DISCUSSION PAPER SERIES

DP16731

Social Media and Collective Action in China

Yanhui Wu, David Strömberg and Bei Qin

DEVELOPMENT ECONOMICS

ISSN 0265-8003

Social Media and Collective Action in China

Yanhui Wu, David Strömberg and Bei Qin

Discussion Paper DP16731
Published 17 November 2021
Submitted 16 November 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Social Media and Collective Action in China

Abstract

This paper studies how social media affect protest dynamics in China during 2009-2017. Based on 13.2 billion microblog posts, we use tweets and retweets to measure social media communication across cities and exploit its rapid expansion for identification. We find that despite strict government control, Chinese social media have a sizeable effect on the geographic spread of protests and strikes. While the spread effect is short-lived and predominantly between similar events, social media considerably increase the scope of protests. Further empirical results and textual analysis show that the effect is likely to be driven by tacit coordination and emotional reactions rather than explicit coordination and sharing tactics. Our study sheds light on the debate regarding whether social media help strengthen authoritarian regimes.

JEL Classification: N/A

Keywords: N/A

Yanhui Wu - yanhuiwu@hku.hk
University of Hong Kong and CEPR

David Strömberg - david.stromberg@ne.su.se
Stockholm University

Bei Qin - beiqin@hkbu.edu.hk
Hong Kong Baptist University

Social Media and Collective Action in China

Bei Qin, David Strömberg, and Yanhui Wu*

September 2021

Abstract

This paper studies how social media affect protest dynamics in China during 2009-2017. Based on 13.2 billion microblog posts, we use tweets and retweets to measure social media communication across cities and exploit its rapid expansion for identification. We find that despite strict government control, Chinese social media have a sizeable effect on the geographic spread of protests and strikes. While the spread effect is short-lived and predominantly between similar events, social media considerably increase the scope of protests. Further empirical results and textual analysis show that the effect is likely to be driven by tacit coordination and emotional reactions rather than explicit coordination and sharing tactics. Our study sheds light on the debate regarding whether social media help strengthen authoritarian regimes.

*Qin: Economics Department, Hong Kong Baptist University, Hong Kong, beiqin@hkbu.edu.hk; Strömberg: Economics Department, University of Stockholm, SE-106 91 Stockholm, Sweden, david.stromberg@ne.su.se; Wu: HKU Business School, University of Hong Kong, Hong Kong, yanhuiwu@hku.hk. This study was approved by the Regional Ethics Review Board of Stockholm, number 2017128-31. The project has received funding from the European Research Council (ERC project number 742983) and the Swedish Research Council (diary number 2016-03220).

1 Introduction

Social media are transforming political communication and interaction, and it is likely that they do so more strongly in nondemocratic regimes with little pre-existing media diversity. Because social media can be used both by and against the regime, it has been debated whether social media weaken or strengthen authoritarian regimes (e.g., Shirky 2011; Morozov 2012). The debate is fueled by the recent development of information technology and big data techniques, which may have increased the ability of authoritarian governments to exploit social media in their favor. A focal question in this debate is whether social media facilitate the formation and dynamics of anti-regime collective action. In this paper, we study whether and how social media affect protests and strikes in China, the largest autocracy in the world, which is known for extensive censorship of social media and the substantial use of artificial intelligence technology for information and social control.

Unraveling the political effect of social media in China requires a deeper understanding of how social media affect protest dynamics–whether social media cause events to cross geographical boundaries and thus generate far-reaching social movements. In China, protests of limited scale and scope are allowed in part because they can inform the regime about unpopular policies and local corruption (e.g., Cai 2008; Lorenzten 2017). An obvious risk is that these protests may snowball into larger social movements that target national problems and challenge the regime.[1] This risk is likely to increase with the flow of information on social media, which enable rapid transmission of information over long distances.

China also provides an opportunity to investigate the mechanisms through which social media affect protests in a strictly-controlled information environment. Since Olson (1965), it has been well recognized that information to organize events is crucial for collective action. Several recent studies show that the organization of events and the explicit provision of information that coordinates people's behavior, such as when and where to meet, is a key driver behind the effect of social media on protests (e.g., Tucker et al. 2016; Steinert-Threlkeld 2017; Enikolopov et al. 2020). This type of information, however, is likely to be censored in China (e.g., King et al. 2013, 2014; Roberts 2018). Without such explicit information, will social media lose their efficacy in spurring protests? They may not, because simply reporting on a protest or complaining about social problems can help people coordinate their protest activities or arouse people's emotions so that they join the protest, as argued by a number of theoretical studies (e.g., Edmond 2013; Little 2016; Pasarelli and Tabellini 2017; Barberà and Jackson 2020). This type of general information, abundant on Chinese social media,

[1] A cautionary example is Solidarity in Poland, which originated from limited demands for workers' rights and better economic conditions but quickly turned into pervasive resistance that proved fatal to the regime. More recent examples involve social media. In 2013, small groups of environmentalists protested the proposed destruction of Gezi Park in Istanbul; the protest exploded to attract millions of tweets on social media and eventually the participation of hundreds of thousands of people in Turkey. In 2013-2014, after the Ukrainian president Viktor Yanukovych failed to sign an agreement with the EU, opposition leaders posted calls for protests on Twitter and Facebook. The protests then grew rapidly and spread widely, eventually culminating in the ouster of Yanukovych.

allows us to assess the importance of the less-explicit mechanisms that may drive the effect of social media on protests.

We primarily examine the effect of social media on the dynamics of protests and strikes, rather than on event incidence. To clarify, by effect on incidence, we mean the impact of an exogenous shock on the probability that an event will take place at a particular location and time. By effect on dynamics, we mean the impact on the extent to which an event increases the probability of a subsequent event in another location within a certain time window. To estimate the causal effect on event dynamics, we use the rapid expansion of the information network across cities (i.e., prefectures in the Chinese context) based on large-scale textual data from Sina Weibo (Weibo for short)—the leading microblogging platform in China. We then investigate the mechanisms that drive the effect by analyzing the content of Weibo posts.

The first empirical question is whether information about protests and strikes circulates in Chinese social media at all, or whether this content is uniformly censored. In Qin et al. (2017), we examine the content of 13.2 billion posts published on Weibo between 2009 and 2013 and find that there were approximately 4 million microblog posts mentioning protests or strikes. These posts predicted real-world protests and strikes. Moreover, users who published these sensitive posts continued to post similar content. This suggests that the Chinese government permits the posting of sensitive content useful for surveillance while maintaining a generally strict information environment.

To avoid protest waves, a potential control strategy is to block the visibility of sensitive content to people in other regions and delete posts that start to spread geographically. To investigate this possibility, we take advantage of our unique data on retweets (forwards). The fact that a user retweets a message indicates that the user has seen it. Thus, retweets are an effective measure of information diffusion (e.g., Kwak et al. 2010). We identify approximately 40 million retweets of the protest and strike posts. In a subset of 3 million first retweets for which we have precise timing and location information, we find that information about protests and strikes spreads quickly and widely. Approximately 30% of the retweets occur within one hour of the posting of the original messages, and 80% within one day. After one hour, the mean distance between the user who posts a message and the user who retweets it is more than 800 km. Evidently, Chinese social media are not only abundant in information about protests and strikes; the rapid diffusion of this information implies that social media may actively promote wider protest activities.

Estimating the effect of social media on the dynamics (spread) of real-world events is challenging. First, social media are likely to increase the observability of events to researchers. Second, cities with strong informational ties through social media typically also have strong ties through other information channels, such as phone calls and face-to-face meetings. Third, cities with strong social media ties tend to be similar in, for example, industry structure and hence exposed to similar shocks affecting the propensity to protest and strike. Therefore, it is difficult to separately identify the informational effect of social media on the spread of events. Related identification challenges have been extensively discussed in studies of social networks, as pioneered by Manski (1993). Proposed solutions include exploiting the network structure to identify instruments (e.g., Bramoulle et al. 2009), using instruments that are uncorrelated with the error terms and the network (e.g., Acemoglu et al. 2015; Konig et al. 2017), matching (Aral et al. 2009), and explicit randomization (e.g., Bakshy et al. 2012).

Unlike previous studies, we exploit the time-series variation in communication capacity arising from the rapid expansion of social media networks to identify the causal effect. Specifically, we use retweets on topics other than protests and strikes from one city to another city within the last few months to measure how information diffusion between cities increases with social media. We use two research designs. First, we use an estimator that allows for arbitrary time-invariant heterogeneity in the spread of events across city pairs. Second, we investigate how the spread of events has evolved across cities that eventually become closely connected through social media over time. Monte Carlo simulations validate our methods. We also show that, even if social media allow researchers to observe more events, our estimation of the effect of social media network expansion on event spread is not affected. The intuition is that increased observability per se does not affect the temporal clustering of events across cities.

The main findings of our empirical analysis are as follows. First, our two research designs consistently show that protests and strikes spread rapidly (within two days) through the social media network. We estimate that, due to the flow of information on Weibo, the occurrence of a protest in one city in the last two days increases the probability of a protest taking place in any of the other cities by 25%. Relative to the mean event probability, the estimated effects amount to a 68% increase for protests and a 32% increase for strikes. Both protests and strikes start to spread in 2010-2011 across cities that eventually become highly connected through Weibo. For strikes, which are politically less sensitive, the spread continues at high levels throughout our sample period (ending in 2017). In contrast, the spread of protests across cities connected by social media covaries with changes in political sensitivity and the strictness of censorship. The spread effect is lower in 2012, when Xi Jinping replaced Hu Jintao as General Secretary of the Chinese Communist Party (CCP), and falls significantly after 2014, when censorship became stricter as measured by the Press Freedom Index published by Freedom House.

Third, we find that social media increase the scope of protests and strikes. Not surprisingly, the spread of events is predominantly narrow in scope—protests spread to other protests with the same cause and strikes spread to other strikes in the same industry. However, we find that social media also induce significant, albeit smaller, spread of events across categories. Importantly, despite the smaller effect of each individual event, the aggregate effect of cross-category spread is greater than within-category spread, because the total number of events across all categories is much larger than within the same category.

We also investigate the mechanisms that drive the spread of events via social media.

Previous studies generally focus on how protesters use social media to organize events. For instance, Enikolopov et al. (2020) report that two thirds of the Russian cities in their sample had social media communities created to organize protest demonstrations. Similar content is found in studies of protests during the Arab Spring and in Turkey and Ukraine (Tucker et al. 2016; Steinert-Threlkeld, 2017). However, our inspection of post content shows that Weibo posts that contain logistical information (e.g., specifying where and when to meet) are extremely rare, most likely due to extensive censorship of this type of content. Another potential mechanism is that social media might carry information about protest tactics (e.g., what to demand and how to deal with the government), from which others can learn how to protest more effectively (e.g., Cai 2008; Little 2016; Chen and Suen 2016). If such a tactic-learning mechanism plays a significant role, we should observe a persistent effect on event spread through the social media network, because effective tactics are unlikely to change in the short term and relevant information remains accessible for a long time. However, we find the estimated effects on both strikes and protests to be short-lived, with the strong effect appearing within a week after the initial event and dying out after a month. Moreover, posts discussing protest tactics or government responses are scarce.

Our empirical findings are consistent with other potential mechanisms, such as tacit coordination and emotional reactions. Tacit coordination arises when the mere visibility of events on social media induces people to protest simultaneously so that protesters have a higher chance of achieving their objectives and expect a lower risk of repression (Edmond 2013; Little 2016; Barberà and Jackson 2020). Protests can also spread through emotional reactions if angry people spontaneously protest against social injustices (Pasarelli and Tabellini 2017). These two mechanisms are likely to generate short-term effects. Furthermore, inspection of a random sample from the dataset shows that, among the posts that mention strikes and protests, a large number of them discuss the causes of events, such as corruption or wage arrears, and criticize the government for bad policies and misbehavior. Among the most retweeted posts, many express anger and sympathy for protesters.

Our findings suggest that China's censorship strategy is faced with a fundamental tradeoff between utilizing bottom-up information and maintaining political control. The Chinese method of censoring social media appears to be geared toward achieving the dual goals of using bottom-up information and mitigating the risk of collective action. Information about social problems that lead to protests and emotional reactions is allowed to be posted and spread widely, making it possible for the regime to gauge whether the sentiments are merely local or national. In contrast, content that is useful for protesters but not for the regime, such as calls for action and discussion of protest tactics, is absent. Nevertheless, the remaining content is sufficient to spread protests and strikes. To limit the spread of collective action, it seems that the government would have to silence the discussion of causes and emotional reactions to social problems. The cost is the loss of bottom-up information that is crucial for surveillance and monitoring.

This paper contributes to the emerging literature on how information and communication technology (ICT)–in particular social media–affects collective action and regime changes in nondemocratic countries. A number of papers have examined the political effects of the expansion of communication infrastructure, notably high-speed Internet or mobile phone networks. For example, Manacorda and Tesei (2020) and Christensen and Garfias (2018) find that cell phone access affects protests in a panel of African countries. Guriev et al. (2021) find that 3G mobile networks reduce government approval. A few studies aim to identify the causal effects of social media on protests (Zhuravskaya et al. 2020). In particular, Enikolopov et al. (2020) show that penetration of the dominant Russian online social network led to protests against election fraud in Russia. Ferguson and Molina (2020) find a positive effect of Facebook on collective action on a global scale. Together with other research (e.g., Tucker et al. 2016; Steinert-Threlkeld 2017; Acemoglu et al. 2018), these studies highlight the explicit coordination role of social media, as they find extensive social media content containing logistical communication and tactical advice-sharing among protesters on the ground.

Our paper advances this literature in several ways. First, we study how social media affect political collective action in China, which is a considerably more strictly controlled environment than Russia and some African countries. Logistical communication and tactical advice are absent on Chinese social media. Second, we shift focus from cross-sectional incidence to the dynamic interaction between geographically dispersed protests. This dynamic interaction is key to understanding whether local protests turn into widespread social movements that have important consequences for an authoritarian regime, as emphasized in the theoretical literature on nondemocratic politics.[2] Finally, our examination of various mechanisms furthers the understanding of the role of social media in collective action. In this last regard, our paper is related to recent research on how information affects participants' behavior in social movements (e.g., Cantoni et al. 2019; Bursztyn et al. 2020).

Our paper also relates to the theoretical literature on protests, government policy, and regime change. Lohmann (1993) models how costly protests may help politicians avoid bad policies. Battaglini (2017) extends the analysis to the case where policy makers and protesters have partially divergent interests, arguing that social media may reduce the informativeness of protests because better coordination means that they no longer act independently. Barberà and Jackson (2020) argue that social media help coordinate uprisings. Our findings suggest that accounting for endogenous censorship of social media, where the regime trades off the benefits and costs of information control (e.g., Egorov et al. 2009; Lorentzen 2014), would be a useful extension of these models.

Finally, our paper contributes to the literature on media control strategies in nondemo-

[2] In their survey of the theory of nondemocratic politics, Gelbach et al (2016) note: "A number of models examine the dynamics of collective action in anti-regime protests and the means by which dictators can prevent their success. An important intuition in the study of protest and revolutions concerns cascades: the possibility that a protest today spurs more protests tomorrow by revealing information about the degree of popular support for the regime."

cratic countries. Several studies take the revealed-preference approach and infer authoritarian governments' goals from their observed strategy of social media censorship (e.g., King et al. 2013, 2014; Qin et al. 2017). We extend this line of research by examining the effect of social media on real-world events after the imposition of censorship, from which we infer that there is an unavoidable trade-off between the regime's dual goals. This has direct implications for the debate over whether social media help strengthen authoritarian regimes.

The remainder of this paper proceeds as follows. Section 2 provides a brief description of the institutional background. Section 3 describes our data and provides descriptive statistics. Section 4 presents the empirical analysis and results. Section 5 concludes the paper.

2 Background

2.1 Social Media in China

After an intense period of explosive growth, social media in China today are as vibrant and extensive as in any Western country. Our study, however, focuses on the earlier period of rapid social media expansion. Because the Chinese government blocked Twitter and Facebook and strictly controlled domestic microblogging services, the use of social media was limited until Sina Weibo appeared in August 2009. Weibo is a hybrid of Twitter and Facebook. It allows users to tweet and retweet short messages with embedded pictures or videos, send private messages, and write comments. Between 2009 and 2012, Weibo use increased exponentially. By 2010, Weibo had 50 million registered users, and this number doubled in 2011, reaching a peak of over 500 million at the end of 2012 (China Internet Network Information Center 2014). Weibo adoption was faster in some areas, notably, those with higher pre-existing levels of mobile phone use. During our main sample period (ending in 2013), Weibo was the dominant microblogging platform in China. Since then, it has lost ground to WeChat, a cellphone-based social networking service, but remains an influential platform for public communication.

The extensive use of social media in China is coupled with extreme government control. To suppress the posting of unwanted content, tens of thousands of information officers and internet monitors police the internet to punish users who post such content. These efforts also induce self-censorship (Chen and Ang 2011). Censorship of Chinese social media is regulated by the national Propaganda Department of the CCP and is implemented largely by private service providers registered in Beijing. The estimated extent of censorship of Sina Weibo ranges from 0.01% of posts published by a sample of VIP users (Fu et al. 2013) to 13% of posts on collective action events such as protests and strikes (King et al. 2013). Although censorship is pervasive, it is far from complete even on these topics (Qin et al. 2017).

The Chinese government also controls content through active posting on Weibo. Qin et al. (2017) estimate that there are 600,000 government-affiliated accounts contributing 4% of all posts regarding political and economic issues on Weibo. The share of government users is larger in areas with more extensive censorship of social media and in areas where newspapers are more strictly controlled. Local governments also employ a large number of internet trolls to interfere with the major social media platforms in China (King et al. 2017).

2.2 Protests and Strikes

Although a lack of opportunities to protest is a hallmark of authoritarian regimes, there have been numerous instances of collective resistance in China in the last two decades. The Chinese government's response to protests is multifaceted. Protests are often met with violent repression, and their leaders are taken into custody (Lorentzen 2017). Even if some protest demands are accommodated, organizers and active participants risk losing their jobs and being arrested or placed under close watch. At the same time, concessions are frequent; protests viewed as sufficiently innocuous are even ignored (Cai 2010; O'Brien and Li 2006; Su and He 2010; Lee and Zhang 2013). As numerous public statements by top CCP leaders make clear, the Chinese central government requires local officials to handle collective action events strategically rather than simply suppressing all such events with police force (Steinhardt 2017). In response to many anti-corruption protests, high-ranking CCP officials were eventually sent to converse with protesters to re-establish the public's trust in the government. In terms of task division, local governments are typically responsible for repression while central leaders make concessions (Cai, 2010).

In China, strikes are often triggered by firms' violation of labor laws, such as wage arrears and illegal work conditions. Due to weak law enforcement and inadequate government intervention, strikes are often the most effective way for workers to claim their rights and express their disapproval. In this sense, strikes are similar to protests, and it is not surprising that many strikes are associated with protest activities. Government reactions to strikes range from repressive acts (e.g., police arrests) to mediation to concessions. The Chinese government does not regard strikes as politically sensitive unless they escalate into violent events that create social unrest.

There are several reasons for the Chinese regime's relative tolerance of protests and strikes. First, China is large and diverse, and most political and economic decisions are decentralized to local governments. Protests are a costly and hence credible way for citizens to communicate their concerns, which may help the regime identify and correct policy oversights, gauge public sentiment, and monitor local officials (Lorentzen, 2017). Second, the absolute suppression of collective action may generate distrust and undermine the legitimacy of the regime. Finally, some collective actions such as strikes may improve welfare and even enhance productivity if they result in fairer competition and better working conditions (Cai 2010; O'Brien and Li 2006; Su and He 2010; Lee and Zhang 2013).

As long as protests remain within a restricted scope and location, they pose little threat to the regime. The regime is only threatened when local protests evolve into larger political action and social movements that shift the focus from local to national policies and leaders, thereby diminishing the legitimacy of the regime and trust in the CCP (Cai 2008).

3 Data

We assemble a unique dataset combining detailed information on thousands of collective action events from 2006 to 2017 together with posts published on Sina Weibo from 2009 to 2013. Data on protests and strikes in China are not available from any official sources, and media coverage of such events in mainland China is rather limited. Hence, we collect data from sources outside of mainland China. Below, we explain how we collect the data and provide a description of our dataset.

3.1 Protests

The data on protests are collected manually from the website of Radio Free Asia (RFA), a private non-profit international broadcasting corporation based in Washington DC. We extract relevant information from the Chinese version of the website, which is widely used by Chinese news portals outside of mainland China. The news reported on the RFA website comes from RFA's special correspondents, from media outlets in mainland China, Hong Kong, and Taiwan, and from Western media outlets such as the New York Times and the BBC. One advantage of using RFA as the information source is its hiring of correspondents on the ground to verify information authenticity. To the best of our knowledge, the RFA website is the most reliable and well-structured data source for protest events in mainland China.[3] We search for keywords related to "protest" and "demonstration" (in Chinese) on the RFA website and download the relevant news reports. Several research assistants were hired to verify the news source and purge duplicate information. Next, they extracted relevant information from each news report and coded the date, location, cause, and scale (number of participants) of each event.

Our dataset contains 1,153 protests between July 2006 and December 2013, which is the focus of our main analysis, and an additional 1,576 protests between 2014 and 2017. Although many of these events are small and confined to certain localities, some of them are large-scale and disruptive, spreading across regions. For instance, an event widely reported by Western media is the 2011 Wukan protest, when thousands of villagers in a village in Guangdong province protested against the corruption of local officials. The event led to direct

[3] One potential data source of protests is the GDELT project, which contains information on massive events collected from the world's news media. However, we find the data related to China extremely noisy. Another data source of protests is the Mass Incident Dataset constructed by the Chinese Academy of Social Sciences (e.g., Miao et al. 2021). But this dataset relies on media outlets in mainland China and covers only events with more than 100 participants. Recently, some researchers have applied machine-learning methods to Chinese social media data to collect collective action events (e.g., Zhang and Pan 2019). However, these data lack event specifics and are difficult to verify. Moreover, this method is likely to overestimate the number of events due to repeated counting of the same event that lasted more than one day or was discussed by social media users in different locations.

confrontation between the villagers and local officials, violent conflicts between protesters and police, and demonstrations in multiple cities in support of the villagers.

As shown in Table 1, many protests concern the government (policy, police and courts, and housing and land reforms) and livelihood issues (employment, environment, and health). In terms of scale, more than 70% of the events involve hundreds or thousands of people. Geographically, these protests span 224 cities (prefectures), accounting for two thirds of all cities in China. Although the distribution is right-skewed, many locations experience more than 10 events during our sample period. Beijing is an outlier with 95 protests. Other provinces with frequent occurrences of protests include coastal Shanghai and Guangdong and inland Hebei, Shaanxi, Chongqing, and Sichuan.[4]

3.2 Strikes

We collect data on strikes mostly from the China Labor Bulletin (CLB), a non-governmental organization based in Hong Kong that supports the development of trade unions in China and the enforcement of China's labor laws. The CLB has collected data on strikes in mainland China since 2007. From 2011 onward, this dataset contains detailed information on the timing, location, employers involved, industry, scale, worker action, and government responses for each event. For earlier strikes, we extract similar information from the annual reports published by CLB and supplement them with data from Boxun, a widely read political news website in Chinese based in the US.

The CLB draws on information from overseas Chinese media outlets, labor movement activists in China, and internet searches including social media. An important source for the 2013-2016 period is Wickedonna, a mass event tracking blog that searched Sina Weibo and other Chinese social media platforms.

Our dataset contains 1,558 strikes between January 2007 and December 2013, and an additional 7,946 strikes between 2014 and 2017. As shown in Table 2, strikes occur in a wide range of economic sectors, with a concentration in the manufacturing and transportation industries (including taxi services). Over 70% of the strikes involve more than 100 people. The most common cause is demands for payment of wage arrears. Geographically, these strikes cover 242 cities, approximately 71% of all cities in China. The developed coastal areas are over-represented, notably Guangdong, Shanghai and Jiangsu, but a significant number of strikes occur in some inland areas such as Chengdu and Chongqing.[5]

[4] Figure A1 in the appendix shows the distribution of the number of protests by city (prefecture). The upper right panel of Figure A2 shows the total number of protests by city.

[5] Figure A1 in the appendix shows the distribution of the number of strikes by city. The upper left panel of Figure A2 shows the total number of strikes by city.

3.3 Social Media

Our social media data come from a database including 13.2 billion posts published on Sina Weibo from 2009 to 2013. The database was created by Weibook Corp, which conducted a massive data collection project of downloading blogposts from more than 200 million active users. They categorized users into six tiers based on the number of followers. They downloaded the microblogs of the top tier users at least daily, the second and third tiers every 2–3 days, and the lowest tier on a weekly basis. For each post, the data provide the content, posting time, and user information (including self-reported location). According to our estimates, the Weibook dataset contains approximately 95% of the total posts published on Sina Weibo before 2012 when we have an alternative measure of the total post volume (Qin et al. 2017).

From this Weibook database, we obtain two datasets. The first dataset contains the aggregate information on the number of posts per city and month, based on the entire 13.2 billion posts available. We use this aggregate measure, labelled as Weibo penetration, to measure how the popularity of Weibo changes over time and across cities.[6]

The second dataset contains individual microblog posts that mention any of approximately 5,000 keywords related to various social and political topics.[7] This dataset comprises 202 million original posts and 133 first retweets of them. We only have the direct retweets of the original posts, not the retweets of the retweets. For each original post and retweet, we obtain the text and information on the posting time, times of being retweeted, and the location of its author. We use this dataset to construct a network of social media information flows across cities and measure its expansion over time. We also use posts referencing protests and strikes to examine how information about protests and strikes spreads on social media.

Figure 1 depicts the total number of protests and strikes per month along with Weibo use per month. Clearly, there is a positive correlation between the incidence of protests/strikes and Weibo penetration over time. The number of strikes per month was approximately six in the 2007-2010 period and increased rapidly to over 52 in 2013. The pattern for protests is similar, with around three protests per month until 2010, followed by a rapid increase to around 54 in 2013. The green dots show the number of Weibo posts per capita, which increased significantly after the start of 2010. This trend of increasing protests and strikes has not gone unnoticed. It has been commented on in numerous news sources including the BBC, CNN, the New York Times and the Washington Post.[8] Of course, this relationship between social media and events may not be causal. The increase in strikes could be driven by other factors, such as a slowdown in Chinese exports (Campante et al. 2019) and by

[6] The geographical distribution of social media posts is shown in the bottom panel of Figure A2.

[7] Details on our selection of keywords and extraction of posts can be found in Qin et al. (2017).

[8] "Can China keep its workers happy as strikes and protests rise?," Mukul Devichand, BBC, December 15, 2011; "China on strike," James Griffiths, CNN, March 30, 2016; "Strikes by Taxi Drivers Spread Across China," Andrew Jacobs, January 14, 2015; "Strikes and workers protests multiply in China, testing party authority," Simon Denyer, Washington Post, February 25, 2016.

increased observability of the events through social media.

3.4 Information Diffusion of Protests and Strikes

The Chinese government's censorship is a crucial determinant of what content about protests and strikes is available on social media in China. Three censorship strategies that the Chinese government potentially can implement are: (1) blanket censorship that removes all posts related to a subject (e.g., protests); (2) localized censorship that allows the posting of messages but not their spread (retweets) across geographical boundaries; and (3) strategic censorship that keeps content useful for surveillance in a controlled environment. Below, we provide evidence ruling out the first two strategies. We discuss the third strategy later when we investigate the mechanisms that drive our estimated effects.

Posting and retweeting on protests and strikes. As described in Qin et al. (2017), there is extensive coverage of protests and strikes on Weibo. From the original post dataset, we find 2.5 million posts containing keywords related to protests and 1.3 million posts related to strikes. To characterize these posts, we inspect the content of a random sample of 1,000 posts in each category. Around 30% of the posts are indeed about protests and strikes. These posts predict and identify real-world protests and strikes. Hence, regarding protest and strike posts, the Chinese government does not implement a blanket censorship strategy.

Moreover, we find that these original posts referencing protests and strikes indeed reach many other users. The post data contain a variable measuring the total number of times each post is retweeted. According to this variable, the 3.8 million protest and strike posts generate 37 million retweets, amounting to an average of 10 retweets per original post. This average masks a huge disparity: some posts are retweeted hundreds of thousands of times and others none. Conditional on being retweeted at least once, there are on average 50 retweets per strike post and 100 retweets per protest post.

Patterns of information diffusion. In the above description of retweeting, we use the number of retweets of an original post without knowing the specifics of an individual retweet. To show the spatial and temporal patterns of the original posts and retweets, we exploit the timing and location information of the 3 million direct retweets (among the 37 million total retweets) of the protest and strike posts. As seen from the left panel of Figure A4, for posts about protests and strikes, approximately 30% of the retweets occur within one hour after the original posts are published, and 80% within one day. Retweets within the first hour are more likely to be geographically close. After that, distance plays no role: the average distance between the user who posted the original post and the user who retweeted it is the same as the average distance between Weibo users. In general, information about protests and strikes on social media disperses rapidly and widely across China, as shown in the right panel of Figure A4.

Documenting the rapid and wide spread of protest and strike information is important in its own right. Even if protests and strikes are only local, the accumulation of widespread information about them could be detrimental to the legitimacy of the regime. Nevertheless, posting and retweeting of such information is permitted. This rules out the abovementioned localized censorship strategy and suggests that the Chinese government uses a strategic censorship strategy, as we investigate later.[9]

Social media information network. Empirically, we use retweets across all topics (other than protests and strikes) to construct a time-variant social media information network, through which information flows across cities. The network changes with the expansion of Sina Weibo. We use the number of posts from users in city i that are retweeted by users in city j as a proxy for the flow of information from city i to city j through the Weibo network. A post from city i being retweeted in city j implies that someone in city j has seen it and decided to retweet it. Of course, many others in city j may see the post without retweeting it. Therefore, our retweeting measure is a conservative measure of information spread.[10] Note that the information network is built on the 133 million first retweets across all topics except the 3 million retweets of the protest or strike posts.

We find that cities that are more strongly connected through the Weibo network are also close in other dimensions. This pattern conforms to an important network property, namely, homophily, which means that individuals tend to form links with similar others. As seen in Figure A3, geographical proximity is the most significant homophily factor, and cities with similar population density and GDP share of the tertiary sector are also more closely connected. For this reason, information from other sources is likely to be correlated with information through the social media network. In addition, errors may be correlated across cities that are more connected through the social media network. For example, strikes are likely to erupt at similar points in time in areas that share a similar industry structure, which are also closely connected through the social media network. Our estimation of the social media effect is designed to address this homophily concern.

3.5 An Example

We end this section with one concrete example. In Guangzhou, Guangdong, a male worker wearing a bomb vest staged a protest against wage arrears in a company in the afternoon of January 18, 2013. Later, he detonated the bomb, killing one person and severely injuring seven people including himself. Immediately after the tragedy, many Weibo posts, including some from direct witnesses, described and commented on the event. We identify 374 posts mentioning our protest-related keywords on that day from Guangzhou, 261 of them indeed

[9] One may suspect that the absence of the localized-censorship strategy is due to the lack of technology to implement this strategy. However, it is well known that the visibility of posts on Sina Weibo differs for accounts registered in and outside of mainland China. Our communication with experts working for the major Chinese social media platforms suggests that it is technologically feasible to block the visibility of a post to a certain type of user.

[10] A much less conservative measure is to use followers, because many followers will not read each blog post. A widely cited study on Twitter (Kwak et al. 2010) probes whether the number of followers or the number of retweets is a better measure of influence and settles on retweets.

talking about this event. These posts were first retweeted by Weibo users in nine cities closely connected to Guangzhou according to our retweet-based measure.[11] Most retweets express sympathy for the worker, condemn employers who default on wages, and decry the local government for disregarding citizens' rights and condoning wage arrears.

Among the nine cities that first retweeted the event posts, three experienced protests in the subsequent two days. In Shanghai, thousands of workers from a Sino-Japan joint venture protested the unfair new labor rules, detaining 18 senior managers, and the situation lasted until over 400 policemen stormed the factory two days later. In Shenzhen, hundreds of people took to the street, protesting against the construction of a heavily polluting LCD factory in their neighborhood. In Shanwei, thousands of villagers protested in front of the city government, demanding the return of lands taken by the government without appropriate compensation. These protests occurred at the same time in different provinces. Although targeting different causes, they all involved people who had suffered enough and protested against perceived injustices. It is possible that the later protests were inspired by the earlier ones, but there is no direct evidence of explicit organization or coordination across events.

Our empirical task is to investigate whether such episodes are isolated random incidents or part of a large-scale systematic pattern. Specifically, do strikes and protests spread more after the expansion of Weibo use and across cities that are closely connected through Weibo?

4 Empirical Analysis

We describe our baseline specification in Section 4.1 and the corresponding results in Section 4.2. Section 4.3 analyzes how the spread of events through Weibo evolves over time. Section 4.4 examines whether social media increase the scope of event waves by spreading protests across causes and strikes across industries. Section 4.5 evaluates various mechanisms that may explain the spread. In Section 4.6, we aggregate the data at the city-by-month level to study how the incidence of protests and strikes is associated with Weibo penetration.

4.1 Specification

We analyze whether information on Sina Weibo affects the spread of protests and strikes across Chinese cities (prefectures) using a panel of N cities at daily frequency, t. Because the specification is the same for both types of events, here we use protests for illustrative purposes. Let y_{it} be a dummy variable indicating the occurrence of a protest. Suppose that the probability of a protest in city i on day t, $\Pr(y_{it})$ depends on the number of people who are informed about protests y_{jt-1} in another city j at time $t-1$. Let f_{ijt} be the number of people who learn about the protest through social media and c_{ij} be the number of people

[11] Controlling for city and time fixed effects, the ranking percentile of the retweets between these cities and Guangzhou are 1.19% (Shanwei), 2.3% (Shenzhen), 4.36% (Wuhan), 5.03% (Shanghai), 5.3% (Hangzhou), 5.77% (Chengdu), 6.25% (Xi'an), 6.98% (Zhengzhou), and 8.62% (Qingdao).

who know the protest from other time-invariant sources. We then use the following model to characterize how protests spread from city j to city i:

$$\Pr(y_{it}) = (\beta f_{ijt} + \gamma c_{ij})\, y_{jt-1}. \tag{1}$$

Intuitively, when f_{ijt}, the flow of information on the Weibo network between cities i and j, increases, protests will spread more from city j to city i. The term c_{ij} captures the time-invariant propensity for events to spread between cities.

Based on Equation (1), we specify the following econometric equation as our baseline model:

$$y_{it} = \alpha h(y_{it-1}) + \beta h(\sum_{i \neq j} f_{ijt-1} y_{jt-1}) + \gamma h(\sum_{i \neq j} d_{ij} y_{jt-1}) + \eta h(\sum_{i \neq j} f_{ijt-1}) + \beta_0 w_{it} + \theta' x_{it} + \delta_t + \delta_i + u_{it}. \tag{2}$$

Here, y_{it} is a binary event dummy, y_{jt-1} is the number of events in city j in the previous period (within two days before t). The key variable for our study of spread is $\sum_{i \neq j} f_{ijt-1} y_{jt-1}$, which measures the diffusion of information on social media at time $t-1$. The variable $\sum_{i \neq j} d_{ij} y_{jt-1}$ measures the geographical diffusion of events, with d_{ij}, the inverse geographic distance between cities i and j, being a proxy for c_{ij}. The function $h(\cdot)$ is used to capture the possibility that protests can spread simultaneously from multiple previous protest locations with a nonconstant marginal effect. We also control for the sum of retweets, that is, $\sum_{i \neq j} f_{ijt-1}$. Below, we provide more details on how we measure f_{ijt-1}, specify $h(\cdot)$, and handle c_{ij}.

In Specification (2), we include Weibo penetration, w_{it}, to address the concern that social media may directly affect the incidence of protests through channels other than spread from other locations.[12] Practically, w_{it} is measured by the logarithm of 1 plus the number of Weibo posts per capita based on the aforementioned 13.2 billion posts in the Weibook dataset. Note that w_{it} is the number of all Weibo posts, most of which are non-sensitive personal communication and hence is only marginally affected by censorship or posting about protests and strikes. Fu et al. (2013) estimate that 0.01 percent of general posts are censored. In our dataset, posts referencing protests and strikes only account for 0.03 percent of the entire pool of Weibo posts related to social and economic issues. We add a set of other controls, x_{it}, including the logarithm of population, GDP, the shares of the industrial and tertiary sectors, and the numbers of cell phone and landline users. We also include an auto-regressive term y_{jt-1} as well as time and city fixed effects, δ_t and δ_i, respectively.

We measure f_{ijt-1}, by the logarithm of 1 plus the number of retweets by users in city i of posts from users in city j on all subjects other than protests and strikes in the past six

[12] For example, with greater Weibo penetration, people are better informed about social problems and thus protest more frequently. Another possibility is that Weibo penetration increases the visibility of protests to upper-level governments, which in turn increases protesters' expectation for favorable outcomes.

months up until one week before day t.[13] We exclude retweets less than one week before day t and all retweets about protests and strikes to avoid reverse causality, which may occur if people who plan to participate in protests and strikes are more likely to retweet posts about them. In matrix form, social media information flows can be represented by F with the element f_{ijt-1}, which is arranged as T stacked $N \times N$ matrices F_t, and hence has $N \times T$ rows and N columns. For interpretative convenience, we normalize this matrix so that the average sum of all elements in a row of a weighting matrix is 1.[14] Similarly, we normalize the distance matrix D with each element d_{ij}. Later, instead of using d_{ij}, we allow for arbitrary time-constant spread c_{ij} across cities.

We use two functional forms $h(x) = x$ and $h(x) = \ln(5x + 1)$. In our data, the values of x are mostly small, and the two functions yield similar values and thus estimated marginal effects (see Figure A5). However, the log transformation implies smaller marginal effects when x is large, which ensures non-exploding paths in the Monte Carlo simulations. In contrast, the linear function produces exploding paths that contradict the bounded increase in event occurrence observed in the real data. For this reason, we use the log transformation in the Monte Carlo simulations and present the results using the log-transformation as a baseline model, while the results from the linear model are presented in the appendix.

Several econometric issues require consideration. First, logistic and probit models are prone to bias with rare events data (King and Zeng 2001) and do not work well in panel data with a large set of fixed effects. Thus, we estimate a linear probability model, which is immune to these problems. Second, our model includes location-fixed effects and lagged dependent variables. In general, the estimates in this type of model are inconsistent when T is fixed (e.g. Arrelano 2003; Baltagi 2008; Hsiao 2014). In the model we use, T is large and the bias is likely to be small. We show that the bias is indeed small using Monte Carlo simulations.[15] Third, we check whether the estimated process is stationary.[16] Fourth, consistency requires no serial auto-correlation of errors. Serial correlation in the error term implies that the error term in the current period will be correlated with y_{it-1}. We test for serial autocorrelation in

[13] Using all cumulative retweets in a 4-month window produces qualitatively similar results. Although we have millions of retweets, some averaging over time is necessary because we estimate over 90,000 city-pair connections. A longer time window increases accuracy but reduces time variation in connectedness across cities.

[14] A common form of normalization is row-normalization in which all coefficients in a row sum up to 1. In our case, however, this normalization would imply that all locations would be equally affected (on average) by protests in other locations, which is clearly incorrect (Elhorst [2001] argues against the row normalization of distance matrices for analogous reasons). Our weighting maintains the relative magnitude between all elements in the weighting matrix. In a constant $N \times N$ matrix, such as the distance matrix, the total sum of all elements is N. In an $(N \times T) \times N$ matrix, such as the retweeting matrix, this sum is $N \times T$.

[15] If the bias was large, one could address this issue by using the GMM estimators of Arrelano and Bond (1991) or Blundell and Bond (1998). However, instrumenting rare events, such as protests and strikes in our setting, with lagged differences and levels is likely to perform poorly.

[16] Sufficient analytical conditions for stationarity are that the errors are not auto-correlated and that $|\beta\omega_{F_t,\max}| + |\gamma\omega_{D,\max}| < 1$, where $\omega_{D,\max}$ and $\omega_{F_t,\max}$ are the largest eigenvalue of the matrices D and F respectively (largest negative if $\gamma < 0$ or $\delta < 0$). We evaluate this criterion and check whether the process is stationary in the Monte Carlo simulations.

the first-differenced residuals. Finally, the error term in (2) may be correlated across both time and spatial units. We use two-way clustering of errors in the temporal and spatial dimensions.

4.2 Results

Baseline estimation We estimate Equation (2) in a panel of cities with at least one protest or strike for the respective outcomes. Table 3 reports the regression results: the first two columns for protests and the last two for strikes. Table A1 shows the summary statistics of the key variables.

Our main interest is the first row, the coefficient of the variable with lagged events, weighted by the number of retweets from city j by users in city i. This coefficient is positive and statistically significant at the 1% level across all specifications. The magnitude is barely affected by the inclusion of controls. Note that the total number of retweets does not affect the incidence of the event, as shown in the fourth row of the table.

The estimate in column 2 implies that a protest in a given city in the last two days increases the expected number of cities with protest incidence by 0.25, while the corresponding number for strikes, based on the estimate in column 4, is 0.17.[17] Relative to the mean event probability, these numbers represent an increase of 68% for protests and 32% for strikes.[18] The reported average effects mask considerable underlying heterogeneity, because the marginal spread effect is proportional to retweeting, f_{ijt-1}, which is 0 before the entry of Weibo in 2009 and increases rapidly afterwards. The marginal effects at a high level in the later sample period are much larger than the averages, as we discuss in Section 4.3.

The results in the third row of Table 3 show that both protests and strikes follow an autoregressive process. The occurrence of an event in the last two days significantly increases the predicted probability of an event occurring in the same city. These within-city dynamics are much less important for the overall dynamics of protest and strike waves than the spread across cities. As reported in the second row, the distance-weighted effect of lagged strikes is positive and statistically significant, but the protest counterpart is not. This result suggests that strikes, but not protests, spread to nearby cities even without spreading through the social media network. As expected, the fifth row shows that the incidences of protests and strikes are both increasing with Weibo penetration.

A dynamic model with fixed individual effects can generate a so-called Nickell bias (i.e., inconsistent estimates) when the number of time periods is kept fixed while the number

[17] For the case $h(x) = \ln(5x + 1)$, the marginal effect is calculated by scaling the estimated coefficients by a factor equal to 4.63 for protest and 4.56 for strikes. Therefore, the marginal effect is $4.63 * .055 = .25$ for protests and $4.56*0.038 = 0.17$ for strikes. These effects are somewhat larger than the corresponding estimates from the linear model reported in columns 2 and 4 of Table A2, which are .17 for protests and .12 for strikes. See Appendix A.1 for more details.

[18] The mean of protest incidence is .0015, and there are 246 cities other than the one where the first protest took place. Thus, the effect is .25/(246*.0015)=.68. For strikes, the corresponding number is .17/(281*.0019)=0.32.

of individuals tends to infinity (Nickell 1981). To assess the Nickell bias in our coefficient estimates, we run a set of Monte Carlo simulations. The bias is very small—the difference between the true and simulated mean estimates is in the magnitude of the third digit for both protests and strikes. See Appendix A.3 for more details.

Alternative specifications The results from the linear model, with $h(x) = x$, are reported in Table A2. The effects are equally significant and somewhat smaller than the corresponding effects in Table 3. For example, the estimate in Column 2 of Table A2 implies that a protest in a given city in the last two days increases the expected number of cities with protest incidence by 0.17, compared to 0.25 based on the corresponding column in Table 3.

Table A3 shows the results from a probit regression of Equation 2, where the time fixed effects are replaced by a quadratic time trend to avoid the problem of incidental parameters. The estimates of β are statistically significant across all specifications. Again, the implied marginal effects on event probabilities are slightly smaller than those in Table 3.

Tables A5 and A6 show how our key estimates are affected when the variables are sequentially added to the regression, starting from the most parsimonious specification with only the cumulative forwards weighted events and ending with the specification in the first column of Table 3. The coefficients on cumulative-forwards-weighted events (first rows of Tables A5 and A6) are stable across specifications, ranging from 0.052 to 0.055 for protests and from 0.039 to 0.052 for strikes.

Observability Social media can be used as a source to discover protests and strikes. This is likely to bias the estimation of the effect of social media penetration on event incidence. However, observability is less likely to affect the estimated effect on the spread of events across cities. Consider the following example in which events do not spread across cities that are connected by social media, so that the probability of an event in city i is independent of whether there was an event in a connected city j in the previous two days. Suppose that only a portion of the events in cities i and j is observed. Reporting on social media increases the share of observed events, but it is unlikely that the newly observed events depend on whether or not there was an event in city j in the past two days. Therefore, social media may increase the number of observed events, but not the probability that one observed event occurs just two days after another event, relative to other days. In this case, a regression of event incidence in city i on incidence in city j within two days will yield a coefficient of 0. Bias could arise if the number of newly observed events increases and falls precisely within the 2-day interval, which, if it occurs, is unlikely to be driven by a general increase in observability.

The real-world situation is more complicated than the above illustrative example because events may spread across cities for other reasons, such as geographic proximity. Therefore, we use Monte Carlo simulations to verify that we can consistently test the null hypothesis that there are no effects of social media on event spread, even if event observability increases with Weibo penetration. In particular, we simulate the event data, y_{it}, using the estimated parameters from Equation 2, except that we set $\beta = 0$ so that there is no spread of events through the social media network. We assume that the probability of observing a simulated event, p_{it}, increases linearly with Weibo penetration, w_{it}. We adjust the size of the effect such that Weibo penetration increases observability by 30%, which is larger than what would be implied if the estimated coefficient on Weibo penetration, $\widehat{\beta}_0$, only reflects increased observability. We then draw a set of observed events $\widetilde{y}_{it}$ with probability p_{it} from the simulated events y_{it}. Finally, we estimate the model in Equation 2 on the observed simulated events, $\widetilde{y}_{it}$. Figure 2 shows that observability induced by social media does not generate any spurious effect on event spread. Hence, we conclude that we can consistently test the hypothesis $\beta = 0$, even if social media affect observability.

Homophily and correlated shocks Another identification concern lies in the difficulty of separately identifying the effect of social media in spreading events if information from other communication channels or the underlying shocks are correlated with the social media network. As shown previously, the social media network we construct exhibits some homophily: cities similar in other dimensions are also closely connected through the social media network. Consequently, strikes and protests could seemingly spread more across cities that are closely connected in our social media network, even if social media play no role. Econometrically, shocks induced by homophily factors would mean that the error term is correlated with the lagged outcome in connected cities y_{jt-1}. Pioneered by Manski (1993), a vast literature discusses this issue and the related identification problems (e.g., Bramoulle et al. 2009; Aral et al. 2009; Bakshy et al. 2012; Acemoglu et al. 2015; König et al. 2017).

Our identification strategy departs from these studies by leveraging time-series variation in the rapidly expanding network for identification. In particular, we allow strikes and protests to spread through channels that are correlated with the network, but assume that these correlations are constant during our short sample period. In other words, although the average spread of protests and strikes between cities may be related to the average strength of the network, the exact timing of network expansion is unrelated to other changes in the spread of events across cities.

In practice, we use a difference-in-differences approach to estimate network interactions. We explore three different methods. We first estimate a model of dyads of locations with time-constant dyad-fixed coefficients. However, the Monte Carlo simulations show that the estimate based on this method produces a strong Nickell bias. Thus, this solution is not reliable, and we turn to the other two solutions, as explained below.

Time-invariant heterogeneity in event spread across city pairs Our second method is to use a model that allows for arbitrary time-invariant heterogeneity in event spread across cities. We implement this method by including $c_{ij}h(y_{j,t-1})$ for each $j \neq i$ in Equation (2), where c_{ij} is a constant coefficient. This term implies that an event in city j is allowed to increases the probability of an event in city i by an arbitrary time-invariant amount. Table 4 reports the results. The pure auto-regressive and distance-weighted effects are not identified. The estimates of the spread effect, β in the first row, remain highly significant. For protests, the coefficient slightly decreases from 0.055 to 0.047; for strikes, the magnitude of the coefficient drops from 0.037 to 0.024. This suggests that factors correlated with the social media network, such as correlated shocks or communication through other channels, are more important for the spread of strikes than for the spread of protests.

With this specification, two potential technical issues deserve further attention. The first concerns the presence of Nickell bias. The Monte Carlo simulations show that a Nickell bias, albeit small, is present. The second issue is that the auto-correlation test we use is not well suited for this case. See Appendix A.3 for a detailed discussion of these two issues.

4.3 Effects over Time

Our third method is analogous to an event-study analysis. In our baseline specification (2), we assume that the spread of events increases with time-variant communication on social media, measured by f_{ijt-1}, and estimate the constant parameter β. We now replace f_{ijt} with $\overline{f}_{ij}$, which is the average retweets in city i of posts published by users in city j, and estimate a time-varying parameter β_p, where p indexes time periods.[19] By fixing $\overline{f}_{ij}$, we let the data reveal when protests and strikes start to spread across cities that eventually develop strong ties through social media. We also replace the constant parameter γ in Equation (2) with γ_p.

In this specification, the estimate of β_p contains both the causal effect of the expanding social media network and any potential bias created by spurious correlations through pre-existing communication channels and shocks correlated with the network. Before August 2009, Sina Weibo did not exist, so for this period, β_p only contains the bias induced by spurious correlations. We can then estimate the causal effect by the increase in β_p after the pre-Weibo period. The key identification assumption is that the correlation between the average network ($\overline{f}_{ij}$) and the communication channels and correlated shocks (c_{ij}) does not change over time.

This method has several advantages. First, it allows us to more directly measure how the effects of social media on event spread change over time. Second, we can verify whether the events in our data spread across cities even before the entry of Weibo. Third, importantly, we can investigate the effect of social media after 2013, when Weibo data are not available but a significant number of events are added. The key assumption is that the average number of retweets between cities in 2013 is a good proxy for the average number of retweets during the 2014-2017 period. This assumption is arguably reasonable. Weibo passed its rapid expansion phase in 2012 and stabilized in 2013. Formally, a regression of the number of retweets between

[19] To make the coeffcients comparable with those in the previous estimation, we normalize the relevant retweeting matrix such that the sum of all elements $\overline{f}_{ij}$ in the 2010-2013 sample is equal to the sum of all elements of f_{ijt} for this period.

cities during the past six months before the last day of 2013 on the same variable before the last day of 2011 yields an R^2 of 0.96.

We implement this method for two time frequencies. We first estimate separate coefficients on event spread across cities for each biannual period. We then divide our full sample into three periods, and estimate a more demanding model, which includes city fixed effects and other control variables, for each period. The results are reported and discussed below.

Biannual periods Figure 3 plots the biannual estimated marginal effects of the spread of events across cities that eventually become closely connected through social media.[20] For both protests and strikes, the spread effect increases rapidly in 2010-2011 when the use of social media took off (recall Figure 1). For protests, the spread effect drops significantly in 2012, bounces back to its highest level in 2013, and then trends downward after 2014. This pattern coincides with changes in political sensitivity and the strictness of censorship in China. In 2012, protests were politically sensitive because the 11th National People's Congress (March 2012) and the 18th CCP National Congress (November 2012) took place and Xi Jinping replaced Hu Jintao as the new government and party leader. In addition, the Chinese government started to tighten the regulation of social media in 2012, for example, by announcing "the real name reform," which required users to register with their real identity and caused a drop in Weibo use. After a short period of reduced political restrictions, China established the Central Cyberspace Affairs Commission in February 2014 to enhance its political control over the internet and media. This change in media control is also reflected by the Media Freedom Index published by the Freedom House. During our early sample period between 2006 and 2013, the Index for China was at a level close to the average between 1990 and 2016. However, after 2013, this Index fell to a historically low level, comparable to the level in the early 1990s after the Tiananmen Square Protests in 1989.

In contrast to the fluctuating effect for protests, the spread effect for strikes continues at a constant and high level after 2011. A likely reason is that the spread of strikes is not regarded by the Chinese central government as politically threatening (e.g., Kuruvilla and Zhang 2016; CLB 2018). The magnitude of the marginal effect for strikes during the later period of our sample reaches a high level, while the number of event days falls drastically. This suggests that there are other factors (such as stronger suppression or less frequent wage arrears) that counteract the strong feedback through social media in this period and limit the occurrence of such events.

Note that the substantial increase in the number of reported events during the 2013-

[20] In the logarithmic case with $h(x) = \ln(5x+1)$, the reported marginal effect is

$$\beta_b \widetilde{E}\left(\frac{5\sum_{j\neq i}\overline{f}_{ij}}{5\sum_{j\neq i}\overline{f}_{ij}y_{kt-1}+1}\right),$$

where $\widetilde{E}()$ is the sample mean. The marginal effects for the linear model with $h(x) = x$, plotted in Figure A8, are simply the estimated β_b-coefficients, which are very similar to those demonstrated in Figure 3.

2016 period, probably partly caused by increased observability through social media, has little impact on the estimated marginal effects. This provides additional evidence that our estimation is robust to increased observability via social media.

Three periods We split the full sample into three periods: pre-Weibo (before 2010), rapid expansion (2010-2013), and stabilization (2014-2017). The first period (2006-2009 for protests and 2007-2009 for strikes) is before the entry of Weibo. The second period is the period during which Weibo data are available. The last period is the period during which we do not have Weibo data. The last two periods coincide with rapid expansion and stabilized use of Weibo in China. We estimate the effect on event spread in each of these three periods (longer than the previous biannual period) in a model controlling for city fixed effects, day fixed effects, and a large number of time-variant city characteristics.

Table 5 presents the estimated marginal effects. In the pre-Weibo period (period 0), protests spread across cities that eventually became closely connected through social media. However, this is not the case for strikes. In the rapid-expansion period (period 1), the spread effect induced by social media increased significantly for both protests and strikes. If we subtract the pre-period β-coefficient, the average spread through social media for the second period is approximately 0.6 for both protests and strikes. Finally, in the stabilization period (period 2), the spread of protests falls to a level that is not significantly different from that in the pre-Weibo period, whereas the spread of strikes remains highly significant, as shown in the F-tests in the last row.

Another important message from Table 5 is that the spread of events induced by social media does not just pick up a general time trend. The coefficients of the distance-weighted past events show that spread induced by geographical proximity does not move in tandem with spread induced by social media. Specifically, the spread of protests to geographically close areas is small and not significant, and the spread of strikes remains roughly the same in the first two periods and drops to a low and nonsignificant level in the last period.

4.4 Do Social Media Increase the Scope of Protests and Strikes?

We now investigate whether social media increase the scope of event waves by spreading protests across causes and strikes across industries. As previously discussed, events that only spread within the same cause and industry are bounded in size and influence and unlikely to be perceived as a threat by the regime. Thus, it seems likely that strikes predominantly spread within industries and that protests predominantly spread within causes. For example, the wave of school teacher strikes in 2014 and 2015, documented by Chang and Hess (2018), spread within the education sector, and the protests against corruption that originated from Wukan and spread among farmers (recall Section 3.3) were for the same cause. However, the example presented in Section 3.4 suggests that strikes and protests may also spread across industries and causes.

To study this scope-of-spread issue, we separately estimate the spread of events through social media within and across the protest causes listed in Table 1 and the strike industries listed in Table 2. Table 6 shows that the estimated spread effect through social media is six to eight times higher within protest cause and strike industry than the cross-category counterpart effect. This result confirms that the spread of events through social media predominantly occurs within categories.

Nevertheless, the cross-category spread induced by social media is statistically significant. Although the effect of an individual event is smaller across than within categories, the total effect of spread across categories is larger because there are many more events across all categories. The means of the social media weighted events within and across categories are reported in the last two rows of Table 6. The aggregate effect of protests spreading across protest causes is 60% higher than within causes. Similarly, the aggregate effect across strike industries is 20% higher than within industries. Not surprisingly, the distance-related spread effect does not appear across categories. We thus conclude that social media break down the boundaries of protest causes and strike industries, leading to a greater overall impact on event waves than through within-category spread.

4.5 Mechanisms

In this section, we assess four mechanisms that may drive the effects of event spread through the social media network: learning tactics, explicit coordination trough logistic information, tacit (implicit) coordination, and emotional response. These mechanisms require the circulation of certain types of content in Weibo posts to support them and have different empirical implications about, for instance, the duration of the effects.

Our goal is to assess which mechanisms are likely to be a major force behind the large estimated effects of social media on event spread in an information environment as strictly controlled as China. Because our assessment of mechanisms hinges, to a large extent, on the content of social media posts, it is useful to further clarify what information is in our dataset and what is visible to the public under the Chinese censorship scheme. As described in Section 3.3, our Weibo-post data were downloaded at various frequencies—those of the top tier users were downloaded at least once a day and those of the other users at a lower frequency. As for the speed of censorship, King et al. (2013) show that Weibo posts containing sensitive keywords are held for review before publication on the Weibo platform. In this case, censorship is immediate, and the censored posts are neither visible to the public nor to researchers. Posts that do not contain sensitive keywords are reviewed after posting. Zhu et al. (2013) show that these posts are quickly reviewed and released to the platform (if not censored), 30 percent within half an hour and 90 percent within one day. This review process and censorship timing imply that our dataset may miss some posts that were visible to the public for a short while (usually less than a few hours) and then censored. On the other hand, our dataset also contains posts that were downloaded before censorship. Given such speedy censorship, posts of these two categories are unlikely to be seen by a large audience.

Content Multiple studies show that protesters use social media to organize events and provide logistical information, such as calls for action or at least mentions of where and when a protest is to take place. Enikolopov et al. (2020) report that the vast majority of protest participants in Russia learned about upcoming protests from social media and that social media were also widely used for organizing protest activities. Tucker et al. (2016) find that, during the Turkish protests in Gezi Park in 2013 and the Maidan protests in Ukraine in 2013-14, tweets provided information instructing demonstrators to converge on particular locations and warning them when the police were approaching these locations. Similarly, Steinert-Threlkeld (2017) describes posts providing logistical information during the Arab Spring. We refer to the mechanism based on such explicit information to coordinate action as explicit coordination.

The situation in China is rather different. Based on a random sample of 1,000 posts about protests, only 15 posts involve a call for action and only two explicitly state a time and location of action. These 15 posts are rarely retweeted, with only one retweet per post on average. In our random sample of 1,000 posts about strikes, none calls for action. Hence, explicit organization and coordination is unlikely to be an important channel for social media effects on collective action in China.

Social media may also cause protests and strikes to spread because they carry information about effective protest/strike tactics (e.g., Little 2016; Chen and Suen 2016). Examples of such information are how to counteract the effects of tear gas (Tucker et al., 2016) and whether it is effective to complement strikes with other tactics such as demonstrations and petitions to local administrators, as described by Chang and Hess (2018). Some posts in our dataset, albeit few, discuss yet other tactics, such as whether violent or peaceful protests are more effective and what to demand from and how to deal with the government. These posts, if existent in large quantities, would help protesters learn tactics and spread protests.

However, few posts in our random sample—five protest posts and 13 strike posts—mention tactics. Similarly, only a few posts discuss the outcome of an event, such as whether the protest was met with concessions or not. The posts that indeed talk about tactics receive only two retweets on average. Given the scarcity and limited circulation of posts mentioning tactics, the tactic-learning mechanism seems unlikely to be a driver of the spread of protests and strikes. This is further confirmed by the short-term effects discussed below.

Most posts in our random samples, however, report other facts about the events, coupled with emotional reactions such as the protesters' anger and sympathy. Many posts—164 about protests and 223 about strikes—discuss what caused the event and the underlying social problems such as corruption or persistent wage arrears. These posts are among the most retweeted. This type of information is important for event spread because it helps build common ground between insiders (people involved in a specific event) and outsiders (people suffering from similar problems) and thus has the potential to cause people's actions to converge.

A significant number of posts—137 about protests and 42 about strikes—criticize and question existing social institutions. The content of these posts ranges from criticism of the legal system to complaints about the lack of free speech. Due to their general nature, these posts have the potential to spread events across causes and industries.

For a certain mechanism to play a role, the content supporting this mechanism must circulate widely. We inspect a sample of the 100 most retweeted posts about protests and strikes. After removing those that are repeated or irrelevant, the sample contains 91 posts. Of them, 55 talk about ongoing events, almost all mentioning the cause of the event. The remaining 36 posts comment on past events, government policies, or social problems. The majority of these popular posts convey certain emotional elements. Many posts, likely published by insiders, express anger after a description of government repression and persecution of protesters. Another common type of emotional content is outsiders' reaction to events, with posts stating that the protesters were unfairly treated and expressing sympathy and moral support.

Effect duration Alternative mechanisms differ in their implications for the duration of the spread effect. For instance, the mechanism by which protesters learn about protest tactics, as in the examples given above, is likely to have an effect that lasts more than a couple of days. The usefulness of a tactic is unlikely to change abruptly, and people who read the posts are unlikely to forget the information quickly (even if they do, the posts remain available online).[21] In contrast, a short-term effect is more likely if people respond to some posts emotionally and take to the street spontaneously, or if people view concurrent protests from various locations as a more effective way to achieve favorable outcomes and move simultaneously as if there is coordination—referred to as tacit coordination in our discussion.

In the previous analysis, we focused on short-term responses (within a 2-day window) because posts about protests and strikes spread quickly and widely (recall Section 3.4). To examine how the effects persist over time, we estimate the same model with the time window extended to various durations. Figure 4 shows the estimated spread effect of events occurring within 1-2 days, 3-7 days, 8-30 days, 31-90 days, and 91-180 days. For both strikes and protests, the effects are greatest in the 2-day window, then drop drastically and become small and nonsignificant after 3-7 days. The lack of long-term effects is evidence against the tactic-learning mechanism and in favor of the emotional-response and tacit-coordination mechanisms.

Summary The above analysis shows that learning tactics and explicit coordination are unlikely to be the main mechanisms underpinning the effect of social media on the spread of

[21] There might be information about protest tactics that is valuable only in the short run. However, we have not found any discussion about such tactics in the literature or in our protest and strike posts.

protests and strikes in China. Consistent with the short-lived effect, the infrequency of content talking about tactics limits the function of the tactic-learning mechanism. Furthermore, we find little content related to explicit coordination and organization of protests. The occasional posts that mention logistics are barely retweeted. This evidence, consistent with the finding that the Chinese government relentlessly censors posts that have the potential to spur collective action (King et al. 2013, 2014), excludes explicit coordination as an important mechanism.

The mechanisms consistent with our findings are tacit coordination and emotional reactions. We find a large amount of widely circulated content describing ongoing events, discussing their causes, and expressing anger or sympathy. This content can spark a short-term response as a consequence of an emotional reaction that motivates people to protest (Pasarelli and Tabellini 2017). It can also result in effective tacit coordination when people anticipate that simultaneous protests may increase the chance of pressuring the parties involved to deal with the problem and reduce the risk of punishment (Edmond 2013; Little 2016; Barberà and Jackson 2020).

4.6 Event Incidence and Social Media Penetration

Although the focus of this paper is on the effect of information diffusion via social media on the spread of events, it is worth examining how social media penetration may affect the incidence of protests and strikes. From our previous analysis (recall Tables 3 and 4), social media penetration, as measured by the number of Weibo posts published in a city per month, is indeed associated with the number of events. Here, we use a more direct specification to study the effect of Weibo penetration on event incidence at the city-month level. Specifically, we estimate the following linear probability model:

$$y_{im} = \alpha_i + \alpha_m + \beta_0 w_{im} + \beta' x_{im} + \varepsilon_{im}, \tag{3}$$

where i indicates the city and m the month. The dependent variable y_{im} is an indicator for an event taking place in city i in month m. The key independent variable w_{im} measures the extent of Weibo penetration, defined as before. We use the same set of controls x_{im} as in Equation (2) except that we drop the total number of retweets by users in city i. The reason is that the number of retweets is strongly correlated with Weibo penetration, and we wish to have one single measure of social media penetration for interpretative convenience.

Table 7 shows the results. As expected, both protest and strike incidences are positively associated with the number of Weibo posts. Adding controls barely affects the estimates. The magnitudes of the estimated coefficients are large. In 2012, the variable Weibo posts has an average of 0.3. Thus, the estimated coefficients imply that an increase in the number of Weibo posts by 0.3 is associated with an increase in the total number of protests by $13 (= .3 * .186 * 224)$ per month and an increase in strike incidence by $14 (= .3 * .188 * 242)$ per month. Figure 5 shows the dynamic response in a regression with quarterly leads and lags of Weibo penetration. There is no pre-trend, and the effects appear with a lag of six months for strikes and immediately for protests.

As mentioned before, these correlations between Weibo penetration and event incidence could be driven by increased observability due to social media. If so, we would expect to see a larger correlation in places where observability prior to social media was lower. Our interviews with industrial experts suggest that RFA, our source of protest data, had better information in certain coastal provinces (Guangdong, Fujian, Zhejiang, and Jiangsu) than in inland provinces before the entry of social media in China. To investigate this heterogeneous effect, we include an interaction term between the Weibo penetration variable and a dummy variable for areas other than these four coastal provinces. The results from columns 3 and 6 of Table 7 show that the effect on strike incidence is not significantly different in coastal and inland areas, but the effect on protests is significantly greater in inland areas. This latter result is consistent with the RFA having fewer news sources there.

It is also likely that observability has a greater impact on the incidence of small events, which are less likely to attract media attention. The lower panel of Figure 5 shows the estimated coefficients by the number of event participants (less than 100, hundreds, thousands, or tens of thousands). The estimated effects are larger for small strikes and medium-sized protests and strikes. These results suggest that observability explains part of the correlation between social media penetration and protests in inland areas and smaller strikes.

5 Conclusion and Discussion

In response to new ICT, authoritarian countries, most notably China, have made substantial efforts to strategically control online information flows. We find evidence consistent with the view that China's strategy aims at achieving the dual goals of utilizing bottom-up information and mitigating the risk of grassroots anti-regime protests. Chinese social media are devoid of content supporting the mechanisms identified in previous studies as the main drivers of the effect of social media on protests. Specifically, very few posts contain explicit information about how to organize protests and protest logistics and tactics, likely due to the Chinese government's censorship. However, Chinese social media massively diffuse the type of information about protests and strikes that can be useful for government surveillance and gauging public sentiment. In particular, a large number of posts discuss the causes of the events, criticize the government and policies, and express anger and sympathy for protesters. These posts spread quickly and widely, as is evident from the patterns of their retweets.

Despite the absence of information explicitly useful for organizing protests, we find that social media posting and retweeting has a sizeable effect on the spread of both protests and strikes across cities. The conclusion is that the diffusion of content that the regime can use to maintain regime stability is sufficient to forcefully spread protests and strikes. One reason, consistent with our empirical findings, is that people can use social media information for tacit coordination (e.g., simultaneous protests), which increases their chances of achieving their goal and reducing the risk of punishment. Another possible reason is that social media information triggers a widespread emotional response, leading to spontaneous protests. The lesson is that, to limit the spread of collective action, an authoritarian government must shut down discussion of causes and emotional reactions to problems and thus bear the cost of losing bottom-up information that is valuable for identifying unpopular policies, monitoring local officials, and tracking social unrest and anti-regime sentiment.

Another important finding of our study is that social media increase the scope of protests. The spread of events induced by social media is rapid and predominantly local, but the spread across event categories is significant and of greater magnitude in aggregate. Although the effect of social media on event spread is short-lived, the effect caused by a sequence of events may lead to waves of social movements.

A natural question arises as a result of our study: do advances in ICT enable an authoritarian regime to adopt an alternative media control strategy of collecting truthful information but avoiding the risk of event spread? This question is highly relevant given that some authoritarian governments such as China's have invested heavily in surveillance and AI technology to improve social monitoring. For example, the government could implement a localized-censorship strategy to fine tune the spread of information across regions. The cost of this strategy depends on how much more information about public attitudes toward the regime and national problems can be learned from knowing how people in other regions respond to a local event. Another strategy is to closely monitor people's actions using advanced surveillance technology while allowing a free social media so as to collect information without worrying about widespread protests—people dare not take to the street anyway. The feasibility of such a strategy hinges on the complementary relationship between online information and offline action. One reason why protests are informative is that they are costly to the protesters. Without offline protests, online protests may be merely cheap talk, rumors, and noise. We leave investigation of this issue to future research.

Our study also has implications for the effect of ICT on political accountability. Information is indispensable for holding politicians accountable to the public in democracies and to higher-level leaders in autocracies. In a large autocracy like China, where traditional media are operated by sub-national levels of government, local officials have an information advantage over citizens and the central government. This creates severe agency problems in the political system. Social media substantially reduce the information asymmetry among the central government, local officials, and citizens. In particular, citizens can easily make their information public, while the central government has the technological capability to collect and aggregate this information. Therefore, when the goal of the central government is aligned with the interests of citizens, social media may help solve agency problems and hold local officials accountable. One caveat is that the low cost of posting complaints and allegations online may reduce the informativeness of social media. Consequently, the informational value of social media relies critically on real events that are costlier and thus more informative. Investigating how this nuanced complementarity between online allegations and offline protests affects local accountability would be an interesting extension of our research.

References

[1] Acemoglu, Daron, Camilo Garcia-Jimeno, and James A. Robinson. "State capacity and economic development: A network approach." American Economic Review 105, no. 8 (2015): 2364-2409.

[2] Acemoglu, Daron, Tarek A. Hassan, and Ahmed Tahoun. 2018. "The Power of the Street: Evidence from Egypt's Arab Spring." The Review of Financial Studies 31, no. 1(January 2018): 1-42

[3] Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya. 2015. "Radio and the Rise of the Nazis in Prewar Germany." The Quarterly Journal of Economics, 130(4): 1885-1939.

[4] Aral, Sinan, Lev Muchnik, and Arun Sundararajan. "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks." Proceedings of the National Academy of Sciences 106.51 (2009): 21544-21549.

[5] Arellano, Manuel. Panel data econometrics. Oxford university press, 2003.

[6] Arellano, Manuel, and Stephen Bond. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." The review of economic studies 58.2 (1991): 277-297.

[7] Bakshy, Eytan, tamar Rosenn, Cameron Marlow, and Lada Adamicet. 2012. "The role of social networks in information diffusion." Proceedings of the 21st international conference on World Wide Web. ACM, p. 16-20.

[8] Baltagi, Badi. Econometric analysis of panel data. John Wiley & Sons, 2008.

[9] Barberà, Salvador, and Matthew O. Jackson. "A model of protests, revolution, and information." Quarterly Journal of Political Science, 15 (3), 297-335, 2020.

[10] Battaglini, Marco. "Public protests and policy making." The Quarterly Journal of Economics 132.1 (2017): 485-549.

[11] Blundell, Richard, and Stephen Bond. "Initial conditions and moment restrictions in dynamic panel data models." Journal of econometrics 87.1 (1998): 115-143.

[12] Born, Benjamin and Jorg Breitung. 2016. "Testing for Serial Correlation in Fixed-Effects Panel Data Models," Econometric Reviews, 35:7, 1290-1316.

[13] Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. "Identification of peer effects through social networks." Journal of econometrics 150.1 (2009): 41-55.

[14] Brauer, Alfred, and Ivey C. Gentry. "Bounds for the greatest characteristic root of an irreducible nonnegative matrix." Linear Algebra and its Applications 8.2 (1974): 105-107.

[15] Bursztyn, L., Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2020). Persistent political engagement: Social interactions and the dynamics of protest movements. American Economic Review: Insights.

[16] Cai, Yongshun. "Power structure and regime resilience: contentious politics in China." British Journal of Political Science (2008): 411-432.

[17] Cai, Yongshun. Collective resistance in China: Why popular protests succeed or fail. Stanford University Press, 2010.

[18] Campante, Filipe R., Davin Chor, and Bingjing Li. The Political Economy Consequences of China's Export Slowdown. No. w25925. National Bureau of Economic Research, 2019.

[19] Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2019). Protests as strategic games: experimental evidence from Hong Kong's antiauthoritarian movement. The Quarterly Journal of Economics, 134(2), 1021-1077.

[20] Chang, Shengping, and Steve Hess. "The Diffusion of Contention in Contemporary China: An investigation of the 2014–15 wave of teacher strikes." Modern Asian Studies 52.4 (2018): 1172-1193.

[21] Chen, Heng, and Wing Suen. 2016. "Falling Dominoes: A Theory of Rare Events and Crisis Contagion." American Economic Journal: Microeconomics, 8(1):228-255

[22] Chen, Xiaoyan and Peng Hwa Ang. 2011. "The Internet Police in China: Regulation, Scope and Myths." In Online Society in China: Creating, Celebrating, and Instrumentalising the Online Carnival, edited by David Kurt Herold and Peter Marolt, 40–52. New York: Routledge.

[23] China Internet Network Information Center. 2014. "The 34th Statistical Report on Internet Development in China." January 2014, Beijing.

[24] China Labor Bulletin. 2015. Search for the Union: the Workers' Movement in China 2011-2013, Research Report by China Labor Bulletin.

[25] China Labor Bulletin. 2018. The Workers' Movement in China: 2015-2017, Research Report by China Labor Bulletin.

[26] Christensen, Darin, and Francisco Garfias. "Can you hear me now? How communication technology affects protest and repression." Quarterly journal of political science 13.1 (2018): 89.

[27] Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." Review of Economic Studies 80(4): 1422–1458.

[28] Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. 2009. "Why Resource-Poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." American Political Science Review 103(4): 645–68.

[29] Elhorst, J. Paul. "Dynamic models in space and time." Geographical Analysis 33.2 (2001): 119-140.

[30] Elhorst, J. Paul. "Spatial panel data models." Spatial econometrics. Springer, Berlin, Heidelberg, 2014. 37-93.

[31] Enikolopov, Ruben, Alexey Makarin, and Maria Petrova. "Social media and protest participation: Evidence from Russia." Econometrica 88.4 (2020): 1479-1514.

[32] Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. "Media and Political Persuasion: Evidence from Russia." American Economic Review, 111(7): 3253-85

[33] Fergusson, Leopoldo and Carlos Molina, 2020. "Facebook Causes Protests," Documentos de Trabajo LACEA 018004, The Latin American and Caribbean Economic Association - LACEA

[34] Fu, King-wa, Chung-hong Chan, and Marie Chau. 2013. "Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and the Real-Name Registration Policy." IEEE Internet Computing 17(3): 42–50.

[35] Gehlbach, Scott, Konstantin Sonin, and Milan W. Svolik. "Formal models of nondemocratic politics." Annual Review of Political Science 19 (2016): 565-584.

[36] Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya. "3g internet and confidence in government." The Quarterly Journal of Economics (2020).

[37] Hsiao, Cheng. Analysis of panel data. No. 54. Cambridge university press, 2014.

[38] King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism But Silences Collective Expression." American Political Science Review 107(2): 1–18

[39] King, Gary, Jennifer Pan, and Margaret E Roberts. 2014. “Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation.” Science 345(6199): 1–10.

[40] King, Gary, Jennifer Pan, and Margaret E. Roberts. "How the Chinese government fabricates social media posts for strategic distraction, not engaged argument." American political science review 111.3 (2017): 484-501.

[41] King, Gary, and Langche Zeng. "Logistic regression in rare events data." Political analysis 9.2 (2001): 137-163.

[42] König, Michael D., Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti. "Networks in conflict: Theory and evidence from the great war of Africa." Econometrica 85, no. 4 (2017): 1093-1132.

[43] Kuruvilla, Sarosh and Hao Zhang. 2016. "Labor Unrest and Incipient Collective Bargaining in China," Management and Organization Review, Vol 12(1): 159-187.

[44] Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon. "What is Twitter, a social network or a news media?." 2010. Proceedings of the 19th international conference on World wide web. AcM, p. 591-600.

[45] Lee, Ching Kwan, and Yonghong Zhang. "The power of instability: Unraveling the microfoundations of bargained authoritarianism in China." American Journal of Sociology 118.6 (2013): 1475-1508.

[46] Little, Andrew T. "Communication technology and protest." The Journal of Politics 78.1 (2016): 152-166.

[47] Lohmann, Susanne. "A signaling model of informative and manipulative political action." American Political Science Review (1993): 319-333.

[48] Lorentzen, Peter. 2014. "China's Strategic Censorship." American Journal of Political Science 58(2): 402–414.

[49] Lorentzen, Peter. "Designing contentious politics in post-1989 China." Modern China 43.5 (2017): 459-493.

[50] Manacorda, Marco, and Andrea Tesei. "Liberation technology: Mobile phones and political mobilization in Africa." Econometrica 88.2 (2020): 533-567.

[51] Manski, C., 1993. Identification of endogenous social effects: The reflection problem. Review of Economic Studies 60 (3), 531 542.

[52] Miao, M. J. Ponticelli, and Y. Shao. 2021. "Eclipses and the Memory of Revolutions: Evidence from China," working paper, Northwestern University.

[53] Morozov, Evgeny. 2012. The Net Delusion: The Dark Side of Internet Freedom. Reprint edition. Public Affairs.

[54] Nickell, Stephen. 1981. Biases in Dynamic Models with Fixed Effects. Econometrica, Vol. 49, No. 6, pp. 1417-1426

[55] O'brien, Kevin J., and Lianjiang Li. 2006. Rightful resistance in rural China. Cambridge University Press.

[56] Olson, Mancur. 1965. The Logic of Collective Action. Harvard University Press, Cambridge, Mass.

[57] Passarelli, Francesco, and Guido Tabellini. "Emotions and political unrest." Journal of Political Economy 125.3 (2017): 903-946.

[58] Qin, Bei, David Strömberg, and Yanhui Wu. "Why does China allow freer social media? Protests versus surveillance and propaganda." Journal of Economic Perspectives 31.1 (2017): 117-40.

[59] Qin, Bei, David Strömberg, and Yanhui Wu. "Media bias in China." American Economic Review 108.9 (2018): 2442-76.

[60] Roberts, Margaret E. Censored: distraction and diversion inside China's Great Firewall. Princeton University Press, 2018.

[61] Shirky, Clay. 2011. "The Political Power of Social Media: Technology, the Public Sphere, and Political Change." Foreign Affairs, January/February.

[62] Steinhardt, H. Christopher. 2017. "Discursive Accommodation: Popular Protest and Strategic Elite Communication in China." European Political Science Review, 9(4):539-560.

[63] Su, Yang, and Xin He. "Street as courtroom: state accommodation of labor protest in South China." Law & Society Review 44.1 (2010): 157-184.

[64] Tucker, J. A., Nagler, J., MacDuffee, M., Metzger, P. B., Penfold-Brown, D., & Bonneau, R. (2016). Big data, social media, and protest. Computational social science, 199.

[65] Yanagizawa-Drott, David. 2014. "Propaganda and Conflict: Evidence from the Rwandan Genocide." The Quarterly Journal of Economics, 129(4): 1947-1994.

[66] Zhang, Han and Jennifer Pan. 2019. "CASM: A Deep-Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media" Sociological Methodology 49(1): 1-57.

[67] Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. "Political effects of the internet and social media." Annual Review of Economics 12 (2020): 415-438.

Tables and Figures

Table 1. Protests by cause, 2006-2013

	size					
Cause	<100	100s	1,000s	10,000s	unknown	Total
Government policy or corruption	58	105	93	6	51	313
Housing and land	23	112	79	5	62	281
Labor	7	52	45	1	15	120
Police and court	12	44	77	7	19	159
Environment and health	3	36	51	5	26	121
Firm and finance	0	18	12	3	3	36
Ethnic and religion	9	21	8	3	9	50
School and education	5	17	12	0	5	39
Social conflicts	0	5	9	2	1	17
International relations	0	0	1	0	1	2
Transportation	0	0	0	0	1	1
Human right	2	2	2	0	5	11
Unknown	0	0	0	1	2	3
Total	119	412	389	33	200	1,153

Table 2. Strikes by industry, 2007-2013

	size					
Industry	<100	100s	1,000s	>=10,000s	unknown	Total
Manufacturing	99	400	216	5	72	792
Taxi	26	58	14	2	19	119
Transportation	138	83	13	0	23	257
Services	63	60	4	0	11	138
Education	36	32	12	0	6	86
Construction	50	29	12	2	3	96
Retail	15	13	2	0	5	35
Mining	1	8	3	0	3	15
Other	3	8	4	0	2	17
Unknown	0	0	3	0	0	3
Total	431	691	283	9	144	1,558

Table 3. Event spread across cities: baseline model with h(x)=ln(5x+1)

		(1) Protest	(2) Protest	(3) Strike	(4) Strike
Number events 1-2 days prior, retweet weighted	$h(\sum f_{ijt} y_{jt-1})$	0.055***	0.055***	0.039***	0.037***
		(0.013)	(0.012)	(0.008)	(0.007)
Number events 1-2 days prior, distance weighted	$h(\sum d_{ij} y_{jt-1})$	-0.008	-0.007	0.031**	0.030**
		(0.008)	(0.008)	(0.013)	(0.012)
Number events 1-2 days prior	y_{it-1}	0.009**	0.009**	0.017***	0.017***
		(0.004)	(0.004)	(0.004)	(0.004)
Total number retweets	$h(\sum f_{ijt})$	-0.000	-0.000	-0.001	-0.001
		(0.000)	(0.000)	(0.001)	(0.001)
Weibo posts		0.006***	0.006***	0.011***	0.010***
		(0.002)	(0.002)	(0.004)	(0.003)
Controls		No	Yes	No	Yes
Observations		668,260	668,260	713,702	713,702
R-squared		0.017	0.017	0.027	0.027
QPtest		0.07	0.15	0.27	0.32

Notes: Results are from a linear regression of an event dummy. The unit of observation is city by day. The function is h(x)=ln(5x+1). The regression includes city and day fixed effects. Controls are population, GDP, shares of the industrial and tertiary sectors in GDP, and the number of cell phone users . The QP statistic reports the p-value of the test for serial correlation in fixed-effects model of Born and Breitung (2016). Standard errors are two-way clustered by date and city. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Event spread across cities: extended model with h(x)=ln(5x+1)
allowing for arbitrary time-invariant heterogeneity in the spread across city pairs

	(1) Protest	(2) Protest	(3) Strike	(4) Strike
Number events 1-2 days prior, retweet weighted	0.049***	0.047***	0.025***	0.024***
	(0.013)	(0.013)	(0.008)	(0.008)
Total number retweets	0.000	0.000	-0.000	-0.001
	(0.000)	(0.000)	(0.001)	(0.001)
Weibo posts	0.006**	0.005**	0.009***	0.009***
	(0.002)	(0.002)	(0.003)	(0.003)
Observations	668,260	668,260	713,702	713,702
R-squared	0.223	0.223	0.239	0.239
Controls	No	Yes	No	Yes

Notes: Results are from a linear regression of an event dummy. The unit of observation is city by day. The function is h(x)=ln(5x+1). The regression includes city and day fixed effects and city-pair fixed effects. Controls are log of one plus the sum of all retweets by users in city i by posts from all users j in all other cities, population, GDP, shares of the industrial and tertiary sectors in GDP, and the number of cell phone users . Standard errors are two-way clustered by date and city. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Three-period analysis

	Period	(1) Protest	(2) Protest	(3) Strike	(4) Strike
Number events 1-2 days prior, average retweet weighted	0	0.260***	0.270***	0.025	0.020
		(0.085)	(0.092)	(0.214)	(0.235)
Number events 1-2 days prior, average retweet weighted	1	0.796***	0.895***	0.622***	0.632***
		(0.200)	(0.206)	(0.135)	(0.133)
Number events 1-2 days prior, average retweet weighted	2	0.379***	0.400***	0.752***	0.779***
		(0.111)	(0.117)	(0.096)	(0.104)
Number events 1-2 days prior, distance weighted	0	0.004	0.003	0.119*	0.116
		(0.017)	(0.018)	(0.070)	(0.071)
Number events 1-2 days prior, distance weighted	1	-0.032	-0.029	0.113*	0.106*
		(0.040)	(0.040)	(0.061)	(0.058)
Number events 1-2 days prior, distance weighted	2	0.024	0.022	0.025	0.031
		(0.026)	(0.026)	(0.036)	(0.035)
Number events 1-2 days prior		0.009***	0.009***	0.010***	0.010***
		(0.002)	(0.002)	(0.002)	(0.001)
Controls		No	Yes	No	Yes
Observations		1,131,840	1,023,893	1,224,880	1,119,858
R-squared		0.019	0.020	0.048	0.049
Pr(b period: 1=0)		0.009	0.003	0.028	0.027
Pr(b period: 2=0)		0.379	0.380	0.001	0.001

Notes: Results are from a linear regression of an event dummy. The coefficients are scaled to measure marginal effects. The unit of observation is city by day. The function is h(x)=ln(5x+1). The regression includes city-by-period and day fixed effects. Controls include interaction terms between period-fixed effects and the following variables: population, GDP, shares of the industrial and tertiary sectors in GDP, and the number of cell phone users. The sum of weights, $\sum \bar{f}_{ij,}$, is constant and perfectly collinear with city fixed effects. Standard errors are two-way clustered by date and city. *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Event spread, within and across categories

	(1)	(2)
	Protest	*Strike*
Within		
Number events 1-2 days prior, retweet weighted	0.0642***	0.0606***
	(0.0132)	(0.0140)
Number events 1-2 days prior, distance weighted	-0.0089	0.0943***
	(0.0107)	(0.0266)
Across		
Number events 1-2 days prior, retweet weighted	0.0086***	0.0064***
	(0.0029)	(0.0020)
Number events 1-2 days prior, distance weighted	-0.0017	0.0056
	(0.0026)	(0.0040)
Observations	8,687,380	7,137,020
R-squared	0.0078	0.0162
Controls	Yes	Yes
Category	cause	industry
Mean within	0.0013	0.0020
Mean across	0.0152	0.0179

Notes: Results are from a linear regression of an event dummy. The unit of observation is city by day. The function is h(x)=h(5x+1). The regression includes city and day fixed effects. Controls are h($\sum f_{ijt}$), population, GDP, shares of the industrial and tertiary sectors in GDP, and the number of cell phone users . Standard errors are two-way clustered by date and city.

Table 7. Effects on event incidence

	(1)	(2)	(3)	(4)	(5)	(6)
		Strike			Protest	
Weibo posts	0.203***	0.188***	0.179***	0.199***	0.186***	0.136***
	(0.045)	(0.045)	(0.045)	(0.043)	(0.042)	(0.030)
Weibo posts, inland			0.038			0.236***
			(0.053)			(0.047)
Observations	23,475	23,463	23,463	21,993	21,981	21,981
R-squared	0.185	0.187	0.187	0.146	0.148	0.154

Notes: Results are from a linear regression of an event dummy. Unit of observation is city by month. The regression includes city and month fixed effects. Controls are population, GDP, shares of the industrial and tertiary sectors in GDP, and the number of cell phone users. Standard errors two-way clustered by city and month in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1. Number of events and Weibo posts per capita per month

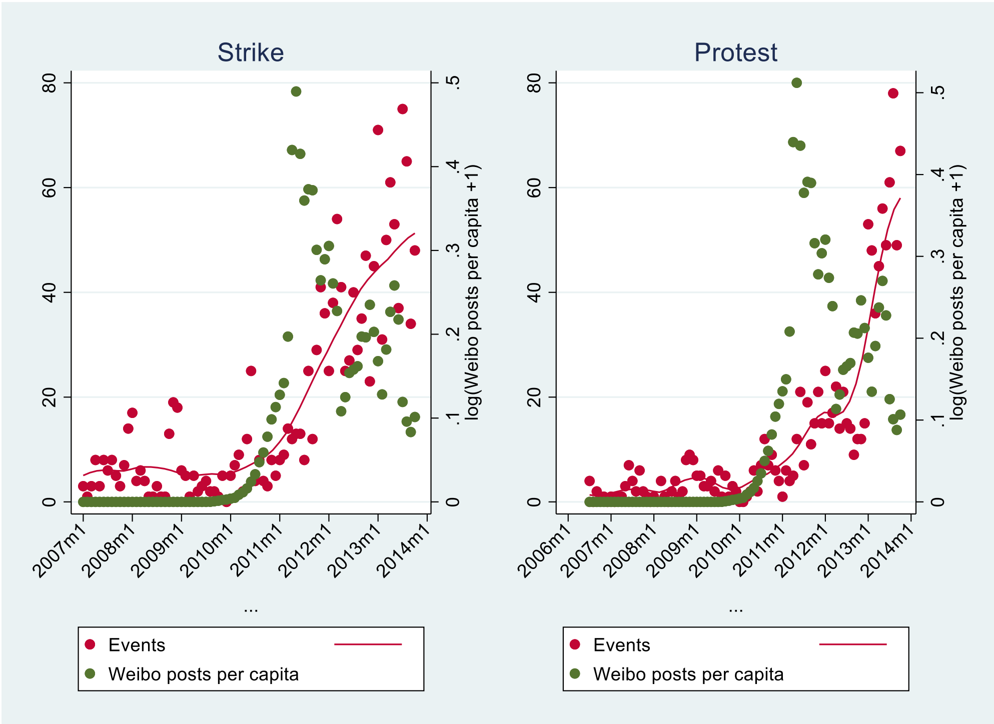


Figure 2. Monte Carlo simulations with observability driven by Weibo and no network spread effect

Notes: The blue line is at the beta-coefficient of the DGP. The red line is at the mean estimated coefficient using the simulated data.

Figure 3. Estimated marginal effects using constant retweeting matrix

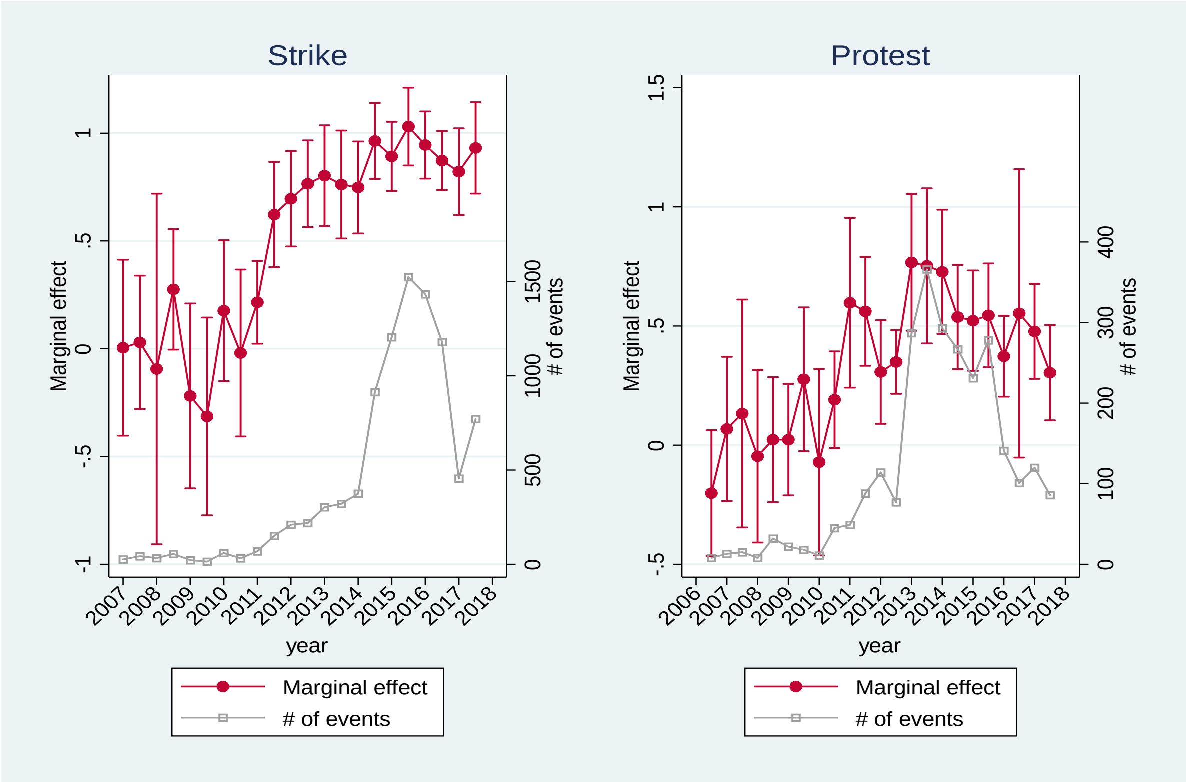
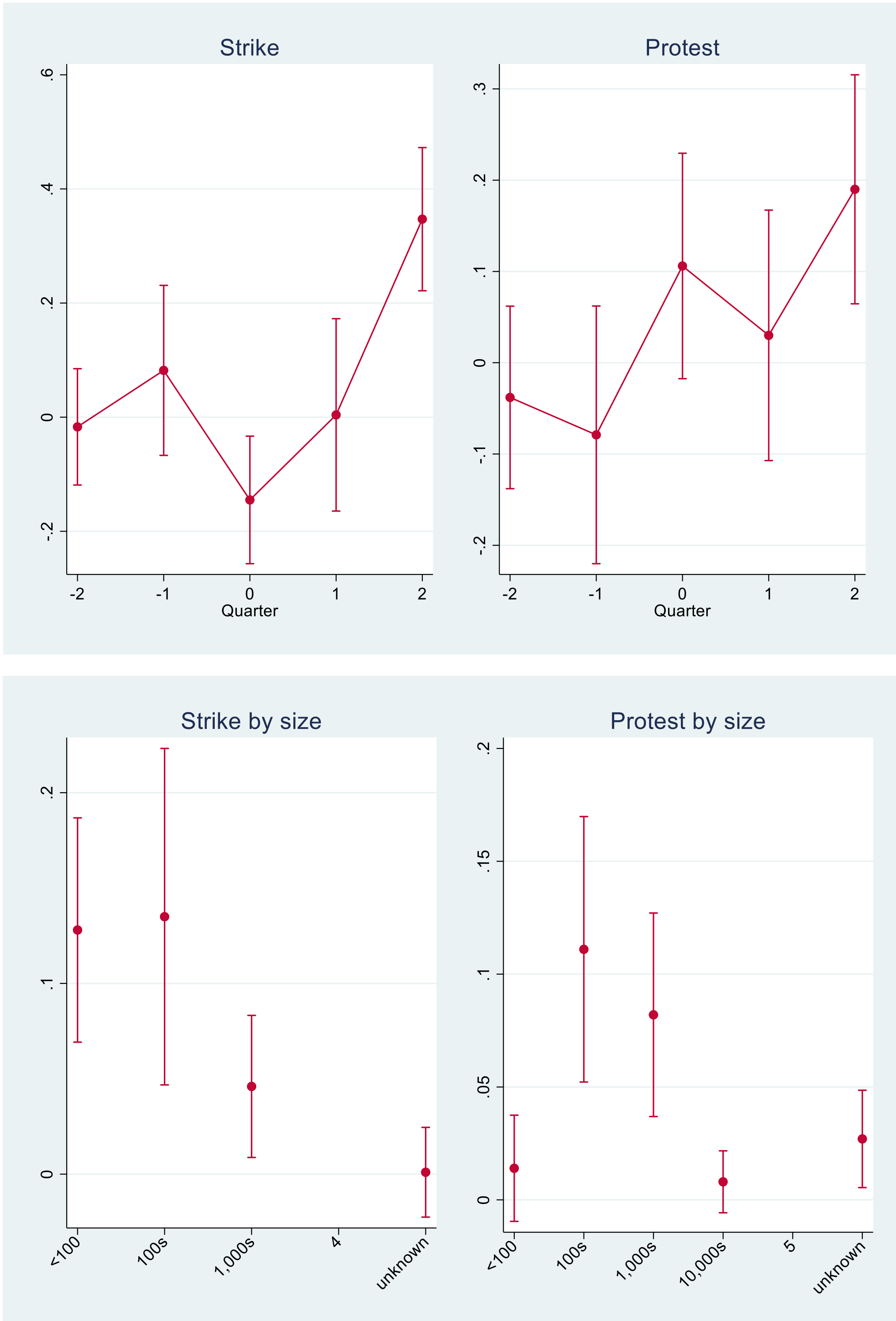


Figure 4. Effect duration

Figure 5. Dynamic and heterogenous effects on event incidence



Notes: These figures plot the coefficients of Weibo penetration estimated from Equation (3). The upper panel shows the dynamic effects. The lower panel shows heterogenous effects by size.

A Online Appendix to Social Media and Collective Action in China

A.1 Marginal effects

To interpret β in Equation (2), note that the average marginal effect of y_{jt-1} across locations $j \neq i$ on $\Pr(y_{it})$ is

$$\beta h'(x_{it}) \frac{1}{N} \sum_{j \neq i} f_{ijt-1}.$$

Then, the sum of these effects across all locations is

$$\beta \sum_{i} h'(x_{it}) \frac{1}{N} \sum_{j \neq i} f_{ijt-1}.$$

For the linear function, $h'(x_{it}) = 1$. Hence, the above expression equals the multiplication of β, and the average row sum of f_{ijt-1} is:

$$\beta \frac{1}{N} \sum_{i} \sum_{j \neq i} f_{ijt-1}.$$

This, averaged across time, equals

$$\beta \frac{1}{N} \frac{1}{T} \sum_{t} \sum_{i} \sum_{j \neq i} f_{ijt-1} = \beta \widetilde{E}(\sum_{j \neq i} f_{ijt-1}) = \beta,$$

where $\widetilde{E}$ is the sample mean. The last step follows since the average row sum of f_{ijt-1} is normalized to one.

For the logarithmic function,

$$\beta \sum_{i} h'(x_{it}) \frac{1}{N} \sum_{j \neq i} f_{ijt-1} = \beta \frac{1}{N} \sum_{i} \frac{5 \sum_{j \neq i} f_{ikt-1}}{5 \sum_{j \neq i} f_{ijt-1} y_{kt-1} + 1}.$$

The above expression, averaged across time, equals

$$\beta \widetilde{E}(\frac{5 \sum_{j \neq i} f_{ikt-1}}{5 \sum_{j \neq i} f_{ijt-1} y_{kt-1} + 1}) = s\beta,$$

with s being defined as a scale factor. Evaluated at the sample means and given the normalization of the average row sum of f_{ijt-1} to 1,

$$s = \frac{5}{5\overline{\sum_{i \neq j} f_{ijt-1} y_{jt-1}} + 1}.$$

For strikes, the sample mean value is

$$\overline{\sum_{i \neq j} f_{ijt-1} y_{jt-1}} = 0.0193.$$

Hence, $s_t = 4.56$, and the estimate $\widehat{\beta} = 0.037$ (column 4 of Table 3) should be multiplied by 4.56 to be comparable with the estimate of the linear model in column 4 of Table A2. That is, the marginal effect is $4.56 * 0.037 = 0.17$. For protests, the scaling factor is $s_t = 4.63$, and the marginal effect is $4.63 * 0.055 = 0.25$. These effects are larger than the corresponding estimates from the linear model reported in columns 2 and 4 of Table A2, which are 0.17 for protests and 0.12 for strikes.

A.2 Stationarity

In our setting, stationarity of the protest and strike processes is not simply an econometric issue. Whether these processes are stable or exploding is likely to be a core concern of an authoritarian regime (the Chinese central government in our context). Stationarity in a dynamic spatial panel data model depends on the parameters of the model as well as on the spatial weight matrix, which determines the amount of feedback in the process. For the location's own autoregressive term and the distance weighted term, this feedback is constant over time (since α, γ and the distance matrix D are constant). However, more intensive use of social media will increase the feedback because each individual row in the forwarding matrix, F, does not sum up to one. This implies that the marginal effect of a change in y_{t-1} on the probability of a protest or strike differs across localities and time. In the linear model, the average effect on a particular date t equals $\beta\overline{f}_{t-1}$, where $\overline{f}_{t-1}$ is the average row sum across locations on that date. The maximum of such a row sum equals 10.5, an order of magnitude larger than the average row sum. This implies that the sufficient conditions for stationarity are not fulfilled for the linear model.[22]

However, the actual data do not exhibit the explosive path suggested by the estimated linear model. One reason could be that the government, after observing that the process was exploding, stepped in and struck down protests and strikes. But we do not have direct evidence to support this explanation. A more likely reason is that the model is mis-specified in that the marginal effects of protests and strikes are assumed linear in the number of cross-sectional events, that is, in the extent of spread of events.

The linearity assumption may be incorrect for several reasons. First, if the mechanism is through learning, then it is likely that the marginal value of information from an additional

[22]Formally, sufficient conditions for stationarity are $|\beta\omega_{F_t,\max}| + |\gamma\omega_{D,\max}| < 1$, where $\omega_{D,\max}$ is the largest real characteristic root of the matrix D (largest negative if $\gamma < 0$) and $\omega_{F_t,\max}$ correspondingly for the F_t matrix (Elhorst 2014). The greatest characteristic roots z of a irreducible non-negative matrix A with maximum row-sum $R(A)$ satisfy $|z| < R(A)$ (Brauer and Gentry 1970). Hence, the second criterion is fulfilled if $|\gamma R(D)| + |R(\beta F_t)| < 1$. It is clear that this condition is not fulfilled in our case with $h(x) = x$. The estimated process is explosive.

event is falling. Second, if the mechanism is through information about the number of protests induced by a common cause, the incentive to organize an additional protest so as to increase government awareness is likely to decrease. Third, it may be the case that there are limited locations where people are so emotionally responsive to a particular issue and protest when they see other people protesting.

Regardless of the specific reason, the total number of protests is likely to be capped, and the process will be eventually concave. This motivates us to model the spread process by a concave function $h(x) = \ln(5x + 1)$. For x involving lagged events in Equation 2, we insert $\ln(5\sum_{i\neq j} f_{ijt-1}y_{jt-1} + 1)$ instead of $\sum_{i\neq j} f_{ijt-1}y_{jt-1}$ ($x = \sum_{i\neq j} f_{ijt-1}y_{jt-1}$). This function is chosen to be sufficiently concave to stabilize the process in the simulations. Figure A5 shows that the linear model and the concave model have similar slopes in the region where the data density is high. Therefore, the estimated effects are very similar. However, some simulations generate high levels of $\sum_{i\neq j} f_{ijt-1}y_{jt-1}$. In these cases, the concave model is stable whereas the linear model explodes. For this reason, we use the concave model in the simulations and report the results of it in the main text while reporting the results of the linear model in the appendix.

Tables A2 and A4, the counterparts of Tables 3 and 4, report the results using the linear transformation $h(x) = x$. To interpret the coefficient magnitudes, the estimates in column 2 implies that a protest occurring in a given location in the previous two days increased the expected number of locations with protests this day by 0.163.[23] The equivalent number of strikes is 0.093. Relative to the mean event probability, this is an increase of 48 percent for protests and 28 percent for strikes.[24]

A.3 Monte Carlo simulations

A.3.1 The baseline model

We run a set of Monte Carlo simulations to assess the Nickel bias in the estimated coefficients of our baseline model. We first estimate the parameters $\alpha, \beta, \gamma, \delta_t$ and δ_i from a regression specified as in Equation 2 without Weibo penetration and controls (for the sake of simplicity). We then generate data using the estimated parameters, adjusted such that $\delta_t + \delta_i \geq 0$ and estimate the model on the generated data. We repeat this procedure 100 times.

Figure A6 plots the distribution of t-statistics of coefficients α, β and γ in Equation (2) against the standard normal density. The bias is very small, as evident from the negligible

[23] In the case $h(x) = x$, the marginal effect of a strike at $y_{j,t-1}$ on the strike probability $\Pr(y_{it})$, through the retweets-weighted term, equals $\beta f_{ijt\ 1}$. Our weighting matrices, D, and F, are normalized so that the average row-sum equals one. Hence, β measures the average increase in strike probability if there was a strike in all locations on the previous day. Since the average column-sum of F_t is one, β also measures the expected total increase in strike probability across all locations at date t due to one strike at a random locality on the previous day.

[24] The mean of the protest incidence is 0.0015, and there are 224 prefectures other than the one where the first strike took place. Hence, the effect is 0.16/(224*.0015)=0.48. For strikes, the corresponding number is 0.09/(218*.0015)=0.28.

difference (in a magnitude of the third digit) between the true and the mean estimated β for both protests and strikes; see the left panels of Figure A7.

A.3.2 Time-invariant heterogeneity in event spread across locations

In Section 4.2, we report the estimates of a model including fixed effects for arbitrary time-constant spread across locations. We run a set of Monte Carlo simulations to assess the Nickel bias in this model. Specifically, we use the baseline model for the data generating process and then estimate the interaction-fixed-effects model using the simulated data.

Figure A7 shows the distribution of β-estimates from the Monte Carlo simulations. The graphs to the left show the results from estimations without interaction-fixed effects (Equation 2), corresponding to the results in Table 3. The graphs to the right are based on the regressions that include interaction-fixed effects, corresponding to the results in Table 4. The blue (green) line shows the coefficients from the estimations without (with) interaction-fixed effects, while the red line shows the mean coefficients from the simulated data. These graphs show a clear, albeit small, bias.

We also test the possibility of autocorrelated errors on the simulated data in which the autocorrelation in errors is absent by construction. The test for autocorrelation in the baseline model verifies this, although it slightly over-rejects the no-autocorrelation hypothesis. The model with interaction-fixed effects can control for the same pattern as the baseline model, but many interaction-fixed effects are imprecisely estimated. Parcelling out slightly incorrect autocorrelated terms generates autocorrelated errors. Hence, it is not surprising that the autocorrelation test for the model with interaction-fixed effects rejects no autocorrelation.

Presumably, autocorrelation would disappear as the sample size grows to infinity, because the coefficients are consistently estimated and would converge to the data generation process. The lack of autocorrelation in the errors in Table 3 shows that there is no significant spread of protests other than what is captured in the baseline model. Although the specification used in Table 4 will generate autocorrelated errors in small samples, this will not severely bias the estimated coefficients as shown in Figure A7.

A4. Additional Empirical Results

Table A1. Summary of statistics for the main variables

Protest (2006-2013)						
		count	mean	sd	min	max
Having event on the day:	y_{it}	668260	0.002	0.039	0.000	1.000
Number events 1-2 days prior:	y_{it-1}	668260	0.005	0.099	0.000	2.398
Number events 1-2 days prior, distance weighted:	$h(\sum d_{ij}y_{jt-1})$	668260	0.016	0.034	0.000	0.546
Number events 1-2 days prior, retweet weighted:	$h(\sum f_{ijt}y_{jt-1})$	668260	0.067	0.149	0.000	1.498
Total number of retweets:	$h(\sum f_{ijt})$	668260	1.265	1.156	0.000	2.916
Weibo posts:		668260	0.060	0.185	0.000	2.647
# of prefectures		247				
Strike (2007-2013)						
Having event on the day:	y_{it}	713702	0.002	0.045	0.000	1.000
Number events 1-2 days prior:	y_{it-1}	713702	0.007	0.114	0.000	2.398
Number events 1-2 days prior, distance weighted:	$h(\sum d_{ij}y_{jt-1})$	713702	0.020	0.038	0.000	0.773
Number events 1-2 days prior, retweet weighted:	$h(\sum f_{ijt}y_{jt-1})$	713702	0.081	0.158	0.000	1.476
Total number retweets:	$h(\sum f_{ijt})$	713702	1.313	1.115	0.000	2.863
Weibo posts:		713702	0.059	0.180	0.000	2.647
# of prefectures		282				

Notes: Variables are in the form of ln(5x+1). “Weibo posts” is defined as log(weibo posts per capita+1).

Table A2. Event spread across cities: baseline model with h(x)=x

	(1)	(2)	(3)	(4)
	Protest	*Protest*	*Strike*	*Strike*
Number events 1-2 days prior, retweet weighted	0.170***	0.169***	0.122***	0.120***
	(0.044)	(0.043)	(0.026)	(0.025)
Number events 1-2 days prior, distance weighted	-0.031	-0.030	0.143**	0.137**
	(0.037)	(0.037)	(0.058)	(0.055)
Number events 1-2 days prior	0.015**	0.015**	0.030***	0.030***
	(0.007)	(0.007)	(0.007)	(0.006)
Total number retweets	0.002***	0.002***	0.002**	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Weibo posts	0.006***	0.006***	0.010***	0.009***
	(0.002)	(0.002)	(0.004)	(0.003)
Controls	No	Yes	No	Yes
Observations	668,260	668,260	713,702	713,702
R-squared	0.017	0.017	0.027	0.027
QPtest	0.05	0.13	0.14	0.34

Notes: Results from a linear regression of an event dummy. The unit of observation is city by day. The function is h(x)=x. The regression includes city and day fixed effects. Controls are population, GDP, shares of the industrial and tertiary sectors in GDP, and the number of cell phone users . The QP statistic reports the p-value of the test for serial correlation in fixed-effects model of Born and Breitung (2016). Standard errors are two-way clustered by date and city. *** p<0.01, ** p<0.05, * p<0.1.

Table A3. Event spread across cities: probit model

VARIABLES	(1) Strike	(2) Strike	(3) Protest	(4) Protest
Number events 1-2 days prior, retweet weighted	0.195***	0.199***	0.195**	0.181**
	(0.058)	(0.059)	(0.084)	(0.079)
Number events 1-2 days prior, distance weighted	0.548***	0.541***	0.393	0.444
	(0.151)	(0.151)	(0.330)	(0.322)
Number events 1-2 days prior	0.190***	0.191***	0.103**	0.097**
	(0.027)	(0.027)	(0.048)	(0.047)
Total number retweets	0.049*	0.051*	-0.094**	-0.095**
	(0.026)	(0.026)	(0.037)	(0.037)
Weibo posts	0.068**	0.070**	0.244***	0.236***
	(0.029)	(0.032)	(0.054)	(0.057)
Observations	581,509	581,509	558,906	558,906
Controls	No	Yes	No	Yes

Notes: Results from probit regressions of an event dummy. The unit of observation is city by day. All regressions include city fixed effects and quadratic time trends. Standard errors clustered by city in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Event spread across cities, extended model with h(x)=x
allowing for arbitrary time-invariant heterogeneity in the spread across city pairs

	(1) *Protest*	(2) *Protest*	(3) *Strike*	(4) *Strike*
Number events 1-2 days prior, retweet weighted	0.217***	0.216***	0.109***	0.107***
	(0.061)	(0.061)	(0.039)	(0.038)
Total number retweets	0.003***	0.003***	0.003***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Weibo posts	0.004**	0.004**	0.008**	0.007**
	(0.002)	(0.002)	(0.004)	(0.003)
Observations	668,260	668,260	713,702	713,702
R-squared	0.223	0.223	0.239	0.239
Controls	No	Yes	No	Yes

Notes: Results from a linear regression of an event dummy. The unit of observation is city by day. The function is h(x)=x. The regression includes city and day fixed effects, and all city pair fixed effects. Controls are population, GDP, shares of the industrial and tertiary sectors in GDP, and the number of cell phone users . Standard errors are two-way clustered by date and city: *** p<0.01, ** p<0.05, * p<0.1.

Table A5. Event spread across cities with sequentially added regressors: protests
Dependent variable: protest dummy; log specification h(x)=ln(5x+1)

VARIABLES	(1)	(2)	(3)	(4)	(5)
Number events 1-2 days prior, cumulative forwards weighted	0.052***	0.053***	0.054***	0.054***	0.055***
	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)
Number events 1-2 days prior		0.008**	0.008**	0.008**	0.009**
		(0.004)	(0.004)	(0.004)	(0.004)
Number events 1-2 days prior, distance weighted			-0.009	-0.009	-0.008
			(0.008)	(0.008)	(0.008)
Total number retweets				0.001**	-0.000
				(0.000)	(0.000)
Weibo posts					0.006***
					(0.002)
Observations	739,260	738,720	738,720	738,720	668,260
R-squared	0.014	0.015	0.015	0.015	0.017
Controls	No	No	No	No	No

Notes: The unit of observation is city by day. The regressions include city and day fixed effects. Standard errors are two-way clustered by date and city: *** p<0.01, ** p<0.05, * p<0.1.

Table A6. Event spread across cities with sequentially added regressors: strikes
Dependent variable: strike dummy; log specification h(x)=ln(5x+1)

VARIABLES	(1)	(2)	(3)	(4)	(5)
Number events 1-2 days prior, cumulative forwards weighted	0.051***	0.052***	0.047***	0.046***	0.039***
	(0.011)	(0.011)	(0.009)	(0.009)	(0.008)
Number events 1-2 days prior		0.017***	0.017***	0.017***	0.017***
		(0.004)	(0.004)	(0.004)	(0.004)
Number events 1-2 days prior, distance weighted			0.031**	0.031**	0.031**
			(0.013)	(0.013)	(0.013)
Total number retweets				0.000	-0.001
				(0.000)	(0.001)
Weibo posts					0.011***
					(0.004)
Observations	779,885	779,275	779,275	779,275	713,702
R-squared	0.023	0.025	0.025	0.025	0.027
Controls	No	No	No	No	No

Notes: The unit of observation is city by day. The regressions include city and day fixed effects. Standard errors are two-way clustered by date and city: *** p<0.01, ** p<0.05, * p<0.1.

Figure A1. Distribution of collective action events across cities (prefectures)

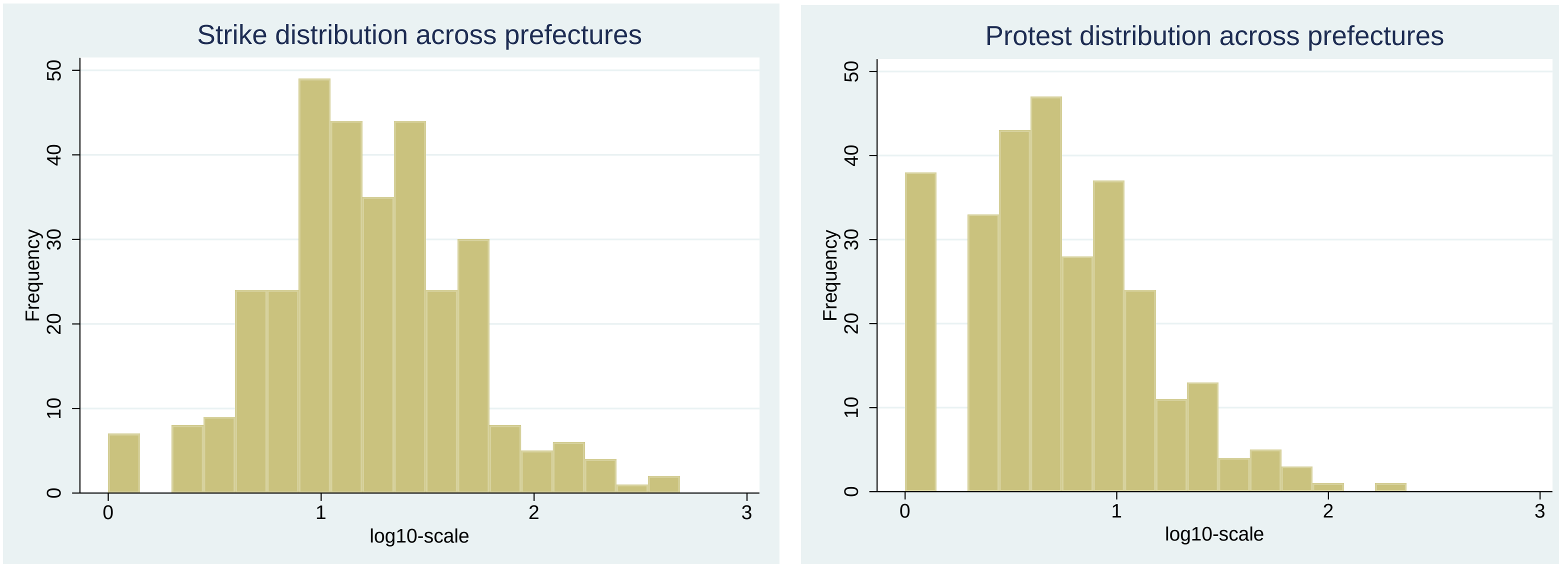


Figure A2. Strike count by city (prefecture) 2007-2013 and protest count by city (prefecture) 2006-2013

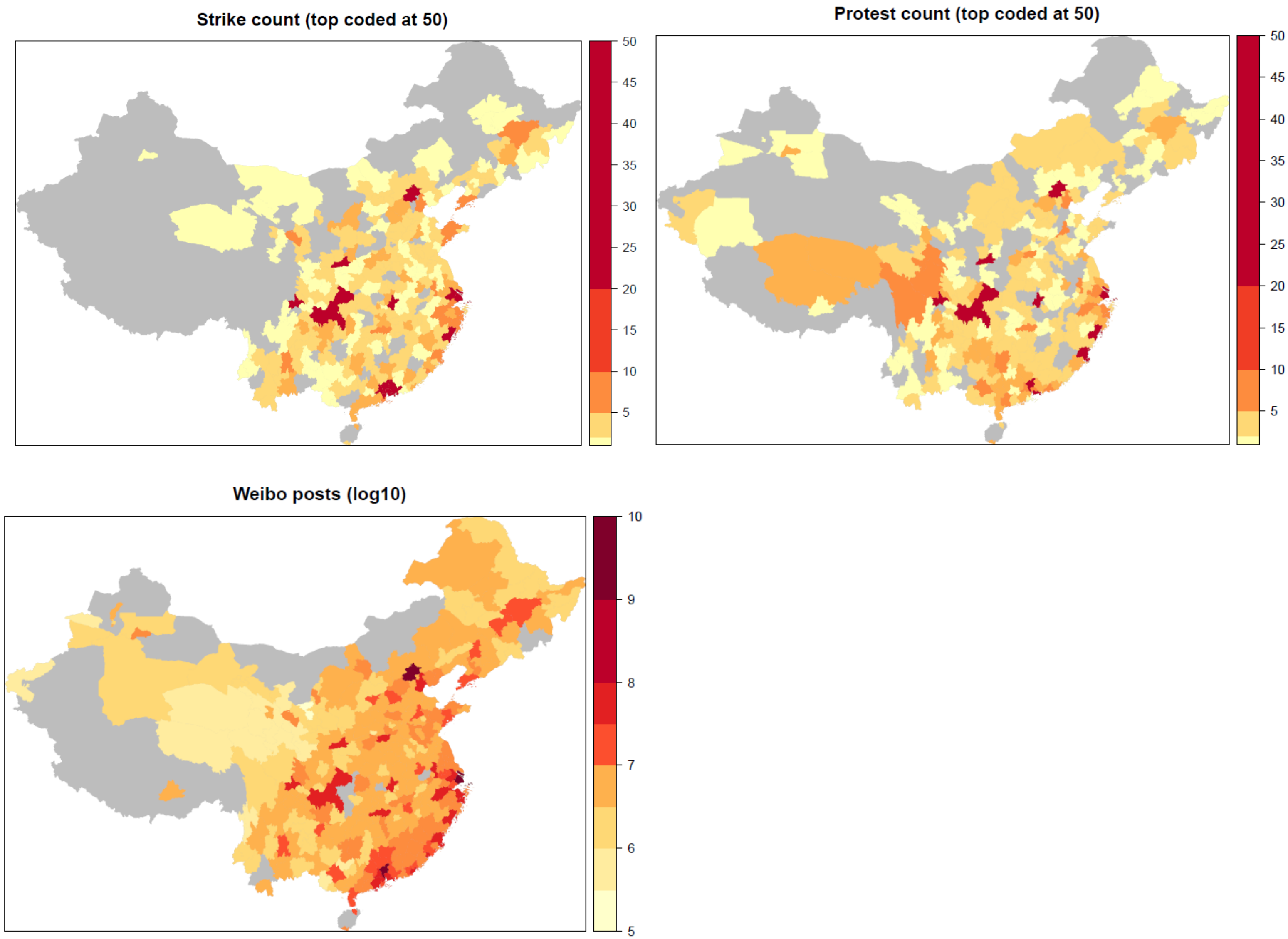


Figure A3: Predictors of relative retweeting (forwarding) between cities

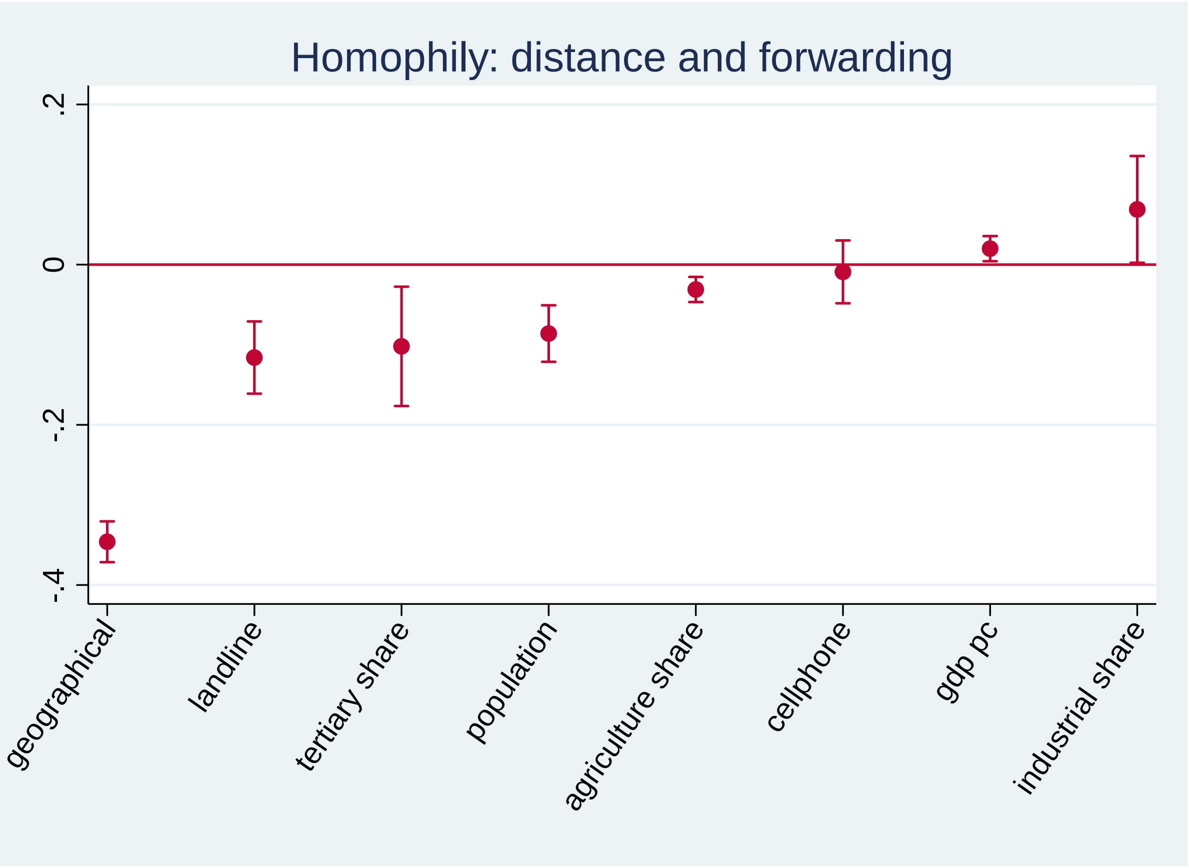
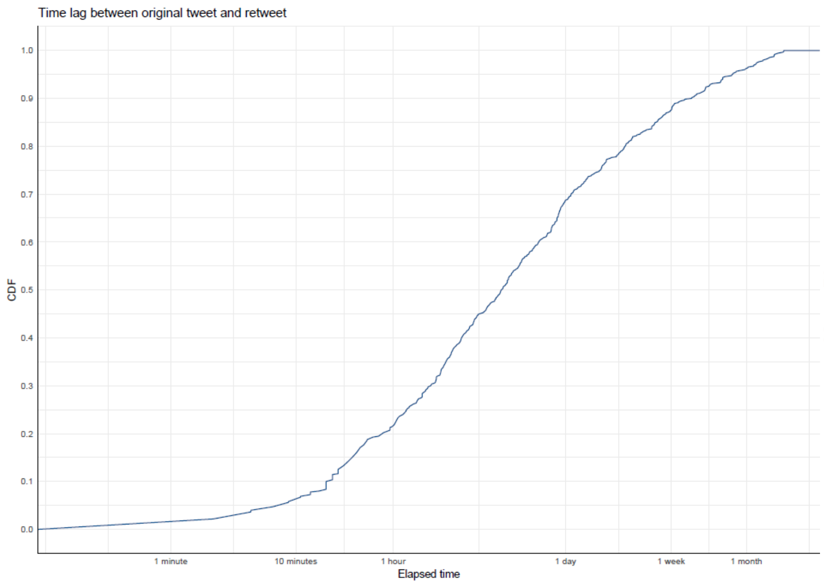


Figure A4. Retweet quantity and distance over time



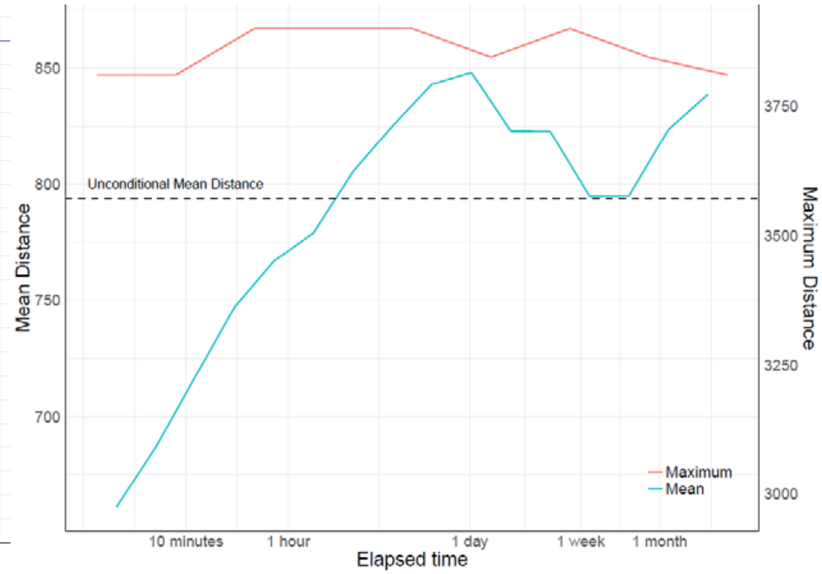
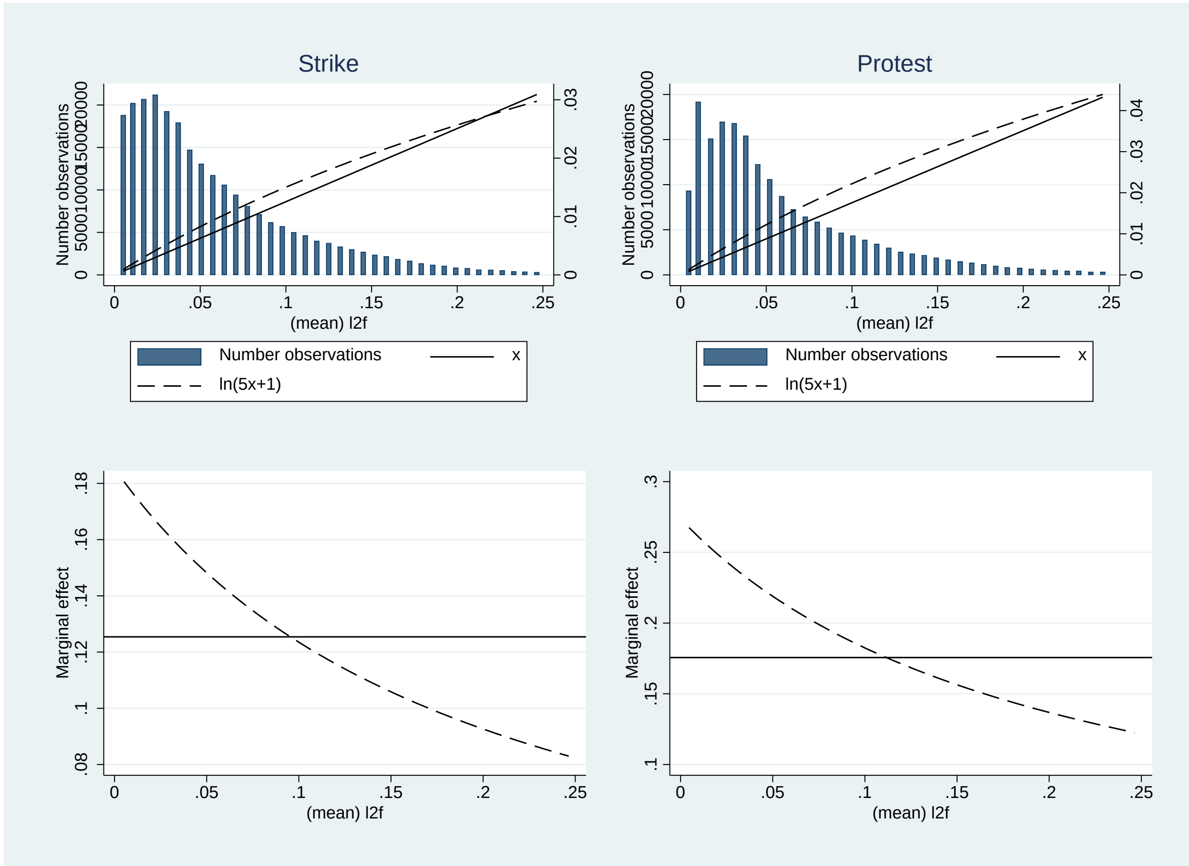


Figure A5. Functional-form comparison in sample range



Notes: The x-axis contains the sum of the retweet-weighted lagged events. The upper panel shows a histogram of the number of observations in each bin, together with beta*h(x) for the linear and the logarithim h(x)-functions. The lower panel shows the implied marginal effects from the two models. Compared to the marginal effect of the linear model (the solid line), the estimated marginal effect using the logarithmic function (the dashed line) is higher for low values and lower for high values of x.

Figure A6. Monte Carlo Simulations: distribution of t-statistic of estimated parameter = true parameter

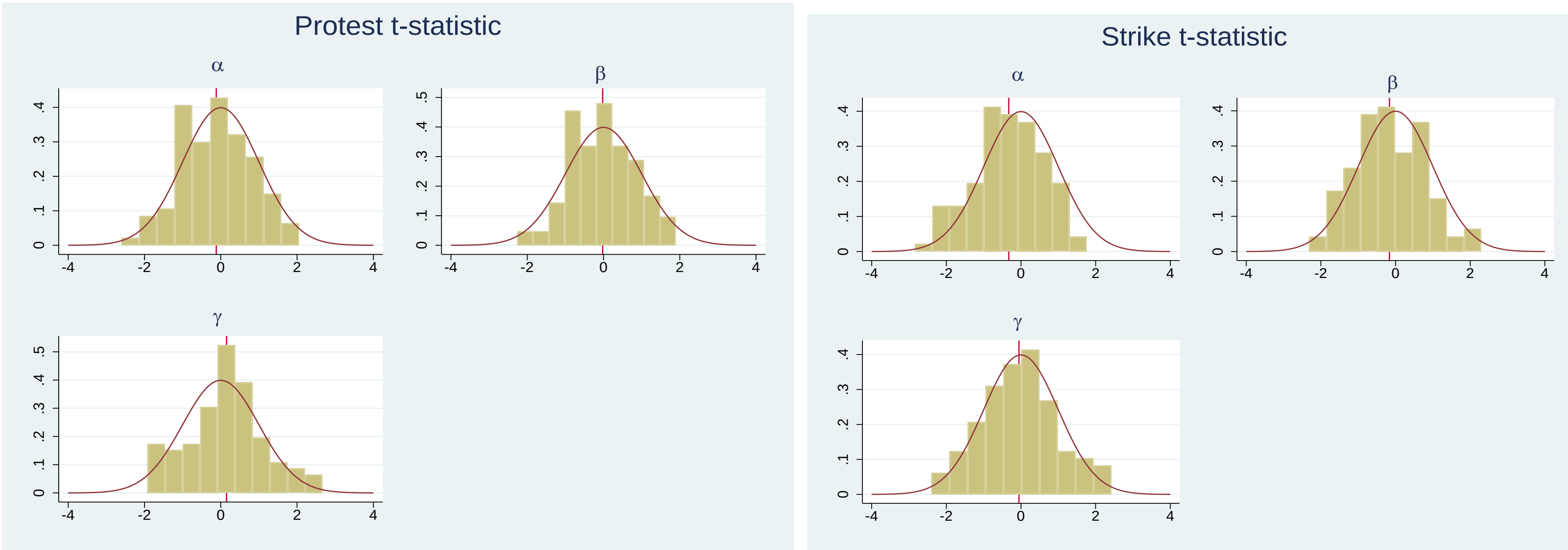


Figure A7. Monte Carlo Simulations: distribution of coefficients around DGP parameter values.

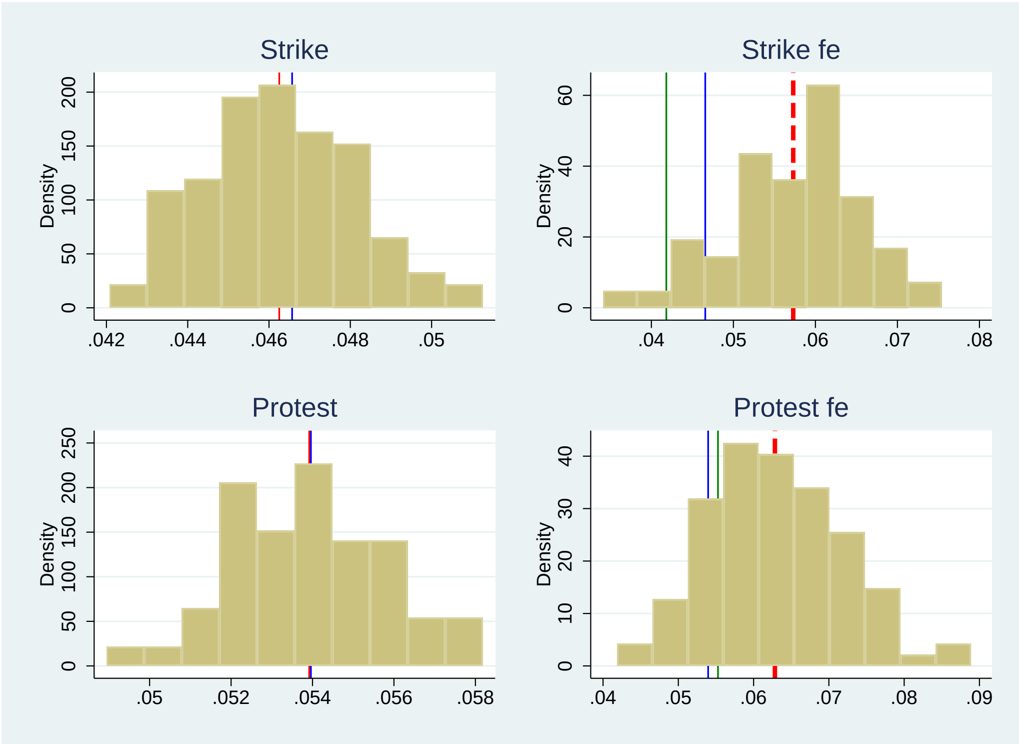


Figure A8. Estimated marginal effects using constant retweeting matrix, h(x)=x