Article:

Flückiger, Matthias orcid.org/0000-0002-2242-0220 and Ludwig, Markus (2025) The structure of online social networks and social movements:Evidence from the Black Lives Matter protests. JOURNAL OF PUBLIC ECONOMICS. 105373. ISSN 0047-2727

https://doi.org/10.1016/j.jpubeco.2025.105373

The Structure of Online Social Networks and Social Movements: Evidence from the Black Lives Matter Protests*

Matthias Flückiger† Markus Ludwig‡

March 28, 2025

Abstract

This paper documents that online social networks—Facebook in particular—can facilitate the spread of social movements across space and time. Focusing on the largest protest movement in recent history, the wave of Black Lives Matter protests sparked by the killing of George Floyd on 25 May 2020, we show that protests are more likely to spill over between US counties when they are more closely connected within the Facebook network. To identify causal effects, we develop an instrumental variable approach that exploits local Facebook outages as a source of exogenous variation in the structure of the online network.

JEL classification: D71, D83, L82, Z13

Keywords: Collective action, online social networks, Black Lives Matter

*We are grateful for the valuable comments received from the editor, three anonymous referees, and participants at various seminars. We also thank Josef Bräutigam, Daniel Burge, Henning Deux, Frederic Heise, and Precious Odusanya for their excellent research assistance.

†University of York. Email: matthias.flueckiger@york.ac.uk

‡TU Braunschweig. Email: markus.ludwig@tu-braunschweig.de

1 Introduction

Social movements are key drivers of economic, political, and cultural change. Online social networks are increasingly thought to be a key facilitator of such movements. These networks allow for easy coordination and the sharing of information between regions, regardless of geographic distance. This suggests that the spatial structure of online networks influences the diffusion of social movements across space and time. In this paper, we address this question, focusing on the largest protest movement in recent years.

The killing of George Floyd on the evening of 25 May 2020 by a Minneapolis police officer triggered a wave of protests that was unprecedented in US history, both in terms of its spatial spread and the number of participants (e.g. Buchanan et al., 2020). The protests were part of the wider Black Lives Matter (BLM) movement, which demands an end to racially motivated police brutality against Black people, a lack of police accountability, and, more generally, discrimination.[1] Online social networks are widely viewed as having played a central role in the rapid diffusion of protests across the country, facilitating the sharing of information and sentiments as well as helping to organise and coordinate collective action (e.g. Pew Research Center, 2020b; Buchanan et al., 2020; Hamzelou, 2020; Ovide, 2020; Tillery, 2019; Anderson and Hitlin, 2016; Freelon et al., 2016). This paper's contribution is to document that the spatial structure of online networks plays a crucial role in explaining when and where protests occur in democracies.

To analyse whether protests in the wake of George Floyd's murder spread across regions that are highly connected within online networks—Facebook in particular—we compile a county×day-level dataset on protest incidence for the period May–August 2020. Data on BLM protests are drawn from the Crowd Counting Consortium (CCC). To measure the extent to which a given county is exposed to these protests within its online social network, we average protest incidence across all other counties, using the number of bilateral Facebook friendships as weights (Bailey et al., 2018a). Because people in the US typically connect on Facebook only with those they know personally (i.e., real-world friends and contacts), Facebook data are uniquely suited to providing a large-scale representation of the US friendship network (see Bailey et al., 2018a).

The resulting index captures the online network-proximity-weighted average protest incidence. The intuition behind this measure is as follows: Protests that occur in counties well-connected

[1] The Black Lives Matter (BLM) movement is a global activist movement that aims to end systemic racism and violence towards Black individuals. The movement was founded in 2013 in response to the acquittal of George Zimmerman, a white neighbourhood watch volunteer who was charged with the murder of a 17-year-old Black teenager (Black Lives Matter, 2023). It gained significant attention in 2014 after the killing of Michael Brown, an unarmed Black man, by a white police officer in Ferguson, Missouri. This event led to large-scale protests in Ferguson and beyond. Within a month of Brown's death, around 200 BLM protests had been recorded nationwide (`elephrame.com`). BLM has since grown into a global movement, with protests and demonstrations sporadically taking place in several countries before 2020 (Mazumder, 2019).

to the home county increase the visibility and awareness of, as well as engagement with, the movement more than protests occurring in areas to which (online) ties are weak. Greater exposure to the movement, in turn, raises the probability that protests will occur in the home county.

The main empirical challenge in testing whether exposure to protests within the online network influences the spatiotemporal diffusion of the protest movement is the existence of unobserved factors that are correlated with our protest exposure measure. Of particular concern are latent networks. To address this issue, we leverage local Facebook outages as a source of exogenous variation. We argue that outages reduce the intensity of online interaction between counties, implying that a home county is less exposed to the protests in locations where Facebook is down. The closer two counties are connected, the stronger the information-disrupting effect of the outage. Based on this reasoning, we use the network-proximity-weighted outages as an instrument for the online exposure to the protest movement. Variation in this measure is solely driven by the timing of Facebook outages in network neighbours and the (predetermined) strength of their Facebook link to the home county.

Our benchmark 2SLS-IV panel regression model accounts for county and date fixed effects. The latter absorb general time-varying factors that influence the protest probability, such as changing news coverage or weekday effects. The former control for differences in county-level characteristics that could influence the baseline risk of collective action (e.g., socio-economic factors). The resulting estimates show that greater exposure to protests within a county's online network raises the likelihood of collective action. Specifically, we find that the probability that a protest takes place in the home county rises on the first and second day after an increase in protest exposure within the online social network. For both lags, the point effects are very similar. A one percentage point increase in network-proximity-weighted average protest incidence raises the protest risk by around one percentage point. This implies that the risk of protest surges on average by around 15 percentage points in our sample if the closest neighbour experiences a protest. Evaluated at the baseline risk of 3.26, this constitutes almost a tripling. An alternative way of quantifying our results is to look at the effects of Facebook outages in the network neighbours directly. The likelihood that a BLM march occurs in the home county decreases by 0.7 percentage points in the two days after its closest neighbour experienced a Facebook outage. This corresponds to a drop of 21% when evaluated at the mean. While the magnitude of the effect is large, it aligns well with those found in other studies that investigate the effects of online networks on protest activity (e.g. Qin et al., 2024; Enikolopov et al., 2020). These sizeable effects illustrate that the structure of the Facebook network strongly influenced the spread of the BLM protests over space and time. Given that the protests have led to local legal and cultural reforms (e.g. Ebbinghaus et al., 2021), our findings imply that online social networks can play an important role in shaping the geography of social change.

There are two main threats to the validity of our identification strategy. The first worry is that Facebook outages in the network neighbours are not exogenous with respect to protest activity in the home county. For example, outages in the home county and network neighbours could be correlated and—at the same time—directly influence the protest risk in the home county. In this case, our estimates would conflate spillover effects that are due to changes in the connectivity within the Facebook network and effects that are driven by Facebook outages in the home county. We show that such effects are unlikely to bias our results. Controlling for outages in the home county or dropping counties that themselves experience outages at any point in time leaves our estimates unaffected. A second natural worry is that our estimates are picking up (correlated) latent spatial networks. To mitigate this risk, we condition our regressions on exposure within the distance network throughout. Reassuringly, including this control has little effect on the size of our main point estimate. Relating to the validity of the exclusion restriction, the primary concern is that network-proximity-weighted outages may influence the local protest risk through mechanisms other than the exposure channel that we have in mind. For example, it is conceivable that outages in important network neighbours reduce the relevance of the Facebook feed in the home county and consequently reduce Facebook usage. This, in turn, could change the protest risk. While we cannot directly test for the relevance of alternative mechanisms, we provide some suggestive evidence indicating that they are unlikely to be of primary importance.

We run a range of robustness checks to document the stability of our results. Inter alia, we show that our findings remain unaltered when we include additional county-level controls, vary the sample period, manipulate parameters in the construction of our main instrumental variable, and employ an alternative instrumental variable approach. Finally, we also run a falsification exercise in which we show that randomly assigning outages across space and time produces point estimates that are centred around zero and orders of magnitude smaller than our main coefficients.

Our paper builds on and contributes to several strands of literature. Most directly linked to our study are papers that investigate the effects of online social networks on political protests. Ample qualitative evidence indicates that online networks play an important role in the emergence and diffusion of protest movements (e.g., The Economist, 2021; Mortensen et al., 2018; Mundt et al., 2018). However, there is a relative scarcity of studies that quantitatively evaluate the impacts. Closely related to our work are the recent studies by Qin et al. (2024) and Enikolopov et al. (2020) which document that online social networks influence the spatio-temporal spread of protest movements in China and Russia, respectively. Our paper expands the literature by looking at the dynamics of spillover effects within the largest online social network and a democratic setting.

Also directly related to our study is the branch of literature that investigates how the structure and the content of social networks (mostly in-person) influence political participation and polarisation (Azzimonti and Fernandes, 2022; Caprettini et al., 2021; Fujiwara et al., 2023; González, 2020; Zhuravskaya et al., 2020; Cantoni et al., 2019; Campante et al., 2017; Halberstam and Knight, 2016), the diffusion of hate crimes (Cao et al., 2023; Müller and Schwarz, 2023; Bursztyn et al., 2024), the intensity of economic ties and investment decisions (Bailey et al., 2021, 2018b), as well as the spread of diseases (e.g. Kuchler et al., 2022). In historical settings, García-Jimeno et al. (2022) and Aidt et al. (2022) show that railway and trade networks influenced the diffusion of political movements in the more distant 19th century. Our results highlight that the structure of digital networks can influence the location and timing of protests and, consequently, the dynamics of mobilisation.

This paper also contributes to the literature on the effects of media and technology on collective action. There is ample evidence that traditional media—newspapers, radio, and television—can influence attitudes and shape collective action (e.g. Adena et al., 2015; Yanagizawa-Drott, 2014; DellaVigna and Kaplan, 2007). Similarly, a range of studies document that new types of technologies and media, the internet and social media in particular, influence the risk of protests (e.g. Manacorda and Tesei, 2020; Bursztyn et al., 2024; Larson et al., 2019; Barberá et al., 2016).[2] In that respect, the papers by Fergusson and Molina (2021) and Müller and Schwarz (2021) are most closely related to our study. They show that the availability and content of Facebook increase the risk of protests and hate crimes. We complement these findings by illustrating that the risk of protests not only varies with local accessibility of the Facebook platform, but also changes depending on how regions are interconnected within the network.

Specific to the causes of BLM protests, narrative evidence suggests that online networks have played an important role in the coordination of activism, pushing the BLM agenda, and growing the movement (Pew Research Center, 2020b; Buchanan et al., 2020; Ovide, 2020; Tillery, 2019; Anderson and Hitlin, 2016; Freelon et al., 2016). Furthermore, qualitative and quantitative evidence indicates that the movement, and protests in particular, can shift the public discourse towards the movement's agenda (Dunivin et al., 2022), increase the salience of racial inequalities (Drakulich and Denver, 2022; Mutz, 2022), reduce racial prejudice (Boehmke et al., 2023; Mazumder, 2019), and change policing approaches (Campbell, 2024; Wang, 2022; Ebbinghaus et al., 2021; The Economist, 2022).[3]

The remainder of this paper is structured as follows: We begin by presenting our data in Section 2. In Section 3, we outline the regression methodology. The resulting estimates, along with threats to identification, are discussed in Section 4. Finally, we offer some conclusions in Section 5.

[2]See Aridor et al. (2024) and Zhuravskaya et al. (2020, chapter 3) for more extensive reviews of the literature on internet, social media and protests.

[3]However, Engist and Schafmeister (2022) find no evidence that the BLM protests increased political mobilisation in the U.S., as measured by voter registration.

2 Data

Three types of data form the basis of our empirical analysis: (i) the location and timing of BLM protests, (ii) online social connectedness between counties, and (iii) local Facebook outages. We combine these data sources into a county×day-level dataset that links protest activity in the home county to exposure to protests in counties within its online social network. The data sources and construction processes are described below.

Protests

Data on BLM protests come from the Crowd Counting Consortium (CCC) which collates publicly available information on political protests from a range of sources, including newspapers, television sites, and law enforcement.[4] We restrict our analysis to protests that are clearly assignable to the Black Lives Matter movement.[5] We geocode these protests and aggregate them at the county×day level. Our main outcome variable is a dummy variable that takes the value of one if at least one protest takes place in a given county and day.[6]

The solid black line in Panel (a) of Figure 1 illustrates the temporal dynamics of the BLM protest movement across the US in 2020. During the first five months of the year, protest activity remained extremely limited.[7] This changed dramatically with the killing of George Floyd on 25 May 2020. After the first protest march in Minneapolis on 26 May 2020, protests quickly spread to other locations. Just five days later, 478 counties reported at least one manifestation. Protest intensity culminated on the 6^{th} of June, when a total of 721 BLM protests were reported. An estimated 500,000 individuals took to the streets on this day (Buchanan et al., 2020). After this peak, the number of protests declined continuously, with surges periodically observable. Spikes typically coincide with weekends.

Given the rapidly dissipating nature of the protest movement, we restrict our analysis to the time period 26 May 2020 – 25 July 2020. This cut-off represents the last date on which more than 50 BLM protests were staged on a single day in the US.[8] The dotted line in Figure 1, Panel (a), represents the cumulative share of total protests during 2020 and illustrates that more than 80% of all BLM marches occurred within our two-month sample period. During this period, 44% of counties saw at least one protest. The location of these counties is shown in Panel (b) of Figure 1.

[4] `sites.google.com/view/crowdcountingconsortium`, last accessed 17 March 2022.

[5] The CCC assigns claims to each protest. We retain all protests where the claims contain at least one of the following keywords: 'Black Lives Matter', 'antiracism', 'against police brutality', and 'against police shooting'.

[6] Unfortunately, detailed descriptions of many protests are unavailable. This prevents us from stratifying them by characteristics such as protest size or whether they were peaceful or violent.

[7] Between 1 January 2020 and 25 May 2020, only 45 protests occurred on 43 different days. Most of these took place in Los Angeles as regular demonstrations against police brutality by the LAPD, held in front of the Hall of Justice in downtown LA.

[8] The choice of time period is somewhat arbitrary. In robustness checks we show that our results remain stable when we manipulate the cut-off date.

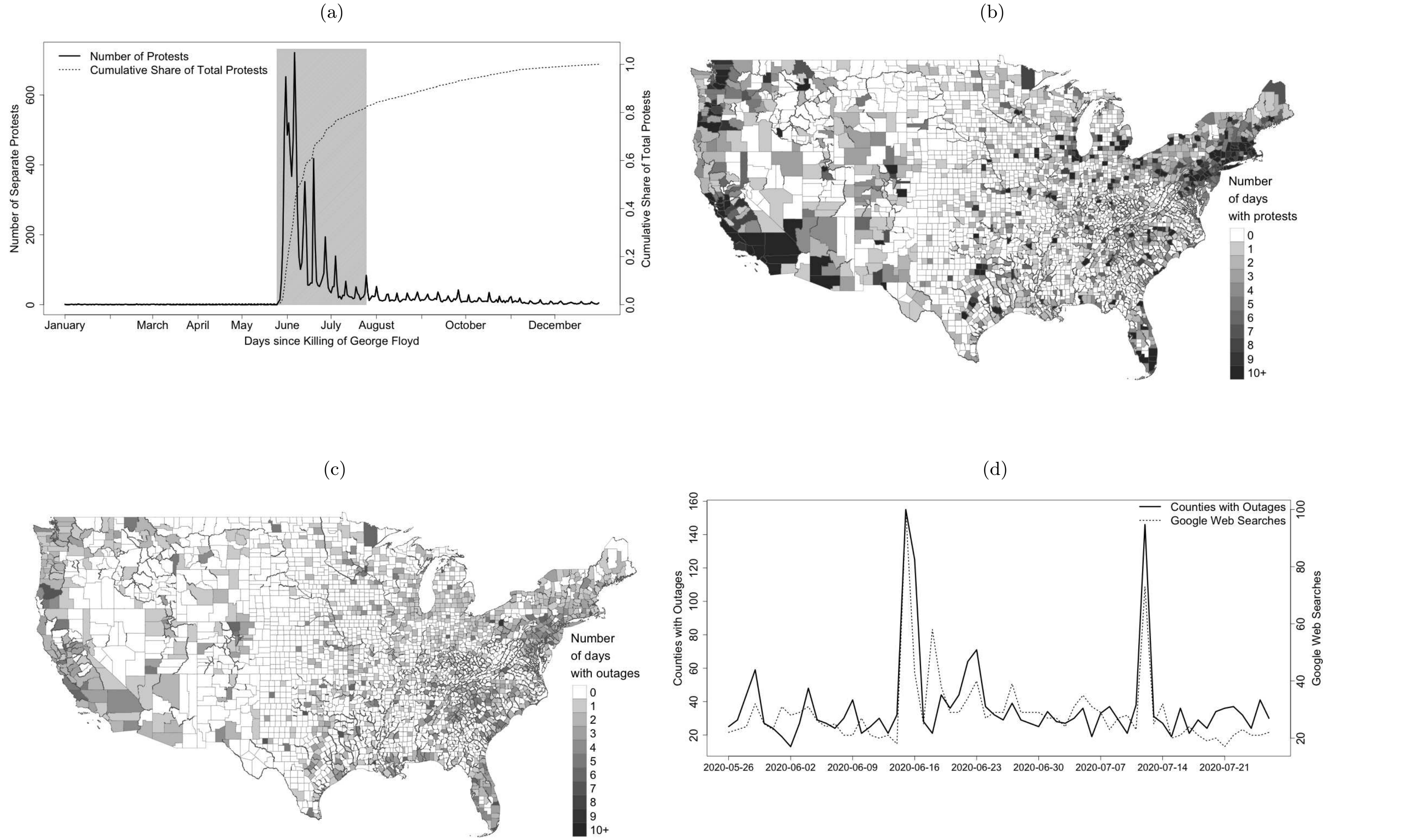

Figure 1: Protests and outages over space and time

Panel (a) depicts the number of BLM protests (solid line) and cumulative share of total protests (dotted line) in 2020. The grey shading delineates our sample period (26 May 2020–25 July 2020). Panel (b) llustrates the spatial distribution of BLM protests. Darker shadings indicate a higher number of protests. Panel (c) illustrates the spatial distribution of Facebook outages. Darker shadings indicate a higher number of outages. Panel (d) depicts the number of Facebook outages (solid line) and Google web searches for Facebook connection issues (dashed line).

Facebook stability

We construct a county×day level indicator variable that captures local stability in access to the Facebook platform. Data on Facebook connection issues comes from `Downdetector.com`, which collects status reports from a range of sources (e.g., Twitter and reports submitted via their online and mobile apps). These reports are analysed and validated in real-time.

Information on the daily number of reported outages are commercially available at the city/town level. We geocode these data and aggregate the number of reports at the county×day level.[9,10] We then create a dummy variable that represents the local stability of the Facebook platform. This dummy is equal to one if there are no major Facebook accessibility issues, and zero if there is a serious outage.[11] We define outages as days on which the number of reports is two standard deviations or more above the county-level mean.[12,13] In total, we observe 2,285 county×day outages over the course of our sample period; the average likelihood of an outage is 1.193%. The frequency of Facebook outages is very similar to that observed in Müller and Schwarz (2021).[14] Panel (c) of Figure 1 shows which counties were affected by outages while the solid line in Panel (d) represents the daily number of counties that experience a Facebook outage. The locations of outages are scattered across space and occur throughout the period of our analysis. Both the spatial and temporal correlations of outages are low.[15]

To validate our outage measure, we draw on Google Trends data which provide information on the daily number of search requests made to Google. The dashed line in Panel (d) of Figure 1 traces out the number of Google searches related to Facebook outages over time.[16] Reassuringly, there is a very high degree of co-movement between the number of outages

[9]For larger cities and metropolitan areas, the geographical information is typically granular enough to assign the outage reports to one specific county. For example, information for New York City is available at the borough level. Similarly, information for the Los Angeles metro area is provided at the level of the individual cities. When a location is not clearly assignable to a county, we assume that the outage occurs in the county that encompasses the centroid of the named city or town.

[10]Within our sample period, Downdetector recorded 126,261 validated reports on Facebook connection issues. Aggregated at the county×day level, the average number of reports is 10.05 (conditional on any report being recorded).

[11]See Müller and Schwarz (2021) for a similar use of Facebook outages in the context of Germany.

[12]Formally, the Facebook stability indicator $s_{j,d}^{\text{Facebook}}$ for county j on day d is given by:

$$s_{j,d}^{\text{Facebook}} = \mathbf{I}\left(r_{j,d} \leq \mu_{r,j} + 2 \times \sigma_{r,j}\right)$$

where $r_{j,d}$ represents the number of outage reports, $\mu_{r,j}$ the average number of outages, and $\sigma_{r,j}$ the local standard deviation.

[13]In robustness checks, we illustrate that our results remain stable when we manipulate the outage threshold or employ structurally different outage definitions.

[14]Müller and Schwarz (2021) construct a week×municipality-level dataset for Germany and find that the average weekly risk of outages is around 7.2 percent. When we collapse our data at the county×week level, the mean likelihood of an outage is 6.99 percent.

[15]The raw correlation between outages in the home county and the network-weighted outages is 0.143 (see equation (3)). The raw autocorrelation between today's outage and yesterday's outage is 0.079.

[16]Specifically, we focus on web searches containing the following keywords: 'Facebook down', 'Facebook issues', 'Facebook problems', 'Facebook connection', 'Facebook access', and 'Facebook outage'.

and Google searches.[17] The daily search volumes can be further stratified by regions. This allows us to test if local Facebook outages lead to a surge in outage-related searches by conditioning on region and date fixed effects. The result document a strong and statistically significant relationship (see Appendix C for more details). To further illustrate that our measure captures Facebook-specific connection issues, we examine search requests related to outages of other online social media networks. Specifically, we look at the five most popular platforms in addition to Facebook: YouTube, Instagram, TikTok, Snapchat, and Twitter (see doofinder.com). We do not observe any statistically significant relationship between our Facebook outage measure and outage-related searches for other platforms. The exception is Instagram, where we find a positive, albeit weaker, effect. This is not surprising, given that Facebook acquired Instagram in April 2012. It is therefore plausible that Instagram is also affected when Facebook experiences issues.

Exposure to protests in online social network

We map online social connectedness across space using the Social Connectedness Index (SCI).[18] The SCI–developed and described in detail in Bailey et al. (2018a)—captures the link strength between two counties within the Facebook network.[19,20] The SCI is constructed using the universe of friendship links between all Facebook users and is available for all US county pairs. In the United States, Facebook mainly serves as a platform for connecting with real-world friends and acquaintances online. Since people tend to add only those they know personally, Facebook data provides a representation of US friendship networks (p. 261 Bailey et al., 2018a). For the purposes of our study, our interpretation of the SCI is that more socially connected counties ultimately exchange more information, as individuals in these counties are more likely to engage with and share content from their network ties. Furthermore, the information shared is potentially more relevant and trusted, as it originates from friends and personal connections.

Based on the SCI, we construct a measure for protest exposure within a county's social network in two steps: For each county c, we first compute the proportion of total Facebook links it shares with county j:

$$\omega_{c,j} = \frac{SCI_{c,j}}{\sum_{k=1}^{K} SCI_{c,k}}, \tag{1}$$

where $SCI_{c,j}$ is the SCI between county c and j. The weight $\omega_{c,j}$ lies between zero and one, where higher values indicate greater social connectedness.[21] For our analysis it is im-

[17] The raw correlation is 0.784.

[18] This data is publicly accessible at `data.humdata.org/dataset/social-connectedness-index`.

[19] User locations are determined based on their information and activity on Facebook, which include the city listed on their profile, as well as device and connection data.

[20] Note that the data released by Facebook are scaled by an (unknown) factor. Therefore, the absolute values of the SCI are not meaningful, as they cannot be directly interpreted as the number of Facebook user links between counties.

[21] The normalisation to one facilitates interpretation and also implies that the size of the county does not

portant to note that the structure of the network—and therefore $\omega_{c,j}$—is time-invariant and predetermined.[22]

In the second step, we use the connectedness between individual county pairs ($\omega_{c,j}$) and compute the social-network-proximity-weighted average protest incidence, $P^N_{c,d}$, for county c on day d as:

$$P^N_{c,d} = \sum_{j \notin c}^{J} \omega_{c,j}\, p_{j,d}\, s_{j,d}, \tag{2}$$

where $p_{j,d}$ is an indicator equal to one if a BLM protest takes place in county j on day d. Local Facebook stability for county j on day d is represented by $s_{j,d}$. The intuition behind this measure is that a protest occurring in another county is more likely to spark demonstrations in the home county, the more closely the two counties are connected within the online social network. However, the extent to which protest activity is visible varies with Facebook outages.

Connection stability within the online social network

Analogous to the social-network-proximity-weighted protest incidence, we compute a measure of connection stability within a county's online network. The only difference from equation (2) is that the protest indicator is omitted (see García-Jimeno et al. (2022) for a similar IV approach). Consequently, the measure contains no information on protest activity, but simply captures the network-proximity-weighted Facebook connection stability in the other counties ($j \neq c$). Formally, the measure for county c on day d, $S^N_{c,d}$, is given by:

$$S^N_{c,d} = \sum_{j \notin c}^{J} \omega_{c,j}\, s_{j,d}, \tag{3}$$

Analogous to the protest exposure measure, higher values of $S^N_{c,d}$ imply better online connection to other counties.

Our final estimation dataset consists of 192,758 observations, covers the period 26 May 2020–25 July 2020 and includes 3,109 counties, all located within the contiguous United States or Hawaii. Summary statistics of key variables are presented in Table A.1.

3 Empirical Strategy

To investigate if the probability of BLM marches increases with online-social-network exposure to protests, we use the following linear probability model:

$$p_{c,d} = \sum_k \gamma_{d-k}\, P^N_{c,d-k} + \pi_c + \tau_d + \theta_{s,w} + \sum_k \beta_{d-k}\, \mathbf{X}_{c,d-k} + \varepsilon_{c,d}, \tag{4}$$

matter; the weights capture the relative importance of link intensity.

[22] The SCI captures the structure of the network as of March 2020.

where outcome $p_{c,d}$ is an indicator variable that takes the value of one if a BLM protest occurs in county c on day d. The main regressors of interest are $P^N_{c,d-k}$, the network-proximity-weighted average protest incidence in counties $j \neq c$ on days $d-k$. As outlined above, we expect γ to be positive: An increase in protest activity in socially closely connected counties raises the likelihood that a protest takes place in the home county. However, we do not have a prior regarding the appropriate lag structure. In the first, preliminary, step of our empirical analysis, we will therefore take a hands-off approach to identifying the relevant lag(s), by running regression model (4) separately for each day within the 10-day time window centred around day d.

All regressions control for county fixed effects π_c. These absorb any time-invariant characteristics that influence the likelihood of protests. Such aspects include population density, demographic makeup, or past prevalence of police brutality.[23] General time-varying shocks are absorbed by date fixed effects τ_d. We further include state×weekend fixed effects ($\theta_{s,w}$) to allow for the possibility that spikes in protest activity observed during the weekend differ by state (cf. Panel (a) of Figure 1). The vector $\mathbf{X}_{c,d-k}$ represents a varying set of county-level characteristics, and the idiosyncratic error term is symbolised by $\varepsilon_{c,d}$. Throughout, we cluster standard errors at the county level.

The validity of the regression model (4) hinges on the assumption that $P^N_{c,d-k}$ only picks up the effects on collective action that are due to changes in exposure to protest activity within a county's online social network. A main worry is that there are omitted variables, and in particular, latent (and correlated) networks through which protest activity in other counties influences local collective action. A natural concern, for instance, is that $P^N_{c,d-k}$ simply captures the effects of geographic proximity. As an initial step to mitigate this concern, we compute a network exposure measure in analogy to equation (2), using the (inverse) geodesic distance between counties as weights.[24] In the regressions below, we include this distance-based network exposure measure as a control.

To more rigorously account for the possibility that omitted variables bias our estimates and to isolate effects that operate via the Facebook network, we implement an instrumental variable strategy akin to García-Jimeno et al. (2022). Specifically, we predict $P^N_{c,d}$ using network-proximity-weighted Facebook connection stability (cf. equation (3)). Variation in this instrument is only generated by (network-proximity-weighted) local Facebook outages, which

[23]The panel structure of our analysis also implies that the reflection problem associated with the estimation of social interactions is not a concern (García-Jimeno et al., 2022; Manski, 1993).

[24]We compute the weights of the geographical distance network in analogy to our SCI weights. Specifically, the distance weight is given by: $\zeta_{c,j} = \dfrac{1/dist_{c,j}}{\sum_{k=1}^{K} 1/dist_{c,k}}$, where $dist$ represents the geodesic distance between county c and j. The correlation between the SCI and the distance weights is 0.585. Table A.3 illustrates that this correlation is higher than alternative definitions of the distance weights (e.g., inverse distance or inverse log distance). However, robustness tests show that our results are robust to the inclusion of distance exposure measures constructed using alternative weights.

are exogenous with respect to changes in protest activity.[25] Formally, the first stage of the instrumental variable approach is represented by:

$$P^N_{c,d} = \sum_k \lambda_{d-k}\, S^N_{c,d} + \mu_c + \eta_d + \chi_{s,w} + \sum_k \Omega\, \mathbf{X}_{c,d-k} + \psi_{c,d}. \tag{5}$$

As before, $P^N_{c,d}$ represents the network-proximity-weighted average protest incidence. The outage-based instrument is given by $S^N_{c,d}$, while county and time effects are captured by μ_c, η_d, and $\chi_{s,w}$. Finally, $\mathbf{X}_{c,d}$ represents a set of controls. In analogy to the OLS setup, this set includes the (geodesic) distance-weighted connection stability to account for effects of the outages working through latent networks. The error term is denoted by $\psi_{c,d}$.

The main identification assumptions underlying our 2SLS-IV approach are that Facebook outages in network neighbours are exogenous with respect to protest activity in the home county and only influence the latter via reduced interaction within the Facebook network. We discuss the plausibility of these assumptions in more detail in Section 4.3. Furthermore, we conduct a number of robustness checks to show that our findings are not driven by specific assumptions or choices made during data construction. Among other things, we show that our estimates remain unchanged if we condition on Facebook outages in the home county. Furthermore, we obtain qualitatively equivalent results if we disregard Facebook outages in the construction of the endogenous variable, i.e. in equation (2). This addresses the concern that the strength of the (first-stage) results may be driven by a mechanical link between our endogenous and instrumental variables.

4 Results

In this section, we first identify the relevant lag(s) to be included in our regression analysis. We then document that increased exposure to the BLM movement in the Facebook network raises the local protest risk. In the final part, we illustrate that our estimation strategy is unlikely to produce biased results.

4.1 Temporal dynamics: Identifying relevant lags

We use a hands-off approach to investigate the temporal dynamics of the relationship between network exposure and protest activity. The main aim is to identify the relevant lags to be included in the subsequent analysis. To this end, we estimate the 2SLS-IV regression model (4) separately for all leads and lags of network-proximity-weighted protest incidence within a 10-day time window centred around day d, i.e., for each day $d-k$, where $k = [-5, 5]$. Figure 2 shows the resulting point estimates along with the 95% confidence intervals. The grey squares

[25] We provide empirical evidence for this claim in the subsequent sections.

represent the values of the first-stage F-statistics for the excluded instrument; these are very high throughout.[26]

Figure 2 reveals two important findings. First, the spread of the BLM protests is influenced by the spatial structure of the Facebook network. The likelihood of a protest taking place in the home county statistically significantly increases in the first and second day after a surge in online network protest exposure to the BLM movement.[27] After the second day, the protest risk returns to its baseline. This implies that the relevant lag structure is 1–2 days.

The identified lag structure aligns well with the results of other studies that investigative the effects of online networks on protest activity. For example, Qin et al. (2024) find that protests in China spread from city to city 1–2 days after exposure within social media network Sina Weibo. Similarly, Müller and Schwarz (2021) show that anti-refugee Facebook posts trigger hate crime in the same week of the posts.

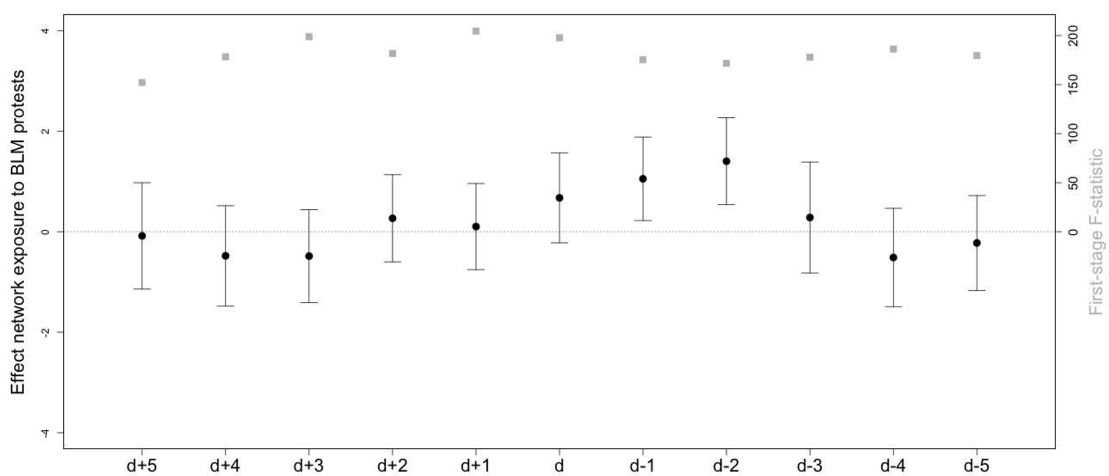

Figure 2: Temporal dynamics of network exposure
Figure depicts 2SLS point estimates and 95% confidence intervals of the effect of network exposure on protest probability. Point estimates obtained by separately estimating regression model (4) for each day within the 10-day time period centred around day *d*. Standard errors are clustered at the county level. Dependent variable equals 100 if a protest takes place in a given county and day, and zero otherwise. 'Network exposure to BLM protests' is the online network-proximity-weighted average protest incidence as defined in equation (2). The grey squares represent the values of the first-stage Kleibergen Paap F-statistics for the excluded instrument (right axis).

Second, the figure documents the absence of pre-trends. That is, variation in future network-weighted protest incidence does not influence today's protest risk. This provides a first piece of evidence that our estimates specifically capture the effects of increased exposure to the BLM movement in the online social network. We discuss the validity of our estimates in more detail below.

[26] Table B.1 presents the estimates shown in Figure 2 in tabular format.

[27] We observe the same pattern of results when we look at the reduced-form estimates (see Figure B.1) or estimate the coefficients jointly in one single regression (see Figure B.2).

4.2 Main analysis

Based on the temporal dynamics revealed in Figure 2, we subsequently focus on analysing the effects of the first and second lag of online network exposure. As a reference point, we start by estimating a parsimonious OLS version of regression equation (4), which only accounts for county and date fixed effects. Column (1) of Table 1 documents that there is a strong positive relationship between exposure to protests within the social network and collective action in the home county. The higher the proximity-weighted incidence of protests during the previous two days, the greater the likelihood of BLM protests. The point estimates indicate that the probability of observing a protest in the home county rises by 0.65 percentage points on the first day after a one percentage point increase in exposure to protests within the online network, and by 0.34 percentage points in the second day after the increase.

As outlined in previous sections, the OLS estimates are potentially biased due to unobservable factors. As a first step in addressing this issue, we control for exposure to protests within the distance network. This represents a catch-all variable for latent spatial networks. Column (2) shows that the inclusion of the additional control reduces the size of the point estimates slightly.

To isolate Facebook-specific diffusion effects, we turn to our 2SLS-IV strategy in column (3). In this approach, we use variation in proximity-weighted Facebook stability across network nodes as the instrument. Looking first at Panel B, we see that the instrument successfully predicts network-proximity-weighted protest incidence. For both, lag 1 and lag 2 the first-stage F-statistic is very high. The joint first-stage F-statistic for the validity of the instruments is 45. As expected, the sign of the coefficients is positive. Greater Facebook stability within a county's social network increases corresponding protest activity.[28] The results of the first-stage regressions further document that we can separate the effects of the first and second lag. That is, the two instruments only have predictive power for their respective endogenous counterparts.

We next focus on the second-stage estimates reported in Panel A. The point estimates in column (3) are larger than their OLS counterparts.[29] The difference may stem from correlated latent networks or from unobserved spatial and temporal factors that influence protest risk. The downward bias in the OLS estimates is consistent with patterns found in other studies (e.g., García-Jimeno et al., 2022). The magnitude of our 2SLS-IV coefficients implies that a one percentage point increase in protest exposure within the online network raises the likelihood of a BLM protest occurring in the home county by approximately one percentage point in both subsequent days. Relative to the average protest probability of 3.261, this corresponds to a 31% increase for both lags. As an alternative way of quantifying the results, we can look

[28] Table B.2 reports the reduced-form estimates.

[29] Statistically, however, the OLS and IV estimates are indistinguishable. The p-values for the tests of coefficient equality are 0.199 (lag 1) and 0.131 (lag 2).

Table 1: Effects of exposure to BLM protests in online social network

	Any BLM Protest			
	(1)	(2)	(3)	(4)
	Panel A: Second stage			
Lag 1 network exposure to BLM protests	0.651***	0.502***	1.045**	0.978**
	(0.035)	(0.043)	(0.421)	(0.414)
Lag 2 network exposure to BLM protests	0.339***	0.374***	1.003**	0.891**
	(0.025)	(0.032)	(0.415)	(0.419)
Lag protest incidence				0.124***
				(0.012)
	Panel B: First stages			
	Lag 1 network exposure to BLM protests			
Lag 1 Facebook connection stability			0.049***	0.048***
			(0.003)	(0.003)
Lag 2 Facebook connection stability			0.005	0.004
			(0.004)	(0.004)
	Lag 2 network exposure to BLM protests			
Lag 1 Facebook connection stability			-0.006	-0.006
			(0.005)	(0.005)
Lag 2 Facebook connection stability			0.049***	0.048***
			(0.004)	(0.003)
County fixed effects	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes
State × weekend fixed effects	Yes	Yes	Yes	Yes
Distance-weighted exposures	No	Yes	Yes	Yes
Observations	192,758	192,758	192,758	192,758
Mean dependent variable	3.261	3.261	3.261	3.261
First-stage F-statistic lag 1			101.899	101.494
First-stage F-statistic lag 2			96.712	96.516
Joint first-stage F-statistic			45.141	44.625

Notes: Panel A reports the second-stage estimates of equation(4) using 2SLS-IV. Panel B reports the corresponding first-stage estimates (equation (5)). Standard errors clustered at the county level are reported in parentheses. 'Any BLM Protest' is a dummy equal to one if at least one BLM protest takes place in a given county and day. 'Network exposure to BLM protests' is the online network-proximity-weighted average protest incidence as defined in equation (2). 'Facebook connection stability' is the online network-proximity-weighted average Facebook connection stability. Variable is defined according to equation (3). All variables multiplied by hundred to facilitate interpretation. Values of the Kleibergen-Paap F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

at the effects of Facebook outages in the network neighbours directly (i.e. the reduced-form). In the two days after the closest neighbour experienced a Facebook outage, the likelihood of a BLM march occurring in the home county decreases by 0.7 percentage points.

In column (4), we additionally include lagged protest incidence (in the home county) as a control. This addresses worries related to the possibility that our regression model may not accurately capture potential persistence of protest activity. The lagged protest indicator is statistically significant and positive, indicating that current protest activity increases the probability of future protests. Crucially, however, the main coefficients of interest remain stable, increasing confidence in the validity of our estimates.

The magnitude of the estimates presented in Table 1 is large. However, as summarised in Table B.3, it aligns well with the results from other studies that investigate the role of online networks in the propagation of protests. Qin et al. (2024), whose study is very closely related to ours in terms of research question and methodology, find that a one percentage point increase in protest exposure within the Chinese Sina Weibo network raises the local protest risk—evaluated at the sample mean—by 34%. This is remarkably similar to our estimate of a 31% increase. This shows that online social networks are similarly powerful in facilitating protest movements in two very different institutional contexts.Effect sizes are also comparable when we juxtapose our estimates with studies that focus on the effects of local penetration of online networks rather than on spatio-temporal spillovers. For Russia, Enikolopov et al. (2020) find that, evaluated at their sample mean, a 10% increase in VKontakte user numbers raises protest risk by 35%. Similarly, our estimates imply a 21% increase in protest risk relative to the baseline when protest exposure increases by 10%. Fergusson and Molina (2021) find that a one-standard-deviation increase in a country's share of people for whom Facebook is available in their mother tongue increases protest counts by up to 0.11 standard deviations. Our data suggest that a one-standard-deviation increase in online exposure to protests raises the number of protests by 0.26 standard deviations in the two subsequent days. While Fergusson and Molina (2021)'s estimates are half the size of ours, our 95% confidence interval comfortably includes their point estimate.

A study that explicitly incorporates the structure of the network—analogue rather than online—is García-Jimeno et al. (2022). Focussing on spillover effects within the railway network during the Temperance movement in the late 1800s, they find that protests in the two closest railway neighbours increase the probability of local collective action by a factor of around 5.6 relative to the baseline risk. In our setting, the likelihood of collective action multiplies by 7.3 when the two closest network neighbours experience protests. One possible reason for the larger effects in our study is that information and sentiment diffuse more efficiently within online networks compared to analogue technologies. This is likely a key explanation for the important role of social networks in spreading collective action documented in our and other papers.

In the context of our study, an additional factor may contribute towards explaining the large effect sizes: There is a strong demographic overlap between BLM protest participants and Facebook users. Notably, Blacks and Hispanic individuals, individuals aged 18–49, and college-educated individuals were significantly over-represented among BLM protesters (Pew Research Center, 2020a). Among these over-represented groups, Facebook usage was particularly prevalent. In the year 2021, Facebook was the most popular social networking platform, with around 74% of 18 to 49-year-olds using it (Pew Research Center, 2021). This share was similarly high among Blacks (74%) and Hispanics (72%). College-educated individuals had the highest usage rate at 73%. Furthermore, social media—Facebook among others—played a crucial role in sharing information and mobilising support for the George Floyd protests, especially among Black users, who actively used the platform to raise awareness and encourage action (Pew Research Center, 2020b). Together, this suggests that changes in exposure to BLM protests within the Facebook network may have particularly amplifying effects because they influence groups that were already highly engaged with the movement.

Overall, the results presented above document that online social networks—and Facebook in particular—play an important role in explaining the spread of one of the largest social movements in recent history. This implies that the structure of online networks shapes the geography of social change. We next discuss the internal validity of our analysis.

4.3 Threats to identification and robustness

Threats to identification

Our identifying variation is generated by local Facebook outages, weighted by the connectedness of the county in which the outage occurs. In order for our 2SLS-IV estimates to capture the effects of protest exposure within the online social network, we require that an outage in a network neighbour is exogenous with respect to protest risk in the home county and only influences this risk by changing the degree of visibility within the online network. One immediate worry is that outages are correlated across space and directly influence the protest risk. In this case, our instrument would not only change the degree of online interaction between counties, but also change the likelihood of protest in the home county. As a consequence, our estimates would conflate effects transmitted through the online network with 'direct' effects. In Table 2, we investigate to what extent such effects are present and, consequently, bias our results.

We start by adding contemporaneous and lagged Facebook outages in the home county to our preferred 2SLS-IV specification. Column (1) shows that contemporaneous outages are negatively associated with protest risk, while past outages are not statistically significant. This outage-induced reduction in the protest risk is consistent with the narrative that Facebook may play an important role in coordinating local protests (Karduni and Sauda, 2020). Location and timing of gatherings, for example, are often disseminated via online platforms. The size of the point estimate implies that an outage in the own county decreases the likelihood of collective action by 1.8 percentage points. Crucially, however, controlling for own outages barely changes the size of our main coefficients of interest. In columns (2)–(4), we provide additional evidence against (spatially) correlated outages influencing our estimates. First, we drop observations when a home county experiences an outage. Second, we exclude all counties that ever experience an outage over our sample period. Third, we construct our instrument using only counties that are at least 500 kilometres away from a given home county. Reassuringly, these three modifications produce results that are very similar compared to our

Table 2: Effects of exposure to BLM protests in online social network

	Any BLM Protest					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Second stage					
Lag 1 network exposure to BLM protests	1.054**	0.957**	0.636**	1.048**	0.969**	2.143***
	(0.422)	(0.425)	(0.314)	(0.496)	(0.433)	(0.504)
Lag 2 network exposure to BLM protests	1.009**	1.090***	0.599*	0.908*	1.048**	
	(0.415)	(0.410)	(0.305)	(0.491)	(0.425)	
Own Facebook outage	-0.018***					
	(0.005)					
Lag 1 own Facebook outage	-0.001					
	(0.005)					
Lag 2 own Facebook outage	0.004					
	(0.005)					
	Panel B: First stage					
	Lag 1					
Lag 1 Facebook connection stability	0.048***	0.048***	0.038***	0.048***	0.047***	
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	
Lag 2 Facebook connection stability	0.005	0.005	-0.001	0.004	0.004	
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	
Lag 1 internet connection stability						0.031***
						(0.006)
	Lag 2					
Lag 1 Facebook connection stability	-0.006	-0.005	-0.003	-0.004	-0.007	
	(0.005)	(0.005)	(0.007)	(0.006)	(0.005)	
Lag 2 Facebook connection stability	0.049***	0.048***	0.040***	0.049***	0.047***	
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	
Instrumental variable	Main	Main	Main	500km	Main	Internet outages
Sample	Full	No outage	Never outage	Full	Full	Full
Similarity-weighted exposure	No	No	No	No	Yes	No
State × weekend fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Distance-weighted exposures	Yes	Yes	Yes	Yes	Yes	Yes
Observations	192,758	190,458	124,434	192,758	192,758	192,758
Mean dependent variable	3.261	3.224	0.643	3.261	3.261	3.261
First-stage F statistic lag 1	101.575	99.509	50.395	83.350	99.080	31.385
First-stage F statistic lag 2	96.526	95.335	56.896	82.170	95.688	
Joint first-stage F-statistic	44.890	44.270	18.335	37.158	43.928	

Notes: Panel A of this table reports the second-stage estimates of equation(4) using 2SLS-IV. Panel B reports the corresponding first-stage estimates (equation (5)). Standard errors clustered at the county level are reported in parentheses. 'Any BLM Protest' is a dummy equal to one if at least one BLM protest takes place in a given county and day. 'Network exposure to BLM protests' is the online network-proximity-weighted average protest incidence as defined in equation (2). 'Own Facebook connection stability' is a dummy equal to one if there are no outages in a given county and day. 'Facebook connection stability' is the online network-proximity-weighted average Facebook connection stability. 'Internet connection stability' is the online network-proximity-weighted average internet connection stability. Variables are defined according to equation (3). See Appendix D.1 for more details. All variables multiplied by hundred to facilitate interpretation. 'Similarity-weighted exposure' represents the similarity-weighted measures of exposure to protest activity. Included are measure of similarity along the following county characteristics: share of population with a graduate degree, share of population aged 15–25, share of black population, income per capita, and Democratic vote share. Values of the Kleibergen-Paap F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

main estimates.

Another threat to our identification strategy is that our connectedness measure—i.e., the relative frequency of Facebook friendships—may reflect the degree of overlap in the socio-demographic profiles of counties. For example, more similar counties in terms of age, political attitudes, and ethnicity are likely to be more connected. This raises the concern that our estimates could capture the diffusion of the protest movement across socio-economically similar counties rather than diffusion driven by frequency of individual connections. To address this, we construct a range of socio-economic similarity measures at the bilateral county level and use these to compute similarity-weighted protest exposure, following the procedure described in equation (2). Specifically, we construct similarity measures using the share of the population with a graduate degree, the share of Black residents, the share of population aged 15–25, income per capita, and the Democratic vote share. Appendix D.2 provides further details on the construction of these measures. It also shows that correlation is low across all measures, indicating that similarities in county characteristics alone are not particularly predictive of the Facebook connectedness. Reflecting this, column (5) shows that the point estimates for both lags of our Facebook exposure measure change only marginally when accounting for similarity-weighted exposure to the BLM movement.

A further concern is that outages are endogenous to protest activity. For example, more people could be trying to access Facebook on protest days. This could increase the number of reports sent to Downdetector, thereby biasing our estimates. The fact that outages are negatively correlated with local protest activity makes this unlikely (see column (1) of Table 2). To additionally mitigate worries, we construct an alternative instrument which is based on internet outages. Compared to our main instrument, the only difference being that we exploit local internet outages rather than Facebook outages as the source of variation. The outage measure does not rely on reported issues, but rather on measured internet stability. However, in the context of our analysis, this alternative instrument is associated with two drawbacks. First, the internet-based instrument captures the effects of online networks generally. This includes Facebook and other (correlated) online networks. Second, there is a substantial degree of autocorrelation in the internet outages.[30] As a consequence, the various lags of the instrument are considerably correlated. This implies that we are no longer able to cleanly separate the effects of different lags. In column (6) we therefore restrict our regression setup to only include the first lag.[31] The internet-outage-driven approach produces estimates that

[30]The correlation between the contemporaneous and lagged internet outage is 0.79, while the corresponding correlation for Facebook outages is 0.06.

[31]We describe the construction of this instrument in more detail in Appendix D.1. Empirically, the correlation between the first lag of the internet and Facebook outage driven instruments is very limited: the unconditional correlation is -0.069 while the correlation is 0.010 after partialling out county and date fixed effects. Table D.1 cross-tabulates the Facebook and internet stability indicators. The overlap in outages is minimal: A county simultaneously experiences a Facebook and internet outage in only 0.003 percent of all cases (i.e. observations).

are qualitatively equivalent to our main results, providing further evidence that online social networks have played an important role in the spread of the BLM protests.

The finding that Facebook outages reduce the local protest risk (cf. column (1) of Table 2) raises a concern related to the plausibility and interpretation of the exclusion restriction. Specifically, the dominant mechanism linking variation in our instrument to local BLM activity could simply be an outage-induced reduction in protest incidence among the network neighbours. Rather than reducing the visibility of protests that take place in neighbouring counties, outages could reduce the probability that they occur in the first place. While we cannot directly test for the relevance of this mechanism, we do not believe it is the primary driver underlying our results. Even when Facebook outages in neighbouring counties reduce the protest risk there, they still induce a Facebook-connectedness-specific shock in the home county due to the SCI-based weighting. General effects of lower protest activity and Facebook outages in neighbouring counties should be absorbed by the distance-weighted exposure controls.

Another worry—also pertaining to the validity of the exclusion restriction—is that Facebook outages in network neighbours could reduce the relevance of the Facebook feed for users in the home county. This could lead them to spend less time on Facebook which, in turn, could impact protest organisation and participation. In this case, the mechanism linking outages to protest activity would operate via reduced Facebook usage rather than reduced exposure to BLM protests. Unfortunately, we do not have data with sufficient granularity to test the relevance of this alternative channel. However, the fact that outages in the network neighbours impact protest activity in the home county with a one- to two-day lag provides some (limited) evidence against the usage mechanism being of central importance. As documented in column (1) of Table 2, an outage in the home county—which directly reduces Facebook usage—reduces the protest risk on only the day of the outage. This suggests that effects driven by reduced usage would likely materialise contemporaneously.

A final, more nuanced question relating to the mechanisms underlying our results is the extent to which the diffusion of protest between counties is driven by person-to-person exchange of content or by the Facebook algorithm serving individuals more content from more closely connected counties (i.e. without individuals interacting directly). While we lack data that would allow us to disentangle different channels of diffusion, our estimates will capture Facebook-specific effects as long as the diffusion process—whether inter-personal or algorithm-driven—is dependent on the connectedness of the counties.

Taken together, the results presented in Table 2 provide support for the validity of our identifying assumption, i.e., that connection issues reduce the degree of online interaction between counties and thereby the exposure to protests within the online network. As a final way of substantiating this interpretation, we use Google search data to show that online search behaviour depends on exposure to protests within the online network. The volume of web searches for BLM-related keywords in a county surges when exposure to protest within the online network increases. This suggests that awareness and interest in the social movement increases. This, in turn, helps explain the rise in protest probability. Appendix C provides more details on the Google search data and the results.

Robustness

A number of robustness checks show that our findings are not driven by specific assumptions or data construction choices. The results of the subsequently discussed exercises are presented in Appendix D.

Table D.3, column (1) shows that estimates are very similar if we drop counties in the bottom 10 percent of the population size distribution. When we drop the top 10 percent of counties in column (2), we also obtain qualitatively equivalent results. However, due to the fact that around 57% of all protests take place in these counties, we lose statistical power. As a consequence, the first lag no longer is statistically significant at conventional confidence levels, the second lag is borderline statistically significant at the 95% level. It is worthwhile noting, though, that evaluated at the mean of the outcome variable, the size of the point estimates is extremely similar compared to our main estimates. Columns 3–4 of Table D.3 show that results remain qualitatively unchanged if we shorten or extend our sample period by one week. To provide further evidence that our coefficients are not biased due to unobserved time-varying factors, we illustrate that our results remain stable if we control for a range of (time-interacted) county-level characteristics (column 5).[32] This is also the case when we account for state×week effect to account for the possibility that (Covid-19-induced) gatherings allowance was varying by state over time (column 6). Our main coefficients of interest are also left virtually unchanged if we control for contemporaneous and lagged internet outages in the home county in column (7). Our estimates further remain unaffected if we control for a wide range of alternative geodesic distance-based protest exposure measures (column 8).[33] The last three columns illustrate that our results are not dependent on a particular clustering approach. In column (9), we employ a two-way clustering approach that allows for correlation along the county and state×week dimension to account for potential correlation across counties within the same state and time period (e.g., due to state-level COVID-19 measures). In column (10), we adjust standard errors to account for spatial (Conley, 1999) and temporal (Newey and West, 1994) autocorrelation.[34] In the last column, we adapt the Conley (1999) approach to the structure of the Facebook social network and allow for correlation to decay in network distance (see Acemoglu et al., 2015 for a similar approach).[35]

[32] These include demographic, economic and political county-level characteristics.

[33] Table A.3 lists the different weighting approaches we control for.

[34] We allow for spatial autocorrelation within a radius of up to 1,000 km and temporal correlation within a window of up to 5 days. Results are stable when varying these cut-off values.

[35] Specifically, we use the bilateral connectedness weights (as defined in equation (1)) as weighting kernel

Table D.4 directly speaks to the worry that the results may be specific to the construction of our main explanatory variable and the instrument. Columns (1)–(2) indicate that our findings are not contingent on the precise choice of the threshold value when defining Facebook outages. We obtain similar coefficients when varying the threshold between 1.5 and 2.5 standard deviations. As documented in column (3), this is also the case when we use the number of reports relative to the county population to identify severe connection issues. Akin to Müller and Schwarz (2021), we define an outage as a situation in which the ratio of reported Facebook problems to population lies in the top decile.

Another concern related to the construction of our main explanatory variables is that the insights might depend on the (relative) weighting of the bilateral Facebook connections, $\omega_{c,j}$ (see equation 1). To mitigate this, we use the absolute (log) SCI index—without normalising it by the county's total number of connections—in the construction of the protest exposure measure and the instrumental variable.[36] The resulting 2SLS-IV estimates are reported in column (4). Consistent with our main results, we find that past exposure to protests within the Facebook network statistically significantly increases the protest risk in the home county. However, the log transformation results in a higher degree of cross-correlation between the variables, implying that we can no longer cleanly separate the effects of the two lags.

A further concern is that the strength of our (first-stage) results may be driven by a mechanical link between our endogenous and instrumental variables. For both variables, part of the variation is driven by Facebook outages. To alleviate this worry, we document that we obtain qualitatively equivalent results when we disregard Facebook outages in the construction of the endogenous variable. In this case, variation is only driven by the (time-invariant) structure of the Facebook network and changing protest incidence in neighbouring counties.[37] The last column of Table D.4 shows that our first-stage relationship is not due to a mechanical relationship. The resulting first-stage as well as the second-stage point estimate of the 2SLS-IV procedure is positive and highly statistically significant.

As a final exercise, we illustrate that our findings are not the result of chance. To this end, we randomly permutate the Facebook outages across counties (where, for each day, the number of randomly assigned outages corresponds to the actual number of outages observed in the data). We then construct our instrument as described in Section 2 and estimate the reduced-form model (4). We repeat this exercise 1,000 times and present the results in Figure D.1. Point estimates for both lags are centred around zero and orders of magnitude smaller than the true size of the reduced-form coefficient.

and additionally allow for temporal correlation within a window of up to 5 days.

[36] Again, it is worth noting that absolute values of the SCI are not meaningful in the sense that they cannot be directly interpreted as the number of Facebook user links between counties.

[37] Formally, the network exposure measure becomes: $P^N_{c,d} = \sum_{j \notin c}^{J} \omega_{c,j}\, p_{j,d}$, where $\omega_{c,j}$ is the time-invariant connectedness between individual county-pairs and $p_{j,d}$ is an indicator taking the value one if a BLM protest takes place in county j on day d.

5 Conclusion

This paper empirically assesses the role of online social networks—Facebook in particular—in explaining the diffusion of social movements. In the context of the largest wave of protests in recent years—the BLM protests triggered by George Floyd's death—we show that exposure within the online network influences the spatio-temporal dynamics of the movement. Protests are more likely to spill over between counties the closer they are connected within the Facebook network. This implies that online social networks play an important role in shaping the geography of social change.

References

Acemoglu, Daron, Camilo García-Jimeno, and James A. Robinson (2015) 'State capacity and economic development: A network approach.' *American Economic Review* 105(8), 2364–2409

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya (2015) 'Raido and the rise of the Nazis in prewar Germany.' *The Quarterly Journal of Economics* 130(4), 1885–1940

Aidt, Toke S, Gabriel Leon, and Max Satchell (2022) 'The social dynamics of collective action: Evidence from the Captain Swing riots, 1830-31.' *The Journal of Politics* 84(1), 209–225

Anderson, Monica, and Paul Hitlin (2016) 'The hashtag #blacklivesmatter emerges: Social activism on Twitter.' Technical Report, Pew Research Center

Aridor, Guy, Rafael Jiménez-Durán, Ro'ee Levy, and Lena Song (2024) 'The economics of social media.' *Journal of Economic Literature* 62(4), 1422–74

Azzimonti, Marina, and Marcos Fernandes (2022) 'Social media networks, fake news, and polarization.' *European Journal of Political Economy* p. 102256

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert J Richmond, and Johannes Stroebel (2021) 'International trade and social connectedness.' *Journal of International Economics*

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong (2018a) 'Social connectedness: Measurement, determinants, and effects.' *Journal of Economic Perspectives* 32(3), 259–80

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel (2018b) 'The economic effects of social networks: Evidence from the housing market.' *Journal of Political Economy* 126(6), 2224–2276

Barberá, Pablo, Richard Bonneau, Megan Macduffee Metzger, Jonathan Nagler, Duncan Penfold-Brown, and Joshua A. Tucker (2016) *Big Data, Social Media, and Protest: Foundations for a Research Agenda* Analytical Methods for Social Research

Black Lives Matter (2023) 'Black lives matter: About'

Boehmke, Frederick J., Samuel M. Avery, Marissa S. Good, Thomas C. Dainty, and Hyein Ko (2023) 'Did Black Lives Matter Protests Change Public Opinion?' *American Politics Research* 51(6), 683–700

Buchanan, Larry, Quoctrung Bui, and Jugal K. Patel (2020) 'Black Lives Matter May Be the Largest Movement in U.S. History'

Bursztyn, Leonardo, Georgy Egorov, Ruben Enikolopov, and Maria Petrova (2024) 'Social media and xenophobia: Evidence from Russia.' Working Paper 26567, National Bureau of

Economic Research, December

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

Campbell, Travis (2024) 'Black Lives Matter's effect on police lethal use of force.' *Journal of Urban Economics* 141, 103587

Cantoni, Davide, David Y Yang, Noam Yuchtman, and Y Jane Zhang (2019) 'Protests as strategic games: Experimental evidence from Hong Kong's antiauthoritarian movement.' *The Quarterly Journal of Economics* 134(2), 1021–1077

Cao, Andy, Jason M. Lindo, and Jiee Zhong (2023) 'Can social media rhetoric incite hate incidents? evidence from Trump's "Chinese Virus" tweets.' *Journal of Urban Economics*

Caprettini, Bruno, Marcel Caesmann, Hans-Joachim Voth, and David Yanagizawa-Drott (2021) 'Going viral: Propaganda, persuasion and polarization in 1932 Hamburg.' Technical Report, University of Zurich

Conley, T. G. (1999) 'GMM estimation with cross sectional dependence.' *Journal of Econometrics* 92(1), 1–45

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

Drakulich, Kevin, and Megan Denver (2022) 'The Partisans and the Persuadables: Public Views of Black Lives Matter and the 2020 Protests.' *Perspectives on Politics* pp. 1–18. Publisher: Cambridge University Press

Dunivin, Zackary Okun, Harry Yaojun Yan, Jelani Ince, and Fabio Rojas (2022) 'Black Lives Matter protests shift public discourse.' *Proceedings of the National Academy of Sciences* 119(10), e2117320119. Publisher: Proceedings of the National Academy of Sciences

Ebbinghaus, Mathis, Nathan Bailey, and Jacob Rubel (2021) 'Defended or defunded? Local and state policy outcomes of the 2020 Black Lives Matter protests'

Eichenauer, Vera Z., Ronald Indergand, Isabel Z. Martínez, and Christoph Sax (2022) 'Obtaining consistent time series from google trends.' *Economic Inquiry* 60(2), 694–705

Engist, Oliver, and Felix Schafmeister (2022) 'Do political protests mobilize voters? Evidence from the Black Lives Matter protests.' *Public Choice* 193(3), 293–313

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

Fergusson, Leopoldo, and Carlos Molina (2021) 'Facebook Causes Protests.' Documentos CEDE 18002, Universidad de los Andes, Facultad de Economía, CEDE, April

Freelon, Deen, Charlton D. McIlwain, and Meredith Clark (2016) 'Beyond the hashtags: #ferguson, #blacklivesmatter, and the online struggle for offline justice.' *Center for Media &*

Social Impact

Fujiwara, Thomas, Karsten Müller, and Carlo Schwarz (2023) 'The effect of social media on elections: Evidence from the United States.' *Journal of the European Economic Association* 22(3), 1495–1539

García-Jimeno, Camilo, Angel Iglesias, and Pinar Yildirim (2022) 'Information networks and collective action: Evidence from the women's temperance crusade.' *American Economic Review* 112(1), 41–80

González, Felipe (2020) 'Collective action in networks: Evidence from the Chilean student movement.' *Journal of Public Economics* 188, 104220

Halberstam, Yosh, and Brian Knight (2016) 'Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter.' *Journal of Public Economics* 143, 73–88

Hamzelou, Jessica (2020) 'In the decade since the Arab Spring, social media has changed how people protest.' *The Telegraph.* Accessed: February 22, 2023

Karduni, Alireza, and Eric Sauda (2020) 'Anatomy of a protest: Spatial information, social media, and urban space.' *Social Media + Society* 6(1), 2056305119897320

Kuchler, Theresa, Dominic Russel, and Johannes Stroebel (2022) 'Jue insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by Facebook.' *Journal of Urban Economics* 127, 103314. JUE Insights: COVID-19 and Cities

Larson, Jennifer M., Jonathan Nagler, Jonathan Ronen, and Joshua A. Tucker (2019) 'Social networks and protest participation: Evidence from 130 million Twitter users.' *American Journal of Political Science* 63(3), 690–705

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

Manski, Charles F. (1993) 'Identification of endogenous social effects: The reflection problem.' *The Review of Economic Studies* 60(3), 531–542

Mazumder, Soumyajit (2019) 'Black Lives Matter for Whites' racial prejudice: Assessing the role of social movements in shaping racial attitudes in the United States'

MIT Election Data and Science Lab (2018) 'County Presidential Election Returns 2000-2020'

Mortensen, M., C. Neumayer, and T. Poell (2018) *Social Media Materialities and Protest: Critical Reflections* (Taylor & Francis)

Müller, Karsten, and Carlo Schwarz (2021) 'Fanning the Flames of Hate: Social Media and Hate Crime.' *Journal of the European Economic Association* 19(4), 2131–2167

Müller, Karsten, and Carlo Schwarz (2023) 'From hashtag to hate crime: Twitter and anti-minority sentiment.' *American Economic Journal: Applied Economics* 15(3), 270–312

Mundt, Marcia, Karen Ross, and Charla M Burnett (2018) 'Scaling Social Movements

Through Social Media: The Case of Black Lives Matter.' *Social Media + Society* 4(4), 2056305118807911. Publisher: SAGE Publications Ltd

Mutz, Diana C. (2022) 'Effects of changes in perceived discrimination during BLM on the 2020 presidential election.' *Science Advances* pp. 1–9

Newey, Whitney K., and Kenneth D. West (1994) 'Automatic lag selection in covariance matrix estimation.' *The Review of Economic Studies* 61(4), 631–653

Ovide, Shira (2020) 'How social media has changed civil rights protests.' *The New York Times.* Accessed: February 22, 2023

Pew Research Center (2020a) 'Recent Protest Attendees Are More Racially and Ethnically Diverse, Younger than Americans Overall'

_ (2020b) 'Activism on social media varies by race and ethnicity, age, political party'

_ (2021) 'Social Media Use in 2021'

Qin, Bei, David Strömberg, and Yanhui Wu (2024) 'Social media and collective action in China.' *Econometrica* 92(6), 1993–2026

Quan, Lin, John Heidemann, and Yuri Pradkin (2013) 'Trinocular: Understanding internet reliability through adaptive probing.' In 'Proceedings of the ACM SIGCOMM Conference' ACM Hong Kong, China pp. 255–266

Smith, Braeton, Matthew Riddle, Amanda Wagner, Lesley Edgemon, Carmella Burdi, and Iain Hyde (2021) 'County Economic Impact Index.' *Argonne National Laboratory*

The Economist (2021) 'Digital media fuel global protests but can be used against them'

_ (2022) 'Since George Floyd's murder, new ways of policing have been spreading'

Tillery, Alvin B. (2019) 'What kind of movement is Black Lives Matter? The view from Twitter.' *The Journal of Race, Ethnicity, and Politics* 4(2), 297–323

Wang, Xiaohong (2022) 'The Effect of Black Lives Matter Protests on Racial Disparities in Nonfatal Police-Civilian Interactions'

Yanagizawa-Drott, David (2014) ' Propaganda and Conflict: Evidence from the Rwandan Genocide.' *The Quarterly Journal of Economics* 129(4), 1947–1994

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov (2020) 'Political effects of the internet and social media.' *Annual Review of Economics* 12(1), 415–438

Appendices

A Data sources and construction

Table A.1: Descriptive Statistics key variables

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
Any BLM Protest	3.261	17.760	0	100	192,758
Lag network exposure to BLM protests	2.071	4.384	0	63.254	192,758
Lag Facebook connection stability	99.131	1.948	57.111	99.996	192,758
Own Facebook outage	1.193	10.858	0	100	192,758

Table A.2: Additional variables: definition and sources

Variable	Definition	Sources
Share black population	Percent of total population	US Census Bureau
Police violence	Police shootings of non-white individuals 2016–2019 (normalised by population)	`mappingpoliceviolence.us`
Poverty rate	Percent of total population	DataUSA
Population size	Total population	US Census Bureau
Covid cases per capita	Cumulative cases up to 25 May 2020 (normalised by population)	New York Times
Economic impact	County Economic Impact Index	Smith et al. (2021)
Vote share Dem vs Rep	Vote share Democrats vs vote share Republicans in 2016 presidential election	MIT Election Data and Science Lab (2018)
Share graduates	Percent of population with graduate degree	American Community Survey
Share population aged 15–25	Percent of population aged 15–25	American Community Survey
Income per capita	Income per capita	Census Quick Facts

Table A.3: Alternative distance weights

Measure	Correlation with SCI
Normalised inverse distance	0.585
Inverse distance	0.566
Inverse log(distance)	0.332
$e^{-\text{distance}}$	0.029
Population-weighted inverse distance	0.016

Notes: Table reports bilateral correlation between various geographic distance measures and Facebook connectedness weight, as defined in equation (1).

B Additional results

Table B.1: 2SLS-IV regressions separate by lag

Lag	Point estimate γ_{d-k}	First-stage F-statistic	Observations
$d-5$	-0.083 (-0.540)	152.157	192,758
$d-4$	-0.480 (0.508)	178.351	192,758
$d-3$	-0.486 (0.471)	198.840	192,758
$d-2$	0.267 (0.442)	181.732	192,758
$d-1$	0.102 (0.437)	204.585	192,758
d	0.674 (0.457)	197.782	192,758
$d+1$	1.052** (0.423)	175.333	192,758
$d+2$	1.402*** (0.441)	171.796	192,758
$d+3$	0.282 (0.563)	177.920	192,758
$d+4$	-0.513 (0.499)	186.232	192,758
$d+5$	-0.224 (0.481)	179.678	192,758

Notes: Table reports the 2SLS-IV estimates of regression equation (4). Each row represents a separate regression for each day within the 10-day time period centred around day d, i.e., for each day $d-k$, where $k = [-5, 5]$. The dependent variable is a dummy equal to 100 if at least one BLM protest takes place in a given county and day, and zero otherwise. The coefficient reported (γ_{d-k}) is the second-stage point estimate of online network-proximity-weighted average protest incidence, as defined in equation (2). Standard errors clustered at the county level are reported in parentheses. Values of the Kleibergen-Paap F-statistics are reported. All regressions account for county fixed effects, date fixed effects, state $\times$ weekend fixed effects, and distance-weighted exposure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

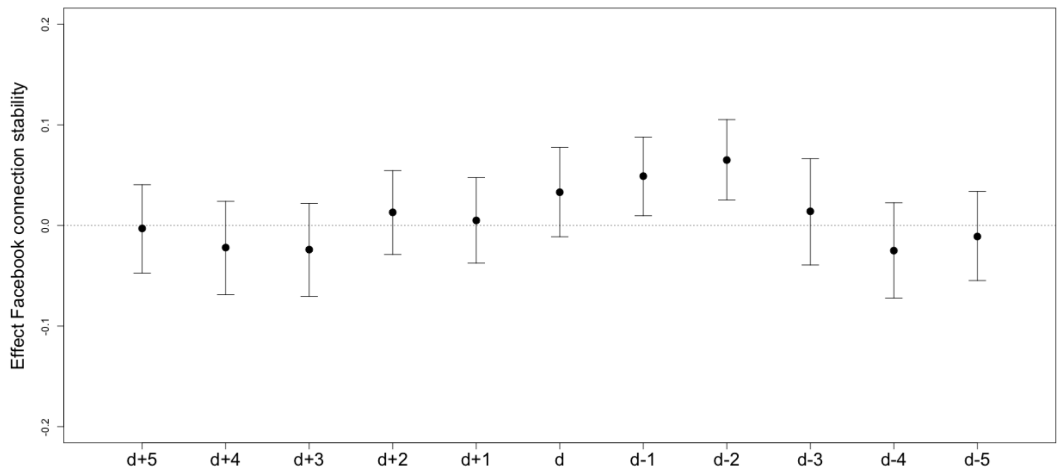

Figure B.1: Temporal dynamics of network connection stability
Figure depicts reduced-form point estimates and 95% confidence intervals of the effect of Facebook connection stability on protest probability. Point estimates obtained by separately estimating the reduced-form version of regression model (4) for each day within the 10-day time period centred around day *d*. Standard errors are clustered at the county level. Dependent variable equals 100 if a protest takes place in a given county and day, and zero otherwise. 'Facebook connection stability' is the online network-proximity-weighted connection stability, as defined in equation (3).

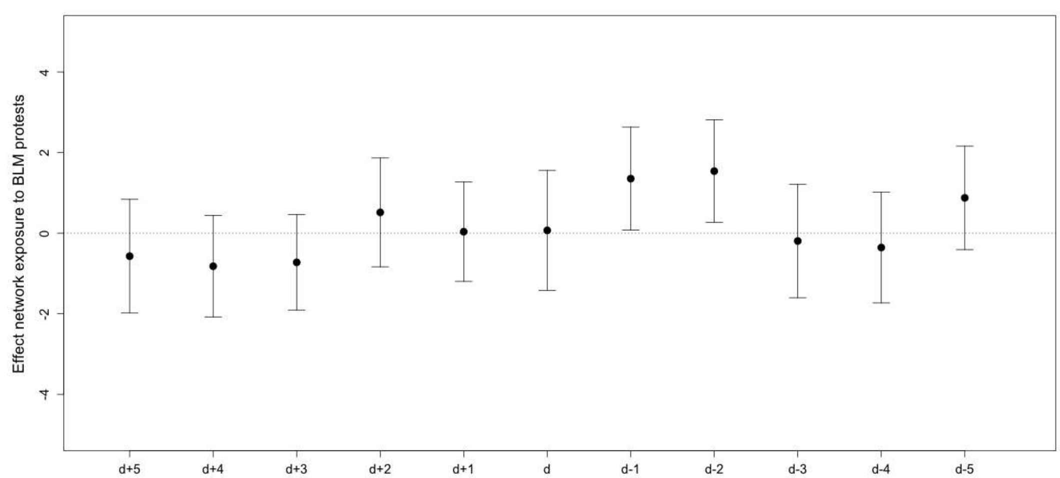

Figure B.2: Temporal dynamics of network exposure
Figure depicts 2SLS point estimates and 95% confidence intervals of the effect of network exposure on protest probability. Point estimates obtained by estimating regression model (4) and simultaneously include all leads and lags within the 10-day time period centred around day *d*. Standard errors are clustered at the county level. Dependent variable equals 100 if a protest takes place in a given county and day, and zero otherwise. 'Network exposure to BLM protests' is the online network-proximity-weighted average protest incidence as defined in equation (2).

Table B.2: Reduced-form effects of exposure to BLM protests in online social network

	Any BLM Protest			
	(1)	(2)	(3)	(4)
	Reduced form corresponding to Table 1			
Lag 1 Facebook connection stability			0.045**	0.042**
			(0.020)	(0.020)
Lag 2 Facebook connection stability			0.054***	0.047**
			(0.020)	(0.019)
County fixed effects	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes
State × weekend fixed effects	Yes	Yes	Yes	Yes
Distance-weighted exposures	No	Yes	Yes	Yes
Observations	192,758	192,758	192,758	192,758
Mean dependent variable	3.261	3.261	3.261	3.261

Notes: Table reports the reduced-form estimates corresponding to the 2SLS-IV estimates presented in Panel A of Table 1. Standard errors clustered at the county level are reported in parentheses. 'Any BLM Protest' is a dummy equal to one if at least one BLM protest takes place in a given county and day. 'Facebook connection stability' is the online network-proximity-weighted average Facebook connection stability. Variable is defined according to equation (3). All variables multiplied by hundred to facilitate interpretation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Comparison of effect sizes across studies

Study	**Effect Size**	**What We Find**
Qin et al. (2024) China - Sina Weibo	One percentage point increase in protest exposure → 34% increase in protest incidence relative to the mean.	Our results closely align: one percentage point increase in protest exposure → 31% increase in protest incidence relative to the mean. The confidence intervals from both studies overlap.
Fergusson and Molina (2021) Worldwide - Facebook	1 SD increase in Facebook availability → 0.11 SD increase in protest counts.	Our estimated effect is about two times larger, suggesting a stronger influence of network exposure. However, confidence intervals from both studies overlap.
Enikolopov et al. (2020) Russia - VKontakte	10% increase in VK users → 4.5 percentage points increase in protest likelihood equivalent to 6% evaluated at the mean.	Our estimated effect is similar in size: 10% increase in network exposure → 0.2 percentage points increase in protest activity equivalent to 6% evaluated at the mean.
García-Jimeno et al. (2022) U.S. - Temperance Movement	Protests in two closest railway network neighbours → 5.6× increase in local protest likelihood	Our estimate is slightly larger (7.3×), indicating stronger network effects within modern digital networks.

C Google trends

We use Google Trends data to (a) validate our outage measure and (b) investigate if network exposure to protests increases the volume of BLM-related web searches. The Google trends database provides information on the daily number of search requests made to Google, which can be stratified by keywords and regions. Google does not report search volumes at the county level but rather for Designated Market Areas (DMA), a higher level of aggregation. We therefore conduct the subsequent analysis at the DMA$\times$day level. The data construction process, empirical methodology, and results are described below.

Data

For our sample period, we extract daily search volumes for each of the 204 DMAs located in the contiguous US.[38] We use the following keywords to proxy for searches related to Facebook outages: 'Facebook down', 'Facebook issues', 'Facebook problems', 'Facebook connection', 'Facebook access', and 'Facebook outage'. We follow the same procedure to identify searches related to outages on the five other most popular social media platforms in the US: YouTube, Instagram, TikTok, Snapchat, and Twitter (see doofinder.com).

To check for changes in the volume of BLM-related searches, we focus on the following keywords: 'Black Lives Matter', 'George Floyd', 'antiracism', 'decolonization', and 'defund the police'. The search volumes represent a largely unfiltered sample of actual search requests, where Google normalises search volumes for each query such that the highest value is equal to 100 and the lowest value is zero. To make the data consistent across queries and reduce sampling noise, we apply the procedure developed in Eichenauer et al. (2022). The result is a DMA$\times$day-level index that consistently captures changes in searches for the BLM keywords.

Because the DMAs are larger spatial units than the counties, we need to aggregate the bilateral SCI index to the DMA-pair level. We use the DMA-county matching provided by Jacob Schneider to map counties into DMAs.[39] We then aggregate the SCI (i.e., Facebook connectedness) to the DMA-pair level. In analogy to the procedure described in Section 2, we subsequently compute the social-network-proximity-weighted average protest incidence as well as the outage-based instrumental variable for each county and day.

The final dataset—combining the Google trends, Facebook outages, and protest exposure data—consists of 12,648 observations.

[38] On average, a DMA encompasses 15 counties.

[39] See https://sites.google.com/view/jacob-schneider/resources.

Empirical framework

To validate our main outage measure, we analyse if local Facebook outages trigger corresponding web searches using the following OLS model:

$$G^{F}_{a,d} = \theta\, O_{a,d} + \delta_a + \zeta_d + \chi_{a,d}, \tag{C.1}$$

where $G^{\mathrm{F}}_{a,d}$ is the Google trends search index for outage keywords for a given online social network platform F in Designated Market Areas a on day d. The variable $O_{a,d}$ represents the share of counties located within a given DMA a that experience a Facebook outage on day d. DMA fixed effects are symbolised by δ_a, date fixed effects by ζ_d. Error terms $\chi_{a,d}$ are clustered at the DMA level.

Turning to the investigation of spillover effects, we test if searches for BLM keywords vary with online network exposure to protest employing our usual 2SLS-IV approach. The second stage can be formally written as:

$$G^{\mathrm{BLM}}_{a,d} = \gamma\, P^{N}_{a,d-1} + \pi_a + \tau_d + \varepsilon_{a,d}, \tag{C.2}$$

where $G^{\mathrm{BLM}}_{a,d}$ is the Google trends search volume index for BLM keywords in DMAs a on day d. The main regressor of interest is $P^{N}_{a,d-1}$, the network-proximity-weighted average protest incidence in DMAs $j \neq a$ on the previous day $d-1$.[40] All regressions control for DMA fixed effects, π_a, and date fixed effects, τ_d.

The first-stage regression is given by:

$$P^{N}_{a,d} = \lambda\, S^{N}_{a,d} + \mu_a + \eta_d + \Omega\, \mathbf{X}_{a,d} + \psi_{a,d}. \tag{C.3}$$

In analogy to the main part, $S^{N}_{a,d}$ is the network-proximity-weighted Facebook access stability.

Results

Column (1) of Table C.1 reports the results of regression equation (C.1) using the search volume for Facebook outages as dependent variable. The estimate suggests that the searches related to Facebook outages increase by around 19.8 index points when moving from a day on which no outages occur within a DMA to a day on which the entire DMA experiences disruptions in access to the Facebook platform. In columns (2)–(6), we run the same regression, now using the search volumes related to outages of other popular social media platforms as outcomes. We do not observe that Facebook outages trigger searches about outages for connection issues of other platforms. Point estimates are statistically non-significant and small in size. The exception is Instagram, where we do find a positive effect, though weaker compared

[40] In keeping with the main analysis, we lag the explanatory variable. However, using contemporaneous values results in very similar estimates.

to column (1). This is not surprising, given that Facebook acquired Instagram in April 2012. It is therefore plausible that Instagram is also affected when Facebook experiences issues. Together, columns (1)–(6) suggest that our outage measure specifically captures connection issues of the Facebook platform.

In the last column of Table C.1, we investigate the relationship between exposure to protests in the online network and BLM-related web searches. The point estimate implies that a one percentage point increase in online exposure to protests raises the volume of BLM-related web searches by around 1.993 index points. This translates to an increase of around 42.45 points when the closest network neighbour experiences a protest.[41]

[41] On average, the proportion of total Facebook links a DMA shares with its closest network neighbour amounts to 21.3 percent.

Table C.1: Facebook outages and web searches

	Search volume Facebook outage	Search volume YouTube outage	Search volume Instagram outage	Search volume TikTok outage	Search volume Snapchat outage	Search volume Twitter outage	Search volume BLM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Second stage						
Share Facebook outage	0.198***	-0.040	0.136*	0.033	0.042	0.018	
	(0.063)	(0.063)	(0.070)	(0.031)	(0.047)	(0.035)	
Lag network exposure to BLM protests							1.993***
							(0.752)
	Panel B: First stage						
Lag Facebook connection stability							0.355***
							(0.077)
DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,648	12,648	12,648	12,648	12,648	12,648	12,648
Mean dependent variable	29.442	24.147	17.967	6.085	13.068	11.082	53.661
First-stage F-statistic							21.224

Notes: Panel A of this table reports the OLS estimates of equation (C.1) in column (1) and the second-stage estimates of equation (C.2) using 2SLS-IV in column (2). Panel B reports the corresponding first-stage estimate. Standard errors clustered at the DMA level are reported in parentheses. 'Search volume Facebook outage' is the Google search volume index for Facebook outage keywords in a given DMA and day. 'Search volume Instagram outage' is the Google search volume index for Instagram outage keywords in a given DMA and day. 'Search volume YouTube outage' is the Google search volume index for Youtube outage keywords in a given DMA and day. 'Search volume TikTok outage' is the Google search volume index for TikTok outage keywords in a given DMA and day. 'Search volume Snapchat outage' is the Google search volume index for Snapchat outage keywords in a given DMA and day. 'Search volume Twitter outage' is the Google search volume index for Twitter outage keywords in a given DMA and day. 'Search volume BLM' is the Google search volume index for BLM keywords in a given DMA and day. 'Share Facebook outage' represents the share of counties located within a given DMA that experience a Facebook outage. Variable is multiplied by 100 to facilitate interpretation. 'Network exposure to BLM protests' is the online network-proximity-weighted average protest incidence as defined in equation (2). Variable is multiplied by 100 to facilitate interpretation. 'Facebook connection stability' is the online network-proximity-weighted average Facebook connection stability. Variable is defined according to equation (3) and multiplied by 100 to facilitate interpretation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Robustness

D.1 Alternative instrumental variable

We construct our internet-outage-based instrument analogously to our main instrument (see Section 2). The only difference is that we use local internet outages (rather than Facebook outages) as a source of variation. Data on internet connection stability comes from the Center for Applied Internet Data Analysis (CAIDA). Specifically, we draw on hourly county-level data on internet connectivity. CAIDA derives this measure using active probing and the trinocular measurement and inference technique (Quan et al., 2013).[42] Trinocular (the outage detection system) sends pings to # /24 IPv4 network blocks to determine their activity. The system then measures if a block is either up, down, or uncertain. We observe the share of active blocks within a county. We aggregate the hourly data to the day level and create a dummy variable that equals one if there are no major internet connection issues, and zero if there is a serious outage. In analogy to Facebook stability, we define major issues as situations where values of the metric drop at least two standard deviations below the local (i.e., county) average.[43]

Using the internet stability indicator, we then construct the alternative IV as defined in equation (3). The only difference is that we exchange local Facebook stability in equation (2) with local internet outages (i.e., $s^{\text{CAIDA}}_{j,d}$).

Table D.1 below cross-tabulates Facebook and internet outages. There is extremely limited overlap between the two.

Table D.1: Facebook and Internet Outages (%)

		Internet Outage	
		0	1
Facebook Outage	0	98.511	0.296
	1	0.003	1.190

[42] See https://www.caida.org/projects/ioda/ for more information.

[43] Formally, the internet connection stability index $s^{\text{CAIDA}}_{j,d}$ is defined as:

$$s^{\text{CAIDA}}_{j,d} = \mathbf{I}\left(m_{j,d} \geq \mu_{m,j} + 2 \times \sigma_{m,j}\right),$$

where $m_{j,d}$ is the internet connection metric for county j on day d. The county-specific mean and standard deviations are represented by $\mu_{m,j}$ and $\sigma_{m,j}$

D.2 Similarity weights

We construct a range of socio-economic similarity measures at the bilateral county level and use these to compute similarity-weighted protest exposure, analogously to the procedure described in equation (2). Specifically, we construct similarity measures using the share of the population with a graduate degree, the share of Black residents, the share of population aged 15–25, the Democratic vote share, and income per capita. For all but income per capita, we compute the similarity between county c and j in characteristic Γ as $1 - |\Delta\Gamma_{c,j}|$. For income per capita, we use the absolute difference as the measure of (dis-)similarity, i.e. $|\Delta\Gamma_{c,j}|$. Table D.2 below lists the correlations between our Facebook-connectedness weight (as defined in equation (1)) and the respective similarity weight. Generally speaking, the correlations are low, documenting that similarities in county characteristics alone are not particularly predictive of the Facebook connectedness.

Table D.2: Similarity measures

Measure	**Correlation with SCI**
Share graduates	0.027
Share black	0.030
Vote share Democrats	0.031
Share population aged 15–25	0.004
Income per capita	-0.024

Notes: Table reports the bilateral correlations between the similarity measures and the Facebook connectedness weight (as defined in equation (1)). Data sources are listed in Table A.2.

Table D.3: Robustness effects of exposure to BLM protests in online social network (1)

	Any BLM Protest										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Second stage											
Lag 1 network exposure to BLM protests	1.013**	0.435	1.230***	0.949**	0.861*	0.917**	1.045**	0.890**	1.045**	1.045**	1.045**
	(0.433)	(0.386)	(0.415)	(0.458)	(0.452)	(0.453)	(0.421)	(0.452)	(0.469)	(0.474)	(0.426)
Lag 2 network exposure to BLM protests	1.041**	0.742*	1.056***	0.812*	1.113**	0.911**	1.003**	0.951**	1.003**	1.003**	1.003**
	(0.436)	(0.380)	(0.395)	(0.442)	(0.433)	(0.428)	(0.415)	(0.449)	(0.464)	(0.455)	(0.420)
Panel B: First stages											
Lag 1 network exposure to BLM protests											
Lag 1 Facebook stability	0.049***	0.044***	0.056***	0.041***	0.046***	0.046***	0.049***	0.045***	0.049***	0.049***	0.049***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)	(0.008)	(0.004)
Lag 2 Facebook stability	0.005	-0.002	0.011**	0.003	0.004	0.004	0.005	0.004	0.005	0.005	0.005
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.009)	(0.004)
Lag 2 network exposure to BLM protests											
Lag 1 Facebook stability	-0.003	-0.007	-0.004	-0.008*	-0.007	-0.008	-0.006	-0.008	-0.006	-0.006	-0.006
	(0.004)	(0.005)	(0.006)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.005)
Lag 2 Facebook stability	0.049***	0.044***	0.057***	0.041***	0.047***	0.047***	0.049***	0.044***	0.049***	0.049***	0.049***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.008)	(0.008)	(0.004)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance-weighted exposures	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	173,538	173,538	170,995	214,521	192,758	192,758	192,758	192,758	192,758	192,758	192,758
Mean dependent variable	3.617	1.552	3.556	3.008	3.261	3.261	3.261	3.261	3.261	3.261	3.261
First-stage F statistic lag 1	96.358	83.310	108.955	91.974	89.828	90.294	101.910	22.030	101.899		
First-stage F statistic lag 2	91.647	82.326	109.431	89.426	86.644	91.641	96.739	20.500	96.712		
Joint first-stage F-statistic	46.537	29.109	53.486	36.612	39.856	39.239	45.166	40.204	9.485	8.853	41.842
Robustness	Drop bottom 10% population	Drop top 10% population	−7 days	+7 days	County characteristics	State×Week FE	Internet outage	Distance controls	Double clustering	Conley distance	Conley network

Notes: Panel A reports the second-stage estimates of equation(4) using 2SLS-IV. Panel B reports the corresponding first-stage estimates (equation (5)). Standard errors clustered at the county level are reported in parentheses. 'Any BLM Protest' is a dummy equal to one if at least one BLM protest takes place in a given county and day. 'Network exposure to BLM protests' is the online network-proximity-weighted average protest incidence as defined in equation (2). 'Facebook connection stability' is the online network-proximity-weighted average Facebook connection stability. Variable is defined according to equation (3). All variables multiplied by hundred to facilitate interpretation. Values of the Kleibergen-Paap F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.4: Robustness effects of exposure to BLM protests in online social network (2)

	Any BLM Protest				
	(1)	(2)	(3)	(4)	(5)
	Panel A: Second stage				
Lag 1 network exposure BLM protests	0.691*	1.235***	1.338**	0.039***	2.040**
	(0.389)	(0.419)	(0.645)	(0.012)	(0.894)
Lag 2 2etwork exposure BLM protests	0.743*	1.090***	0.367	-0.009	1.688**
	(0.404)	(0.407)	(0.512)	(0.012)	(0.840)
	Panel B: First stages				
	Lag 1 network exposure to BLM protests				
Lag 1 Facebook connection stability	0.050***	0.051***	0.039***	0.301***	0.025***
	(0.003)	(0.003)	(0.004)	(0.020)	(0.004)
Lag 2 Facebook connection stability	0.007	0.014***	0.008*	0.296***	0.005
	(0.004)	(0.004)	(0.005)	(0.014)	(0.004)
	Lag 2 network exposure to BLM protests				
Lag 1 Facebook connection stability	-0.009*	0.051***	-0.006	0.041**	-0.004
	(0.005)	(0.003)	(0.006)	(0.017)	(0.005)
Lag 2 Facebook connection	0.049***	0.014***	0.044***	0.388***	0.026***
	(0.003)	(0.004)	(0.004)	(0.017)	(0.004)
County fixed effects	Yes	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes	Yes
Distance-weighted exposures	Yes	Yes	Yes	Yes	Yes
Observations	192,758	192,758	192,758	192,758	192,758
Mean dependent variable	3.261	3.261	3.261	3.261	3.261
First-stage F statistic lag 1	113.498	107.263	45.309	603.87	22.224
First-stage F statistic lag 2	104.813	100.152	52.669	1632.62	23.072
Joint first-stage F-statistic	51.887	98.367	28.501	729.058	11.562
Robustness	1.5SD outage threshold	2.5SD outage threshold	Top decile reports/population	Log absolute SCI weights	No outages endogenous variable

Notes: Panel A reports the second-stage estimates of equation(4) using 2SLS-IV. Panel B reports the corresponding first-stage estimates (equation (5)). Standard errors clustered at the county level are reported in parentheses. 'Any BLM Protest' is a dummy equal to one if at least one BLM protest takes place in a given county and day. 'Network exposure to BLM protests' is the online network-proximity-weighted average protest incidence as defined in equation (2). 'Facebook connection stability' is the online network-proximity-weighted average Facebook connection stability. Variable is defined according to equation (3). All variables multiplied by hundred to facilitate interpretation. Values of the Kleibergen-Paap F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

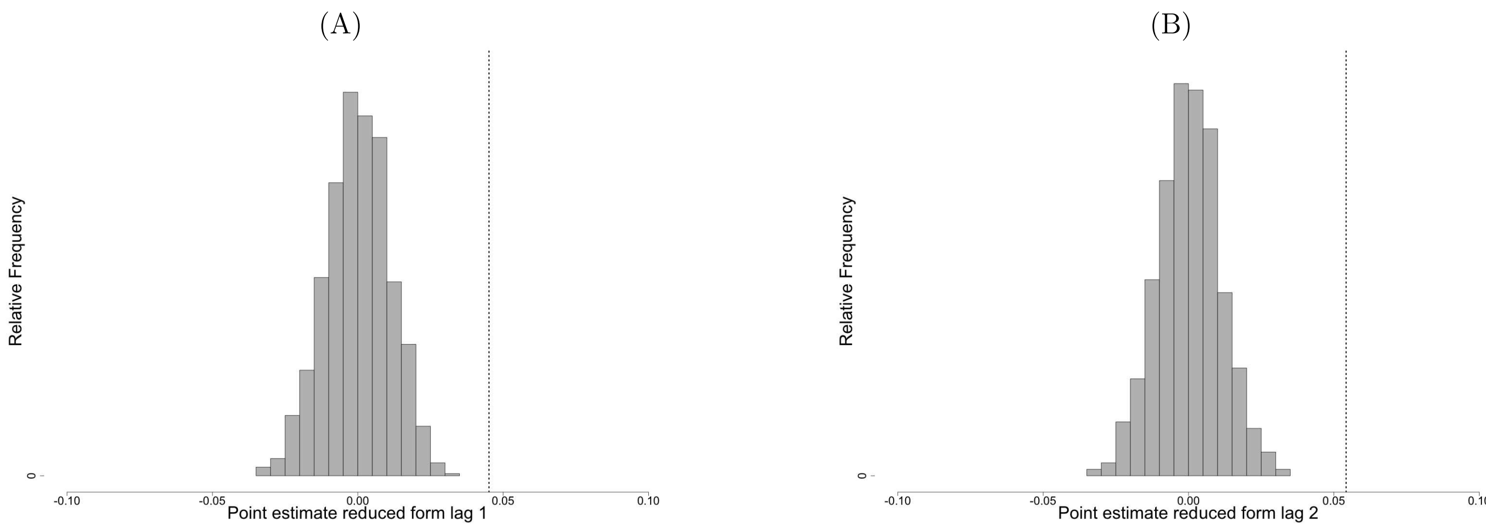

Figure D.1: Random outages
Figure depicts the distribution of the reduced-point estimates obtained from 1,000 random permutation of Facebook outages across counties. Panel A reports the estimates for lag 1 of Facebook stability. Panel B reports the estimates for lag 2 of Facebook stability. The dashed black vertical lines represent the point estimates obtained using the actual outage data (see Table B.2, column (3)).