure, social media take-up and BLM protest. In this section, we explore additional (non-exclusive) mechanisms for the broadening of the BLM protests in response to the pandemic, considering *i*) a scattering rather than a broadening of protest *ii*) pandemic-induced salience of racial inequality *iii*) lower opportunity costs of protesting and *iv*) increased overall agitation and propensity to protest.

6.1 Broadening versus Scattering of Protest

In this section, we discuss the possibility that spatial spillovers from BLM protest (i.e., from the cities to the suburbs) are driving our results. There are several ways this can happen. First, the pandemic may have changed the scope and structure of BLM protests (smaller but more numerous). Second, neighboring counties may inspire future protest in close proximity. If SSEs and BLM protests themselves have spill-over effects, we may falsely attribute an increase in protest to the pandemic. Third, the pandemic and its restrictions on mobility may have led to a geographic spread of the protest movement, substituting large protests in cities with smaller protests in suburbs.

If the observed increase in the number of counties hosting a BLM event for the first time after George Floyd's murder is driven by a substitution of protest across space (e.g. re-location of protesters or creation of multiple smaller protest events), we should observe that the number of protests increases while the number of participants should decrease. In Section 4.3, we have shown that neither is the case and that – on the contrary – the scope of BLM protest in places with prior BLM protest even increases.

In addition, we construct a dummy variable that indicates whether or not one of the neighboring counties has observed a BLM protest before the pandemic. We use this variable in two ways. First, we include it as an additional control (column 1 of Table 8) and second we interact this dummy variable with COVID-19 deaths per 1000 population (column 2 of Table 8). Results show that having a traditional protester as a neighbor does not increase the probability of protesting overall within the sample of counties that had never protested before.[29] More importantly, the interaction term between exposure to COVID-19 and having a traditional protester as a neighbor in column 2 is not significant, and if anything reduces the likelihood of protesting in response to the pandemic. This seems to indicate that the displacement

[29]The results for the full sample and for the counties with at least one BLM protest before George Floyd's murder can be found in Appendix Table D7 and Table D8 correspondingly

effect is not a driver of our results.

It is also possible that protests in one county could inspire protests in neighboring counties during the outbreak of protests. To test this hypothesis we construct an indicator similar to the one described above ("traditional protester as a neighbor") but for the period after George Floyd's murder. In other words, this dummy variable indicates whether the county has a neighboring county that protested just *before* and during the same wave of protests. This allows us to show that there are no spillovers over time.

Lastly, we analyze the geographic diffusion of protest. In columns 5 and 6 of Table 8, we investigate whether proximity to the earliest and largest protest hub (Minneapolis, the location of the murder of George Floyd) affected the likelihood of observing a protest. We use the distance and squared distance to Minneapolis and find no significant impact of proximity to Minneapolis. If anything, counties further away may respond slightly more to COVID-19 exposure, with the caveat that the first stage of the interaction term becomes weak in column 6.

Overall, we take these results as evidence that the observed spread of the BLM protest is not driven by the spread of pre-existing protesters. We also find no evidence for learning or imitation through time and space. We argue that this is consistent with the social media mechanism because online exposure to the protest trigger is much less dependent on learning over time or through geographic proximity.

6.2 Salience of Racial Inequality

The second alternative mechanism is a rise in the salience of racial inequality due to the pandemic itself and not through exposure to BLM-related content online. For instance, an a-priori indiscriminate virus should affect whites and Blacks equally but if there are racial disparities in death rates, then people may be more inclined to believe that there are systemic disadvantages afflicting the Black community. We test this mechanism in two ways.

First, we hypothesize that if this mechanism is at play, counties facing a higher proportion of Black deaths due to COVID-19 would be more likely to protest after the trigger of George Floyd's death. Column 1 of Table 9 shows the estimate of the interaction term between COVID-19 death per 1000 population and the Black death burden. The Black death burden is computed as the ratio of the Black COVID-19 deaths per 1000 Black population over the total COVID-19 deaths per 1000 population. Results show that the effect on COVID-19 on protest is not higher in counties with relatively more COVID-19 related deaths among Blacks.[30]

[30]The results for the full sample and for the counties with at least one BLM protest before George Floyd's murder

Additionally, we test whether the results are driven by an increase in the awareness and sympathy towards BLM-related issues during the pandemic but *before* the murder of George Floyd. We hypothesize that if people are empathizing with problems faced by the Black community because of the pandemic itself, we would observe an increase in interest towards BLM already before the protest trigger. We test this in column 2 of Table 9, where we interact the relative popularity of BLM search terms on Google in the month leading up to George Floyd's murder with pandemic exposure. We provide more information on the BLM search terms in the Data Appendix Section A. We do not find that an increased interest in racial injustice before the protest trigger increased the probability of a demonstration.

6.3 Opportunity Cost of Protesting

Next, we test whether the results can be explained by a decrease in the opportunity cost of protesting. It is possible that new people joined the movement because they had lower (social and economic) opportunity costs of protesting during the pandemic. We test this in two ways. First, a decrease in the overall opportunity cost of protesting can be owed to a decrease in employment and economic opportunities during the pandemic. According to Bureau of Labor Statistics (2020): "in June 2020, 40.4 million people reported that they had been unable to work at some point in the last 4 weeks because their employer closed or lost business due to the coronavirus pandemic —that is, they did not work at all or worked fewer hours" which "represented 16 percent of the civilian non institutional population". We proxy the decrease of economic opportunity cost using the unemployment rate in the month before the murder of George Floyd. Column 3 of Table 9 shows the interaction between unemployment and COVID-19 deaths per 1000 population. Results show that the effect of COVID-19 on protest is not higher in counties with higher unemployment rate.

Second, we consider the decrease of the *social* opportunity costs as a possible channel. Individuals may trade off leisure activities and joining a protest. In the absence of alternative leisure activities due to lockdown and social distancing measures, the social opportunity cost of protesting decreased and made protesting relatively more attractive. We proxy the decrease of social opportunity cost with the stringency of social distancing measures at the state level. Column 4 of Table 9 shows the interaction between the stringency of social distancing measures and pandemic exposure. Results show that effect of COVID-19 deaths on protest is not higher in counties having stricter lock-down and social distancing rules.

are shown in Appendix Table D9 and Table D10 correspondingly

6.4 Agitation and Propensity to Protest

Lastly, we investigate whether the pandemic has agitated the general public. It is possible that our results are driven by overall discontent and are not linked to BLM. In order to address this concern, we look at the effect of COVID-19 on other protests, using the ACLED US Crisis Monitor protest data. We exclude BLM-related protests from this data set and expand the observation period to 3 months after George Floyd's murder to make sure we do not capture a substitution effect between BLM protests and other protests immediately after the BLM protest trigger. We report the results in column 5 of Table 9; we do not find an effect of COVID-19 on other protests. We also consider the possibility that pandemic exposure increased polarization and sparked counter-movements. In this case, it is not exposure to BLM related content on social media that makes citizens more sympathetic to the movement but the pandemic is riling up resentments towards the political opponent. To test this, we consider only COVID-19 related protests in column 6. These protests are largely comprised of anti-mask protests that were sympathizers of the Trump administration. We do not find evidence that anti-mask protests increased in response to pandemic exposure. Additionally, we verify whether the pandemic also mobilized the counter-movement to BLM. Two of the most popular hashtags in opposition to BLM were #AllLivesMatter and #BlueLivesMatter. Again, we scrape and geo-localize the universe of tweets in the three weeks following the murder of George Floyd and show in columns 7 and 8 that the pandemic did not lead to a counter-mobilization on Twitter.

While we cannot fully rule out that other mechanisms are at play simultaneously, we believe that they are unlikely to be the main driver of our results.

7 Conclusion

In this paper, we shed light on the role of social media in mobilizing new segments of society to protest against racial injustice. We begin with the stylized fact that half of the counties that observe a protest in response to the murder of George Floyd have never participated in a BLM related protest before. We then establish a causal link between pandemic exposure and BLM protest. In line with the idea of a broadening of the BLM coalition, we find that these counties are whiter, more affluent, more rural and exhibit a low baseline protest probability to protest for BLM. We interpret this as the mobilization of new segments of society that had previously not been affected by the movement's grievances.

Next, we show that the pandemic-induced take up of social media is the main driver of the reduced form effect. We develop a novel index of social media penetration at the county level to show that the pandemic acted as a demand shock to social media. We also show that pandemic exposure increased protest participation only in counties with higher social media engagement during the pandemic. In addition, we consider various alternative explanations and show that none of these dimensions can explain the increase in BLM protest in response to pandemic exposure.

Our research highlights the important role of social media for modern social movements. Exogenous changes in the use of social media may increase political mobilization in the presence of a protest trigger, notably among people not directly impacted by the movement's grievances. It is yet to be shown whether these effects are lasting and whether movements on the other side of the political spectrum were also able to take advantage of the pandemic-induced shift of attention to the digital space.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge, "When should you adjust standard errors for clustering?," Technical Report, National Bureau of Economic Research 2017.

Abouk, Rahi and Babak Heydari, "The Immediate Effect of COVID-19 Policies on Social-Distancing Behavior in the United States," *Public Health Reports*, 2020, p. 0033354920976575.

Acemoglu, Daron, Tarek A Hassan, and Ahmed Tahoun, "The power of the street: Evidence from Egypt's Arab Spring," *The Review of Financial Studies*, 2018, *31* (1), 1–42.

Adam, Dillon, Peng Wu, Jessica Wong, Eric Lau, Tim Tsang, Simon Cauchemez, Gabriel Leung, and Benjamin Cowling, "Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong," 2020.

Anderson, Monica and Paul Hitlin, "The hashtag# BlackLivesMatter emerges: Social activism on Twitter," *Pew Research Center*, 2016, *15.*

Arthur, W Brian, "Competing technologies, increasing returns, and lock-in by historical events," *The economic journal*, 1989, *99* (394), 116–131.

Askitas, Nikos, Konstantinos Tatsiramos, and Bertrand Verheyden, "Lockdown strategies, mobility patterns and covid-19," *arXiv preprint arXiv:2006.00531*, 2020.

Bates, Robert, Avner Greif, and Smita Singh, "Organizing violence," *Journal of Conflict Resolution*, 2002, *46* (5), 599–628.

Bazzi, Samuel and Christopher Blattman, "Economic shocks and conflict: Evidence from commodity prices," *American Economic Journal: Macroeconomics*, 2014, *6* (4), 1–38.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig, "This Mine Is Mine! How Minerals Fuel Conflicts in Africa," *American Economic Review*, June 2017, *107* (6), 1564–1610.

Besley, Timothy and Torsten Persson, "The logic of political violence," *The quarterly journal of economics*, 2011, *126* (3), 1411–1445.

Black Lives Matter, "Home - Black Lives Matter," https://blacklivesmatter.com/ 2020. Accessed: 2020-10-31.

Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro, "Greater Internet use is not associated with faster growth in political polarization among US demographic groups," *Proceedings of the National Academy of Sciences*, 2017, *114* (40), 10612–10617.

Bureau of Labor Statistics, "Supplemental data measuring the effects of the coronavirus (COVID-19) pandemic on the labor market," https://www.bls.gov/cps/effects-of-the-coronavirus-covid-19-pandemic.htm 2020. Accessed 2022-04-19.

Bursztyn, Leonardo, Davide Cantoni, David Y Yang, Noam Yuchtman, and Y Jane Zhang, "Persistent political engagement: Social interactions and the dynamics of protest movements," *American Economic Review: Insights*, 2021, *3* (2), 233–50.

_ , Georgy Egorov, Ruben Enikolopov, and Maria Petrova, "Social media and xenophobia: evidence from Russia," Technical Report, National Bureau of Economic Research 2019.

Butts, Kyle and John Gardner, "did2s: Two-stage difference-in-differences," *arXiv preprint arXiv:2109.05913*, 2021.

Campante, Filipe R, Ruben Durante, and Andrea Tesei, "Media and Social Capital," Working Paper 29230, National Bureau of Economic Research September 2021.

Campante, Filipe, Ruben Durante, and Francesco Sobbrio, "Politics 2.0: The multifaceted effect of broadband internet on political participation," *Journal of the European Economic Association*, 2018, *16* (4), 1094–1136.

Cantoni, Davide, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang, "Protests as Strategic Games: Experimental Evidence from Hong Kong's Anti-Authoritarian Movement," *Quarterly Journal of Economics*, May 2019, *134* (2).

Casanueva, Annalí, "Can Chants in the Street Change Parliament's Tune? The Effects of the 15M Social Movement on Spanish Elections," *The Effects of the 15M Social Movement on Spanish Elections (April 22, 2021)*, 2021.

CDC, "Risk Assessment and Management," Technical Report, Centers for Disease Control and Prevention March 2020.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–96.

Chenoweth, Erica, Barton H Hamilton, Hedwig Lee, Nicholas W Papageorge, Stephen P Roll, and Matthew V Zahn, "Who Protests, What Do They Protest, and Why?," Technical Report, National Bureau of Economic Research 2022.

Cinelli, Carlos, Andrew Forney, and Judea Pearl, "A crash course in good and bad controls," *Sociological Methods & Research*, 2021, p. 00491241221099552.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig, "Inference with Arbitrary Clustering," IZA Discussion Papers 12584, Institute of Labor Economics (IZA) August 2019.

Dave, Dhaval M, Andrew I Friedson, Kyutaro Matsuzawa, Joseph J Sabia, and Samuel Safford, "Black Lives Matter Protests, Social Distancing, and COVID-19," Technical Report, National Bureau of Economic Research 2020.

_ , Drew McNichols, and Joseph J Sabia, "Political Violence, Risk Aversion, and Non-Localized Disease Spread: Evidence from the US Capitol Riot," Technical Report, National Bureau of Economic Research 2021.

Dube, Oeindrila and Juan F Vargas, "Commodity price shocks and civil conflict: Evidence from Colombia," *The review of economic studies*, 2013, *80* (4), 1384–1421.

Eisinger, Peter K, "The conditions of protest behavior in American cities," *The American Political Science Review*, 1973, *67* (1), 11–28.

Ellis, C.J. and J. Fender, "Information cascades and revolutionary regime transitions," *Economic Journal*, 2011, *121* (553), 763–792.

Endo, Akira, Sam Abbott, Adam J Kucharski, Sebastian Funk et al., "Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China," *Wellcome Open Research*, 2020, *5* (67), 67.

Engist, Oliver and Felix Schafmeister, "Do political protests mobilize voters? Evidence from the Black Lives Matter protests," *Public Choice*, 2022, pp. 1–21.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova, "Social media and protest participation: Evidence from Russia," *Econometrica*, 2020, *88* (4), 1479–1514.

_ , Maria Petrova, and Konstantin Sonin, "Social media and corruption," *American Economic Journal: Applied Economics*, 2018, *10* (1), 150–74.

Freelon, D., C. McIlwain, and M. Clark, "Beyond the hashtags: #Ferguson, #Blacklivesmatter, and the online struggle for offline justice," Technical Report, Center for Media & Social Impact 2016.

Friedson, Andrew I, Drew McNichols, Joseph J Sabia, and Dhaval Dave, "Did California's shelter-in-place order work? Early coronavirus-related public health effects," Technical Report, National Bureau of Economic Research 2020.

Fujiwara, Thomas, Karsten Müller, and Carlo Schwarz, "The effect of social media on elections: Evidence from the United States," Technical Report, National Bureau of Economic Research 2021.

Galvani, Alison P and Robert M May, "Dimensions of superspreading," *Nature*, 2005, *438* (7066), 293–295.

García-Jimeno, Camilo, Angel Iglesias, and Pinar Yildirim, "Information Networks and Collective Action: Evidence from the Women's Temperance Crusade," *American Economic Review*, 2022, *112* (1), 41–80.

González, Felipe and Mounu Prem, "Police Repression and Protest Behavior: Evidence from Student Protests in Chile," *Available at SSRN 3705486*, 2020.

Guriev, S. and D. Treisman, "How modern dictators survive: an informational theory of the new authoritarianism," Technical Report, National Bureau of Economic Research 2015.

Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya, "3g internet and confidence in government," *Available at SSRN 3456747*, 2019.

Hager, Anselm, Lukas Hensel, Johannes Hermle, and Christopher Roth, "Political activists as free-riders: Evidence from a natural field experiment," 2020.

Hale, Thomas, Samuel Webster, Anna Petherick, Toby Phillips, and Beatriz Kira, "Oxford covid-19 government response tracker," *Blavatnik School of Government*, 2020, *25.*

Heckman, James J, *The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process and some Monte Carlo evidence*, University of Chicago Center for Mathematical studies in Business and Economics, 1987.

Ince, Jelani, Fabio Rojas, and Clayton A Davis, "The social media response to Black Lives Matter: how Twitter users interact with Black Lives Matter through hashtag use," *Ethnic and racial studies*, 2017, *40* (11), 1814–1830.

Jinjarak, Yothin, Rashad Ahmed, Sameer Nair-Desai, Weining Xin, and Joshua Aizenman, "Accounting for Global COVID-19 Diffusion Patterns, January-April 2020," Technical Report, National Bureau of Economic Research 2020.

Kutlaca, Maja, Helena RM Radke, and Julia C Becker, "The impact of including advantaged groups in collective action against social inequality on politicized identification of observers from disadvantaged and advantaged groups," *Political Psychology*, 2022, *43* (2), 297–315.

Lancaster, Tony, "The incidental parameter problem since 1948," *Journal of econometrics*, 2000, *95* (2), 391–413.

Lasry, Arielle, Daniel Kidder, Marisa Hast, Jason Poovey, Gregory Sunshine, Kathryn Winglee, Nicole Zviedrite, Faruque Ahmed, Kathleen A Ethier, CDC Public Health Law Program et al., "Timing of community mitigation and changes in reported COVID-19 and community mobility—four US metropolitan areas, February 26–April 1, 2020," *Morbidity and Mortality Weekly Report*, 2020, *69* (15), 451–457.

Leclerc, Quentin J, Naomi M Fuller, Lisa E Knight, Sebastian Funk, Gwenan M Knight, cmmid covid 19 Working Group et al., "What settings have been linked to SARS-CoV-2 transmission clusters?," *Wellcome open research*, 2020, *5.*

Lelkes, Yphtach, Gaurav Sood, and Shanto Iyengar, "The hostile audience: The effect of access to broadband internet on partisan affect," *American Journal of Political Science*, 2017, *61* (1), 5–20.

Lipsky, Michael, "Protest as a political resource," *The American Political Science Review*, 1968, *62* (4), 1144–1158.

Manacorda, Marco and Andrea Tesei, "Liberation technology: Mobile phones and political mobilization in Africa," *Econometrica*, 2020, *88* (2), 533–567.

Mangrum, Daniel and Paul Niekamp, "JUE Insight: College Student Travel Contributed to Local COVID-19 Spread," *Journal of Urban Economics*, 2020, p. 103311.

Marcus, Michelle and Pedro HC Sant'Anna, "The role of parallel trends in event study settings: An application to environmental economics," *Journal of the Association of Environmental and Resource Economists*, 2021, *8* (2), 235–275.

Mazumder, Soumyajit, "Black Lives Matter for Whites' Racial Prejudice:Assessing the Role of Social Movements in Shaping Racial Attitudes in the United States," Technical Report, Harvard University May 2019.

McCarthy, John D and Mayer N Zald, "Resource mobilization and social movements: A partial theory," *American journal of sociology*, 1977, *82* (6), 1212–1241.

McKersie, Robert B, "The 1960s Civil Rights Movement and Black Lives Matter: Social Protest from a Negotiation Perspective," 2021.

Melnikov, Nikita, "Mobile Internet and Political Polarization," *Available at SSRN 3937760*, 2021.

Miller, Danielle, Michael A Martin, Noam Harel, Talia Kustin, Omer Tirosh, Moran Meir, Nadav Sorek, Shiraz Gefen-Halevi, Sharon Amit, Olesya Vorontsov et al., "Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel," *medRxiv*, 2020.

Morris, Aldon D, *The origins of the civil rights movement*, Simon and Schuster, 1986.

Müller, Karsten and Carlo Schwarz, "From hashtag to hate crime: Twitter and anti-minority sentiment," *Available at SSRN 3149103*, 2020.

_ and _ , "Fanning the flames of hate: Social media and hate crime," *Journal of the European Economic Association*, 2021, *19* (4), 2131–2167.

Olson, Mancur, "Collective action," in "The invisible hand," Springer, 1989, pp. 61–69.

Ostrom, Elinor, *Governing the commons: The evolution of institutions for collective action*, Cambridge university press, 1990.

Passarelli, Francesco and Guido Tabellini, "Emotions and political unrest," *Journal of Political Economy*, 2017, *125* (3), 903–946.

Porta, Donatella Della and Mario Diani, *The Oxford handbook of social movements*, Oxford University Press, 2015.

_ and _ , *Social movements: An introduction*, John Wiley & Sons, 2020.

Rosenbaum, Pail R. and Donald B. Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 04 1983, *70* (1), 41–55.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe, "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *arXiv preprint arXiv:2201.01194*, 2022.

Rupasingha, Anil, Stephan J Goetz, and David Freshwater, "The production of social capital in US counties," *The journal of socio-economics*, 2006, *35* (1), 83–101.

Sanderson, Eleanor and Frank Windmeijer, "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of Econometrics*, 2016, *190* (2), 212–221. Endogeneity Problems in Econometrics.

Sangnier, Marc and Yanos Zylberberg, "Protests and trust in the state: Evidence from African countries," *Journal of Public Economics*, August 2017, *152*, 55–67.

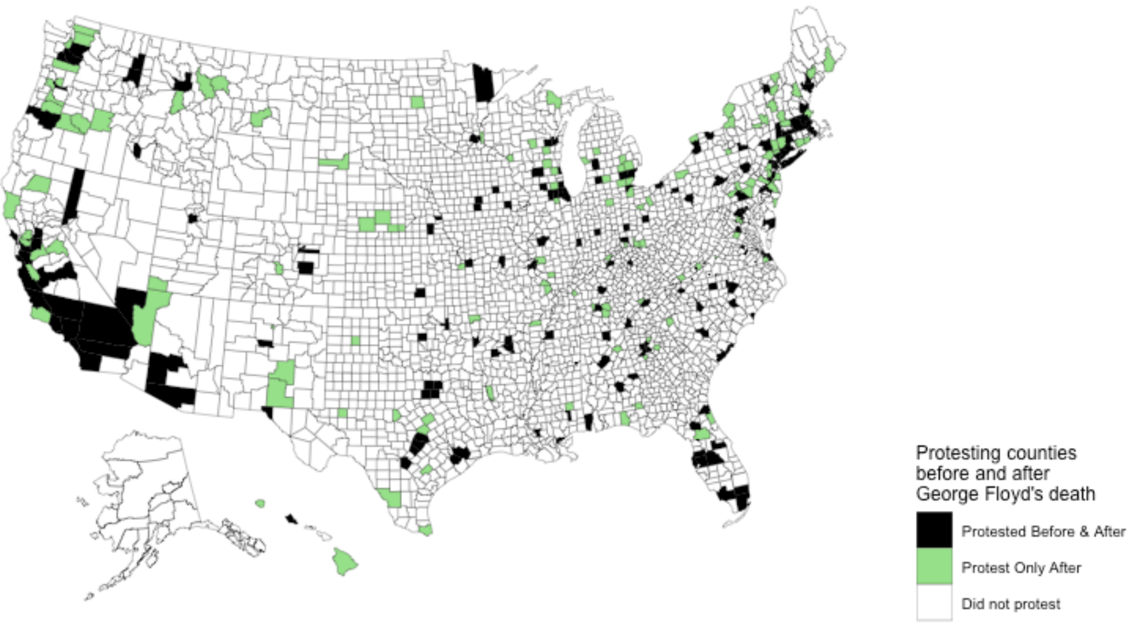
Takhteyev, Yuri, Anatoliy Gruzd, and Barry Wellman, "Geography of Twitter networks," *Social networks*, 2012, *34* (1), 73–81.

Tibshirani, Robert, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, *58* (1), 267–288.

Wooldridge, Jeffrey M, *Introductory econometrics: A modern approach*, Cengage learning, 2015.
Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov, "Political effects of the internet and social media," *Annual Review of Economics*, 2020, *12*, 415–438.

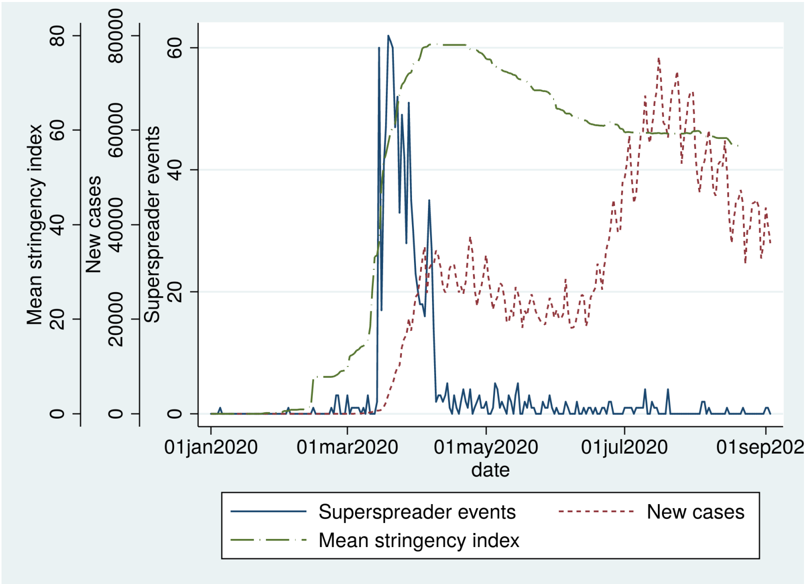
Figures and Tables

Figure 1: Traditional and new counties with BLM protest



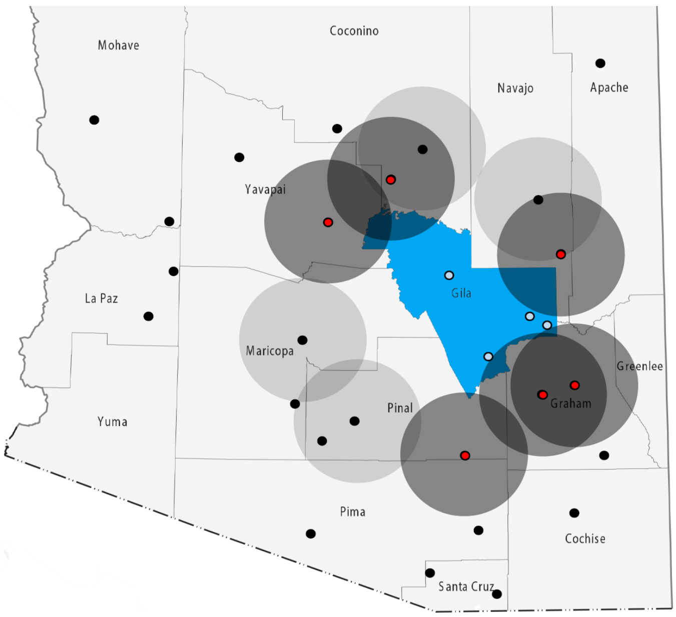
Note: Map based on data from *Elephrame*. This map represents whether US counties that protested in the three weeks following the murder of George Floyd (May 25 to June 14, 2020) already held a BLM protest before the murder of George Floyd or not. Counties in Black protested both before and after the murder of George Floyd. Counties in green are counties whose first BLM protest was after George Floyd's murder. Counties in white did not protest after the murder.

Figure 2: Window of opportunity for SSEs



Note: Solid (blue) line represents the number of daily total SSEs over time (January 2020 to September 2020). Dashed (green) line shows the daily average stringency index across all US states, as measured by the Oxford COVID response tracker. Dotted (red) line shows the number of daily new COVID-19 cases as recorded by the New York Times.

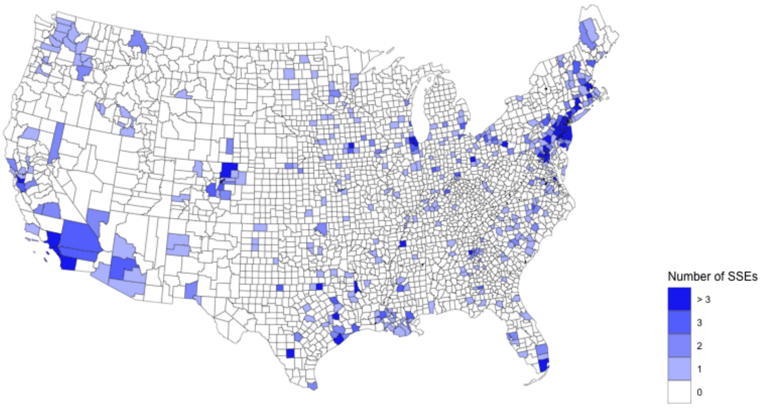
Figure 3: Construction of the super-spreading events instrument (example)



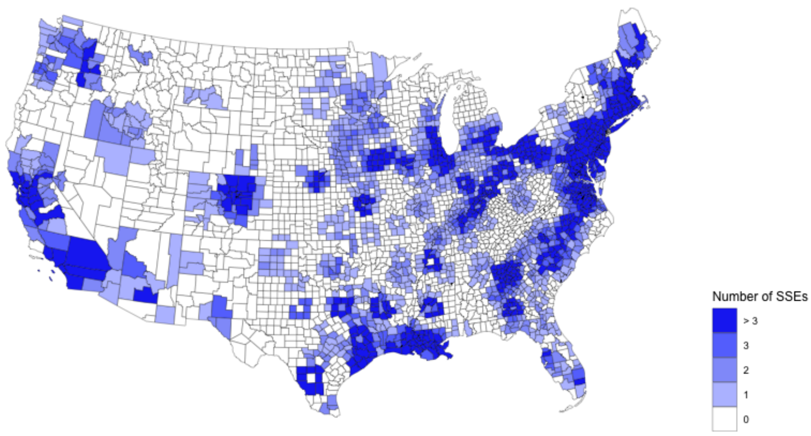
Note: Example of the construction of the instrument we use for COVID-19 death per 100k population: the total number SSEs in neighbouring counties during a certain period. Red point are the super-spreader events assigned to the blue county. Gray shaded area represents the 50km radius around each super-spreader event. Black points represent super-spreader event that are not assigned to the blue county because are too far away from the border. White points represents super-spreader events that are inside the county and therefore not assigned to the county (to increase exogeneity).

Figure 4: Geographic distribution of super-spreader events (SSEs)

(a) Number SSEs at county level 6 weeks prior to the murder of George Floyd



(b) Number of SSEs in 50 km radius outside of own county, 6 weeks prior to the murder of George Floyd



Note: Geographic distribution of two measures related to super-spreader events. Sub-figure (a) shows the cumulative number of COVID-19 super-spreader events at the county level up to 6 weeks prior the murder of George Floyd. Sub-figure (b) shows the number of super-spreader events in neighbouring counties up to 50km around the county border in the same period of time. The variable shown in sub-figure (b) correspond to the measure of the main instrument used to instrument the number of COVID-19 deaths per one hundred thousand population.

Table 1: Summary statistics: Counties without prior BLM event

From 25th of May to 14th of June 2020:	N	Mean	SD	Min	Max
Presence of BLM events	2768	0.048	0.213	0.000	1.000
Number of BLM events	2768	0.064	0.322	0.000	5.000
Participants in BLM events	2768	21.026	172.090	0.000	5500.000
Participants per event	132	355.115	621.538	0.000	5500.000
Tweets mentioning BLM	2768	322.706	4818.930	0.000	243596.000
New users tweeting about BLM	2768	1.809	25.763	0.000	1306.000
Followers of @BlkLivesMatter created during the pandemic	2768	0.441	2.167	0.000	78.000
Tweets mentioning #AllLivesMatter	2768	47.488	326.063	0.000	15659.000
Tweets mentioning #BlueLivesMatter	2768	6.125	36.206	0.000	1647.000
Neighbor protested first	2768	0.334	0.472	0.000	1.000
On the 25th of May 2020:					
COVID deaths (total)	2768	8.366	46.396	0.000	1025.000
COVID cases (total)	2768	164.485	663.300	0.000	15169.000
COVID deaths (per 1000)	2768	0.099	0.230	0.000	2.935
COVID cases (per 1000)	2768	2.596	5.662	0.000	145.513
Super-spreader events, 6+ weeks ago, neighboring	2768	2.327	7.564	0.000	143.000
Black death burden	2768	1.344	0.985	0.000	4.104
Lockdown stringency index	2768	68.210	8.543	47.220	89.810
Before the 25th of May 2020:					
Google searches for Twitter	2731	60.523	10.941	17.000	100.000
Residential stay	1348	10.633	3.387	3.600	26.286
New Twitter users (sample)	3106	3.062	17.034	0.000	600.000
Later outcomes:					
Followers of @BlkLivesMatter	2768	16.186	94.110	0.000	3304.000
County characteristics:					
Black police-related deaths (2014-2019)	2768	0.207	0.724	0.000	15.000
Black police-related deaths (2020)	2768	0.014	0.131	0.000	3.000
Unemployment rate (year average)	2768	4.713	1.575	0.708	17.442
Black population share	2768	0.093	0.146	0.000	0.875
Non-white population share	2768	0.134	0.160	0.000	0.928
Large cities	2768	0.001	0.027	0.000	1.000
Suburban areas	2768	0.105	0.307	0.000	1.000
Smaller towns	2768	0.201	0.400	0.000	1.000
Rural areas	2768	0.694	0.461	0.000	1.000
BLM events (2014-2019)	2768	0.000	0.000	0.000	0.000
Black poverty rate	2768	0.283	0.236	0.000	1.000
Population share with 3+ risk factors	2768	25.957	5.066	10.685	48.448
Vote share for republicans (2016)	2768	0.656	0.141	0.083	0.960
Vote share for republicans (2012)	2768	0.614	0.140	0.060	0.959
Median household income (2016)	2768	47521.697	12362.349	20170.891	129150.343
Social capital	2768	1.426	0.726	0.000	6.887
Distance to Minneapolis	2768	1192.851	538.680	34.438	6474.706
Notable Deaths	2768	0.001	0.033	0.000	1.000
Log(Twitter users observed in 2019)	2768	1.739	1.215	0.000	8.093
Log(SXSW followers created before March 2017)	2768	0.090	0.228	0.000	1.427
Log(SXSW followers created during March 2017)	2768	0.157	0.312	0.000	1.658

Note: summary of main variables used in our analysis. The sample consists of all US counties with no BLM event before George Floyd's death. Summary statistics for all sub-samples and all counties is provided in appendix table A5. We report the number of observations, the mean, the standard deviation as well as the minimum and maximum value of each of the variables. We present the variable in different panels corresponding to the time at which every variable is measured. The first panel presents the variables measured during the three weeks following the murder of George Floyd; the second, the variables that were measured on the same day of the murder of George Floyd; the third during the period after the start of the pandemic but before the murder of George Floyd; the forth was measured in 2022. The final panel shows county-level controls measured at different times before the murder of George Floyd.

Table 2: Main Result: Effect of COVID-19 on BLM Protest

	BLM Protest				
	(1)	(2)	(3)	(4)	(5)
Panel A: 2SLS estimates					
COVID	0.555***	0.675***	0.404**	0.575***	0.456***
(deaths/1000)	(0.0745)	(0.160)	(0.184)	(0.169)	(0.168)
Panel B: reduced form estimates					
Number of SSEs	0.00548***	0.00595***	0.00303*	0.00431***	0.00343**
(6 weeks prior, $\leq$ 50 km)	(0.000966)	(0.00178)	(0.00161)	(0.00140)	(0.00126)
Panel C: OLS estimates					
COVID	0.0661	0.0503	0.0385*	0.0695**	0.0744**
(deaths/1000)	(0.0445)	(0.0319)	(0.0218)	(0.0319)	(0.0315)
Observations	2,768	2,767	2,767	2,767	2,767
KP F-stat	115.10	44.53	27.04	27.04	27.04
Mean dep. var.	0.0477	0.0477	0.0477	0.0665	0.0795
Mean COVID	0.099	0.099	0.099	0.099	0.099
County controls			Y	Y	Y
State fixed effects		Y	Y	Y	Y
Time frame	3 weeks	3 weeks	3 weeks	6 weeks	9 weeks

Note: estimation of the effect of the COVID-19 pandemic on the presence of BLM protests. The sample consists of counties with no BLM protest before George Floyd's murder. The top panel reports 2SLS results, using the number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument. Panel B presents the reduced form estimates and Panel C, the corresponding OLS results. In columns 1 - 3, we define the outcome as the presence of at least one BLM event during the three weeks following the murder of George Floyd. In columns 4 and 5, we use 6 and 9 weeks, respectively. In column 1 we do not include any control. In column 2 (column 3 respectively) we include state fixed effects (state fixed effects and county-level controls). Control variables include: the share of Black population, urban (category [1-6]), median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience, Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations) and deadly force used by police against Black people. We report Kleibergen-Paap rkWald F statistics for weak instruments. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Heterogeneity by baseline county characteristic

	BLM Protest									
Sample of all counties	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
COVID	0.215*	0.331**	-0.908*	0.276	-0.176	0.256	-0.314	-0.0301	0.295**	0.243**
(deaths/1000)	(0.121)	(0.487)	(0.204)	(0.301)	(0.175)	(0.188)	(0.147)	(0.112)	(0.120)	(0.152)
... × Probability of protest		-2.175***								
		(0.805)								
... × Share non-Black			1.301**							
			(0.548)							
... × Share non-white				-0.163						
				(0.411)						
... × Median household income					0.00417*					
					(0.00243)					
... × Republican vote share						-0.102				
						(0.393)				
... × Not large cities							0.608***			
							(0.155)			
... × Suburban areas								0.321***		
								(0.112)		
... × Smaller towns									0.0391	
									(0.137)	
... × Rural areas										-0.155
										(0.159)
Interacting variable		-0.191	1.695***	-0.111**	0.00224**	-1.056***	-0.566***	-0.0572**	0.0703***	-0.0652***
		(0.185)	(0.0537)	(0.000926)	(0.191)	(0.0865)	(0.0251)	(0.0220)	(0.0223)	(0.596)
Observations	3,106	3,002	3,106	3,106	3,106	3,106	3,106	3,106	3,106	3,106
F COVID	36.05	25.17	20.93	22.58	22.62	24.95	42.35	37.12	20.25	17.36
F interaction		35.45	17.19	46.17	18.79	13.81	110.8	48.84	42.69	9.229
County controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: estimation of the differential effect of COVID-19 deaths per 1000 population on the presence of at least one Black Lives Matter event during the three weeks following the murder of George Floyd depending on different counties characteristics. Table presents 2SLS estimation, using number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument for the full sample (all counties with and without prior BLM protest). Each column presents the interaction with a different county characteristic. The line "interacting variable" presents the coefficient of the variables that is interacted with COVID-19 deaths respectively for each column. It is different for each column. Column (1) is the baseline regression for comparisson. All controls from the preferred specification of Table 2 and state fixed effects are included. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Alternative outcomes

	Any BLM protest	Number of BLM protests	Total participants	Participants per protest	Tweets BLM	Followers @BLM
	(1)	(2)	(3)	(4)	(5)	(6)
COVID	0.404**	0.621**	-81.00	-106.2	2,586*	174.8**
(deaths/1000)	(0.187)	(0.247)	(261.4)	(265.9)	(1,449)	(79.07)
Observations	2,767	2,767	2,767	2,767	2,767	2,767
KP F-stat	27.04	27.04	27.04	27.04	27.04	27.04
Mean dep. var.	0.0477	0.0636	21.03	16.94	322.6	16.18
County controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: estimation of the effect of COVID-19 deaths per 1000 population instrumented by the number of SSEs in neighbouring counties (50km radius) on different outcomes. The sample consists of counties with no BLM protest before George Floyd's murder. Columns 1 to 4 use protest information from *Elephrame* (described in the data section). Column 1 is the baseline result for reference. The outcome of column 2 is the number of BLM events in the three weeks following the murder of George Floyd. Column 3 reports results for the number of participants and column 4 divides the number of participants by the number of events in the county (imputing zero to counties with no BLM event). Column 5 reports the number of geo-located tweets that use at least one of the following hashtags #BlackLivesMatter #BlackLifeMatters #BLM in the three weeks following the murder. Column 6, reports the number of geo-located accounts that follow the official BLM account @BlkLivesMatter. We report Kleibergen-Paap rkWald F statistic for the first stage. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: COVID-19 exposure and social media use

	PC 1	New Twitter accounts	(log) New Twitter accounts	Google search for Twitter	Residential stay
	(1)	(2)	(3)	(4)	(5)
COVID	1.264***	17.88**	1.317***	18.28**	3.885***
(deaths/1000)	(0.380)	(7.871)	(0.339)	(8.838)	(0.931)
Observations	1,014	2,767	2,767	2,730	1,022
KP F-stat	20.21	27.04	27.04	26.03	20.16
Mean dep. var.	-0.237	1.808	0.420	60.52	10.01
County controls	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y

Note: estimation of the effect of COVID-19 deaths per 1000 population instrumented by the number of SSEs in neighbouring counties (50km radius) on the use of social media. The sample consists of counties with no BLM protest before George Floyd's murder. Column 1 shows the standardized first principal component of four outcomes of interest: new Twitter accounts (created during the pandemic but before the murder of George Floyd) tweeting about BLM during the three weeks after the murder of George Floyd (and its log), Google searches for "Twitter" and the level of change of residential stay with respect to the baseline before the pandemic. Table A7 details the construction of the principal component. Column 2 (resp column 3) shows estimates for new Twitter accounts (log of new accounts) created after the beginning of the pandemic but before George Floyd's murder that tweet about BLM in the three weeks following George Floyd's death. Column 4 (resp column 5) shows results for Google searches for "Twitter" (the level of change of residential stay with respect to the baseline before the pandemic) between April 13 to May 24. All specifications include state fixed effects and the standard controls of the main specification. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Effect of Twitter on BLM protest

	BLM Protest					
Measure for Twitter	*New Accounts*		*Baseline Accounts*		*Baseline Accounts (IV)*	
	(1)	(2)	(3)	(4)	(5)	(6)
COVID × Twitter		0.205**		0.245***		0.232*
		(0.0834)		(0.0880)		(0.118)
COVID	0.355*	-0.0444	0.386**	-0.599	0.353**	-0.578
	(0.188)	(0.277)	(0.178)	(0.409)	(0.167)	(0.568)
Twitter	0.0374***	0.0193*	0.0319***	0.0128	0.0677*	0.0406
	(0.00760)	(0.0102)	(0.00540)	(0.00854)	(0.0398)	(0.0453)
Observations	2,767	2,767	2,767	2,767	2,767	2,767
F-stat interaction		60.91		47.35		18.87
F-stat COVID	26.84	15.28	27.40	11.35	30.82	8.530
F-stat Twitter					16.78	19.31
County controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: estimation of the differential effect of exposure to the COVID-19 pandemic on the presence of BLM protest following George Floyd's murder depending on different measure of Twitter usage. The sample consists of counties with no BLM protest before George Floyd's murder. COVID-19 is measured as in the main specification: deaths per 1000 population instrumented by the number of SSEs in neighbouring counties (50km radius). Column 1 and 2 show the effect of the log number of new Twitter users that have created an account during the pandemic, but before the murder of George Floyd and instrumented COVID-19 deaths. Column 2 additionally shows the interaction. Columns 3 and 4 show the effect of Twitter users at baseline (i.e. December 2019) and instrumented COVID-19 deaths. Column 4 additionally shows the interaction. Columns 5 and 6 use instrumented log number of Twitter users at baseline (i.e. December 2019), using the Müller and Schwarz (2020) SXSW instrument. The first stage regression is reported on Table C4. All specifications include state fixed effects and all controls from the baseline specification. Columns 5 and 6 include an additional control: the log number of users before the SXSW festival. First stage F statistics are presented following Sanderson and Windmeijer (2016). Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table 7: COVID-19 and individual-level news consumption and attitudes

	News consumption			**Attitudes towards Blacks, BLM & COVID-19**			**Placebo**
	Follow news about GF	Receive news about GF on social media	Ratio social media to overall GF news	Higher Black COVID hospitaliz. not their fault	Protest because structural racism	Protest because criminal behaviour	Illegal immigration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
COVID-19 deaths per capita (category)	0.0480***	0.0343**	0.0225*	0.0115*	0.0259***	-0.0254**	-0.00641
	(0.00964)	(0.0152)	(0.0134)	(0.00645)	(0.00907)	(0.0109)	(0.00540)
Observations	9,201	9,121	9,111	9,212	9,190	9,183	9,212
Black	Y	Y	Y	Y	Y	Y	Y
Metropolitan area	Y	Y	Y	Y	Y	Y	Y
Female	Y	Y	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y	Y	Y
Education	Y	Y	Y	Y	Y	Y	Y
Income	Y	Y	Y	Y	Y	Y	Y
Democrat	Y	Y	Y	Y	Y	Y	Y

Note: relation between living in a county with different levels of COVID-19 deaths per capita on different outcomes related to news consumption and attitudes towards Blacks, BLM and COVID-19. Columns 1 to 3 present the estimates for outcomes related to news consumption. In particular, column 1, 2 and 3 show respectively: the interest in George Floyd related news, the amount of George Floyd related news received through social media and the ratio of the variable of column 2 over the variable of column 1. Columns 4 to 6 show the results for the outcomes related to attitudes towards BLM and racism awareness. Column 4 corresponds to the likelihood of answering that the higher COVID-19 mortality rate faced by Blacks is due to their disadvantaged circumstances instead of to their personal life style choices. Columns 5 and 6 correspond to the likelihood of answering that the protest following George Floyd's death is related with structural racism or to criminal behaviour respectively. Finally, column 7 shows a placebo result. The exact framing of the questions is as follows: column 1: *"How closely have you been following news about the demonstrations around the country to protest the death of George Floyd, a black man who died while in police custody?"*; column 2: *How much, if any, news and information about the demonstrations to protest the death of George Floyd have you been getting on social media (such as Facebook, Twitter, or Instagram)?*; column 4:*Do you think the reasons why black people in our country have been hospitalized with COVID-19 at higher rates than other racial or ethnic groups have more to do with... Circumstances beyond people's control*; column 5: *How much, if at all, do you think each of the following has contributed to the demonstrations to protest the death of George Floyd? Longstanding concerns about the treatment of black people in the country*; column 6: *Some people taking advantage of the situation to engage in criminal behavior*; column 7: *Which comes closer to your view about how to handle undocumented immigrants who are now living in the U.S.? There should be a way for them to stay in the country legally, if certain requirements are met* All columns include controls for various characteristics of the respondent: race, living in a metropolitan area, gender, age, education, income and leaning towards the democratic party. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Competing Mechanisms: Broadening versus Scattering of Protest

	BLM Protest					
	(1)	(2)	(3)	(4)	(5)	(6)
IV: COVID (deaths/1000)	0.410**	0.306	0.412**	0.175	-0.153	3.307
	(0.189)	(0.343)	(0.191)	(0.257)	(0.806)	(2.464)
× Neighbor protested historically		0.116				
		(0.345)				
× Neighbor protested currently				0.236		
				(0.240)		
× Distance to Minneapolis					0.000371	-0.00770
					(0.000532)	(0.00473)
× Distance to Minneapolis (squared)						3.73e-06*
						(2.11e-06)
Observations	2,767	2,767	2,767	2,767	2,767	2,767
F first stage	27.08	13.75	26.60	13.47	29.44	16.61
F interaction		16.57		32.07	21.22	11.51
F interaction (with squared)						7.753
Mean of dep. var.	0.0477	0.0477	0.0477	0.0477	0.0477	0.0477
Neighbour protested historically	Y	Y				
Neighbour protested currently			Y	Y		
County controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: estimation of the effect of COVID-19 deaths per 1000 population instrumented by the number of SSE in neighbouring counties (50km radius) on the presence of BLM protests. The sample consists of counties with no BLM protest before George Floyd's murder. Column 1 (column 3) shows estimates for instrumented COVID deaths controlling for a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 2 and 4 present heterogeneous effects for the presence of a neighbouring county that protested before. Column 2 (column 4) shows the interaction term with a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 5 and 6 presents results for interaction with distance to Minneapolis and distance to Minneapolis squared. All results are shown for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and the same controls as the baseline specification. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Competing Mechanisms: Salience, Opportunity Cost, and Agitation

	BLM Protest				Other Protests	COVID-19 Protests	Tweets AllLivesMatter	Tweets BlueLivesMatter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
COVID (deaths/1000)	0.119	-0.0466	0.501	-0.624	0.333	0.129	528.1	62.57
	(0.250)	(0.535)	(0.336)	(2.621)	(0.271)	(0.123)	(348.2)	(39.19)
× Black death burden	0.502							
	(0.410)							
× BLM Google searches		0.0135						
		(0.0153)						
× Unemployment			-0.0251					
			(0.0695)					
× Stringency index				0.0141				
				(0.0376)				
Observations	2,767	2,706	2,767	2,767	2,767	2,767	2,767	2,767
F COVID	27.09	13.18	35.93	12.95	27.04	27.04	27.04	27.04
F interaction	3.579	14.95	14.85	12.98				
Mean of dep. var.	0.0477	0.0473	0.0477	0.0477	0.0322	0.00976	47.45	6.123
County controls	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y

Note: estimation of the effect of COVID-19 deaths per 1000 population intrumented by the number of SSE in neighbouring counties on the presence of several types of protest. The sample consists of counties with no BLM protest before George Floyd's murder. Columns 1 to 4 show heterogeneous effects (using the interaction between several county characteristics and COVID-19) on the presence of BLM events. The interacting variables are, respectively: Black death burden one week prior to George Floyd's murder; Google searched for "BLM" 3 weeks prior to George Floyd's murder; unemployment rate and an index capturing the stringency of socially distancing measures (both variables measured at the closest date available prior to the murder). The coefficient for the interacting variable in column 4 is dropped as stringency is measured at the state level and state fixed effects are included. Column 5 (resp column 6) presents results for all other protests besides BLM (protest related to COVID-19; e.g. anti-mask protest) during the 3 weeks following George Floyd's murder. Columns 7 and 8 show as an outcome the number of tweets including the pro-police and anti BLM hashtags #AllLivesMatter and #BlueLivesMatter. All results are shown for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and the same controls as the baseline specification. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Online Appendix

Contents

Appendix A: Data Appendix

A.1 Description of Data Sources

Super Spreader Events. Our identification strategy relies on records of SSE in the early stages of the pandemic. In this section, we discuss the limitations of the SSE data set and how we address these in the empirical section. The data set is collected from various sources by researchers from the London School of Hygiene and Tropical Medicine and published as a free access data base for researchers and the media under the *SARS-CoV-2 Superspreading Events from Around the World* Project.

A main challenge in the construction of this data base is that there is no standard definition of a SSE. The data base mainly refers to "outbreak" and "clusters" for which they use the UK Government Public Health Definition: "two or more test-confirmed cases of COVID-19 among individuals associated with a specific non-residential setting with illness onset dates within a 14-day period." The outbreak definition is expanded to "identified direct exposure between at least 2 of the test-confirmed cases in that setting (for example under one metre face to face, or spending more than 15 minutes within 2 metres) during the infectious period of one of the cases when there is no sustained local community transmission - absence of an alternative source of infection outside the setting for the initially identified cases."

The data base draws from one main source: Leclerc et al. (2020) who performed a systematic review of available literature and media reports to find settings reported in peer reviewed articles and media with "outbreak" or "cluster" characteristics. There were various extensions to this data set, using articles of journalists, expanding that data set to second and third generation events by Swinkels (2020), and including the Western Pacific Region for a project of the World Health Organisation (under the project lead of Fatim Lakha, also from the London School of Tropical Medicine and Hygiene). We will primarily draw from Leclerc et al. (2020), as we focus on SSEs in the United States during the early stages of the pandemic.

There are various limitations in the measurement of SSEs. First, there exists some uncertainty about the exact date of the SSE. If, for instance, there was a COVID-19 cluster at a worker dormitory, the exact date of the transmission event is difficult to narrow down. In these cases, researchers make an approximation based on the timing of tests and overall case numbers. We address this concern by using the cumulative number of SSEs until a certain cut-off date (first week of April in the baseline version of the instrument), thereby not relying on the specific timing of the SSE. Second, for many SSEs it is not known exactly how many people were infected (either directly at the SSE or by somebody who was infected at the initial SSE). The database always uses the lowest number cited in the articles about the SSE but actual numbers can be much higher. The actual detected number of cases will be related to testing capacity and potentially other unobserved factors at the county level. For this reason, we use the most simple version of the instrument, i.e. counting the number of SSEs rather than using the cases associated with the SSE. Third, the GPS coordinates of SSEs are almost always approximate. For instance, when an SSE occurred somewhere in city A, typically the database uses GPS coordinates for a random location within that city, not the for precise location. In a robustness check, we make sure that our results are not sensitive to changing the radius around SSEs to account for potential measurement error. Overall, the measurement error in SSEs would only bias our results if it is somehow related to the counties' overall propensity to protest (and is not captured in the set of controls or state fixed effects). One important exercise, addresses this concern: SSEs do not predict past BLM events. If SSEs were disproportionately recorded in places with a higher likelihood of a BLM event occurring, we should see a systematic relationship to previous BLM protest, which is not the case.

Twitter usage during the protests. Twitter data is an important source when studying social events and protests. Previous work on BLM events has used this data (Ince et al., 2017). We collected tweets using the Twitter Academic Research API. In particular, we collected all tweets that contain the keywords "BLM", "Black Lives Matter", "Black Life Matters" or "George Floyd",[31] including retweets, between May 25 and June 14. For each tweet, we extract the time and text of the tweet, the user, the user's stated location, and account creation date. We present a selection of tweets that are part of our sample in Table A3.

Geo-location of tweets. We follow the literature in assigning the location of a tweet or a user by extracting information on their self-reported location from their Twitter profile (Enikolopov et al., 2020; Takhteyev et al., 2012; Müller and Schwarz, 2020). Not all users report a location and among those who do, not all state a valid location (e.g., "in the heart of Justin Bieber") so we restrict the sample to the users that state a valid location that can be matched to a USA county (in particular, we exclude users whose location only mentions a state). The location is an arbitrary text field which is not meant to be machine-readable. We use the Nominatim geocoding engine (based on the Open Street Map database) to find the coordinates of the most likely match for the location. We then filter out all locations outside the US and all locations that are too vague (i.e. that map the whole country or a whole state). Finally, we map these coordinates to counties using the US Census Bureau cartographic boundary files. Across our different tweet collections, we end up with 23.3 million tweets. This approach has clear limitations as it relies only on self-reported locations and may not be representative of the whole Twitter universe. We report summary stats on the counties for which we were able to assign tweets and compare them to the characteristics of the full set of counties in Table A4. We would be particularly concerned if counties with geolocalizable tweets were substantially different from other counties. Reassuringly, counties without localizable tweets only form a tiny minority: out of the 3106 counties in our universe, only 21 (0.7%) are not attributed any tweet.

Pre-existing Twitter usage and instrument. For the study of mechanisms, we use a proxy of pre-existing Twitter usage measured in December 2019. This is measured by sampling all tweets containing the word "the" during random intervals in one week of December 2019. One million tweets were collected from 765 000 users. Users were attributed to counties using the location in their profile. To study causally the effect of pre-existing Twitter usage on the reaction to COVID-19, we collected data to reproduce the SXSW instrument used by Müller and Schwarz (2020): we collected in November 2021 the locations of all 639 915 followers of the @SXSW Twitter account as well as the date they joined the network.

BLM account followers. As an additional outcome, we use the number of all followers of the official BLM account @Blklivesmatter. We collected the followers and their geolocation in February 2022. This gap between the period of analysis and the date of data

[31] These keywords are considered both in when appearing separated with space, or without spaces as a hashtag (e.g. #BlackLivesMatter)

collection can lead to measurement error because we do not know the starting date of following. Accounts that followed the official BLM account may stop following it and accounts that are computed as followers may start following just a few hours before the collection. Similarly, geolocation of accounts may have changed between the period of study and the date of data collection. Using this data we also compute the number of accounts created between the first COVID-19 death in the USA and the 24th of May (the day before the murder of George Floyd) that are followers of the account @Blklivesmatter.

Google mobility. We use data on mobility, collected through mobile phones that use Google apps (such as Google Maps). This data collects information on the time a person spent on certain mobility tasks like the time spent in parks, being at home, doing groceries, in the transit stations and finally at their workplace (as identified by Google). This information is then aggregated at the county level to measure the aggregate daily mobility.

Google searches. We also use the Google Trends data to analyze patterns of search activity before and after the death of George Floyd. Each variable is a normalized index of search activity for a given search term. The indices are specified on a Nielsen's Designated Market Area (DMA) level. A DMA is a region of the United States that consists of counties and ZIP-codes. There are 210 DMA regions covering the US. Search activity is averaged across the period of interest: each observation is a number of the searches of the given term divided by the total searches of the geography and time range, which is then normalized between regions such that the region with the largest measure is set to 100. The important limitation of the Google Trends data is that an index of search activity is an integer from zero to one hundred with an unreported privacy threshold. The search terms that were used in the analysis are presented in Table A8.

SafeGraph. We rely on two data sets provided by SafeGraph. Both of them are based on anonymized mobile data. SafeGraph aggregates data from around 45 million smartphones on the level of US Census Block Groups. With the help of the first data set, Monthly Patterns (MP), we can answer such questions as: who visited each «point of interest», where they came from and where they go to. The set of «points of interests» consists of millions of places such as hotels, restaurants, public parks, malls and other establishments. The MP data allows us to observe home locations at the level of the US Census Block Group, which we can use to construct our variable of touristic flows out of spring break locations in March 2020. In our alternative identification strategy we employ an instrumental variable based on data provided by the data company SafeGraph. The SafeGraph data is GPS location data that reveal the spatial mobility of population between the points of interest. For the region of interest (three vacation destinations in Florida: Miami Beach, Panama Beach and Fort Lauderdale) the SafeGraph data provide rich set of points of interest, which include more than 3000 places such as restaurants, bars, hotels, gyms, public parks, malls and other establishments. Using this data, we measure the number of devices that "pinged" in each of the point of interest during March, 2020. Using these data we can also observe home locations on the level of the US Census Block Groups (CBG). An individual "home" is defined as a place where user's devices pinged most often in the night time between 6 PM and 7 AM during the baseline 6-week period determined by the SafeGraph.

Elephrame. Elephrame is a crowd-sourced platform that collects data on Black Lives Matter and other protests. It provides information on the place and date of each BLM protest and estimated number of participants, as well as a link to a news article covering the protest. We extracted all protest records from June 2014 to September 2020 and geo-coded their location. The observation period starts with the first BLM demonstration for Eric Garner on 7/19/2014 and consists of any public demonstration or public art installation focused on "communicating the value of a Black individual or Black people as a whole". Each observation is manually collected by the creator of Elephrame, Alisa Robinson, from sources that include press, protest organizers, participants and observers.

Lockdown stringency. We use data from the Oxford COVID-19 Government Response Tracker (Hale et al., 2020) to measure the restrictiveness of the government's pandemic policy. Use of this data is inspired by recent work which shows that stringent policies lead to lower mortality, mobility and consequently spread of infection during the pandemic (Jinjarak et al., 2020; Askitas et al., 2020). This data provides four key indices (i) an overall government response index, (ii) a containment health index, (iii) an economic support index, and (iv) an original stringency index which captures the strictness of lockdown-style policies. Each of this indices reports values between 1 and 100 and varies across states and weeks.

Additional county-level controls. We include unemployment data available on a monthly basis at the county level from the Local Area Unemployment Statistics of the US Bureau of Labor Statistics and the total population, population by ethnicity, income statistics (such as Black poverty rate and median household income (all in 2018), as well as past Republican vote share (in 2012 and 2016) from the American Community Survey. We use a dummy for rural counties which is constructed from the Office of Management and Budget's February 2013 delineation of metropolitan and micropolitan statistical areas.[32] The measure of social capital that we use aggregates the information on the number of local organizations.[33] In addition, we include an index of county resilience towards a pandemic provided by the US Census bureau, which incorporates health and infrastructure indicator and is described in more detail next.

Community resilience. One of the most important COVID-19 related control variables used in our empirical analysis is the ability of counties to cope with the pandemic. This variable comes from the United States Census Bureau. These estimates measure the capacity of individuals and households to absorb, endure, and recover from the health, social, and economic impacts of a disaster such as a hurricane or a pandemic. For each county the population living under each of 11 risk factors is estimated and these factors are aggregated into 3 composite risk factors: (i) population with 0 risk factors; (ii) population with 1-2 risk factors, and (iii) population with 3 or more risk factors. These risk factors are based on households' and individuals' socio-economic and health conditions. Risk factors include: Income-to-Poverty Ratio, single or zero caregiver household, unit-level crowding defined as $>$ 0.75 persons per room, communication barriers (defined as either limited English-speaking households or no one in the household over the age of 16 with a high school diploma), no one in the household is employed full-time, disability posing constraint to significant life activity, no health insurance coverage, being aged 65 years or older, households without a vehicle and households without broadband Internet access. For

[32] 2013 NCHS Urban-Rural Classification Scheme for Counties, Vintage 2012 postcensal estimates of the resident U.S. population. NCHS Urbanization levels are designed to be convenient for studying the difference in health across urban and rural ares. This classification has 6 categories: large "center" metropolitan area (*inner cities*), large "fringe" metropolitan area (*suburbs*), median metropolitan area, small metropolitan area, micropolitan area and non-core (nonmetropolitan counties that are not in a micropolitan area).

[33] This includes: (a) civic organizations; (b) bowling centers; (c) golf clubs; (d) fitness centers; (e) sports organizations; (f) religious organizations; (g) political organizations; (h) labor organizations; (i) business organizations; and (j) professional organizations.

our analysis we look at populations within each county that are classified as living under 1-2 risk factors and 3 or more risk factors.

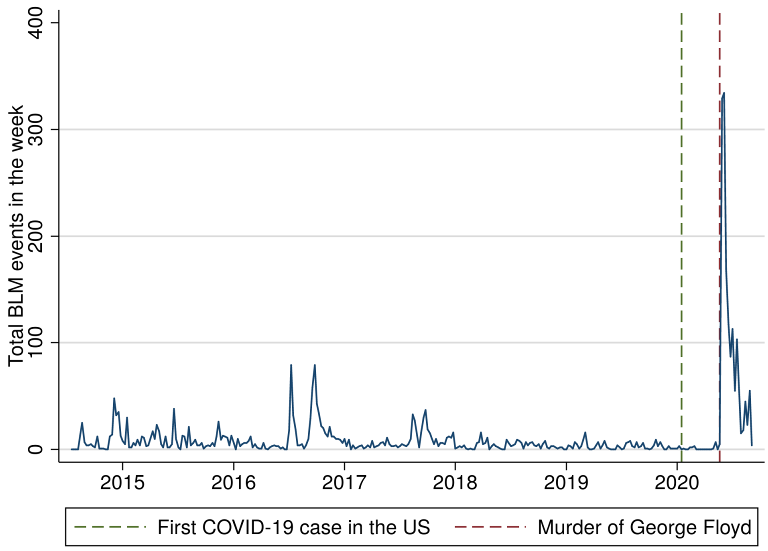
Notable deaths. We collect data on all notable Black deaths that have occurred in the country since 2014. Notable deaths are defined as deaths of Blacks at the hands of a police officer and which are covered in national media and/or have a dedicated Wikipedia page. This data set includes personal information about the victim like their name, age, sex and race. It also has details about the event, like the county and zip code of the place where shooting took place, cause of death, whether the victim was armed, if a video of the incidence was taken by onlookers and if the police officer wore a body camera. We also collect information on date of the shooting, date of the official verdict from this incident and whether the police officer was convicted. From 2014 till 2020, we have 34 notable deaths from all over the country. The list of notable deaths can be found in Table A6. Average age of victim is 34 years, 31 out of 34 are men. All victims in our data are Black.

Use of deadly force by police. We obtain this from the collaborative platform Fatal Encounters. This data is collected by a multi-disciplinary team at the University of Southern California. The results are published as part of the *National Officer-Involved Homicide Database*. The data is available from 2000 onward and contains the name, gender, race, and age of each victim and the specific address where the death occurred, among other variables.

The American Trends Panel survey by Pew Research Center. To zoom in and move from county to individual level analysis, we employ the American Trends Panel survey (ATP), conducted by Pew Research Center. The panel is based on a representative sample of U.S. adults who participate via self-administrated online web-survey. Participants with no internet access were provided with tablets and wireless connection to answer the survey, which is crucial for studying the effects of social media. For our analysis we drawn from the panel wave 68 that took place from June 4 to June 10, 2020. Participants were questioned on a wide range of topics, including Black Lives Matter movement, police brutality, ideologies and politicians, race relations, social issues, coronavirus and president Donald Trump. The survey also contains a group of demographic variables, which are included in our analysis as controls: race, age, gender, education, income, political leanings, level of urbanisation of participant's region. It is important to note that the participants' region of residence is anonymised, therefore the exact data on COVID cases and deaths is not available. However, the panel does include the aggregate version of this data: on which decile of COVID prevalence the respondent's region is in. We can make only associative conclusions based on this limited information.

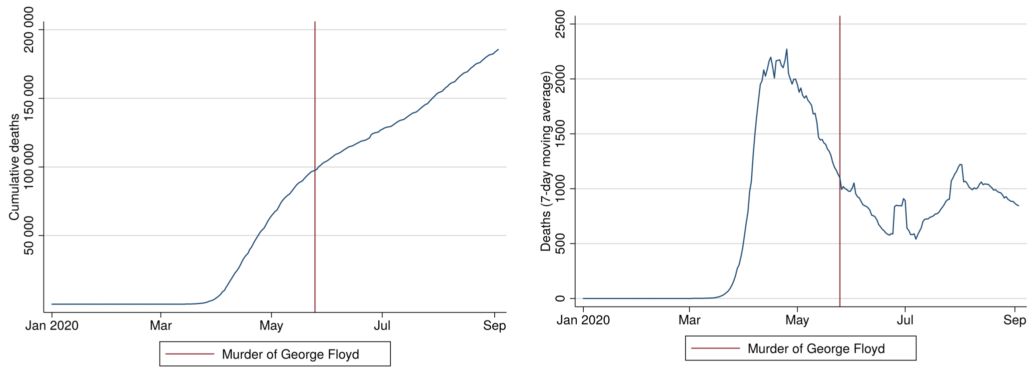
A.2 Data Tables and Figures

Figure A1: BLM events over time



Note: Number of BLM events per week in the US from June 2014 to September 2020. The green vertical line denotes the week of the first confirmed COVID-19 case in the US (January 21, 2020), and the red vertical line denotes the week of the murder of George Floyd (May 25, 2020).

Figure A2: COVID-19 deaths and timing of George Floyd's murder



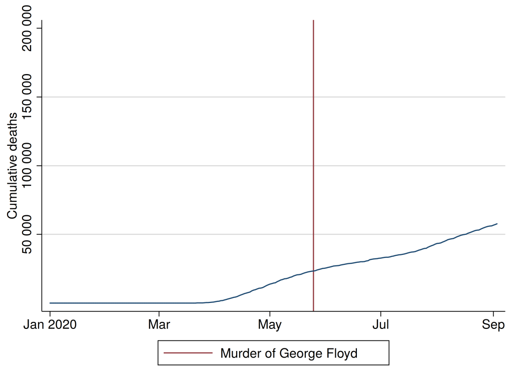
(a) Cumulative deaths

(b) New deaths

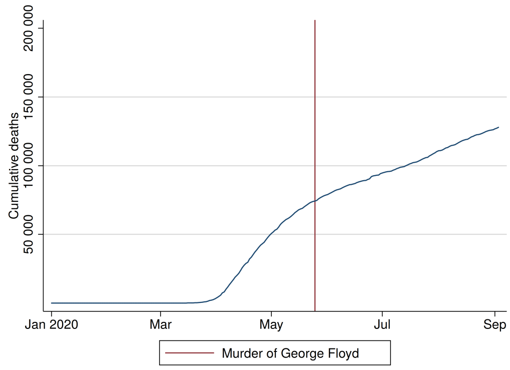
Note: Number of cumulative COVID-19 deaths and daily new COVID-19 deaths in the US between January and September 2020. New COVID-19 deaths are presented as a 7-day moving average. The red vertical line denotes the day of the murder of George Floyd (May 25, 2020).

Figure A3: Evolution of COVID-19 pandemic per sub-sample

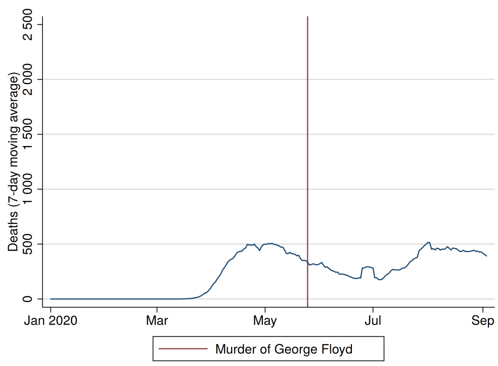
(a) Cumulative deaths. Sub-sample: counties with no BLM before



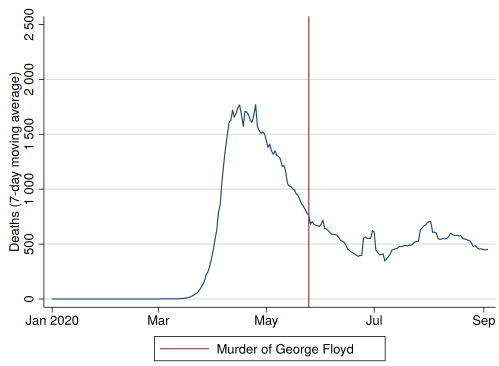
(b) Cumulative deaths. Sub-sample: counties with previous BLM events



(c) New deaths per week. Sub-sample: counties with no BLM before

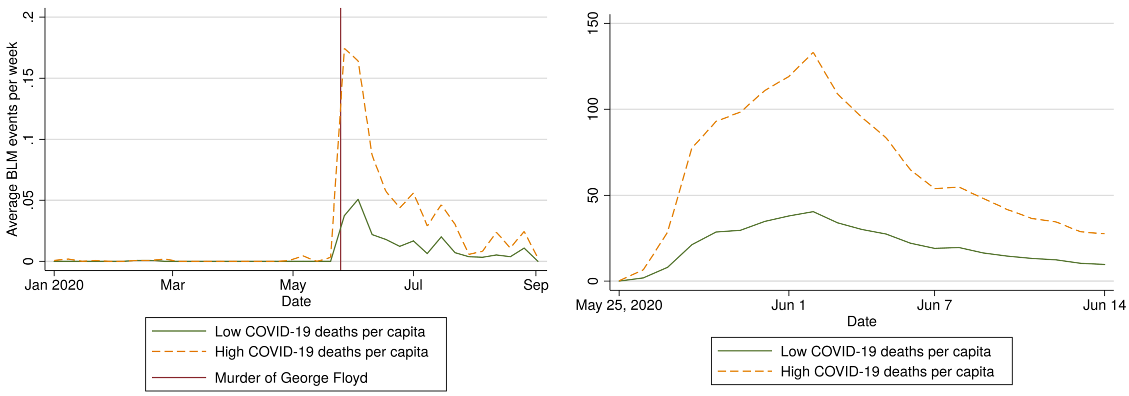


(d) New deaths per week. Sub-sample: counties with previous BLM events



Note: Evolution of two measures of COVID-19 pandemic from January to September 2020 in different sub-samples. Sub-figures (a) and (b) show cumulative deaths. Sub-figures (c) and (d) show new deaths per week. Sub-figures (a) and (c) correspond to the sub-sample of counties with no prior BLM events and sub-figures (b) and (d) correspond to the sub-sample of counties that had BLM events previous to the murder of George Floyd.

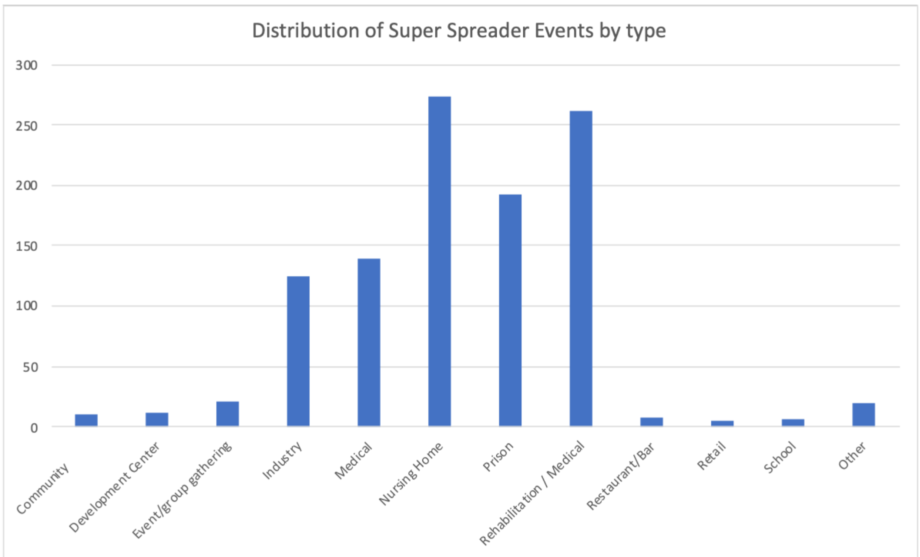
Figure A4: BLM events and tweets in counties with above and below median COVID-19 deaths per-capita



(a) Average BLM protests per week (b) Average tweets mentioning BLM per day

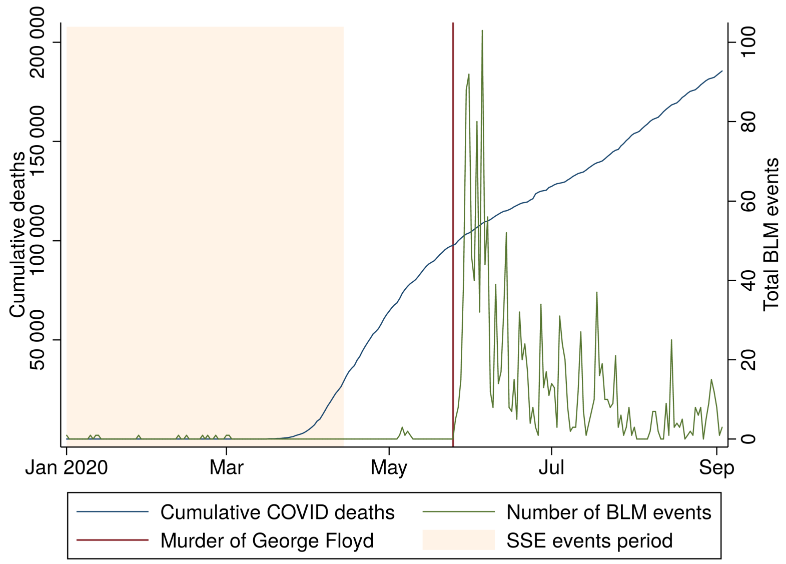
Note: Evolution of two variables over time in counties with below and above median COVID-19 deaths per capita. Sub-figure (a) presents the average number of BLM protests per week between January and September 2020. The red vertical line represents the day of the murder of George Floyd (May 25, 2020). Sub-figure (b) presents the average number of daily tweets mentioning "BLM" or "Black Lives Matter" from May 25 to June 14, 2020.

Figure A5: Distribution of super-spreader events in the US by their type



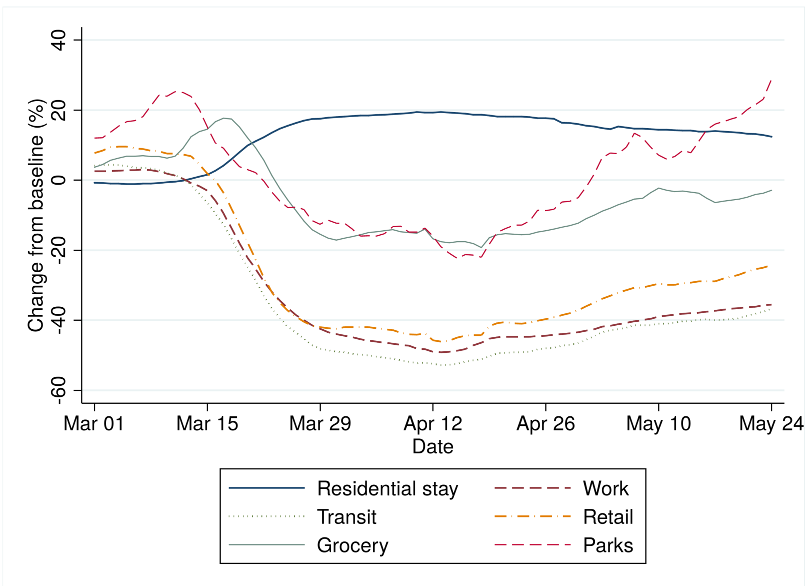
Note: Histogram showing the distribution of the types of place where super spreader events took place.

Figure A6: Timing of SSEs relative to Floyd's murder, protest and COVID-19 deaths



Note: Plot showing the evolution per day of various variables of interest. Blue line show cumulative COVID-19 deaths and BLM events per day from January to September 2020 as reported by the New York Times. The green line shows the evolution of the number of BLM events. The red vertical line denotes the week of the murder of George Floyd (May 25, 2020), and the orange shaded area is the period we consider for super-spreader events.

Figure A7: Evolution of mobility index



Note: This graph represents the components of the Google Community Mobility index: residential stay, and mobility to different types of places, between March 1st and May 24th, 2020. The index is relative to the average mobility to these places in the same day of the week between January 3 and February 6, 2020. The displayed value is an average of the 7 previous days.

Table A1: Description of variables and data sources

Variable	Source	Start of period	End of period
BLM related variables			
Presence of BLM events	Elephrame	June 2014	September 2020
Total participants	Elephrame	June 2014	September 2020
Participants per event	Elephrame	June 2014	September 2020
Number of events	Elephrame	June 2014	September 2020
Use of deadly force	Fatal Encounters	January 2000	January 2022
Notable deaths	Self-collected	July 2014	May 2020
COVID-19			
COVID deaths (per 1000)	NYTimes data	January 2020	May 2020
SSEs event	SSEs Database	January 2020	May 2020
Spring breakers (Visits per device)	SafeGraph data	March 2020	
Black death burden	CDC	January 2020	May 2020
Stringency	OxCGRT	January 2020	May 2020
Internet and Social Media			
New Twitter accounts	Twitter API	5 weeks before George Floyd's death	
Predicted new Twitter accounts	Twitter API	5 weeks before George Floyd's death	
Google searches for Twitter	Google Trends	5 weeks before George Floyd's death	
Google BLM search	Google Trends	5 weeks before George Floyd's death	
Demography, Income, Unemployment and Votes			
Urban	2013 NCHS Urban-Rural Classification	2013	
Rural areas	2013 NCHS Urban-Rural Classification	2013	
Suburban areas	2013 NCHS Urban-Rural Classification	2013	
Not large cities	2013 NCHS Urban-Rural Classification	2013	
Black population	American Community Survey	2018	
Non-black population share	American Community Survey	2018	
White population share	American Community Survey	2018	
Non-white population share	American Community Survey	2018	
Black poverty	American Community Survey	2018	
Median hh income	American Community Survey	2018	
3+ risk factors	U.S. Census	2019	
Unemployment	Local Area Unemployment Statistics	January 2020	May 2020
Republican vote shares	American Community Survey	2012, 2016	
Social capital	Rupasingha et al. (2006), (2014)	2014	
Residential stay	Google mobility report	January 2020	May 2020

Note: Description of source and the period of collection (i.e. the period during which the data is measured that we include for our analysis but not the period at which the data was collected).

Table A2: Summary statistics for super spreading events by their type

Type of SSE event	Total events	Total Events 6 weeks before GF's murder	Mean	Standard Deviation	Total Cases
Community	11	9	1.364	0.505	504
Development Center	12	12	3.833	1.404	1612
Event/group gathering	21	13	3	1.549	1083
Industry	125	87	15.656	8.642	17825
Medical	140	134	36.586	17.037	13731
Nursing Home	273	261	80.597	37.073	26684
Prison	193	187	45.487	19.674	49747
Rehabilitation / Medical	262	251	89.618	41.009	26979
Restaurant/Bar	8	4	1.5	0.535	1306
Retail	5	0	1	0	68
School	7	2	1.286	0.488	218
Other	20	15	2.5	1.051	1592

Note: All SSE in the USA by their type. Total events are total number of SSE event of each type occurring till 29 August. Total Events 6 weeks before George Floyd's murder is sum of all SSE events by their type that occurred 6 weeks before George Floyd's death. Total cases is sum of all reported COVID-19 positive cases attributed to each type of SSE event.

Table A3: Example Tweets

Date	Text
May 29, 2020	While #BlackLivesMatter is raising awareness on Twitter, it shouldn't stop there. While you're inside with your families, talk about racism and discrimination. Especially with older generations who don't use social media and don't see further than the national new's portrayal.
May 30, 2020	This is called UNITY. this is what white america doesn't want. they're afraid of the non racist whites to form patternship and unity with POC bc then they will be out numbered. I stand by my brothers #BlackLivesMatter https://t.co/EPYE9HKkBN
May 31, 2020	Reach out to black friends, peers, and social media connections to LISTEN to them with the understanding that I do not know what their struggles are like as a person that has lived with privilege. #BlackLivesMatter
Jun 2, 2020	If it weren't for Twitter and social media the videos of George Floyd and Ahmed Arbery would have not been seen and murderers would have walked free. Fact. #BlackLivesMatter
Jun 4, 2020	(3/7) We will also be sharing courses made by the Arist community designed to educate allies. The first example: https://bit.ly/anti-racism101 This 20-day text message course will teach you about systemic racism against Black people and how you can practice anti-racist allyship.#BlackLivesMatter
Jun 4, 2020	I made a decision when I came on twitter to keep it strictly for work. I have other social media for expressing personal and political views. However, given the events of the last week, I feel compelled to say something - so here is my bit #BlackLivesMatter #WhitePrivilege
Jun 6, 2020	#IAmASuburbanMom and Black Lives Matter to me! I just went to a rally in a suburb of Atlanta, and there are a lot of us moms who want racial justice and change!
Jun 7, 2020	White privlage means you CAN walk away from #BlackLivesMatter when you get weary and you go back to your regular routine. Our black and coloured allies don't have that privilege to simply walk away. It's their life. Recognizing our white privilege means refusing to walk away. 3/3
Jun 11, 2020	1/ I've been trying to learn more about all the complexities of everything going on lately, and how to be a better ally, better support the #blacklivesmatter movement & simply be an anti-racist. For what it's worth, here's a few things I've found to be especially helpful:
Jun 13, 2020	There was a #BlackLivesMatter car parade in my VERY white, VERY red suburban San Antonio neighborhood today. I was afraid we'd be the only car. There were 50 of us!!!

Note: Example of tweets containing #BLM or #BlackLivesMatter between May 25th and June 14

Table A4: Summary statistics - counties with and without geo-localized tweets

	All counties					Counties with tweets					Counties without tweets				
From 25th of May to 14th of June 2020:	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Presence of BLM events	3106	0.099	0.298	0.000	1.000	3085	0.100	0.299	0.000	1.000	21	0.000	0.000	0.000	0.000
Number of BLM events	3106	0.250	1.348	0.000	36.000	3085	0.252	1.352	0.000	36.000	21	0.000	0.000	0.000	0.000
Participants in BLM events	3106	270.759	5968.521	0.000	323687.500	3085	272.602	5988.765	0.000	323687.500	21	0.000	0.000	0.000	0.000
Participants per event	307	539.141	878.429	0.000	8991.319	307	539.141	878.429	0.000	8991.319	0	.	.	.	.
Tweets mentioning BLM	3106	819.502	7187.496	0.000	243596.000	3085	825.005	7211.615	0.000	243596.000	21	11.048	27.558	0.000	124.000
New users tweeting about BLM	3106	4.586	53.812	0.000	2442.000	3085	4.616	53.994	0.000	2442.000	21	0.143	0.478	0.000	2.000
Tweets mentioning #AllLivesMatter	3106	134.741	833.066	0.000	28943.000	3085	135.658	835.824	0.000	28943.000	21	0.000	0.000	0.000	0.000
Tweets mentioning #BlueLivesMatter	3106	17.753	113.478	0.000	4117.000	3085	17.874	113.854	0.000	4117.000	21	0.000	0.000	0.000	0.000
Neighbor protested first	3106	0.348	0.477	0.000	1.000	3085	0.350	0.477	0.000	1.000	21	0.143	0.359	0.000	1.000
On the 25th of May 2020:															
COVID deaths (total)	3106	24.461	141.132	0.000	3304.000	3085	24.627	141.597	0.000	3304.000	21	0.095	0.301	0.000	1.000
COVID cases (total)	3106	459.678	2438.202	0.000	72010.000	3085	462.751	2446.204	0.000	72010.000	21	8.190	18.878	0.000	66.000
COVID deaths (per 1000)	3106	0.113	0.248	0.000	2.935	3085	0.113	0.249	0.000	2.935	21	0.016	0.052	0.000	0.182
COVID cases (per 1000)	3106	2.791	5.664	0.000	145.513	3085	2.801	5.676	0.000	145.513	21	1.340	3.365	0.000	15.273
Superspreader events, 6+ weeks ago, neighboring	3106	3.070	9.790	0.000	143.000	3085	3.083	9.816	0.000	143.000	21	1.238	4.392	0.000	20.000
Black death burden	3106	1.346	0.963	0.000	4.104	3085	1.350	0.960	0.000	4.104	21	0.682	1.094	0.000	4.104
Lockdown stringency index	3106	68.445	8.508	47.220	89.810	3085	68.502	8.459	47.220	89.810	21	60.007	11.416	49.070	78.700
Before the 25th of May 2020:															
Google searches for Twitter	3056	61.265	11.222	17.000	100.000	3037	61.317	11.217	17.000	100.000	19	53.053	9.132	41.000	65.000
Residential stay	1348	10.633	3.387	3.600	26.286	1347	10.633	3.388	3.600	26.286	1	11.000	.	11.000	11.000
Later outcomes:															
Followers of @BlkLivesMatter	3106	63.198	495.174	0.000	20058.000	3085	63.628	496.830	0.000	20058.000	21	0.000	0.000	0.000	0.000
Followers of @BlkLivesMatter created during the pandemic	3106	1.540	11.207	0.000	453.000	3085	1.550	11.245	0.000	453.000	21	0.000	0.000	0.000	0.000
Street art count	3106	0.703	26.735	0.000	1467.000	3085	0.708	26.825	0.000	1467.000	21	0.000	0.000	0.000	0.000
County characteristics:															
Black police-related deaths (2014-2019)	3106	0.677	3.207	0.000	84.000	3085	0.682	3.217	0.000	84.000	21	0.048	0.218	0.000	1.000
Black police-related deaths (2020)	3106	0.047	0.301	0.000	6.000	3085	0.047	0.302	0.000	6.000	21	0.000	0.000	0.000	0.000
Unemployment rate (year average)	3106	4.691	1.550	0.708	19.650	3085	4.692	1.545	0.708	19.650	21	4.456	2.275	1.642	10.267
Black population share	3106	0.100	0.147	0.000	0.875	3085	0.100	0.147	0.000	0.875	21	0.058	0.156	0.000	0.602
Non-white population share	3106	0.144	0.162	0.000	0.928	3085	0.144	0.160	0.000	0.882	21	0.242	0.302	0.001	0.928
Large cities	3106	0.020	0.140	0.000	1.000	3085	0.020	0.140	0.000	1.000	21	0.000	0.000	0.000	0.000
Suburban areas	3106	0.118	0.323	0.000	1.000	3085	0.119	0.324	0.000	1.000	21	0.048	0.218	0.000	1.000
Smaller towns	3106	0.234	0.423	0.000	1.000	3085	0.235	0.424	0.000	1.000	21	0.095	0.301	0.000	1.000
Rural areas	3106	0.628	0.483	0.000	1.000	3085	0.626	0.484	0.000	1.000	21	0.857	0.359	0.000	1.000
BLM events (2014-2019)	3106	0.617	4.183	0.000	117.000	3085	0.621	4.197	0.000	117.000	21	0.000	0.000	0.000	0.000
Black poverty rate	3106	0.281	0.225	0.000	1.000	3085	0.281	0.224	0.000	1.000	21	0.239	0.370	0.000	1.000
Population share with 3+ risk factors	3106	25.899	5.019	10.685	48.448	3085	25.886	5.017	10.685	48.448	21	27.900	5.028	19.106	39.176
Vote share for republicans (2016)	3106	0.633	0.156	0.083	0.960	3085	0.633	0.156	0.084	0.946	21	0.678	0.214	0.083	0.960
Vote share for republicans (2012)	3106	0.596	0.148	0.060	0.959	3085	0.596	0.147	0.092	0.959	21	0.612	0.203	0.060	0.857
Median household income (2016)	3106	48795.991	13277.575	20170.891	129150.343	3085	48823.327	13286.361	20170.891	129150.343	21	44780.231	11492.168	24306.000	66196.000
Social capital	3106	1.384	0.705	0.000	6.887	3085	1.384	0.698	0.000	6.887	21	1.464	1.424	0.000	6.055
Distance to Minneapolis	3106	1216.679	555.825	11.998	6474.706	3085	1218.933	555.758	11.998	6474.706	21	885.570	468.940	324.576	1898.310
Notable Deaths	3106	0.010	0.116	0.000	3.000	3085	0.010	0.116	0.000	3.000	21	0.000	0.000	0.000	0.000
Log(SXSW followers created before March 2017)	3106	0.114	0.258	0.000	1.474	3085	0.115	0.259	0.000	1.474	21	0.033	0.151	0.000	0.693
Log(SXSW followers created during March 2017)	3106	0.193	0.350	0.000	1.658	3085	0.194	0.351	0.000	1.658	21	0.033	0.151	0.000	0.693

Note: Descriptive statistics of all variables by different sub-samples depending on the presence of tweets containing #BLM or #BlackLivesMatter between May 25th and June 14. Different panels correspond to different moments at which each variable was measured.

Table A5: Summary statistics - counties with and without prior BLM event

	All counties					No BLM event before					Has BLM event before				
From 25th of May to 14th of June 2020:	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Presence of BLM events	3106	0.099	0.298	0.000	1.000	2768	0.048	0.213	0.000	1.000	338	0.518	0.500	0.000	1.000
Number of BLM events	3106	0.250	1.348	0.000	36.000	2768	0.064	0.322	0.000	5.000	338	1.778	3.642	0.000	36.000
Participants in BLM events	3106	270.759	5968.521	0.000	323687.500	2768	21.026	172.090	0.000	5500.000	338	2315.911	17979.700	0.000	323687.500
Participants per event	307	539.141	878.429	0.000	8991.319	132	355.115	621.538	0.000	5500.000	175	677.949	1010.498	0.000	8991.319
Tweets mentioning BLM	3106	819.502	7187.496	0.000	243596.000	2768	322.706	4818.930	0.000	243596.000	338	4887.938	16330.354	0.000	177550.000
New users tweeting about BLM	3106	4.586	53.812	0.000	2442.000	2768	1.809	25.763	0.000	1306.000	338	27.331	143.697	0.000	2442.000
Followers of @BlkLivesMatter created during the pandemic	3106	1.540	11.207	0.000	453.000	2768	0.441	2.167	0.000	78.000	338	10.536	32.057	0.000	453.000
Tweets mentioning #AllLivesMatter	3106	134.741	833.066	0.000	28943.000	2768	47.488	326.063	0.000	15659.000	338	849.290	2224.119	0.000	28943.000
Tweets mentioning #BlueLivesMatter	3106	17.753	113.478	0.000	4117.000	2768	6.125	36.206	0.000	1647.000	338	112.976	312.535	0.000	4117.000
Neighbor protested first	3106	0.348	0.477	0.000	1.000	2768	0.334	0.472	0.000	1.000	338	0.464	0.499	0.000	1.000
On the 25th of May 2020:															
COVID deaths (total)	3106	24.461	141.132	0.000	3304.000	2768	8.366	46.396	0.000	1025.000	338	156.266	382.483	0.000	3304.000
COVID cases (total)	3106	459.678	2438.202	0.000	72010.000	2768	164.485	663.300	0.000	15169.000	338	2877.112	6677.133	0.000	72010.000
COVID deaths (per 1000)	3106	0.113	0.248	0.000	2.935	2768	0.099	0.230	0.000	2.935	338	0.224	0.345	0.000	2.010
COVID cases (per 1000)	3106	2.791	5.664	0.000	145.513	2768	2.596	5.662	0.000	145.513	338	4.391	5.430	0.000	40.048
Superspreader events, 6+ weeks ago, neighboring	3106	3.070	9.790	0.000	143.000	2768	2.327	7.564	0.000	143.000	338	9.154	19.279	0.000	140.000
Black death burden	3106	1.346	0.963	0.000	4.104	2768	1.344	0.985	0.000	4.104	338	1.363	0.755	0.000	4.104
Lockdown stringency index	3106	68.445	8.508	47.220	89.810	2768	68.210	8.543	47.220	89.810	338	70.367	7.969	47.220	89.810
Before the 25th of May 2020:															
Google searches for Twitter	3056	61.265	11.222	17.000	100.000	2731	60.523	10.941	17.000	100.000	325	67.505	11.623	41.000	100.000
Residential stay	1348	10.633	3.387	3.600	26.286	1023	10.006	3.016	3.600	26.286	325	12.606	3.722	4.429	26.143
Later outcomes:															
Followers of @BlkLivesMatter	3106	63.198	495.174	0.000	20058.000	2768	16.186	94.110	0.000	3304.000	338	448.195	1421.137	0.000	20058.000
County characteristics:															
Black police-related deaths (2014-2019)	3106	0.677	3.207	0.000	84.000	2768	0.207	0.724	0.000	15.000	338	4.527	8.589	0.000	84.000
Black police-related deaths (2020)	3106	0.047	0.301	0.000	6.000	2768	0.014	0.131	0.000	3.000	338	0.314	0.783	0.000	6.000
Unemployment rate (year average)	3106	4.691	1.550	0.708	19.650	2768	4.713	1.575	0.708	17.442	338	4.506	1.323	2.492	19.650
Black population share	3106	0.100	0.147	0.000	0.875	2768	0.093	0.146	0.000	0.875	338	0.157	0.142	0.009	0.727
Non-white population share	3106	0.144	0.162	0.000	0.928	2768	0.134	0.160	0.000	0.928	338	0.231	0.150	0.014	0.801
Large cities	3106	0.020	0.140	0.000	1.000	2768	0.001	0.027	0.000	1.000	338	0.178	0.383	0.000	1.000
Suburban areas	3106	0.118	0.323	0.000	1.000	2768	0.105	0.307	0.000	1.000	338	0.228	0.420	0.000	1.000
Smaller towns	3106	0.234	0.423	0.000	1.000	2768	0.201	0.400	0.000	1.000	338	0.506	0.501	0.000	1.000
Rural areas	3106	0.628	0.483	0.000	1.000	2768	0.694	0.461	0.000	1.000	338	0.089	0.285	0.000	1.000
BLM events (2014-2019)	3106	0.617	4.183	0.000	117.000	2768	0.000	0.000	0.000	0.000	338	5.672	11.510	0.000	117.000
Black poverty rate	3106	0.281	0.225	0.000	1.000	2768	0.283	0.236	0.000	1.000	338	0.263	0.099	0.000	0.600
Population share with 3+ risk factors	3106	25.899	5.019	10.685	48.448	2768	25.957	5.066	10.685	48.448	338	25.428	4.600	11.763	39.453
Vote share for republicans (2016)	3106	0.633	0.156	0.083	0.960	2768	0.656	0.141	0.083	0.960	338	0.446	0.143	0.084	0.818
Vote share for republicans (2012)	3106	0.596	0.148	0.060	0.959	2768	0.614	0.140	0.060	0.959	338	0.456	0.131	0.092	0.823
Median household income (2016)	3106	48795.991	13277.575	20170.891	129150.343	2768	47521.697	12362.349	20170.891	129150.343	338	59231.630	15713.952	28625.618	120936.813
Social capital	3106	1.384	0.705	0.000	6.887	2768	1.426	0.726	0.000	6.887	338	1.037	0.336	0.334	2.744
Distance to Minneapolis	3106	1216.679	555.825	11.998	6474.706	2768	1192.851	538.680	34.438	6474.706	338	1411.818	648.904	11.998	6395.519
Notable Deaths	3106	0.010	0.116	0.000	3.000	2768	0.001	0.033	0.000	1.000	338	0.080	0.331	0.000	3.000
Log(Twitter users observed in 2019)	3106	2.034	1.496	0.000	8.806	2768	1.739	1.215	0.000	8.093	338	4.443	1.401	0.000	8.806
Log(SXSW followers created before March 2017)	3106	0.114	0.258	0.000	1.474	2768	0.090	0.228	0.000	1.427	338	0.311	0.378	0.000	1.474
Log(SXSW followers created during March 2017)	3106	0.193	0.350	0.000	1.658	2768	0.157	0.312	0.000	1.658	338	0.489	0.482	0.000	1.619

Note: Descriptive statistics of all variables for different sub-samples depending on the presence or absence of BLM protesting history before the murder of George Floyd. Different panels correspond to different periods where different data is measured.

Table A6: Notable deaths: names of victims of police brutality that received national coverage

Name	Date	Location
Rayshard Brooks	6/12/2020	Atlanta, GA
George Floyd	5/25/2020	Minneapolis, MN
Breonna Taylor	3/13/2020	Louisville, KY
Atatiana Koquice Jefferson	10/12/2019	Fort Worth, TX
Botham Jean	9/6/2018	Dallas, TX
O'Shae Terry	8/1/2018	Arlington, TX
Antwon Rose II	6/19/2018	East Pittsburgh, PA
Stephon Clark	3/18/2018	Sacramento, CA
Jordan Edwards	4/29/2017	Balch Springs, TX
Keith Lamont Scott	9/20/2016	Charlotte, NC
Terence Crutcher	9/16/2016	Tulsa, OK
Paul O'Neal	7/28/2016	Chicago, IL
Philando Castile	7/6/2016	Falcon Heights, MN
Alton Sterling	7/5/2016	Baton Rouge, LA
Akiel Rakim Lakeith Denkins	2/29/2016	Raleigh, NC
Greg Gunn	2/25/2016	Montgomery, AL
Jamar Clark	11/15/2015	Minneapolis, MN
Ricky Javenta Ball	10/16/2015	Columbus, MS
Jeremy McDole	9/23/2015	Wilmington, DE
Christian Taylor	8/7/2015	Arlington, TX
Samuel Dubose	7/19/2015	Cincinnati, OH
Sandra Bland	7/13/2015	Hempstead, TX
Brendon Glenn	5/5/2015	Los Angeles, CA
William Chapman	4/22/2015	Portsmouth, VA
Freddie Gray	4/19/2015	Baltimore, MD
Walter Scott	4/4/2015	North Charleston, SC
Eric Courtney Harris	4/2/2015	Tulsa, OK
Tony Terrell Robinson	3/6/2015	Madison, WI
Rumain Brisbon	12/2/2014	Phoenix, AZ
Tamir E. Rice	11/22/2014	Cleveland, OH
Akai Gurley	11/20/2014	New York, NY
Laquan McDonald	10/20/2014	Chicago, IL
Michael Brown	8/9/2014	Ferguson, MO
Eric Garner	7/17/2014	New York, NY

Note: We collect data on all notable Black deaths that have occurred in the country since 2014. Notable deaths are defined as deaths of Blacks at the hands of a police officer and which are covered in national media and/or have a dedicated Wikipedia page.

Table A7: Principal component analysis of online presence

(a) Correlation between measures

	New Twitter accounts	New Twitter accounts (log)	Google searches for Twitter	Residential stay
New Twitter accounts	1			
New Twitter accounts (log)	0.379 [0.000]	1		
Google searches for Twitter	0.0558 [0.042]	0.234 [0.000]	1	
Residential stay	0.0770 [0.005]	0.355 [0.000]	0.520 [0.000]	1

p-values in brackets

(b) Principal components

	Eigenvalue	Difference	Proportion	Cumulative
PC1	1.845167	.7217762	0.4613	0.4613
PC2	1.123391	.5386709	0.2808	0.7421
PC3	.5847198	.137997	0.1462	0.8883
PC4	.4467228	.	0.1117	1.0000

(c) Factor loadings

	PC1	PC2	PC3	PC4
New Twitter accounts	.3265664	.7302487	.5601388	.2152575
New Twitter accounts (log)	.5339498	.3756387	-.6544048	-.3815068
Google searches for Twitter	.5247868	-.4478813	.484595	-.5377442
Residential stay	.5769323	-.3536025	-.1522054	.7203804

Note: The first table reports the correlation between the online presence measures. The second table reports the eigenvalues of the six principal components. The third table reports the loading of the different components.

Table A8: Search terms used in indices of search activity

Keywords	Start of period	End of period	Duration
Twitter	2020-01-01	2020-05-25	6 months
Twitter	2020-04-20	2020-05-25	5 weeks
BLM	2020-05-25	2020-06-15	3 weeks
Floyd	2020-05-25	2020-06-15	3 weeks
George Floyd	2020-05-25	2020-06-15	3 weeks
BLM + Black Lives Matter + Floyd + George Floyd	2020-05-25	2020-06-15	3 weeks
BLM + Black Lives Matter + Floyd + George Floyd	2020-04-20	2020-05-25	5 weeks

Note: The Google Trends data is generated on a designated market area (DMA) level. Keywords are case-independent. The resulting outcomes are normalised measures generated by Google Trends.

Appendix B: Robustness Checks

Our robustness checks focus on two dimensions: *i*) robustness to changes in the definition and construction of our instrumental variable and *ii*) robustness of our main results to sample composition, spatial correlation and other confounding factors. We present our results in Tables B2 to B5.

B.1 Instrument Robustness and Validity

We present results on the robustness of the instrument in Tables B2 and B3, showing the IV result and first stage coefficient. Our baseline result is always reported in column 1 for reference.

Changing the radius around SSEs. In the baseline specification, we choose the 50km threshold as a distance of the SSE to the county border, as it is approximately two times the average radius of a county in the US.[34] To make sure that this choice is not driving our results, we change the radius of influence to 25 km, 100 km and 200 km (columns 2, 3 and 4 of Table B2 respectively). For all distances the coefficient remains significant and becomes slightly larger in magnitude.

Changing the time window of SSEs. Similarly, in our preferred specification, we take into account the SSEs that occurred in a specific time window that we call "window of opportunity" where there were enough cases to observe SSEs and the social distancing measures were not applied strictly or widely enough. Specifically, we count the number of SSEs between the beginning of the COVID-19 outbreak until April 13th 2020 (e.g., six weeks before Floyd's murder). In columns 6 to 8 of Table B2 we expand and narrow this window to make sure our results are not driven by the specific timing of SSEs. In particular, we count SSEs until April 20th, 5 weeks before the murder of Floyd (column 6), on April 6th, 7 weeks before (column 7) and on March 30th, 8 weeks before (column 8). Results are robust to change in the time window.

Excluding SSEs in prisons. A non-negligible number of SSEs occurred inside prisons. We exclude SSEs in prisons in a robustness check in column 2 of Table B3 for two reasons. First, it is likely that by the nature of prisons, the geographical spread of cases stemming from an SSE in a prison is quite limited and less relevant for the overall population and the protesting population. In this case, we would expect a bigger effect when excluding these SSEs. Second, SSEs in prisons may have an effect on BLM protests other than through overall exposure to COVID, for instance, by raising the salience of the overproportional incarceration of Black people. In this case, we would expect the coefficient to decrease in magnitude when excluding these SSEs. While the salience of racial inequality in prisons may be a possible mechanism, with this exercise we investigate whether our results are indeed solely driven by this subsample of SSEs. We exclude SSEs in prisons in column 2 and find that our results slightly increase in magnitude and precision.

Controlling for SSEs in the county. Our first stage compares the effect of having an SSE outside the county within 50 km of the county border and excluding the effect of SSEs that take place within its border. Therefore, in our analysis a county is "not affected" by an SSE if its border is either further than 50 km from the SSE, or the SSE happened within its

[34]For reference, the average radius of a county is 28 km and the average radius of a state is 220 km.

boundaries. We expect the effect of SSEs to be different between these groups: presumably, counties far away will have no COVID-19 cases from this SSE, while the county where the SSE took place will have a lot of cases and deaths caused by the event. To assuage the concern that correlation of SSEs across counties is driving the variation in SSE exposure, we add as a control the number of SSEs that occurred within the county itself. Estimates are presented in column 3 of Table B3 and show that the results of the baseline specification are robust to the addition of this control.

Weighting SSEs by distance. In our baseline specification, we count any SSE that occurred in a 50 km radius outside the border of a county as an additional SSE affecting the county. However, an SSE 1 km away from the border is likely to have a different level of influence from a SSE 49 km away. To ensure that this simplification is not driving the results, we refine the level of influence in three different ways. First we weight the SSEs by a linear function decreasing with distance (column 4 of Table B3), giving less weight to events that are more distant. Second, we repeat the analysis but with a quadratic function (column 5 of Table B3), weighting distant events less and increasingly so. The results are robust to these distance weighting procedures.

Weighting SSEs by the inverse probability of occurrence. The probability of being near a county that has an SSE is not constant over all counties. For instance, counties neighboring cities have likely a higher probability of being treated by our instrument as their neighbors may be more likely to experience an SSE. This could be a violation of the exclusion restriction because the probability of being treated by our instrument at a certain level is not uniform, and this heterogeneity could be related to certain county characteristics that could in turn be related to the probability of protesting. To address this concern, we weight each observation by the inverse probability of being treated. Using LASSO (a regularized regression procedure that performs variable selection and avoids overfitting, Tibshirani 1996), we select relevant variables predicting (by a logit model) the probability of having a neighbor with an SSE among a set of county characteristics, including a large set of socio-demographic and economic characteristics extracted from the American Community Survey (such as population, population density, race distribution, age groups, poverty rates, among others), indicators for different levels of urbanization, geographical indications (latitude, longitude, and state dummies), as well as the minimum and maximum of these variables for neighboring counties. We use the LASSO selected model to predict the probability of a county having a neighbor with an SSE, then weight the observations by the inverse of this probability. This means that counties with a higher probability of having a neighbor with an SSE that actually had a neighbor with an SSE are weighted less than counties with a lower probability of being treated that are actually treated. Estimates are presented in column 7 of Table B2 and show that our results are robust to this weighting procedure.

Plausibility of exclusion restriction for BLM. If our instrument were to pick up any underlying factors correlated with the overall likelihood of protesting for a BLM-related cause, then this would challenge a causal interpretation of our estimates. To probe the plausibility of the exclusion restriction, we estimate the effect of instrumented COVID-19 on the likelihood of observing past BLM protests, using as a sample the set of all counties instead of counties where no protests has been observed before George Floyd's murder. If our instrument were correlated with the county unobservables that also predict the likelihood of observing BLM protests, then we would expect to see a statistically significant relationship between our instrumented COVID-19 and likelihood of observing a BLM protest in the past. In column 2 of Table B4, we show that exposure to COVID-19 does not predict the presence of BLM events between 2014 and 2019. We take this as additional evidence for the plausibility of our identifying assumption.

Plausibility of exclusion restriction for Twitter. We repeat a similar exercise for Twitter penetration. It is possible that SSEs are correlated with unobserved factors that drive both COVID-19 and Twitter use. For instance, counties that experience SSEs may also generally be more sociable and therefore also use more social media and have more SSEs at the same time. In order to test this, we look at the effect of SSEs on past Twitter use in December of 2019 (we describe the construction of this variable in the previous section). In Table B6, we show that SSEs do not predict past Twitter use, assuaging concerns about the violation of the exclusion restriction in the context of social media use.

B.2 Robustness of Main Results

In this section, we focus on our main results and run robustness checks including changing definitions in treatment and outcome, estimation method, spatial correlation and concerns about the overall propensity to protest. We present these checks in Tables B4 and B5.

Excluding coastal counties and states. Coastal states and counties might behave differently, either with regard to our instrument or to the process of COVID-19 contagion. Coastal regions are generally denser, which increases the chance of having an SSE (Figure 4 shows the density of SSEs). Coastal counties also differ in the construction of the instrumental variable. As defined by our IV, we positively label those counties that have SSEs in neighboring counties. Coastal counties naturally have fewer land neighbors, which decreases its chances of being treated. Traditionally, these counties have also had higher BLM protest activity, so it is also instructive to obtain results for less active landlocked counties. We show that our results are robust to excluding coastal counties (column 3 of Table B4), as well as coastal states (column 4).

COVID-19 cases. In our baseline specification we use the number of COVID-19 deaths per thousand in the county as an explanatory variable for protest. It is possible that COVID-19 deaths may have a different or distinct effect on BLM protest. This could be due to - for instance - different threat perceptions or salience of the pandemic. In column 8 of Table B4, we show that the results hold when using the number of COVID-19 related cases instead of the number of deaths. As expected, the number of COVID-19 related cases exhibits significantly smaller coefficients but continues to significantly and positively affect protest behavior.

Probit estimation. In our baseline specification the effect of COVID-19 is additive. It might be the case that the effect would be multiplicative of some characteristics of the counties. Using a Probit model accounts for this possibility. Non-linear models with many covariates (typically when using fixed effects) suffer from the incidental parameter problem resulting in bias of the estimates (Heckman, 1987; Lancaster, 2000; Wooldridge, 2015). To reduce the extent of this problem we omit the state fixed effects, which significantly reduces the number of covariates. We use an OLS in the first stage, but estimate the second stage with a Probit model. Results are presented in column 2 of Table B5. The Probit model delivers larger and more precisely estimated coefficients.

Controlling for propensity to protest. We add the propensity to protest that we constructed for our matching-based alternative identification (the construction of this propensity measure is detailed in Appendix C.3) as a control in our regression. It yields a continuous measure of ex-ante protest probability even for the sub-set of counties with no prior BLM protest. We first use it directly as a control (column 3 of Table B5). Our results remain robust and are more precisely estimated.

In addition, we include fixed effects for different levels of the propensity to protest. We group observations by groups of 1000, 100 and 10 units with similar propensity to protest and add fixed effects for each group. Results are shown in columns 4 to 6 of Table B5. This is essentially a matching-like strategy, where the fixed effects ensure that observations with similar propensity are compared. Results are robust to the inclusion of fixed effects.

Accounting for spatial correlation. Observations are likely to be spatially correlated for several reasons. For instance, there could be spatially-correlated unobserved factors influencing the decision to protest (such as weather conditions or available TV and radio stations). Clustering by state does not entirely remove these errors because correlation across state borders remains (Colella et al., 2019). To overcome this problem, we use Conley standard errors that allow for spatial correlation within a certain distance. Column 7 of Table B5 shows the estimates when allowing spatial correlation between observations in a 50 km radius. Column 8 of Table B5 shows the estimates when allowing spatial correlation with all neighboring counties. Reassuringly, our results remain robust.

Estimation without clustering. Our preferred specification clusters at the state level and includes state fixed effects (Abadie et al., 2017). Column 9 of Table B5 shows our baseline results when we do not cluster the standard errors.

Table B1: First stage: Super spreading events in neighbouring counties on COVID-19 deaths per 1000 population

	COVID (deaths/1000)		
	(1)	(2)	(3)
First stage			
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.00986***	0.00881***	0.00751***
	(0.000909)	(0.00132)	(0.00144)
Observations	2,768	2,767	2,767
KP F-stat	115.10	44.53	27.04
County controls			Y
State fixed effects		Y	Y

Note: Estimation of the effect of SSE in neighbouring counties (50km radius) six weeks prior to George Floyd's murder on COVID-19 deaths. The sample consists of counties with no BLM protest before George Floyd's murder (and excluding Florida). In column 1 we do not include any control. In column 2 (column 3 respectively) we include state fixed effects (state fixed effects and county-level controls). Control variables include: the share of Black population, urban (category [1-6]), median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience, Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations) and deadly force used by police against Black people. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B2: Instrument robustness - SSE timing and distance

	Presence of BLM event						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IV: COVID (deaths/1000)	0.404**	0.503*	0.499**	0.503**	0.383**	0.440**	0.410**
	(0.187)	(0.266)	(0.191)	(0.225)	(0.188)	(0.193)	(0.187)
First stage coefficient:	0.00751***	0.0126***	0.00304***	0.000901***	0.00738***	0.00770***	0.00926***
	(0.00144)	(0.00331)	(0.000309)	(0.000272)	(0.00139)	(0.00154)	(0.00170)
Observations	2,767	2,767	2,767	2,767	2,767	2,767	2,767
F first stage	27.04	14.40	97.13	10.95	28.12	24.87	29.78
Mean of dep. var.	0.0477	0.0477	0.0477	0.0477	0.0477	0.0477	0.0477
Distance	50 km	25 km	100 km	200 km	50 km	50 km	50 km
Lag	6 weeks	6 weeks	6 weeks	6 weeks	5 weeks	7 weeks	8 weeks
All controls	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y

Note: Results on robustness of the SSE instrument to different time and distance selection. The sample consists of counties with no BLM protest before George Floyd's murder. Column 1 corresponds to our baseline specification. Columns 2 to 4 vary the distance at which SSE are counted from 25 to 200km. Columns 5 to 7 vary the time lag between the murder of Floyd and the last SSE, going back 5, 7 or 8 weeks. All specifications include the whole set of controls and state fixed effects. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table B3: Instrument robustness - SSE definition and weighting

	Presence of BLM event					
	(1)	(2)	(3)	(4)	(5)	(6)
IV: COVID (deaths/1000)	0.404**	0.482**	0.401*	0.515**	0.581**	0.363*
	(0.187)	(0.180)	(0.232)	(0.217)	(0.266)	(0.195)
First stage coefficient:	0.00751***	0.00798***	0.00653***	0.0178***	0.0239***	0.00781***
	(0.00144)	(0.00159)	(0.00130)	(0.00448)	(0.00637)	(0.00164)
Observations	2,767	2,767	2,767	2,767	2,767	2,766
F first stage	27.04	25.03	25.08	15.73	14.09	22.55
Mean of dep. var.	0.0477	0.0477	0.0477	0.0477	0.0477	0.0494
Excluding SSEs in prisons		Y				
Control SSE in county			Y			
Measure				linear	square	
SSE probability weighting						Y
All controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Results on the robustness of the SSE instrument. The sample consists of counties with no BLM protest before George Floyd's murder. Column 1 corresponds to our baseline specification. Column 2 excludes SSEs that took place in prisons. In column 3, a control is added for the number of SSEs within the county 6 weeks before the murder of George Floyd. Columns 4 and 5 weigh the effect of SSEs by distance with smaller weights given to more distant SSEs. Weights are applied linearly (column 5), or quadratically (column 6). In column 6, observations are weighted by the inverse probability of observing a SSE affecting the county if a SSE is observed, no SSE if no SSE is observed. All specifications include the whole set of controls and state fixed effects. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table B4: Robustness of main results - sample composition and variable definition

	Presence of BLM events					
	3 weeks	Past events	3 weeks	3 weeks	3 weeks	3 weeks
	(1)	(2)	(3)	(4)	(5)	(6)
IV: COVID (deaths/1000)	0.404**	-0.0523	0.531**	0.344*		0.451**
	(0.187)	(0.215)	(0.219)	(0.173)		(0.207)
IV: COVID (cases/1000)					0.0312***	
					(0.0116)	
Observations	2,767	3,106	2,616	1,697	2,767	2,767
F first stage	27.04	35.73	45.85	157.2	7.181	30.39
Mean of dep. var.	0.0477	0.108	0.0428	0.0371	0.0477	0.0477
Excluding coastal			counties	states		
All controls	Y	Y	Y	Y	Y	Y
COVID deaths in past 7 days						Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Results on the robustness of our main results. The sample consists of counties with no BLM protest before George Floyd's murder, except for column 2 where the sample consists of all counties. Column 1 correspond to our baseline specification. Column 2 predicts past BLM events (likelihood of observing a BLM event between 2014 and 2019) with (instrumented) COVID-19 deaths just before the murder of George Floyd, and uses all counties instead of only the counties with no BLM protest before George Floyd's murder. Columns 3 and 4 exclude coastal counties and states. Column 5 looks at the effect of COVID-19 cases instead of deaths. Column 6 includes as an additional control the number of new COVID-19 related deaths in the 7 days leading up to Floyd's murder. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table B5: Robustness of main results - estimation method, protest propensity, spatial correlation

	Presence of BLM events								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IV: COVID (deaths/1000)	0.404**		0.423**	0.405**	0.341*	0.338*	0.404*	0.404**	0.404***
	(0.187)		(0.185)	(0.184)	(0.186)	(0.182)	(0.234)	(0.205)	(0.128)
IV Probit: COVID (deaths/1000)		0.878***							
		(0.230)							
Observations	2,767	2,767	2,663	2,767	2,767	2,767	2,767	2,767	2,767
F statistic	27.04	127.7	28.56	28.20	26.64	27.08			72.09
Mean of dep. var.	0.0990	0.0992	0.102	0.0990	0.0990	0.0990	0.0990	0.0990	0.0990
All controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Propensity to protest			Y						
Propensity to protest group: size				1000	100	10			
State clustering	Y	Y	Y	Y	Y	Y			
Spatial clustering							50 km	neighbors	
State fixed effects	Y		Y	Y	Y	Y	Y	Y	Y

Note: Results on the robustness of our main results. The sample consists of counties with no BLM protest before George Floyd's murder. Column 1 correspond to our baseline specification. Column 2 estimates the second stage with an IV Probit model (with an OLS in the first stage) and omits state fixed-effects. Column 3 adds a control for the propensity to protest derived from our LASSO selection model. Columns 4 to 6 add fixed effects for propensity to protest for groups of size 1000, 100 and 10 respectively. Column 7 and 8 replace the state clustering by spatial clustering, allowing correlation in a 50 km radius for column 7, and between neighbors for column 8. Columns 9 omits clustering altogether. All specifications include the whole set of controls, except column 2 where state fixed effects are removed. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level except for columns 7 to 9. *** p<0.01, ** p<0.05, * p<0.1

Table B6: Prediction of Twitter presence from COVID SSE instrument

Outcome:	Log(Preexisting users)		
	(1)	(2)	(3)
Subsamble:	All counties	No BLM events before	Has BLM event before
Cumulative SSE 6 weeks ago, not in county, less than 50km away	0.00134	0.00422	-0.00252
	(0.00379)	(0.00517)	(0.00447)
Observations	3,106	2,767	333
Mean of dep. var	2.034	1.738	4.439
All controls	Y	Y	Y
State fixed effects	Y	Y	Y

Note: This tables shows a regression of pre-existing Twitter presence on the SSE instrument. Column 1 shows the result for all counties, column 2 for the sub-sample of counties that did not hold a BLM protest before the murder of George Floyd, and column 3 for the sub-samples of counties that did. All specifications include state fixed effects and all standard controls. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table B7: Effect of pre-existing Twitter users on new users during the pandemic

Outcome:	Log(New users)					
	(1)	(2)	(3)	(4)	(5)	(6)
Subsample:	All counties		No BLM events before		Has BLM event before	
Panel A: Preexisting users						
Log(Preexisting users)	0.349***	0.510***	0.296***	0.517***	0.676***	-0.391
	(0.0208)	(0.120)	(0.0182)	(0.136)	(0.105)	(4.616)
Observations	3,106	3,106	2,767	2,767	333	333
Mean of dep. var	0.586	0.586	0.420	0.420	1.931	1.931
F users		13.44		13.02		0.119
Panel B: Interaction with COVID						
COVID (deaths/1000)	-0.148	0.222	-0.657	-0.696	-0.236	-2.364
	(0.455)	(1.128)	(0.614)	(0.772)	(1.068)	(9.646)
Log(Preexisting users)	0.332***	0.463***	0.256***	0.413***	0.674***	-0.201
	(0.0208)	(0.0950)	(0.0196)	(0.108)	(0.118)	(2.417)
COVID × Log(Preexisting users)	0.161	0.0806	0.450 ***	0.437***	0.00732	0.370
	(0.102)	(0.248)	(0.128)	(0.159)	(0.209)	(1.916)
Observations	3,106	3,106	2,767	2,767	333	333
Mean of dep. var	0.586	0.586	0.420	0.420	1.931	1.931
F COVID	22.91	8.863	11.35	8.530	22.99	0.859
F interaction	29.68	6.054	11.35	8.530	30.31	0.750
F Twitter		18.70		19.31		0.309
All controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y
SXSW instrument		Y		Y		Y

Note: This table shows the effect of pre-existing Twitter users on new Twitter users during the pandemic. Column 1 and 2 show the result for all counties, column 3 and 4 for the sub-sample of counties that did not hold a BLM protest before the murder of George Floyd, and column 5 and 7 for the sub-samples of counties that did. Panel A present the simple estimates while panel B shows the interaction with (instrumented) COVID. Columns 1, 3, and 5 use the uninstrumented pre-existing Twitter users, and columns 2, 4 and 6 use an IV estimation using the SXSW instrument. The first stage is given in Table C4. All specifications include state fixed effects and all standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Appendix C: Other Estimation Strategies

C.1 Alternative Instrument: Florida Spring Break

In our preferred empirical strategy, we chose smaller and decentralized SSEs to argue for a causal relationship between COVID-19 and BLM protests. Here, we add another cross-sectional instrumental variable: the spatial distribution of touristic flows originating in major Florida Spring Break destinations during March of 2020. Instead of collecting information on multiple independent SSEs as in the previous section, we now focus on one single, large-scale event that is known to have contributed substantially to the spread of COVID-19 (Mangrum and Niekamp, 2020).

Despite the fact that COVID-19 infections had surged in Florida's main spring break destinations and despite the fact that the Center for Disease Control had issued multiple warnings, Florida Governor DeSantis failed to implement social distancing orders until April 1st 2020[35]. We exploit this unique, large scale event to track the diffusion of COVID-19 infections that originated in Florida during spring break and then spread across the United States. To track these movements we benefit from rich data on cell phone mobility provided by SafeGraph. We can identify spring breakers' home counties – locations that they most likely returned to after vacationing in highly infectious spring break locations.

Specifically, we pick three Florida vacation destinations: Miami Beach, Panama Beach and Fort Lauderdale. In early March these three destinations caught the attention of the media, which reported high level of COVID cases among tourists who did not respect social distancing measures (BBC, CNN). We use anonymized mobile data for the period from March, 1, 2020 to April 1, 2020, covering the majority of spring break periods across the country. With the help of the Monthly Patterns data (MP), we measure unique devices that visited specific «points of interest» in one of three popular spring break destinations.

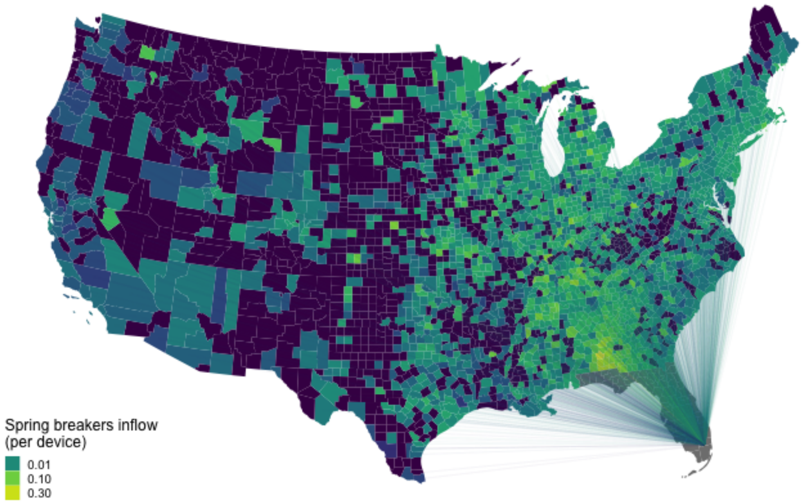
The SafeGraph data provides us with a rich set of points of interest, which include more than 3000 places such as restaurants, bars, hotels, gyms, public parks, malls and other establishments. Using this data, we measure the number of devices that «pinged» in each point of interest during March, 2020. The MP data also allows us to observe home locations on the level of the US Census Block Groups (CBG). An individual "home" is defined as a place where a user's devices pinged most often in the night time between 6 PM and 7 AM during the baseline 6-week period determined by SafeGraph.

Using this information, we calculate the number of unique visitors to points of interest in three cities in Florida and group this number by device home counties. Given that cell phone data is anonymized, each device is counted as many times as it has visited different places (such as restaurants and shops) in a given tourist destination. Therefore, this measure captures both intensity of tourism flow from the county and mobility of these tourists during their spring break. Since higher mobility is associated with higher chances of disease contraction, our variable captures both extensive and intensive margins of COVID-19 spread. We see this variable as an improvement over ones used in literature examining stay at home behaviour (Abouk and Heydari (2020); Lasry et al. (2020); Friedson et al. (2020); Dave et al. (2020); Dave et al. (2021)). The exposure to COVID-19 is therefore instrumented by the number of spring-break tourists.

$$Z_c = \frac{\sum_{POIs} pings_{POI,c}}{devices_c} \tag{4}$$

[35]Local officials had started to close some of the beaches for public access in mid March

Figure C1: Spring Breakers by US counties. Number of visitors per device in home county



We normalise this variable calculating a ratio of the total number of devices detected in spring breakers' home counties at March 1, 2020 to account for differences in population size and differences in resident device coverage between counties in the SafeGraph data. Figure C1 shows our resulting measure of "spring breakers" inflow split into five categories: high flow, moderate-high flow, moderate-low flow, low flow, no flow (missing).

We use the same set of controls and connotations as in our baseline cross-sectional estimation. Our estimating equation is:

$$BLM_c = \beta_0 + \beta_1 \widehat{Covid}_{cs} + \mathbf{X}_c \beta_{\mathbf{X}} + \delta_s + \epsilon_{cs}$$

We present our 2SLS results in Table C1. We use the same set of controls as in the previous cross-sectional estimations, including socio-economic, demographic and political control variables. We also follow the same approach as in our primary identification and restrict sample to counties that had no BLM protests before the murder of George Floyd additionally excluding the state of Florida. For all specifications in the Table C1 we find a positive and statistically significant coefficient for COVID-19 on the presence of a BLM event and the first stage is sufficiently strong.

Table C1: Spring breakers IV: Covid-19 deaths on the presence of BLM events, 2SLS, counties with no prior BLM events

	BLM Protest				
	(1)	(2)	(3)	(4)	(5)
Panel A: 2SLS estimates					
COVID	0.265**	0.774***	0.583*	0.598	0.631*
(deaths/1000)	(0.110)	(0.278)	(0.315)	(0.371)	(0.379)
Panel B: reduced form estimates					
Spring Break Tourists	0.419**	0.647***	0.387**	0.398*	0.420*
	(0.170)	(0.187)	(0.184)	(0.227)	(0.231)
Panel C: OLS estimates					
COVID	0.0668***	0.0510***	0.0385**	0.0744***	0.0769***
(deaths/1000)	(0.0175)	(0.0190)	(0.0193)	(0.0238)	(0.0242)
Observations	2,719	2,718	2,717	2,717	2,717
F-stat	74.56	19.35	13.05	13.05	13.05
County controls			Y	Y	Y
State fixed effects		Y	Y	Y	Y
Time frame	3 weeks	3 weeks	3 weeks	6 weeks	9 weeks

Note: estimation of the effect of the COVID-19 pandemic on the presence of BLM protests. The sample consists of counties with no BLM protest before George Floyd's murder (and excluding Florida). The top panel reports 2SLS results, using the number of Florida visitor's "pings" per number of devices in county during March 2020 as an instrument. Panel B presents the reduced form estimates and Panel C, the corresponding OLS results. In columns 1 - 3, we define the outcome as the presence of at least one BLM event during the three weeks following the murder of George Floyd. In columns 4 and 5, we use 6 and 9 weeks, respectively. In column 1 we do not include any control. In column 2 (column 3 respectively) we include state fixed effects (state fixed effects and county-level controls). Control variables include: the share of Black population, urban (category [1-6]), median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience, Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations) and deadly force used by police against Black people. We report Cragg-Donald Wald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

C.2 Difference in Differences: Notable Deaths Sample

We use data on BLM at the county-week level starting in 2014 and exploit differences in protest behavior following what we call a "notable" death. Deaths of Black people at the hands of the police have been a trigger for BLM protests across the country. Roughly, more than 300 Black people die each year in the US either due to police brutality or under police custody. However, not all of these deaths result in media coverage, which is crucial for generating public discourse or action. Many of these events only received public traction since they were - mostly by chance - recorded on a phone camera. We construct a data set of all police-related Black deaths since July 2014 that were covered in a major national daily newspaper like the Washington Post, received TV coverage by CNN and/or has a dedicated Wikipedia page.

We exploit the variation at the county-week level by interacting our main COVID-19 variable with a dummy variable for a notable death occurring in a certain week. We use information on BLM protests in counties in the 3 weeks after the recorded notable death (we can reduce this to 2 weeks and expand it to 4 weeks without significantly changing the first and second-stage results). With this strategy, we alleviate concerns about time-varying state-level unobserved heterogeneity. Following a difference in differences logic, we then look at whether the reaction following this trigger differs in counties that were more exposed to the COVID-19 pandemic. Again, we use the SSE IV to account for the fact that COVID-19 exposure may be endogenous to past and present protest behavior. We estimate the below model on all counties in the US.[36]

$$Covid_{ct} = \zeta_0 + \zeta_1 Notable\_deaths_{cst} + \zeta_2 Z_{cst} + \zeta_3 Notable\_deaths \times Z_{cst} + \mathbf{X}_{cs}\zeta_{\mathbf{X}} + \gamma_c + \theta_{st} + \eta_{cst}, \quad (5)$$

$$Z_{cst} = \sum SSE_{cst}^{neighbor} \quad (6)$$

The second stage is written as:

$$BLM_{cst} = \beta_0 + \beta_1 Notable\_deaths_{cst} + \beta_2 \widehat{Covid}_{cst} + \beta_3 Notable\_deaths_{cst} \times \widehat{Covid}_{cst} + \mathbf{X}_{cs}\zeta_{\mathbf{X}} + \mu_c + \delta_{st} + \epsilon_{cst}$$

where, $Notable\_deaths_{cst}$ is a dummy variable that takes the value of one in the three weeks following a nationally covered death and zero otherwise. We include county and state-week fixed effects, as well as all police-related deaths of Black people at the county level. This is a crucial control as it allows us to exploit the "extra" trigger that nationally covered deaths create, above and beyond the local level of deadly force used by local police. The key coefficient of interest is β_3 which is the difference in differences estimator.

Table C2 shows the results of this estimation. Columns 1 and 3 report the effect of notable deaths up to 4 weeks after they occurred and columns 2 and 4 report for up to 3 weeks. In both cases we find that the effect of notable deaths in predicting the likelihood of observing a BLM protest is significantly higher in counties with higher pandemic exposure. In columns 3 and 4, we include county-specific time trends to account for unobserved county-specific time trend. The coefficient of Notable Deaths is negative and significant. This implies that the effect of notable deaths on the probability of observing a BLM event

[36] Unlike the main analysis, this estimation relies on the pre-trends of BLM protests in the US, therefore, we include all counties in this estimation.

reduces for counties with average exposure to COVID-19 deaths.

It is important to mention, particularly in the light of new literature on generalized difference in differences - especially the designs that use two-way fixed effects like our estimation model - that the underlying assumption for causal interpretation of β_3 is that the effect of treatment, which in our case is the occurrence of a notable death, is homogeneous across space and time (Roth et al., 2022; De Chaisemartin and d'Haultfoeuille, 2020; Marcus and Sant'Anna, 2021; Butts and Gardner, 2021). The assumption of a homogeneous effect of notable deaths relies on the fact the occurrence of these deaths is random and their location and time cannot be predicted. Therefore, each county of the country has an equally likely probability of being affected by this. While the exposure to COVID-19 is staggered in time across the USA, in this estimation we assume all counties to be equally exposed to the COVID-19 pandemic since it broke out in the US in January 2020.[37] Additionally, in table C2 we estimate robust difference in difference estimator using Butts and Gardner (2021) two-stage estimator which adjusts for negative weighting problem discussed above. As expected our result remains robust to this estimation.

[37]We test for the negative weights in our estimation using the methodology proposed by De Chaisemartin and d'Haultfoeuille (2020) and find that there are no negative weights in our estimator.

Table C2: Notable Deaths Regression

	(1)	(2)	(3)	(4)
	Presence of BLM			
Notable deaths × Covid deaths	1.4926***	2.0714***	1.4935***	2.0707***
	(0.1053)	(0.1095)	(0.1057)	(0.1102)
Covid deaths per thousand	0.0595***	0.0597***	0.0450***	0.0451***
	(0.0166)	(0.0166)	(0.0116)	(0.0116)
Notable deaths	-0.0389***	-0.0391***	-0.0410***	-0.0412***
	(0.0125)	(0.0128)	(0.0127)	(0.0130)
Black police-related deaths	Y	Y	Y	Y
Unemployment	Y	Y	Y	Y
Weeks post Notable Death	4	3	4	3
County FE	Y	Y	Y	Y
State-Week FE	Y	Y		
County Week Trend			Y	Y
Observations	96286	96286	96329	96329
F First Stage (COVID)	18.03	17.92	32.23	32.09
F First Stage (Interaction)	13.05	13.87	14.59	14.97
Two Stage DID			1.4290***	
			(0.2642)	

Note: Estimation of the effect of Notable deaths and COVID-19 deaths on occurrence of BLM events after each notable death. This table presents 2SLS results, using the cumulative number of all super-spreader events in neighbouring counties (50km radius) as an instrument for COVID-19 deaths. Columns (1) and (3) presents the effect of instrumented cumulative number of COVID-19 deaths and notable deaths on the likelihood of having a BLM event in the county within 4 weeks of the notable death. Column (2) and (4) presents the effect of instrumented cumulative number of COVID-19 deaths and notable deaths on the likelihood of having a BLM event in the county within 3 weeks of the notable death. All specifications include county fixed effects and two time varying controls (the number of black police-related deaths and the unemployment rate both at a county level) along with either state-week fixed effects or county week time trend to increase precision. Weekly data by county from year 2014 until the 14th June 2020. We provide robust difference in difference (DID) estimator using Butts and Gardner (2021) two stage DID estimator. We compute the instrumented exposure to COVID-19 by hand for this estimation. The non-instrumented coefficient is 1.0833 with a standard error of 0.0852 which is in line with our main estimation. We report Kleibergen-Paap rkWald F statistic. Standard errors clustered at the county level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

C.3 LASSO Matching: Propensity to Protest

We exploit data on past protests to predict the propensity of a county to protest in response to previous notable deaths using a large set of county characteristics.

More precisely, we start by estimating the following logit model:

$$\log \frac{\Pr(BLM_{ci} = 1)}{1 - \Pr(BLM_{ci} = 1)} = \beta_0 + \mathbf{X}_c \beta_{\mathbf{X}} + \varepsilon_{ci}$$

where BLM_{ci} is a dummy variable indicating whether there was a BLM protest in county c during the three weeks following notable death i, and $\mathbf{X}_c$ is a vector of county-level controls. The sample used for this estimation includes all counties, including those that did have a BLM protest before George Floyd's murder.

We select the most relevant subset of variables with LASSO regression (Tibshirani, 1996). This avoids overfitting and gives confidence in using the model to predict the propensity to react to another notable death. This model is estimated on the subset composed of all counties, and we compute the estimated propensity to protest for each county.

We then perform a propensity score matching-like estimation. To allow the comparison of a "treated" and "control" group, we simplify our continuous treatment into a binary treatment: counties are considered treated if they had at least one COVID-19 related death on or before May 24th. The outcome of interest is still whether these counties held a BLM protest in the 3 weeks following the murder of George Floyd. For all counties that did not protest before the murder of George Floyd, we compute the propensity score using the previously selected model. This is the same sample as our main regression, but only a subsample of the sample used to select the model: since counties in our main sample have, by definition, not experienced a BLM protest before, we need to consider all counties to estimate the propensity-to-protest model. We match counties with similar historical propensities to protest. The results are presented in Table C3. The results are positive and significant; their magnitude is not comparable with our main specification as the treatment is different.

Note that this is not a traditional propensity score matching (Rosenbaum and Rubin, 1983): we are not matching on the propensity to have a COVID death but on the (past) probability to hold a protest. By providing an estimate that does not rely on the SSE instrument, this analysis reinforces our confidence in our result. While this analysis controls much more extensively for the (past) propensity to protest, its estimate of the causal effect may suffer from bias from two sources. First, the propensity to protest in the past might be linked to unobservables that also predict the likelihood of having a COVID-19 death. Second, the LASSO model only predicts the propensity to protest after notable deaths before the murder of George Floyd. There may exist observable characteristics of the counties that influence the probability of treatment and protests after the death of George Floyd, but did not influence the past propensity to protest as much. One such example would be the quality of the health system: it raises both the probability of deaths from COVID, and people are likely more concerned about the quality of the health care system than they were for past protests. These characteristics would also lead to bias, since they are not accounted for in the propensity score.

Table C3: Matching on past propensity to protest

	(1)	(2)	(3)
	Presence of BLM events		
At least one COVID-19 related death	0.0440***	0.0545***	0.0533***
	(0.00867)	(0.0102)	(0.0113)
Observations	2,768	2,768	2,768
Mean of dep. var.	0.0477	0.0665	0.0795
Mean COVID	0.099	0.099	0.099
Propensity to protest	Y	Y	Y
Time frame	3 weeks	6 weeks	9 weeks

Note: Estimation of the effect of having at least one COVID-19 death on presence of BLM protests. The average treatment effect is evaluated by matching on the past propensity to protest after a notable death. The sample consists of counties with no BLM protest before George Floyd's murder. Column 1 uses as outcome the presence of a BLM protest in the 3 weeks following the murder of George Floyd, column 2, 6 weeks, and column 3, 9 weeks. Propensity-to-protest model estimated on the full sample using Logit LASSO regression using all available controls. Standard errors (in parentheses) are not clustered. *** p<0.01, ** p<0.05, * p<0.1

C.4 Baseline Twitter Penetration

This appendix details the construction of the instrument for baseline Twitter penetration used in Section 5.2. We instrument pre-pandemic Twitter penetration in December 2019 with the SXSW (South by Southwest) instrument following Müller and Schwarz (2020).

SXSW is a film, interactive media, and music festival and conference held annually in Austin, Texas. During the March 2007 edition Twitter was heavily promoted, leading to a rapid increase in the social network's popularity. Müller and Schwarz (2020) exploit this event by exploiting the fact that, through network effects, places that had more accounts created by visitors to SXSW continued to have more accounts created later on. It is not possible to directly accounts created by SXSW attendees: instead, Müller and Schwarz (2020) measure the number of followers of the official account of the festival (@SXSW) that joined Twitter during the month of the festival (March 2017). To reproduce this instrument, we collect the location of all followers of the @SXSW account of the South by Southwest festival, the date they joined Twitter, and the location set in their profile. The dataset we end up with is not entirely identical to the one used by Müller and Schwarz (2020): some users created on or before March 2007 might have started or stopped following SXSW later. They might also have changed their location between the time Müller and Schwarz collected their dataset and when we collected ours (2019 versus November 2021). Finally, our geolocation method might be different: we automatically geocode the location given by the user using Nominatim, as described in the Data section. Müller and Schwarz (2020) do not detail their geolocation method. For comparison, we attribute 52% of users to US counties (excluding imprecise locations and locations outside the US). In comparison, Fujiwara et al. (2021) (reusing this instrument) are able to locate 58% of users that joined Twitter between 2006 and 2008 using a similar method.

For each county we compute the number of followers whose account was created in March 2007 and the number of users whose account was created before this date. With our data collection and user localization strategy, we find users that follow @SXSW and joined in March 2007 in 172 counties, only 67 of which did not have BLM events before (Müller and Schwarz (2020) find 155 affected counties). To increase the number of treated counties, and thus the power of our identification, we also consider users in neighboring counties: assuming that Twitter presence diffuses, in part, geographically,[38] these counties should also have a higher number of Twitter users. We find 817 such counties, 618 of which did not have a BLM protest before.

We estimate the log number of observed Twitter users in December 2019 using the number of users that joined during SXSW controlled by the number of SXSW followers that joined before. This variable controls for the interest in SXSW festival and acts as a proxy control for the general interest in Twitter in the county. The specification is as follows:

$$
\begin{aligned}
Twitter_c = & \xi_0 + \xi_1 SXSW\ Users_{sc} + \xi_2 SXSW\ Pre\ Users_{sc} \\
& + \mathbf{X}_c \xi_{\mathbf{X}} + \gamma_s + \eta_{cs}
\end{aligned} \tag{7}
$$

where SXSW Users$_{sc}$ is the logarithm of one plus the number of SXSW followers who created their account in March 2007 in the county and neighboring counties, and Pre SXSW Users$_{sc}$ is the logarithm of one plus the number of SXSW followers in the county and neighboring counties that created their account before March 2007.

[38] This assumption is also made by Müller and Schwarz at the level of a county. Here we just extend it to neighboring counties.

The results of this first stage regression are reported for all sub samples in Appendix Table C4. For the subsample of counties without BLM events before, the coefficient of SXSW users is positive and highly significant, and the first stage is strong (F = 13.02). Unfortunately, for the subsample of counties with BLM events before George Floyd's murder, the F-statistic of the first stage is low and makes the interpretation of the results of the second stage (presented in column 5 and 6, panel B of Table D6) unreliable.

Table C4: Effect of SXSW users on Twitter presence

Outcome:	Log(Preexisting users)		
	(1)	(2)	(3)
Subsample:	All counties	No BLM events before	Has BLM event before
Log(SXSW users)	0.394***	0.373***	0.0447
	(0.108)	(0.103)	(0.130)
Log(Pre-SXSW users)	0.361***	0.382***	0.0722
	(0.0764)	(0.0896)	(0.0802)
Observations	3,106	2,767	333
Mean of dep. var	2.034	1.738	4.439
F statistic	13.44	13.02	0.119
All controls	Y	Y	Y
State fixed effects	Y	Y	Y

Note: This table shows the first stage regression for predicting existing Twitter users at the end of 2019 in the county using SXSW followers that joined Twitter during the festival in the county and its neighboring counties. Column 1 shows the result for all counties, column 2 for the sub-sample of counties that did not hold a BLM protest before the murder of George Floyd, and column 3 for the sub-samples of counties that did. All specifications include state fixed effects and all standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Appendix D: Sub-sample Analysis

Table D1: Main Result - COVID exposure and BLM protest

	Presence of BLM events				
	(1)	(2)	(3)	(4)	(5)
Panel A: ***All counties***					
2SLS estimates					
COVID	0.647***	0.730***	0.589***	0.296**	0.215*
(deaths/1000)	(0.0930)	(0.187)	(0.167)	(0.117)	(0.121)
Reduced form estimates					
Number of SSEs	0.00777***	0.00831***	0.00620***	0.00277**	0.00200
(6 weeks prior, $\leq$ 50 km distance)	0.00145)	(0.00138)	(0.00121)	(0.00120)	(0.00128)
OLS estimates					
COVID	0.203**	0.158**	0.0758*	0.0382	0.0323
(deaths/1000)	(0.0831)	(0.0638)	(0.0435)	(0.0289)	(0.0264)
Observations	3,108	3,107	3,107	3,106	3,106
F first stage	95.03	31.92	27.44	38.38	36.05
Mean dep. var.	0.0994	0.0991	0.0991	0.0988	0.0988
Panel B: ***Counties with BLM protest before***					
2SLS estimates					
COVID	0.277***	0.502**	0.386*	0.116	0.0104
(deaths/1000)	(0.0597)	(0.229)	(0.206)	(0.289)	(0.266)
Reduced form estimates					
Number of SSEs	0.00380***	0.00632***	0.00468**	0.00110	0.000100
(6 weeks prior, $\leq$ 50 km distance)	(0.000760)	(0.00213)	(0.00198)	(0.00270)	(0.00256)
OLS estimates					
COVID	0.252***	0.435***	0.224***	0.0733	0.0682
(deaths/1000)	(0.0494)	(0.0963)	(0.0740)	(0.102)	(0.102)
Observations	340	334	334	333	333
F first stage	105.3	37.56	32.01	29.27	28.09
Mean dep. var.	0.521	0.515	0.515	0.514	0.514
County controls			Y	Y	Y
State fixed effects		Y	Y	Y	Y
Time Frame	3 weeks	3 weeks	3 weeks	6 weeks	9 weeks

Note: Estimation of the effect of COVID-19 deaths per 1000 population on the presence of at least one Black Lives Matter event during the three weeks following the murder of George Floyd. Panel A presents 2SLS estimation, using number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument and OLS results for all US counties. Panel B presents these results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. Each column include sequentially different sets of additional controls. Demographic controls: share of Black population, urban (category [1-6]). Economic controls: median household income, unemployment share, Black poverty rate, 3+ risk factors/community resilience. Political controls: Republican vote share in 2012 and 2016, social capital (number of different types of civic organizations), number of past BLM events between 2014 and 2019, deadly force used by police against Black people. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D2: COVID deaths interacted with county characteristics - Counties without BLM events before

	Presence of BLM events									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample: *Counties with no BLM protest before*										
COVID (deaths/1000)	0.404**	-0.776	0.421	-0.214	1.083*	-0.0160	0.134	0.457**	0.433**	0.389
	(0.187)	(0.920)	(0.294)	(0.257)	(0.618)	(0.117)	(0.221)	(0.201)	(0.195)	(0.309)
... × Non-Black population share		1.305								
		(1.048)								
... × Non-white population share			-0.0523							
			(1.003)							
... × Median household income				0.00703**						
				(0.00294)						
... × Vote Republican 2016					-1.386					
					(1.236)					
... × Not large cities						0.430***				
						(0.0955)				
... × Suburban areas							0.333*			
							(0.176)			
... × Smaller towns								-0.320		
								(0.237)		
... × Rural areas									-0.179	
									(0.187)	
... × Probability of protest										-3.738
										(9.498)
Interacting variable		-0.143	-0.102	0.00131	-0.328*	0.227***	-0.0416	0.0701**	-0.0186	5.784***
		(0.309)	(0.0822)	(0.000856)	(0.183)	(0.0731)	(0.0284)	(0.0303)	(0.0228)	(2.038)
Observations	2,767	2,767	2,767	2,767	2,767	2,767	2,767	2,767	2,767	2,663
F COVID	27.04	15.24	11.86	45.98	14.34	114.1	18.50	23.50	12.70	14.75
F interaction		14.71	11.17	96.93	18.67	159.4	71.80	92.71	6.328	48.55
All controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 inhabitants on first-time BLM protest, interacted with county characteristics. We present results for the sub-sample of counties with no BLM protest before the murder of George Floyd. All specifications include state fixed effects and all standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D3: COVID deaths interacted with county characteristics - Counties with BLM events before

	Presence of BLM events									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample: *Counties with BLM protests before*										
COVID (deaths/1000)	0.0104	-1.368	0.187	-0.610	-0.156	-0.268	-0.115	0.0889	0.113	0.124
	(0.266)	(0.990)	(0.354)	(0.845)	(0.431)	(0.301)	(0.257)	(0.265)	(0.273)	(0.324)
... × Non-Black population share		1.662*								
		(0.986)								
... × Non-white population share			-0.401							
			(0.701)							
... × Median household income				0.00635						
				(0.00681)						
... × Vote Republican 2016					0.512					
					(0.707)					
... × Not large cities						0.324**				
						(0.140)				
... × Suburban areas							0.169			
							(0.107)			
... × Smaller towns								0.122		
								(0.141)		
... × Rural areas									0.245	
									(1.236)	
... × Probability of protest										-0.992
										(1.023)
Interacting variable		-0.645	-0.0956	0.00383	-2.513***	-0.374***	-0.0953	0.0907	-0.226*	-0.0684
		(0.427)	(0.587)	(0.00350)	(0.789)	(0.0709)	(0.109)	(0.0849)	(0.119)	(0.445)
Observations	333	333	333	333	333	333	333	333	333	333
F COVID	28.09	13.42	18.30	16.78	16.74	13.46	12.85	17.61	14.36	13.84
F interaction		15.65	39.02	21.69	13.53	84.91	63.94	27.95	1.246	17.08
All controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 inhabitants on first-time BLM protest, interacted with county characteristics. We present results for the sub-sample of counties that hosted BLM protests before the murder of George Floyd. All specifications include state fixed effects and all standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D4: Alternative outcomes

	Presence of BLM events	Number of BLM events	Total participants	Participants per event	Tweets BLM	Followers @BLM
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: ***All counties***						
IV: COVID	0.215*	0.558	1,344	200.0	2,318	186.1*
(deaths/1000)	(0.121)	(0.662)	(1,403)	(274.4)	(1,948)	(93.47)
Observations	3,106	3,106	3,106	3,106	3,106	3,106
F first stage	36.05	36.05	36.05	36.05	36.05	36.05
Mean dep. var.	0.0988	0.250	270.8	53.29	819.5	63.20
Panel C: ***Counties with BLM protest before***						
IV: COVID	0.0104	1.381	8,187*	560.5*	900.2	480.8*
(deaths/1000)	(0.266)	(1.482)	(4,380)	(308.7)	(4,604)	(275.8)
Observations	333	333	333	333	333	333
F first stage	28.09	28.09	28.09	28.09	28.09	28.09
Mean dep. var.	0.514	1.763	2276	337.1	4903	449.3
All controls	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths on different outcomes. Panel A presents 2SLS estimation, using number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument for all US counties. Panel B presents these results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. Columns 1 to 4 use protest information from *Elephrame* (described in the data section). Column 1 is the baseline result for reference. The outcome of column 2 is the number of BLM events in the three weeks following the murder of George Floyd. Column 3 reports results for the number of participants and column 4 divides the number of participants by the number of events in the county (imputing zero to counties with no BLM event). Column 5 reports the number of geo-located tweets that use at least one of the following hashtags #BlackLivesMatter #BlackLifeMatters #BLM in the three weeks following the murder. Column 6, reports the number of geo-located accounts that follow the official BLM account @BlkLivesMatter. We report Kleibergen-Paap rkWald F statistic for COVID and the respective interaction terms. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D5: COVID-19 exposure and social media use

	PC 1	New Twitter accounts	(log) New Twitter accounts	Google search for Twitter	Residential stay
	(1)	(2)	(3)	(4)	(5)
Panel A: ***All counties***					
IV: COVID	0.767**	-0.709	0.690*	12.94*	3.155***
(deaths/1000)	(0.309)	(20.17)	(0.376)	(6.453)	(0.592)
Observations	1,332	3,106	3,106	3,056	1,348
F first stage	27.65	36.05	36.05	35.71	27.52
Mean dep. var.	0	3.062	0.615	61.27	10.63
Panel B: ***Counties with BLM protest before***					
IV: COVID	0.148	-37.13	-0.374	5.724	2.437***
(deaths/1000)	(0.326)	(62.07)	(0.395)	(6.164)	(0.886)
Observations	312	333	333	320	320
F first stage	25.27	28.09	28.09	26.65	26.13
Mean dep. var.	0.838	18.02	2.007	67.62	12.62
All controls	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on the use of social media. Column 1 shows the standardized first principal component of four outcomes of interest: new Twitter accounts (created during the pandemic but before the murder of George Floyd) tweeting about BLM during the three weeks after the murder of George Floyd (and its log), Google searches for "Twitter" and the level of change of residential stay with respect to the baseline before the pandemic. Table A7 details the construction of the principal component. Column 2 (resp column 3) shows estimates for new Twitter accounts (log of new accounts) created after the beginning of the pandemic but before George Floyd's murder that tweet about BLM in the three weeks following George Floyd's death. Column 4 (resp column 5) shows results for Google searches for "Twitter" (the level of change of residential stay with respect to the baseline before the pandemic) between April 13 to May 24. Panel A presents 2SLS estimation, using number of super-spreader events in neighbouring counties (50km radius) six weeks prior as an instrument for all US counties. Panel B presents these results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. All specifications include state fixed effects and the standard controls of the main specification. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table D6: Effect of Twitter on BLM protest

Outcome:	Presence of BLM events					
Measure for Twitter	*New Accounts/ Pandemic Twitter*		*Baseline Twitter*		*Instrumented Baseline Twitter*	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All counties						
COVID × Twitter		0.000707		0.150***		0.0737
		(0.0622)		(0.0531)		(0.132)
COVID	0.172	0.171	0.207*	-0.527*	0.183	-0.177
	(0.124)	(0.220)	(0.117)	(0.301)	(0.114)	(0.746)
Twitter	0.0625***	0.0624***	0.0585***	0.0441***	0.0749*	0.0649
	(0.00857)	(0.0109)	(0.00758)	(0.00860)	(0.0399)	(0.0450)
Observations	3,106	3,106	3,106	3,106	3,106	3,106
F COVID	36.52	44.29	36.11	22.91	54.22	8.863
F interaction		132.9		29.68		6.054
F Twitter					14.38	18.70
Mean dep. var.	0.0988	0.0988	0.0988	0.0988	0.0988	0.0988
Panel B: *Counties with BLM protest before*						
COVID × Twitter		-0.133		0.0738		-0.0351
		(0.100)		(0.107)		(1.090)
COVID	0.0271	0.368	0.0549	-0.338	-0.191	-0.000256
	(0.273)	(0.341)	(0.242)	(0.649)	(0.642)	(5.753)
Twitter	0.0445	0.0744**	0.169***	0.158***	-0.480	-0.462
	(0.0302)	(0.0306)	(0.0320)	(0.0404)	(1.813)	(1.373)
Observations	333	333	333	333	333	333
F COVID	27.83	43.74	27.37	22.99	0.612	0.859
F interaction		56.90		30.31		0.750
F Twitter					0.251	0.309
Mean dep. var.	0.514	0.514	0.514	0.514	0.514	0.514
Instruments	SSE		SSE		SSE & SXSW	
All controls	Y	Y	Y	Y	Y	Y
Pre-SXSW users					Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Estimation of the differential effect of exposure to the COVID-19 pandemic on the presence of BLM protest following George Floyd's murder depending on different measure of Twitter usage. Panel A presents results for the sample of all US counties. Panel B presents results for the sub-sample of counties with at least one BLM protest before the murder of George Floyd. COVID is measured as before: deaths per 1000 population. Column 1 and 2 show the effect of the log number of Twitter users at baseline (i.e. December 2019) and instrumented COVID deaths on the presence of BLM events in a county. Column 2 additionally shows the interaction. Columns 3 and 4 show the effect of the log number of new Twitter users that have created an account during the pandemic, but before the murder of George Floyd. Column 4 additionally shows the interaction. Columns 5 and 6 use instrumented log number of Twitter users at baseline (i.e. December 2019), using the Müller and Schwarz (2020) SXSW instrument. The first stage regression is reported on Table C4. All specifications include state fixed effects and all controls from the baseline specification. Columns 5 and 6 include an additional control: the log number of users before the SXSW festival. First stage F statistics are presented following Sanderson and Windmeijer (2016). Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D7: Competing Mechanisms: Broadening versus Scattering of Protest (all counties)

	Presence of BLM events					
	(1)	(2)	(3)	(4)	(5)	(6)
Sample: *All counties*						
IV: COVID (deaths/1000)	0.219*	-0.0333	0.219*	-0.0445	-0.154	3.634*
	(0.122)	(0.232)	(0.123)	(0.271)	(0.985)	(2.119)
× Neighbor protested historically		0.252				
		(0.183)				
× Neighbor protested currently				0.254		
				(0.207)		
× Distance to Minneapolis					0.000234	-0.00775**
					(0.000603)	(0.00359)
× Distance to Minneapolis (squared)						3.48e-06**
						(1.47e-06)
Neighbor protested historically	-0.0138	-0.0384**				
	(0.0144)	(0.0177)				
Neighbor protested currently			-0.0108	-0.0320		
			(0.0183)	(0.0193)		
Distance to Minneapolis					3.21e-05	0.000387
					(5.28e-05)	(0.000240)
Distance to Minneapolis (squared)						-1.76e-07*
						(1.03e-07)
Observations	3,106	3,106	3,106	3,106	3,106	3,106
F first stage	35.56	18.67	33.96	17.09	25.67	19.05
F interaction		31.43		37.29	19.74	13.65
F interaction sq						9.791
Mean of dependent variable	0.0988	0.0988	0.0988	0.0988	0.0988	0.0988
All controls from preferred specification	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on the presence of BLM protests for all counties. Column 1 (column 3) shows estimates for instrumented COVID-19 deaths controlling for a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 2 and 4 present heterogeneous effects for the presence of a neighbouring county that protested before. Columns 2 (column 4) shows the interaction term with a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 5 and 6 presents results for interaction with distance to Minneapolis and distance to Minneapolis squared. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D8: Competing Mechanisms: Broadening versus Scattering of Protest (counties with BLM protest before)

	Presence of BLM events					
	(1)	(2)	(3)	(4)	(5)	(6)
Sample: ***Counties with BLM protests before***						
IV: COVID (deaths/1000)	0.00928	-0.322	0.0111	-0.305	-0.236	6.247
	(0.265)	(0.306)	(0.267)	(0.603)	(1.605)	(4.195)
× Neighbor protested historically		0.302				
		(0.202)				
× Neighbor protested currently				0.308		
				(0.493)		
× Distance to Minneapolis					0.000161	-0.0120*
					(0.000911)	(0.00692)
× Distance to Minneapolis (squared)						5.01e-06*
						(2.73e-06)
Neighbor protested historically	-0.0237	-0.0676				
	(0.0682)	(0.0788)				
Neighbor protested currently			-0.0165	-0.0588		
			(0.0930)	(0.130)		
Distance to Minneapolis					0.000231	0.00201
					(0.000273)	(0.00142)
Distance to Minneapolis (squared)						-7.27e-07
						(4.74e-07)
Observations	333	333	333	333	333	333
F first stage	28.19	22.76	27.07	17.70	21.41	16.57
F interaction		169.8		168.6	21.19	16.65
F interaction sq						14.75
Mean of dependent variable	0.514	0.514	0.514	0.514	0.514	0.514
All controls from preferred specification	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on the presence of BLM protests for sub-sample of counties with BLM protests before. Column 1 (column 3) shows estimates for instrumented COVID deaths controlling for a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 2 and 4 present heterogeneous effects for the presence of a neighbouring county that protested before. Columns 2 (column 4) shows the interaction term with a dummy equal to one if at least one neighbouring county protested for BLM at anytime before 2020 (during the 3 weeks after the murder of George Floyd). Columns 5 and 6 presents results for interaction with distance to Minneapolis and distance to Minneapolis squared. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D9: Competing Mechanisms: Salience, Opportunity Cost, and Agitation (all counties)

	Presence of BLM				Other Protests	COVID-19 Protests	Tweets AllLivesMatter	Tweets BlueLivesMatter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample: ***All counties***								
COVID (deaths/1000)	-0.172	0.316	0.0950	-0.117	0.232*	0.277**	357.2**	60.52**
	(0.212)	(0.441)	(0.556)	(2.108)	(0.135)	(0.111)	(174.8)	(23.80)
× Black death burden	0.964**							
	(0.387)							
× Google BLM searches		-0.00203						
		(0.0136)						
× Unemployment			0.0285					
			(0.122)					
× Stringency index				0.00438				
				(0.0282)				
Interacting variable	-0.0101*	0.000114	-0.00710					
	(0.00519)	(0.00147)	(0.00539)					
Observations	3,106	3,032	3,106	3,106	3,106	3,106	3,106	3,106
F COVID	24.40	17.60	24.30	25.76	36.05	36.05	36.05	36.05
F interaction	4.514	17.13	24.87	25				
Mean of dependent variable	0.0988	0.0986	0.0988	0.0988	0.0808	0.0296	134.7	17.75
All controls	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on presence of several types of protest. Columns 1 to 4 show heterogeneous effects on the presence of BLM events for Black death burden weeks prior to George Floyd's murder, Google searched for BLM 3 weeks prior to George Floyd's murder, unemployment and stringency 3 weeks after George Floyd's murder. The coefficient for the interacting variable in column 4 is dropped as stringency is measured at the state level and state fixed effects are included. Column 5 (resp column 6) presents results for all other protests besides BLM (protest related to COVID-19; e.g. anti-mask protest) during the 3 weeks following George Floyd's murder. Columns 7 and 8 show as an outcome the number of tweets including the pro-police and anti BLM hashtags #AllLivesMatter and #BlueLivesMatter. All results are shown for the sample of all counties. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1

Table D10: Competing Mechanisms: Salience, Opportunity Cost, and Agitation (counties with BLM protest before)

	Presence of BLM				Other Protests	COVID-19 Protests	Tweets AllLivesMatter	Tweets BlueLivesMatter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample: *Counties with BLM protest before*								
COVID (deaths/1000)	0.916	1.199	-0.00835	3.090	0.0526	0.188	794.7*	140.2**
	(1.187)	(1.049)	(0.919)	(6.919)	(0.423)	(0.368)	(455.2)	(60.82)
× Black death burden	39.63							
	(39.71)							
× Google BLM searches		-0.0295						
		(0.0308)						
× Unemployment			0.00418					
			(0.165)					
× Stringency index				-0.0395				
				(0.0890)				
Interacting variable	-0.0372	0.00390	0.00509					
	(0.0354)	(0.00662)	(0.0279)					
Observations	333	320	333	333	333	333	333	333
F COVID	15.73	15.19	17.34	17.62	28.09	28.09	28.09	28.09
F interaction	2.824	13.30	10.47	16.94				
Mean of dependent variable	0.514	0.522	0.514	0.514	0.477	0.192	848	113.1
All controls	Y	Y	Y	Y	Y	Y	Y	Y
State fixed effects	Y	Y	Y	Y	Y	Y	Y	Y

Note: Estimation of the effect of COVID-19 deaths per 1000 population on presence of several types of protest. Columns 1 to 4 show heterogeneous effects on the presence of BLM events for Black death burden weeks prior to George Floyd's murder, Google searched for BLM 3 weeks prior to George Floyd's murder, unemployment and stringency 3 weeks after George Floyd's murder. The coefficient for the interacting variable in column 4 is dropped as stringency is measured at the state level and state fixed effects are included. Column 5 (resp column 6) presents results for all other protests besides BLM (protest related to COVID-19; e.g. anti-mask protest) during the 3 weeks following George Floyd's murder. Columns 7 and 8 show as an outcome the number of tweets including the pro-police and anti BLM hashtags #AllLivesMatter and #BlueLivesMatter. All results are shown for the sub-sample of counties with BLM protests before the murder of George Floyd. All specifications include state fixed effects and standard controls. We report Kleibergen-Paap rkWald F statistic. Standard errors (in parentheses) are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1