Social Media and Protests in China

Bei Qin, David Strömberg, and Yanhui Wu[*]

September 2019

Abstract

This paper studies whether the explosive growth of social media in China affects the spread and incidence of protests. We combine a unique dataset of 13.2 billion microblog posts published during 2009-2013 with detailed information on thousands of protests and strikes during 2006-2017. We use retweets to measure the network of social media information flows across cities, and estimate the effects of this rapidly exapanding network. Despite the strict media control in China and the lack of information for explicit coordination, we find that the social media network has a sizeable and significant effect on the spread of both protests and strikes. The spread of events over social media is fast and predominantly local – between events within the same category (e.g., cause and industry); event spread across categories is still significant, albeit weaker. Furthermore, we find that social media networks increase the incidence of protests and strikes. These findings shed light on the recent debate regarding the political role of social media in autocracies.

---

[*]Qin: The University of Hong Kong, beiqin@hku.hk; Stromberg: University of Stockholm, david.stromberg@ne.su.se; Wu: University of Southern California, yanhuiwu@marshall.usc.edu.

1   Introduction

In authoritarian countries, protests are an extreme form of political participation, which imposes a substantial cost on protesters and may threaten the stability of the regime. Nevertheless, in China, protests against unpopular policies or corruption and strikes demanding fair compensation and better work conditions have increased drastically in the last two decades (China Labor Bulletin 2015, 2018). This surge coincides with an exploding use of social media in China. Although there is evidence that social media play an important role in facilitating street protests in some contexts (e.g., Acemoglu et al. 2018; Enikolopov et al. 2019), its effect in China is unclear. The Chinese social media are strictly controlled by the government to the extent that Freedom House has rated China as the world's worst abuser of internet freedom in the last few years. Extensive censorship substantially limits the generation and circulation of protests-related information; advances in information and AI technology enhances the Chinese government's ability to use social media for surveillance and thus to nip protests in the bud (Qin et al., 2017). In this extreme environment, we study the effect of social media on protests. In particular, we use large-scale textual data from Chinese social media to measure high-frequency information networks across cities and to estimate the effect of social media on the geographical spread of protests and strikes in China. Our extensive textual data allow us to rule out simple direct mechanisms, such as explicit organization of protests, and instead point to more subtle mechanisms. This sheds light on the logic and costs of information control in China, which is a leader in the political use of modern information technology and who may stand as a role model for other authoritarian regimes.

The first empirical question is whether information about protests and strikes circulates in Chinese social media at all. To this end, we study the content of a unique social media dataset consisting of 13.2 billion posts published on Sina Weibo—the most prominent microblogging platform in China—during 2009-2013. We find that, during our sample period, there was approximately 4 million microblog posts mentioning protest or strikes and that users dared to post even after censoring. Moreover, these posts predict real-world protests and strikes (see Qin et al. 2017 for more details).

It is not obvious why protests or social media coverage about them are allowed. Some argue that allowing local protests of narrow scope may help the Chinese government monitor local politicians and policies (e.g. Cai, 2008, Lorenzten, 2017, Qin et al., 2017).[1] However, an obvious risk for the regime is that initially small events may snowball into widespread movements that target national problems and politicians. A cautionary example is Solidarity in Poland, which originated from limited demand for workers' rights and better economic conditions but soon developed into pervasive resistance that proved fatal to the regime. Uncontrolled social media likely increases this risk, since its communication is speedy and

---

[1]This is the so-called "mass line" in the political dogma of the Chinese Communist Party. Its essence is to learn about public opinion and policy outcomes from people (the mass) through observing their activities, particularly collective action (Zhao 1998).

long distanced. As the information environment becomes increasingly dynamic, the dynamics of protests and strikes may change as well. Consequently, to understand the effect of social media on regime changes and the issue of media control, it is crucial to study how social media affect protest dynamics—whether they make events break the geographical boundaries and generate far-reaching social movements. This is precisely the focus of our study.

Even though a large number of posts are available, this may be inconsequential if people cannot find them because of search censorship or if the posts that start to spread geographically are systematically censored. Our unique data on retweets (forwards) allow us to study how information about protests and strikes spread across China. We identify approximately 40 million forwards of the protest and strike posts. That a user forwards (retweets) a message indicates that the user has read the message. In a subset of 3 million forwards for which we have precise timing and location information. We find that information about protests and strikes spread far and wide, around 30% of the forwards occur within one hour after 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 forwards 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 suggests that social media have the potential to generate sweeping waves of protest activities.

To study whether this is actually the case, we combine our social media dataset with detailed information on thousands of protests and strikes from 2006 to 2017. Estimating the effect of social media on the spread of real-world events is intricate. Cities with strong informational ties through social media are likely to have other communication channels (e.g., phones, meetings). Moreover, social media users tend to communicate more extensively with users from cities that are more similar to theirs. Events taking place in socially well-connected and similar cities are naturally more closely correlated, regardless of the impact of social media. Therefore, it is challenging to separately identify the informational effect of social media on events. This type of identification challenge has been widely discussed in the studies of social networks, pioneered by Manski (1993). Existing 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).

We propose a novel identification approach that exploits the time-series variation from the rapid network expansion to identify causal effects. We approximate the information flow between cities over social media by the number of forwards in one city of the posts originated from another city– posts about all topics except protests and strikes. We allow the average spread of events between cities to be correlated with the average network strength while assuming that the exact timing of the network expansion is unrelated to other changes in the spread of events. Practically, we employ two research designs. First, we use a differences-in-

differences estimator that allows for arbitrary time-invariant heterogeneity in the interactions between cities. Second, we treat the average connectedness between cities as fixed and then investigate how the average spread of events evolves across cities that are eventually closely connected through social media over time. Monte Carlo simulations justify the validity of our methods. We also show that we can consistently test the null hypothesis that the social media network does not affect event spread even if social media enable researchers to observe more events. The intuition behind the result is that increased observability per se does not affect the temporal clustering of events across cities. In contrast, one cannot accurately estimate the effects of social media penetration on the incidence of events if social media affect observability.

As mentioned, we focus on how social media affect protests spread rather than the direct effects on incidence. To clarify, by direct effects on incidence, we mean an exogenous shock to the probability that an event takes place in a particular place and time. By spread, we mean the event dynamics across locations, that is, to what extent the occurrence of an event in one place increases the probability of a following event in another place. We focus on spread because we think that it is the most interesting outcome, and also because we can consistently estimate the effect on spread even though social media affects the observability of events.

The main findings of this study are summarized as follows. First, our two research designs consistently provide strong evidence that protests and strikes spread rapidly (within 2 days) through the social media network. We estimate that, because of information flows through social media, a protest within one prefecture in the last two days increases the probability of a protest in all other prefectures by 16 percentage points. Relative to the mean event probability, this is an increase of 48 percent for protests and 28 percent for strikes. Second, we find that after a peak in 2013-2014, spread of strikes through the social media network slowed down and even ceased after 2016 and that protests continued to spread but at a substantially lower level. This is likely to be caused by the Chinese central government's more stringent control over the social media and its crackdown on collective action after 2013.[2]

Third, we find that social media help protesters break geographical, industrial, and protest cause barriers and generate widespread social movements. Not surprisingly, we find that the spread of events is predominantly local– events spread more within the same event category (for the same social cause or in the same industry). For example, strikes occurring in the manufacturing industry in a province spread more to other strikes in the same industry within the same province than non-manufacturing strikes outside of the province. However, this strong within-category spread does not mean that events do not spread across categories

---

[2]The Chinese government's stricter control of social media after 2013 is reflected in the declining media freedom index constructed by the Freedom House. Moreover, the Chinese government's procurement of media-based surveillance equipment increased massively from 2013 to 2016. Source: China Government Procurement Website.

at all. Instead, we find that social media also induce weaker but still significant spread of events across event categories. Absent social media, there is no such spread. In addition, the aggregate effect from across-category spread is larger than from within-category spread. This is because the total number of events across all categories is ten times larger than within the same category. While the individual effect of one such event is smaller across than within categories, the larger number of across-category events makes the aggregate effect larger.

We also report results on incidence, although these are likely biased because social media increases observability. Based on a difference-in-differences approach, we find that the penetration of Sina Weibo measured by the number of posts per capita is associated with an increase in the numbers of both protests and strikes. The estimated coefficients are statistically significant and sizeable. If these are interpreted as causal, they would imply that social media would account for one fifth of the total increase in protests and strikes over our sample period. Heterogenous effect across regions with different access to information before the launch of Sina Weibo suggests that the penetration of Sina Weibo still plays an important role in increasing event incidence even after accounting for the bias induced by observability.

Social media may influence the incidence and spread of protests through a number of mechanisms. First, it may carry *information about problems*, such as extensive corruption, that officials might be forced to address if there are protests. This information may affect the rational calculus of whether to protest but also induce emotional responses, create a sense of aggrievement of being unfairly treated (Pasarelli and Tabellini, 2017). Social media may also carry *information about effective protest tactics* (e.g. Little 2016; Chen and Suen 2016) and likely *government responses* to the protests. Second, it may be used for *coordination and organization* of protests (e.g., Acemoglu et al. 2018; Barbera and Jackson, 2018; Enikolopov et al., 2018). Third it may increase the *visibility* of protests. A goal of the protests is frequently to attract the attention of national leaders, who may press local administrators to solve problems. Social media increases visibility and hence the expected benefits from protesting. Finally, social media may by used by the government for *surveillance* to identify and nip protests in the bud (Qin et al, 2017).

We evaluate these mechanisms in two ways. First, some mechanisms imply that certain content must be present in the blogs, for example, the information-driven mechanisms require that this information is present in the blogs and the explicit coordination and organization imply that the posts organizing and explicitly coordinating the protests are found. We describe the frequency of this information in the blogs. Second, we study heterogenous effects. Effects through information are likely to be long lasting, since the information remains available after its posted and since what is, for example, a good protest tactic is not likely go change rapidly, We find that the effects are short lived, with the strongest effect within a week and the effect disappearing after a month. For a subset of the strike data, we also have information on whether the response was concession or repression. We find that strikes do differentially depending on the response. Perhaps this is because we study fast spread and

concessions typically take more time.

This paper primarily contributes to the emerging literature on the effect of social media and more generally information and communication technology (ICT) on collective political action in non-democratic countries. Inspired by the "colour revolution", particularly the Arab Spring, some scholars believe that social media can provide an important means for citizens to organize collective action opposing and even overthrowing the ruler, and that this power released by social media helps hold authoritarian governments accountable (e.g., Shirky 2011; Tucker et al. 2016; Acemoglu et al. 2018). Recently, Manacorda and Tesei (2018) show that mobile phones are instrumental to mass mobilization during economic downturns in Africa. Enikolopov et al. (2018) show that the penetration of a dominant Russian online social network led to more protest activity in Russia. In a field experiment in Hong Kong, Cantoni et al. (2019) show that information on other individuals' participation has an impact on an individual's protest participation. Our study advances this line of research several important aspects. First, our study focusses on the country with the strictest internet control in the world, with a Freedom House score of 88 in 2018. In comparison, Russia's internet is considerably freer with a score of 67 and rank 13. Below, we document the lack of calls for action, explicit organization of protests and even description of where and when protests in Chinese social media, which can be contrasted with the availability of such material in Russia. Second, we study protest dynamics rather than cross-sectional incidence. This dynamic feedback is key to whether protests expand to widespread movements. We do this by developing methods to measure the effect of expanding social networks using time-series variation in the presence of homophily and endogenous event observability. Third, our unique data set allows us to investigate information availability and spread over the social media network.

In a broader context, our study is related to the literature on the role of government-controlled media in autocracies. The media has long been regarded as a powerful propaganda instrument for autocrats to exert political control. This traditional view has gained support in a number of studies (eg., Enikolopov et al. 2011; Yanagizawa-Drott 2014; Adena et al. 2015). Another strand of literature argues that autocratic regimes can use media to monitor local officials (e.g., Egorov et al. 2009; Lorentzen 2014) and social media for surveillance and propaganda (e.g., Morozov 2012; Edmond 2013; Lorentzen 2017). Confirming this empirically, Qin et al. (2017) show that social media in China are very effective for surveillance of protests and strikes. This is a possible reason why regime's extreme control apparatus has allowed millions of posts about protests and strikes to spread across the country, inducing additional protests and strikes, as we show in this paper.

On the other hand, King et al. 2013, 2014, show that the government also strictly monitors and censors coverage of collective action. Even though most posts remain online, the share who are censored is higher than in any other category. This strict monitoring of collective action posts is understandable, given the high impact of social media on the spread of protests and strikes that we document in this paper.

The remainder of this paper proceeds as follows. Section 2 provide a short background of the development of social media, protests and strikes in recent years. Section 3 describes our data and provides descriptive statistics, such as the number of blog posts covering protests and strikes and how rapidly these spread across locations. Section 4 presents the main empirical analysis and results on how the social media network affect the spread of protests and strikes. Section 5 extends the analysis to the period after 2013, and studies how long-lived the effects are and whether social media induces protests and strikes to spread across causes and industries. Section 6 concludes.

2   Background

2.1   Social Media in China

After an intense period of explosive growth, social media in China today is as vibrant and extensive as in any Western country. Since we will examine effects during this period of rapid growth, we now describe it shortly. With the Chinese government blocking Twitter and Facebook and strictly controlling domestic microblogging services, the use of social media was limited until Sina Weibo appeared in August 2009. Sina Weibo is a hybrid of Twitter and Facebook, allowing users to tweet and retweet short messages with embedded pictures or videos and send private messages and write comments. Between 2009 and 2012, Sina Weibo use increased exponentially. By 2010, Sina 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, where pre-existing levels of mobile phone use was higher. Over our main sample period, which ends in 2013, Sina Weibo was the dominant microblogging platform in China. Since then, it has lost ground to WeChat, a cellphone-based social networking service, but has remained an influential platform for public communication.

This 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 posting this content and induce self-censoring (Chen and Ang 2011). To remove unwanted content, censorship on Chinese social media is regulated by the national Propaganda Department of the CCP and is implemented largely by private service providers who are registered in Beijing. The estimated extent of censorship of Sina Weibo ranges from 0.01 percent of posts published by a sample of prioritized users (Fu, Chan, and Chau 2013) to 13 percent of posts on collective action events such as protests and strikes (King, Pan, and Roberts 2013). Even though censorship is pervasive, it is far from complete even on these topics.

The government also controls content by active posting on Sina Weibo. In Qin et al. (2017), we estimate that there are 600,000 government-affiliated accounts contributing 4% of all posts regarding political and economic issues on Sina Weibo. The share of government

users is higher in areas with more extensive censoring of social media posts (Bamman et al. 2012) and in areas where newspapers are more strictly controlled as measured in Qin et al. (2018). Local governments also employ a large number of internet trolls that write posts involving cheerleading for China, the revolutionary history of the Communist Party, or other symbols of the regime (King et al., 2017).

Under these extreme control conditions, can social media still affect protests and strikes, is the relevant material posted, can people find and read it, how quickly does it disseminate? We will use our unique data to answer these questions.

2.2 Protests and Strikes

Although the 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 the protests is multi-faceted. Protests are often met with violent repression by the police or plainclothes thugs, and leaders are arrested (Lorentzen, 2017). Even if some protest demands are accommodated, organizers and active participants risk being arrested, losing their job and being under the close watch of the government. At the same time, concessions are frequent (Cai, 2010; O'Brien and Li, 2006; Su and He, 2010; Lee and Zhang, 2013)), as well as simply ignoring events viewed as sufficiently innocuous (Hoffman and Sullivan, 2015). Central policies require local officials to handle collective action events strategically rather than simply suppressing all events and top CCP leaders have made public statements urging restraint in police handling of protests (Steinhardt 2017). In many anti-corruption protests, high-rank CCP officials were eventually sent to converse with the protesters so as 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 deliver concessions (Cai, 2010).

What explains the regime's relative tolerance of protests? The literature discusses a number of possible benefits to the regime from allowing local protests on narrow issues. 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 and may help the central government to identify and correct policy oversights, gauge public sentiment, and monitor local officials (Lorentzen, 2017). Second, absolute suppression of collective action may generate distrust and undermine the legitimacy of the ruling party. Finally, some collective action events such as strikes may be welfare improving and even productivity enhancing if they result in fairer competition, redistribution and better working conditions (Cai, 2010; O'Brien and Li, 2006; Su and He, 2010; Lee and Zhang, 2013). As long as the protests are contained within scope and geography, they pose no threat to the regime. They may event be useful in identifying important social problems and corruption leaders.

In contrast, costs to the regime may be higher if protest become geographically widespread

and wider in scope. This type of protests naturally shifts the focus from local to national policies and leaders, and may undermine the legitimacy or trust in the CCP. The risk the regime faces when allowing local protests is that they exhibit a scale shift and evolve into better-organized political action and social movements (Cai, 2008). This is the lesson from, for example, the Solidarity movement in Poland, that grew from limited economic demands to widespread resistance that was large in scope and eventually lead to the downfall of the regime. Similarly, the protests in Tianmen square initially involved students in Beijing but spread to include workers in faraway regions.

3   Data

We assemble a unique dataset combining detailed information about thousands of collective action events from 2006 to 2017 together with posts published on Sina Weibo over the period of 2009-2013. We now describe how this data was collected and describe it briefly.

3.1   Sources

Data about protests and strikes in China are not available from any official sources and media coverage of these kinds of events in mainland China is rather limited. Hence, our collection of data must rely on sources outside of mainland China. Below, we explain how we collect these event data, and then provide a description of the data.

The data on protests are collected manually from the website of Radio Free Asia (RFA), which is a private non-profit international broadcasting corporation based in Washington DC. We focus on the news content published in Chinese. The news reported on the RFA website comes from news collected and written by RFA's special correspondents, as well as media outlets in mainland China, Hong Kong, Taiwan and Western media outlets such as the New York Times and the BBC. News produced by the RFA's special correspondents are widely used by Chinese news portals outside of mainland China. To the best of our knowledge, the RFA website is the most comprehensive and well-structured data source for protest events in mainland China. We searched key words related to "protest" and "demonstration" (in Chinese) on the RFA website and downloaded the relevant news reports. Several research assistants were hired to first verify that the news is indeed about events of interest and to purge duplicate information. Then, they extracted relevant information from each news report and coded the date, location, cause, and scale (number of participants) of each event.

We collect data on strikes mostly from the China Labor Bulletin (CLB), a non-governmental organization based in Hong Kong. Since its start in 1994, CLB has evolved into an influential organization 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 and onward, this data contains detailed information on the timing, location, employers involved, industry, scale, worker action, and government responses for each event.

For earlier strikes, we extracted this information from the annual reports published by CLB and supplemented this with data from Boxun, a widely-read Chinese website that is based in the US and specializes in political news.

According to interviews with CLB, their data collection before 2011 relies mostly on overseas Chinese media outlets and occasionally on information provided by labor movement activists in China. Starting from 2011, CLB also collected information on strikes by searching over social media and following accounts on Sina Weibo that specialized in labor dispute settlements and trade union activities in Mainland China. During the period of 2013-2016, CLB also collected strike information from Wickedonna, a mass-event-tracing blog that gathered information on mass demonstrations in China. After the arrest of the founders of Wickedonna by the Chinese government in June 2016 this source stopped updating its data. The CLB instead extended its data collection to other Chinese social media such as WeChat. While admitting that social media help data collection, especially for small-scale events that are less likely to be covered by traditional media, CLB expressed little concern about data quality being affected by censorship. One director of CLB said that strikes in China are typically not regarded as sensitive by the Chinese government unless a strike evolves into a violent event that threatens public security. This suggests that a strike of significant size, once reported, will become known on social media.

Our social media data was extracted from a database including 13.2 billion posts published on Sina Weibo from 2009 to 2013. The database was collected by Weibook Corp, which executed a massive data collection strategy to download blogpost from the 200-300 million active users they identified. 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 downloaded on a weekly basis. Thus, the data include posts that are later censored. For each post, they provided 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 (Qin et al., 2017).

From this Weibook database, we obtain two datasets. One data set contains monthly Weibo penetration across prefectures. This consists of the aggregate number of blogposts per prefecture and month, based on all 13.2 billion posts in the Weibook data. The penetration data will be used to measure how the use of Sina Weibo was spreading across time and prefectures.

The second data set contains individual microblog posts extracted from the 13.2 billion posts in the Weibook data. These posts are of two types: original posts and forwarding posts. The original posts are the 202 million blog posts that mention any of approximately 5,000 key words related to various social and political topics. The forwarding posts are the 133 million retweets of the original posts. We only have the direct forwards of the original post, not the forwards of these forwards. For each original and forwarding post, we obtain the

textual content as well as the posting time, how many times it has been forwarded, and the location of its author.[3] This data will be used to measure how information about protests and strikes spread, and more generally to measure the expansion of communication through social media across Chinese cities.

This unique data base will inform us about how information on protests and strikes was produced and spread across China.

3.2 Data description

We now describe the temporal and spatial variations of events (strikes and protests) and social media penetration, followed by a description of the blog post data. We then explain how we use forwards (retweets) of posts to measure information flows over social media, based on which we depict the pattern of social media communication across cities.

We first describe protests during our sample period. There are in total 1153 protest events in our dataset between July 2006 and December 2013 when our main analysis is conducted. While many of these instances are small-scale protests confined within certain localities, some of them are large scale, spreading across regions, and disruptive. For instance, an event that was widely reported by Western media is the Wukan Event in 2011 when thousands of farmers in a city in Guangdong province protested against corruption of local officials. The event led to direct confrontation between farmers and officials, violent conflicts between protesters and police, and demonstrations in multiple cities in support to the farmers.

As shown in Table 1, many protests concern governments (policy, police and court, 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 protest events span 224 prefectures, slightly over 66% of all prefectures in China. Many locations experience over ten events during our sample period, even though the distribution is right-skewed (see Figure A1in the Appendix). Figure 1 shows the geographical distribution of events and social media penetration. The upper right panel shows the total number of protests by prefecture. Beijing is an outlier with 95 protests, and coastal areas in Shanghai and Guangdong and inland areas in Hebei, Shaanxi, Chongqing and Sichuan witnessed frequent occurrences of protests.

Next, we describe the strikes. In our dataset, there are 1074 strikes between January 2007 and December 2013. As shown in Table 2, strikes occur in a wide range of economic sectors, with concentration in manufacturing and transportation (including taxi) industries. More than 60% of the strikes involve more than 100 people. The most common cause of strikes is to demand payment of wage arrears. Geographically, these strikes cover 218 prefectures, approximately 64% of all prefectures in China. The distribution of strikes across prefectures is less skewed than the distribution of protests (see Figure A1). Geographically, the devel-

---

[3]Details about Weibook's data collection strategy and our selection of keywords and extraction of posts can be found in Qin et al. (2017b).

oped coastal areas are over-represented, notably in Guangdong, Shanghai and Jiangsu, but a significant number of strikes occurred in some inland areas such as Chengdu and Chongqing; see the upper left panel of Figure 1.

The geographical distribution of social media posts is shown in the bottom panel of Figure 1. Over time, there is a positive correlation between the incidence of protests/strikes and the Weibo penetration. Figure 2 shows the total number of protests and strikes per month along with the Weibo use per month. There were around five strikes per month in the period 2007-2010. The number of strikes rapidly rose to over 40 per month in 2013. The pattern for protests is similar, with around 3 protests per month until 2010, followed by a rapid increase to around 55 protests per month in 2013. The green dots show the number of Weibo posts per capita, which increased drastically after early 2010. The figure shows that the increase in protests coincides with the increase in Sina Weibo penetration while the increase in strikes takes place with a lag. 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.[4] Of course, this relationship between social media and events may not be causal. The increase in strikes could be driven by a slowdown in Chinese exports (Campante et al., 2019). In addition, social media may only have affected the observability of the events rather than the underlying event frequency.

In Section 5.1, we also study the period between 2013 and 2017. This adds 1576 protests and 1460 strikes. This extension is of interest since there are clear signs that the regime in China has taken a tougher stance on protest, strikes and media freedom after 2013. During our main sample period, 2006 to 2013, media freedom was at the average level during the period 1990-2016 for which Freedom House provides data However, post 2013, media freedom in China has dropped to an historical low, comparable to the levels just after the Tianmen Square protests. We also see a clear shift in the response to strikes in our data. In the strike data, the CLB have recorded information about the response to collective action for around one third of the observations. Using this data, we code the government response as being repression if the description of the response contains the words *police*, *arrest*, *beaten*, *threatened* or *wounded*. We code the government response as concession if the description contains the words *mediation* or *negotiation*. Before 2013, over three times more strikes where met by concessions than by repression. After 2013, the ratio was inverted.

3.2.1 Tweet content and mechanisms

A first empirical question is whether social media posts about protests and strikes are allowed to exist and spread at all, given the pervasive censorship. As described in Qin et al. (2017), there is extensive coverage of protests and strikes on Sina Weibo. From the original post

---

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

dataset, we find 2.5 million posts mentioning keywords related to protest and 1.3 million posts mentioning keywords related strikes. To characterize these posts, we drew a random sample of 1000 posts in each category and read them manually. Around 30 percent of the posts are indeed about protests and strikes. As shown in Qin et al. (2017), these posts predict and identify the real-world protests and strikes.

We now discuss how the content in these posts are related to potential mechanisms by which social media may affect these events. Social media may be used for *explicit coordination and organization* of protests involving a call for action or at least a mention of where and when a protest or strike is to take place. For example, Enikolopov et al. (2018) report that the vast majority of protest participants in Russia learned about upcoming protest events from the social media and that social media was also widely used for organizing protest activities. The situation in China is very different. In our sample, only 15 in 1000 protest posts involve a call for action and only two explicitly state a time and location of action. These 15 posts are rarely forwarded, with only one forward per post on average. A likely reason is that explicit calls for protest actions are censored. In our random sample of 1000 posts about strikes, none call 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 make protests spread because it carries *information about effective protest tactics* (e.g. Little 2016; Chen and Suen 2016). However, relatively few posts mention tactics (5 protest posts and 13 strike posts in 1000). For example, some posts discuss whether violence is necessary or helpful for a solution while others describe particular tactics. Hence, tactics is reported, but in very few cases. One could imagine that the few posts that introduce new tactics are still very influential. This is not what we see in this data. The posts discussing tactics only receive two forwards per post, on average. It seems unlikely that this is a main channel for the spreading or protests and strikes across events. Similarly, only a few posts discuss the *outcome of the events*, for example, whether the protest was successful or not.

Most of the posts instead report other facts about the protest, coupled with emotional reactions such as anger and sympathy for the protesters. Many posts mention the cause of the event, such as mis-behaviors of officials or persistent wage arrears or (164 in the protest category and 223 in the strike category). These posts are among the most forwarded. Information on social media about the problems causing the protests and strikes is important since it is very difficult for outsiders to get this information in any other way.

It is also worth noting that many posts also question social institutions (137 of the protest posts and 42 of the strike posts). This includes complaints about the legal system and absence of freedom of speech. Because of their general nature, these types of posts may induce protests to spread across causes and industries.

Perhaps a few posts are very influential. We looked specifically at the 100 most forwarded posts. After removing the repeated posts and the posts that are not relevant to protests, we are left with 91 posts. Of these, 55 talk about ongoing events, while the others comment

on past events, government policy, and social problems. Of the posts about ongoing event, almost all discuss the cause of the protest and a majority discuss persecution of protesters posts and government repression. Many of these posts express anger. Another common type of emotional content are reactions by outsiders to the events, with posts stating that the protesters were unfairly treated and expressing sympathy and moral support.

A couple of mechanisms are consistent with the content we find. One is that *information about problems* on social media, such as corruption, bad policies and human rights issues may spark protests, because learning about the significance of a problem may either induce people to protest against it for instrumental reasons or produce an emotional reaction that motivates people to protest (as in Pasarelli and Tabellini, 2017).

Second, the content we find may also facilitate *implicit coordination.* Coordination may be beneficial as the risk of punishment is lower and the pressure to deal with the problem is higher if there are more simultaneous events. Third, this content increases the *visibility* of protests to outsiders and top leaders. A goal of the protests is frequently to attract the attention of higher-level leaders, who may press local administrators to solve problems. Whereas previously outsiders would have no idea about the reason for a strike or protest, now they can read it on Sina Weibo and the affected government or company cannot cover up this information. This increases the pressure on of higher-level leaders to deal with the problem.

Finally, social media may by used by the government for *surveillance* to identify protests before they erupt. We know that social media contains the relevant content to support this mechanism as Qin et al, (2017) show that upcoming protests can be identified one day in advance, based on social media content.

To sum up, we don't find the content necessary to drive substantial social media effects on protests and strikes through explicit coordination and organization of protests, through information about effective protest tactics or outcomes of the events. Possible remaining mechanisms include effects through negative information, implicit coordination, visibility and surveillance.

These different mechanisms have slightly different implications. Effects through information and instrumental motives are likely to be long lasting, since the information remains available after its posted since, for example, what is a good protest tactic is unlikely to change rapidly. Effects through emotional responses are likely to be shorter lived. The surveillance mechanism is likely to particularly depress the incidence of large events. We will investigate these implications below.

3.2.2 Retweet amount and information spread

The post data contain a variable measuring the total number of times each post was forwarded. Based on this variable, we observe that the 3.8 million protest and strike posts are forwarded in 37 million retweets, which implies that on average, there are around ten

forwards per original post. This average masks a huge dispersion: some posts are forwarded millions of times and others none. Conditional on being forwarded at least once, there are on average 50 forwards per strike post and 100 forwards per protest post.

We use the data on individual retweets of protests and strike posts to examine how quickly and widely news about these events spread, exploiting the time and user location of the original post and the retweet. Since we only have the direct forwards of the original post, our dataset contains around 3 million of the 37 million forwards. For protests and strikes, around 30% of these forwards occur within 1 hour of the original post and 80% of the forwards within one day of the original post. Forwards 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 forwarded it is the same as the unconditional distance between users. In general, the forwarding data shows that information about protests and strikes on social media disperses rapidly and widely across China.

We will use forwards across all topics to measure how information flows across prefectures change as the use of Sina Weibo expands. A post being forwarded implies that someone has read the post and decided to forward it. Of course, many others may read the post without forwarding it, so our forwarding measure is a conservative measure of information spread.[5] We will proxy the information flow from city i to city j by the number of posts by users in city i that are forwarded by users in city j. We use the 133 million forwards across all topics except the 3 million forwards of posts discussing protests or strikes.

Figure 3 plots social media information flows from four of the prefectures with most strikes. The area in black is the source of the original posts. For instance, in the upper left panel of Figure 3, the black area is Shenzhen with 144 strikes. The map shows how intensively tweets from Shenzhen are retweeted by users in other prefectures, controlling for the average number of retweets from that prefecture.[6] Geographical proximity clearly matters for forwarding. For example, areas close to Shenzhen are forwarding more than average. However, there is also many areas far from Shenzhen that are closely connected to it through the social media network.

One reason that geographically distant cities are closely connected is that they are close in other dimensions. Cities that are more strongly connected through the social media network are similar in several characteristics measured in 2008 (geographically, population size, population density, tertiary share, FDI, landline and internet penetration; see Figure A2). People tend to form links with people who are like them, a network property called homophily. For this reason, information from other sources is likely correlated with information through the

---

[5]A much less conservative measure is to use followers, since many followers will not read each blog post. One of the most cited Twitter studies (Kwak et al., 2010) asks whether the number of followers or the number of retweets is a better measure of influence and settles on retweets.

[6]We regress the log number of forwards of post from i by users in j on fixed effects for locations i and j and plot the residual.

social media network. In addition, errors are likely to be correlated across cities that are more connected through the social media network. For example, because they share similar industrial characteristics, strikes are likely to erupt at similar points in time in areas which are tightly connected through the social media network, even if there was no such network. Therefore, one challenge in our empirical analysis is to overcome this homophily effect, which is one focus of our empirical design as will be discussed in detail.

3.2.3 Two examples

We end this section with two examples: one where protests seem to spread across causes and regions, another where strikes seem to spread within one industry.

In Guangzhou, Guangdong, a male worker wearing a bomb west staged a protest against wage arrears in a company in the afternoon of January 18, 2013. Later, he detonated the bomb, causing the death of one person and injuring seven people seriously. Immediately after the event, many Weibo posts described and commented the event, including many posts from direct witnesses. We identify in total 374 Weibo posts mentioning our protest-related key words on that day from Guangzhou, Guangdong, and 261 of them are talking about this event. These posts were first forwarded by Weibo users in nine prefectures tightly connected Guangzhou by our forwarding measure.[7] Most retweets express sympathy for the worker, condemn employers who default on wages, and decry the government who disregards citizens' rights and allow wage arrears.

Among the nine prefectures who first forwarded the event posts, three incurred protests in the subsequent two days. In Shanghai, thousands of workers from a Sino-Japan joint venture protested for the unfair new labor rules, detaining 18 senior managerial persons, and the process lasted until over 400 policemen broke into the factory two days later. In Shenzhen, hundreds of citizens went to the street, protesting against a polluting LCD factory being constructed in their neighborhood. In Shanwei, thousands of villagers protested in front of the city government asking for land back, which was taken by the government without appropriate compensation. These protests spread across provinces, and although targeting different causes, they all involve people who have had enough and protest against perceived injustices. Perhaps the later protests were inspired by the earlier ones, but there is no direct evidence of explicit organization or coordination across events.

The second example is a wave of school teacher strikes and demonstrations spanning from the fall of 2014 to the spring of 2015. The strikes stretched across 19 provinces and included 89 separate incidents. Chang and Hess (2018) describe how participants heavily used social media— primarily Sina Weibo— to post images of protests discuss their plights. In this case, they do not find evidence of explicit organization or organization but argue that later

---

[7]Controlling for the location and time fixed effects, the ranking percentile of the forwarding posts between them and Guangzhou Guangdong are 1.19% (Shanwei Guangdong), 2.3% (Shenzhen Guangdong), 4.36% (Wuhan Hubei), 5.03% (Shanghai), 5.3% (Hangzhou Zhejiang), 5.77% (Chengdu Sichuan), 6.25% (Xi'an Shaanxi), 6.98% (Zhengzhou, Henan), 8.62% (Qingdao Shandong).

protesters were inspired by earlier protests.

We now investigate whether these episodes were isolated random incidents or part of a large-scale systematic pattern. Do strikes and protests start to spread more after social media use expands, and exactly across those cities who are closely connected through social media?

4   Empirical analysis

We describe our baseline specification and results in Section 4.1 and then discuss potential identification problems and how to modify the basic specifications to address them in Section 4.2. Section 4.3 contains several extensions, studying effects after 2013, spread effects within and across strike industries and protest causes, effect duration and direct effects of social media penetration on incidence.

4.1   Baseline

4.1.1   Specification

We analyze whether information on Sina Weibo affects the spread of protests and strikes across Chinese prefectures (cities) using a panel of N cities at a daily frequency, t. Since the specification is the same for both types of events, we will in this section use protests as example. Let y_{it} be a dummy variable for protest occurrence. Suppose that the probability of a protest in city i at day t, $\Pr(y_{it})$ depends on the number of people there 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 informed from other sources, depending on time-invariant factors such as geographical distance. The model of protest spread from location j to location i is then

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

As information flows, f_{ijt}, over the social media network between locations i and j increase, protests will spread more from location j to i. In addition, there is the time-invariant propensity for events to spread between location captured by c_{ij}. Protests can spread simultaneously from multiple previous protest locations. For now, we assume that the marginal effect of each such protest is additive (we later relax this assumption). Hence the model becomes

$$\Pr(y_{it}) = \sum_{i \neq j} (\gamma c_{ij} + \beta f_{ijt}) \, y_{jt-1}$$

To capture that social media use may affect the incidence of protests, we add Sina Weibo penetration, w_{it} to the above model. Furthermore, we add x_{it}, a set of controls. We also include an autoregressive term and add time and prefecture fixed effects. We then have the

following baseline specification:

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

Here, y_{it} is a binary event dummy, y_{jt-1} is the number of events in location j within two days prior to t, and, δ_t and δ_i are time and prefecture fixed effects.

The measures of information flows between cities, f_{ijt-1} and c_{ij} are key to our analysis of protest spread. The variable measuring information flows over social media, f_{ijt-1}, is the log of one plus the number of forwards by users in i by posts from users in j on all subjects other than protests and strikes for the past 6 months up until one week before day t. We exclude forwards of posts about protests and strikes and forwards less than one week before day t to avoid reverse causality that may arise from the possibility that people who are planning protests and strikes are more likely to forward posts about protests and strikes. As a first proxy for c_{ij}, we use d_{ij}, the inverse geographic distance between cities i and j. Below, we will allow for arbitrary time-constant c_{ij}. The matrices capturing information flows, F with the element f_{ijt-1} and D with the distance element d_{ij} are normalized to increase interpretability. F is arranged as T stacked $N \times N$ matrices F_t, and hence has $N \times T$ rows and N columns. We normalize these matrices so that the average sum of all elements in a row of a weighting matrix equals one.[8]

Weibo penetration, w_{it}, is measured by the log of one plus the number of Sina Weibo posts per capita. As mentioned, this includes the total number of blog posts per prefecture and month, based on all 13.2 billion blog posts in the Weibook dataset. The set of controls, x_{it}, includes the log of one plus the sum of all forwards by users in location i by posts from all users j in all other locations, population, GDP, tertiary share, industrial share, and the number of cell phone users and landline users.

The specification in Equation (2) assumes that the marginal effect of lagged protest in other location is additive. We also estimate a model with decreasing marginal effects by replacing x with $\ln(5x + 1)$ for each variable x in Equation (2) that contains lagged events, that is, $x = y_{it-1}$, $\sum_{i \neq j} f_{ijt-1} y_{jt-1}$ and $\sum_{i \neq j} c_{ij} y_{jt-1}$. At the small values of x that we see in the data, the two functions are very similar, because the function $\ln(5x + 1)$ is almost linear; see Figure A3. Hence they yield very similar estimated marginal effects. However, the log transformation implies smaller marginal effects at high levels of x which ensures stationarity; see below. We use the log transformation in the Monte Carlo simulations, to avoid exploding time-series paths. We present the results using the linear model here and the results from the $\ln(5x + 1)$ transformation in the appendix.

---

[8]A common form of normalization is row-normalization by which all coefficients in a row sum to one. However, in our case this normalization would imply that all locations would be equally affected (on average) by protests in other locations, which is clearly incorrect (e.g. Elhorst, 2001, argues against 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 and in a$(N \times T) \times N$ matrix, such as the forwarding matrix, this sum is $N \times T$.

Several econometric issues requires consideration. First, logistic and probit models are biased in 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 that is immune from these problems. Second, our model includes location-fixed effects and lagged dependent variables. In general, the estimates in this type of model are inconsistent with T fixed (e.g. Hsiao 2014, Arrelano 2003, Baltagi 2005). In the current model, T is large and the bias is likely to be small. We will show that the bias is indeed small using Mont Carlo simulations.[9] Third, we need to check whether the estimated process is stationary.[10] Fourth, consistency requires no serial autocorrelation in the errors. Serial correlated in the error term implies that ε_{it} will be correlated with y_{it-1}. We will test for serial autocorrelation in the first-differenced residuals. Finally, errors may be correlated across both time and spatial units. We use two-way clustering of errors in the time and spatial dimensions.

4.1.2   Results

We estimate how protests and strikes spread across Chinese cities, using the specification (2) on a panel of prefectures with at least one protest or strike, for these respective outcomes. The first two columns of Table 3 show the results for protests and the last two for strikes.

Our main interest is in the coefficient on the variable with lagged events, weighted by the amount of forwarding of posts from city j by users from city i. This is significant and positive across specifications. The coefficient estimate is not much affected by the inclusion of controls; see columns II and IV.

To interpret the coefficient magnitudes, the estimates in column 2 implies that a protest in a given location the previous two days increased the expected number of locations with protests this day by 0.16.[11] The equivalent number of strikes is 0.9. Relative to the mean event probability, this is an increase of 48 percent for protests and 28 percent for strikes.[12]

In addition, we find that both protests and strikes are autoregressive processes. Having an event during the last two days significantly increases the probability of an event occurring in the same location. Strikes spread to nearby locations as the distance-weighted effect of

---

[9]If the bias was larger, 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 our protests and strikes, with lagged differences and levels is not likely to perform well.

[10]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 will evaluate this criterion and check whether the process is stationary in Monte Carlo simulations.

[11]To assess the magnitude of the estimated effects, consider first 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 forwards-weighted term, equals βf_{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 the previous day.

[12]The mean of the protest incidence is .0015 and there are 224 other prefectures than the one where the first strike took place, .16/(224*.0015)=0.48. For strikes, the corresponding number is .09/(218*.0015)=.28.

19

lagged strikes is positive. However, this is not true for protests. The incidence of observed protests and strikes are both increasing in Weibo penetration.

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 the mean estimated β is in the third value digit for both protests and strikes, see Appendix A4.

Observability Social media was used as a source to identify protests and strikes. This is obviously likely to bias our estimate of , β_0, the effect of social media penetration on events. However, the estimate of β, the effect of social media on the spread of events across prefectures is less likely to be biased by observability. The reason is that while observability increases the number of observed events across cities, it is unlikely to increase the probability that one observed event occurs just the day after an another, relative to other days.

We use Monte Carlo simulations to verify that we can consistently test our null hypothesis of no effects of social media on event spread, even if event observability is increasing in Weibo penetration. In particular, we simulated 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 in Weibo penetration, w_{it}, adjusting the size of the effect so that Weibo penetration increases observability by 30%. This is higher than what would be implied if the estimated coefficient on Weibo penetration, $\widehat{\beta}_0$, only reflected increased observability. We then draw a set of observed events $\widetilde{y}_{it}$ with probability p_{it} from the simulated events y_{it}. We finally estimate the model in Equation 2 on the observed simulated events, $\widetilde{y}_{it}$. Figure 4 shows that no spurious correlation in spread is generated because social media increase observability. Hence, we conclude that we can consistently test the hypothesis $\beta = 0$, even if social media affects observability.

4.2 Identification through time-series variation in the network

4.2.1 Specification

Another identification concern is that it is difficult to identify the role of social media in spreading events if information from other sources or shocks are correlated with the social media network. We showed in the previous section that the social media network exhibits homophily: cities that are similar in different ways are more connected through the social media network. Consequently, strikes and protests may spread more between cities that are closely connected in the social media network even though social media play no role, as these cities may have strong information flows through other channels as well as correlated shocks. Such shocks would make ε_{it} correlated with the lagged outcome in other cities y_{jt-1}. A large literature, pioneered by Manski (1993), has discussed this issue and related identification problems. Solutions include exploiting the network structure to identify instruments (e.g. Bramoulle et al. 2009), of identifying instruments that are uncorrelated with the error terms

and the network (e.g. Acemoglu et al., 2015, and König et al., 2017), matching (Aral et al., 2009) and explicit randomization (e.g. Bakshy et al., 2012).

We instead use time-series variation in the rapidly expanding network to identify effects, by adapting standard difference-in-difference and event study methods to estimate network interactions. We will allow strikes and protest to spread through correlated channels and that errors are correlated with the network but assume that these effects and correlations are constant during our short sample period. Although the average spread of strikes between cities may be related to the average network strength, the exact timing of the network expansion is unrelated to other changes in the spread of strikes between cities.

We explore three different ways of using time-series variation in the network. Our first solution is to use a model with interaction-fixed effects, c_{ij}, that capture any time-invariant heterogeneity in event spread from location j to i. In this specification, β can be consistently estimated even though the errors are correlated with the network, as long as this correlation is time constant. For example, for the error process $\varepsilon_{it} = \kappa w y_{t-1} + v_{it}$, the model with interaction-fixed effects is correctly specified and the coefficient $c_{ij} = \kappa w_{ij}$ is estimated consistently as T goes to infinity.

Our second solution was to estimate a model of dyads of locations with time-constant dyad-fixed coefficients $\beta_{i,j}$. However, Monte Carlo simulations showed that this method produced strongly (Nickell) biased estimates and hence this approach was abandoned.

Our third solution aims to produce an event-study type graph of the spread across cities connected through social media. Consider the spread of protests between location i and j described in equation (1). If we regress this on the average social media information flows between location i and j, $\overline{f}_{ij}$, we get

$$\Pr\left(y_{it}\right) = \gamma c_{ij} y_{jt-1} + \beta_{ijt} \overline{f}_{ij} y_{jt-1},$$

where

$$\beta_{ijt} = \beta f_{ijt} / \overline{f}_{ij}.$$

This β_{ijt} equals zero for the pre-period when f_{ijt} equals zero and then rises as f_{ijt} does. A regression of this model would identify

$$\mathrm{p}\lim \widehat{\beta}_{ijt} = \beta_{ijt} + bias.$$

where the bias arises if the network $\overline{f}_{ij}$ is correlated with c_{ij}, for the reasons discussed above. However, before August 2009, Sina Weibo did not exist and hence $\beta_{ijt} = 0$ and for this period $\widehat{\beta}_{ijt}$ consistently estimates the *bias*. This means that we can estimate β_{ijt} as the increase in $\widehat{\beta}_{ijt}$ after the pre-period.

The key identifying assumption is that the correlation between the average network $\overline{f}_{ij}$ and c_{ij} is constant over time. This may be violated, for example, if people switch from communicating over the phone to communicating over social media. In this case, we would

underestimate the effect through social media because increased spread via social media would be correlated with decreased spread via other communication channels. However, we would still estimate the total effect of the network on event spread between cities, including equilibrium effects through changed spread via other channels. If one is interested in the effect of social media on, for example, protest dynamics or regime stability, this total effect is the relevant parameter. At longer time scale, increased social media interactions may also affect structural features, such as mobility, trade and even industry structure, that in turn affect c_{ij}. However, it seems unlikely that these structural changes would be important over the short duration of our study.

To allow multiple protests in the same day, to spread to location i, we sum the effects across locations j and let coefficient β_b capture the average effect across locations in each half-year period, b, so that

$$\Pr(y_{it}) = \alpha y_{it-1} + \beta_b \sum_{i \neq j} \overline{f}_{ij} y_{jt-1} + \gamma \sum_{i \neq j} c_{ij} y_{jt-1} + \beta_0 w_{it} + \theta' x_{it} + \delta_t + \delta_i. \tag{3}$$

Here, the element $\overline{f}_{ij}$ is proportional to the average of forwards f_{ijt-1} over time and normalized so that the sum of all elements $\overline{f}_{ij}$ equals N.

4.2.2   Results

We first estimate the model according to our first solution, including interaction-fixed effects. The results are shown in Table 4. The pure auto-regressive and the distance-weighted effects are not identified in this specification. The estimates of β remain highly significant and are, in the case of protests, slightly larger than those in columns I and II of Table 3. This slight increase in magnitude is perhaps caused by the so-called Nickell bias as our Monte Carlo simulations show a clear, albeit small, bias in the estimates including interaction-fixed effects.[13]

We next estimate the model according to our third solution, as described in Equation (3). Figure 5 shows the estimated coefficients, relative to the average coefficient size in the pre-period. Starting from the second half of 2010, protest events began to spread significantly more across cities connected through social media, while the timing for strikes is one year later. This provides additional evidence that the Sina Weibo expansion is driving the effects.

---

[13]Figure A5 shows the distribution of β-estimates from Monte Carlo simulations. The graphs to the left show results from estimations without interaction-fixed effects (Equation 2), the graphs to the right include interaction-fixed effects. As previously mentioned, the bias is very small in the graphs to the left. The graphs to the right, show a clear, albeit small, bias. The blue line shows the coefficients from the estimations without interaction-fixed effects (Equation 2) while the red line shows the mean coefficients from the simulated data. The green line shows the interactions-fixed effects estimate in the original (not simulated) data. These estimates are slightly different from those in Table 4 because the simulation does not include the control for Weibo posts. The slightly larger estimates from the interaction-fixed effects model are explained by the Nickell bias. The bias-corrected estimates from the interaction-fixed effects model would be almost identical to those from the model without these effects.

5 Extensions

We now extend our analysis in four ways. First, we study whether strikes and protests continue to spread over social media after 2013, when content control became even stricter. Second, we investigate whether protests and strikes spread across industries and causes and if this type of spread is affected by social media. Third, we study how long-lived the effects are to draw conclusions on the likelihood of different mechanisms. Fourth, we collapse the data at the level that we have penetration data, prefecture-by-month, and describe how the incidence of protests and strikes is associated with micro-blog penetration.

5.1 Spread of protests and strikes after 2013

A key question is whether protests and strikes are still spreading through the social media network in the new environment after 2013, when media freedom was reduced even further. We can answer this question using the model with time-varying coefficients on the average ties from the social media network under the assumption that the average ties in 2013 is a good proxy for average ties post 2013. This is likely the case. By 2013, the social media network was already extensive. In addition, the closeness in terms of geography, size and development that explain social media ties between cities do not change much over a few years. We can evaluate this assumption by investigating how much our measure of social media connectedness change in the last two years of our sample. A regression of the number of forwards between locations for the past six months before the last day of 2013 and the same variable the last day of 2011 has an R2 of 0.96 (and a coefficient greater than one reflecting the increase in forwards).

Figure 6 shows that the spreading of strikes through social media peaked in 2013-2014. After that, this spreading was reduced and by 2016 it had completely stopped. A similar pattern is found for protests, although the peak was earlier and the spread was positive even in 2016. The reason for the reduced spread of protests could, for example, be increased censoring or self-censoring on social media, the lower share of successful examples met with concessions rather than repression. We do not have the data coverage of Weibo posts after 2013, so we cannot evaluate these channels. Perhaps as a consequence of the reduced spread and tougher stance of the regime, the number of protests and strikes have fallen, from a peak of around 40 strikes and 46 protests per month in 2014 and 2015 to around 11 strikes and 18 protests per month in 2017.

5.2 Do social media break industry and cause bounds?

It seems a priori likely that strikes spread within industries and protests within causes. For example, the wave of school teacher strikes in 2014 and 2015 documented by Chang and Hess spread within the same industry, and the spreading protests among farmers against corruption in Wukan were for the same cause. On the other hand, strikes and protests may

also spread across industries and causes, for example, if people learn about general effective protest and strike tactics. We investigate this by splitting the protests into the categories listed in Table 1 (government policy and corruption, housing and land, etc.) and strikes by the industries listed in Table 2.

Table 5 shows the results from regressions where we have investigated separately the spread of events through social media within and across categories. For strikes we split events by industry and for protests by cause. The table shows that the spread through social media is 6-7 times higher within strike industry and protest cause than across, although the spread both within and across categories is statistically significant.

While the effect of an individual event is smaller across than within categories, the total effect of spread across categories is larger. The reason is that there are many more events across all categories. The mean of the social media weighted events within and across categories is reported in the two last rows of the table. The total average effect of protests spreading across protests causes is 60% higher than within. Similarly, for the total effect across strike industries is 20% larger than the total effect within industries. We conclude that social media are breaking the bounds of strikes industries and protest causes and that the aggregate impact is larger than effects through within-category spread.

5.3   Effect duration

Alternative mechanisms differ in their implications for effect duration. Learning about protest tactics or government responses from protests posts is likely to have long-lasting effects, since what is the optimal tactic is not likely to change abruptly, since people who read the posts are not likely to forget the information quickly and since the posts anyway remain available online. A short-term response is more likely if the posts arouse emotional responses or if people see the opportunity to combine efforts across location and protest for a common cause.

So far, we have analyzed short-run responses, within two days, since we saw that posts about protests and strikes spread over long distances within a short time span. However, we now expand the effects window and study how effects persist over time. The upper left panel of Figure 7 shows the estimated spread of events occurring within 1-2 days, 3-7 days, 8-30 days 31-90 days and 91-180 days. Although the effect is largest just after the strike, strikes continue spread through social media (beta) for around one month. The effects for both protests and strikes die out after a month. The effect seems larger in the short run (within a week). The lack of long-run effects is more supportive of mechanisms that go through emotional responses or coordination across cities than learning effects.

5.4   Social media penetration and event incidence

Finally, we show how protest and strike incidence are associated with social media penetration at the prefecture-month level. We already know from the results presented in Tables 3 and

4 that social media penetration is associated with event incidence. However, this alternative specification at the level at which we observe penetration is simpler and we can more easily discuss magnitudes and heterogenous effects on incidence. Specifically, we will study effect by event size, by whether the prefecture is inland or coastal, and by whether a strike results in concessions or repression.

To this end, we now estimate a linear probability model of the form

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

where i indicates prefecture, m indicates month, y_{im} is an indicator variable for an event taking place in prefecture i at month m, w_{im} is Sina Weibo penetration. We use the same set of controls x_{im} as above, with one exception. We drop the total number of forwards by users in location i because it is strongly correlated with Weibo penetration and we wish to have only one measure of the intensity of social media use.

Table 6 shows the results. The first three columns use an indicator variable for a strike as the dependent variable, whereas the last three columns use an indicator variable for protest as dependent variable. The protest and strike incidences are both positively associated with the number of Weibo posts. Adding controls does not much affect the estimates. The magnitudes of the estimated coefficients are large. In 2012, the variable Weibo posts has an average of 0.3. The estimated coefficients imply that an increase in Weibo posts by 0.3 is associated with an increase in the total number of protests by 11 per month (.3*.157*224) and in the total number of strikes by 9 per month (.3*.131*218).

The upper panel of Figure 8 shows the dynamic response. There is no pre-trend and the effects seem to appear with a lag of 6 months for strikes and immediately for protests. This pattern is also visible in the aggregate trends shown in Figure 2.

If observability drives these correlations, then we would expect to see larger correlations where observability prior to social media was lowest. The RFA has better information in certain coastal provinces (Guangdong, Fujian, Zhejiang, Jiangsu) than inland. Table 6, columns III and VI, include an interaction term between our Weibo penetration variable and a dummy variable for areas other than these four coastal provinces. The effects on strike incidence is not significantly different in coastal and inland areas. However, the effect on protests is significantly larger in inland areas, consistent with the RFA having fewer sources there.

It is also likely that prior observability was worse for small events. The lower left panel of Figure 7 shows the estimated coefficients by the number of event participants (less than 100, in the hundreds, thousands or tens of thousands). We have too few observations for protests and strikes with more than ten thousand participants and for protests with less than a hundred participants. For the other event magnitudes, the estimated effects are significant and largest for medium-sized protests and strikes involving hundreds of participants. To conclude, we find some evidence that observability drive effects in inland areas, but little

evidence that small events were affected more.

Finally, we investigate whether concession or repression are more strongly associated with Sina Weibo penetration. The effect of social media on the mode of response is theoretically unclear. The visibility of a strike on social may attract the attention of higher-up leaders who press local leaders to make concessions. If this is true, then increased social media use would be associated with more concessions. Social media also increase the visibility of repression, for example, as people post images of police brutality. This may be costly for the regime, if it reduces its legitimacy or even arouses anger that sparks new protests (see e.g. Goldstone and Tilly 2009, p.181). On the other hand, repression may signal government strength while concessions signal weakness. In this case, the cost of concessions will increase with social media use as the signal of government weakness is spread across the country (and vice versa for repression).

The lower right panel of Figure 8 shows the results. Higher Weibo penetration is more strongly associated with an increase in strikes met with concessions than with repression. This is consistent with Sina Weibo shifting the government response towards concessions. However, some caveats are in place. For example, it could be that the strikes met with repression are more observable absent social media.

6   Conclusion and discussion

Exploiting a large Chinese social media dataset, this paper addresses the heated debate regarding whether social media facilitate citizens' political action in autocracies. In particular, we examine how information diffusion over Sina Weibo—the leading Chinese microblogging platform—affected the incidence and spread of protests and strikes in China from 2006 to 2017. By incidence, we mean the average probability that an event takes place in a certain location over a certain time period (e.g., a month). By spread, we mean the increase in the probability of an event caused by another event taking place just one or a few days before.

Our research places more emphasis on spread than on incidence for two reasons. First, spread influences the dynamics of protests which is more consequential than isolated incident. If protests are local and evenly distributed across time, they are less likely to pose a serious threat to the regime. By contrast, increased spreading of events through social media may cause small local events to snowball into regime-threatening movements. Second, methodologically, we show that we can consistently test the hypothesis that social media does not affect the spread of events even if social media increase the observability of the events. Intuitively, while observability increases the probability that an event is included in our data, it does not increase the temporal clustering of events in which one event is relatively more likely to take place just after another.

The main findings are as follows. Despite the strict control by the Chinese government, we find that social media massively diffused information about protests and strikes. We doc-

ument millions of microblogs discussing these events that were posted and rapidly forwarded across China. We further find that information diffusion through posting and forwarding had a sizeable effect on the spread of both protests and strikes across Chinese cities during the 2009-2013 period. The spread of events induced by social media was fast and predominantly local— between events within the same province and the same social and economic category (e.g., cause and industry). Nevertheless, spread across these categories was still significant, albeit weaker. Over time, the estimated effect of event spread through social media increased gradually after 2009, reached a peak in 2013 and 2014, then declined, and completely stopped in 2016. Finally, we find that the explosive increase in the use of Sina Weibo had a large and significant effect on the incidence of protests and strikes, although part of the effect was likely to be driven by the increased observability of events due to social media.

Our findings cast new light on the mechanisms that drive the political effect of social media in autocracies. In existing studies, the mechanism that has been stressed most is that social media are used for explicit coordination of collective action, through which protesters call for joint action, plan events, and implement certain strategies. However, by exploring Sina Weibo posts discussing protests and strikes, we find little evidence that citizens used social media for explicit coordination of protests and strikes. This is likely due to the Chinese government's strategic censorship of social media.

Another possible mechanism that has drawn scholarly attention is that people learn through social media about protest tactics and government responses, which enables them to better organize similar protests. However, social media posts rarely discuss tactics or outcomes, so the content that could drive these effects is largely absent. In addition, such a learning mechanism is likely to produce a persistent effect, because the posts containing relevant information remain online for a long period and knowledge about protest tactics and government responses is unlikely to be forgotten in days. In contrast, we find that the spread effect through social media depreciates rapidly, with the largest effects being within just two days.

A couple of mechanisms are consistent with our results. It could be that negative information on social media about the causes of protests and strikes spark new events, either for instrumental or emotional motives. It could also be that social media content help people implicitly coordinate and lower the risk of punishment or increase the chance of concessions. These mechanisms would create predominantly short-lived effects and effects that are stronger within cause and industry. In addition, there is abundant content that could drive these mechanisms, as a large number of social media posts report about the events and their causes. This type of content is less likely to be censored not only because it often occurs spontaneously without a clear common goal and is thus unlikely to be viewed as a threat, but also because it does not rely on explicit content such as calls for participation. These mechanisms may be more resilient to government intervention than the explicit coordination mechanism.

Our findings reflect a central trade-off in the media control strategy and the limitation of this strategy in autocracies. On one hand, social media generate a huge amount of information that is useful for surveillance and monitoring. On the other hand, this information may inspire anti-regime collective action. This trade-off motivates a media control strategy in which the government allows for relatively free discussion about local events and politicians but extensively censors explicit coordination of collective action. Our empirical findings demonstrate the limits of such a strategy. To the extent that a regime cannot prevent information flows across regions and groups, the spread of information about local conditions can effectively generate wide-spread protests and strikes, which may diminish people's trust of the government and thus undermine regime stability. Consistent with our finding and as noted by insiders, Sina Weibo's role in spreading these events decreased after 2016, probably as a result of more intense censorship. Instead, the Chinese government allowed freer political discussion in WeChat, a within-group messaging service in which information diffusion by construction is more localized.

Our study also has important implications for the effect of information technology on political accountability. Information is indispensable for holding political leaders 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 governments, local officials have an informational advantage over citizens and the central government. This creates severe agency problems within the Chinese political system. However, social media substantially reduces 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 central government's goal is aligned with the citizens' interest, social media may help solve agency problems and hold local officials accountable. One caveat is that the low cost of posting complaints and allegation online may reduce the informativeness of social media. Consequently, the informational value of social media relies critically on real events that are costlier (e.g., protests and strikes) and thus more informative. Investigating how this nuanced complementarity between online allegation and offline protests affects local accountability will be an interesting extension of our current 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, https://doi.org/10.1093/qje/qjv030

[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, et al. "The role of social networks in information diffusion." Proceedings of the 21st international conference on World Wide Web. ACM, 2012.

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

[9] Bamman, David, Brendan O'Connor, and Noah A. Smith. 2012. "Censorship and Deletion Practices in Chinese Social Media." First Monday 17(3).

[10] Barberà, Salvador, and Matthew O. Jackson. "A model of protests, revolution, and information." Matthew O., A Model of Protests, Revolution, and Information (March 1, 2018) (2018).

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

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

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

[16] 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.

[17] Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2017). Are protests games of strategic complements or substitutes? Experimental evidence from Hong Kong's democracy movement (No. w23110). National Bureau of Economic Research.

[18] 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.

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

[20] 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.

[21] China Internet Network Information Center. 2011. "The 28nd Statistical Report on Internet Development in China." July 2011, Beijing.

[22] China Internet Network Information Center. 2013. "The 32nd Statistical Report on Internet Development in China." January 2013, Beijing.

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

[24] De Mesquita, Ethan Bueno. "Regime change and revolutionary entrepreneurs." American Political Science Review 104.3 (2010): 446-466.

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

[26] 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.

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

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

[29] Enikolopov, Ruben, Alexey Makarin, and Maria Petrova. 2018. "Social Media and Protest Participation: Evidence from Russia." Available at SSRN: https://ssrn.com/abstract=2696236.

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

[31] Epstein, Gady. 2013. "China's Internet: A Giant Cage." The Economist, April 6. http://www.economist.com/news/special-report/21574628-internet-was-expected-help-democratise-china-instead-it-has-enabled.

[32] 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.

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

[34] International Telecommunication Union. 2013. "The World in 2013: ICT Facts and Figures." Geneva. http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf.

[35] 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

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

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

[38] Kleinberg, Jon. 2006. "Complex Networks and Decentralized Search Algorithms." Proceedings of the International Congress of Mathematicians (ICM).

[39] 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.

[40] Kwak, Haewoon, et al. "What is Twitter, a social network or a news media?." Proceedings of the 19th international conference on World wide web. AcM, 2010.

[41] 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.

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

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

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

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

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

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

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

[49] Reporters Without Borders. 2013. "2013 World Press Freedom Index: Dashed Hopes after Spring." https://rsf.org/en/news/2013-world-press-freedom-index-dashed-hopes-after-spring.

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

[51] Steinhardt, H. Christoph. 2017. "Discursive Accommodation: Popular Protest and Strategic Elite Communication in China." European Political Science Review, 9(4):539-560 https://doi.org/10.1017/S1755773916000102

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

[53] Yanagizawa-Drott, David. 2014. "Propaganda and Conflict: Evidence from the Rwandan Genocide." The Quarterly Journal of Economics, 129(4): 1947-1994 https://doi.org/10.1093/qje/qju020.

[54] Zhu, Tao, David Phipps, Adam Pridgen, Jedidiah R. Crandall, and Dan S. Wallach. 2013. "The Velocity of Censorship: High-Fidelity Detection of Microblog Post Deletions." arXiv preprint, arXiv:1303.0597.

A   Appendix

A.1   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 the

top political leadership. Stationarity in a dynamic spatial panel data model depends on the parameters of the model as well as on the spatial weights matrix that determine 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, increased use of social media increases the feedback because each individual row in the forwarding matrix, F, does not sum up to one. This implies that 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 at a particular date t equals $\beta \overline{f}_{t-1}$, where $\overline{f}_{t-1}$ is the average row sum across locations that day. The maximum such row sum equals 10.5, an order of magnitude larger than the average row sum. This implies that sufficient conditions for stationarity are not fulfilled for the linear model.[14]

However, the actual data do not exhibit the explosive path suggested by the estimated linear model. This could be because the government observed that the process was exploding and stepped in and struck down on protests and strikes, but we do not have direct evidence of this. 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, in other words, in how wide spread protests are. The linearity assumption may be incorrect for several reasons. If the mechanism is through information about protest methods, then it is likely that the marginal value of information from an additional event is falling. If the mechanism is through information about the number of protests about an issue, then the incentive to organize an additional protest so as to increase government awareness is likely to be decreasing. It may also be the case that there is a limited number of areas where people are upset enough about a particular issue to potentially protest or strike if they see that other people protesting for this issue. Hence there is a cap to the total number of protests and the process will eventually be concave.

We model this by using the concave function $h(x) = \ln(5x + 1)$, for terms x involving lagged events in equation 2, for example, we insert $\ln(5 \sum_{i \neq j} f_{ijt-1} y_{jt-1} + 1)$ instead of $\sum_{i \neq j} f_{ijt-1} y_{jt-1}$ (in this case $x = \sum_{i \neq j} f_{ijt-1} y_{jt-1}$). This function was chosen to be sufficiently concave to make the process stable and is used in the simulations. Figure A3 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 reach high levels of $\sum_{i \neq j} f_{ijt-1} y_{jt-1}$. In these cases, the concave model is stable while the linear model explodes. For this reason, we use the concave model in the simulations. We prefer to report the linear model results in the main tables since the coefficients are simpler to interpret and report the results from the concave model here.

---

[14]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 (from 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.

Table A3and A4 correspond to Tables 3 and 4, but uses the concave transformation $h(x) = \ln(5x + 1)$ for all terms involving lagged events. Significance levels and magnitudes are very similar. For the case $h(x) = \ln(5x + 1)$, the marginal effect equals $\beta s_t f_{ijt-1}$, where the scaling factor s_t equals 4.67 at the sample mean. The scaling factor equals

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

For strikes, at the sample mean value of

$$\sum_{i \neq j} f_{ijt-1} y_{jt-1} = 0.0138,$$

$s_t = 4.67$ So the estimate $\widehat{\beta} = 0.033$ in column IV should be multiplied by 4.67 to be comparable to the estimate in column II of Table 3: $4.67 * 0.033 = 0.15$. For protests, the scaling factor is $s_t = 4.61$ and the marginal effect is $4.61 * .051 = .23$.

A.2   Probit regression

Table A1 shows the results from a probit regression of Equation 2, where the time fixed effects have been replaced by a quadratic time trend to avoid the incidental parameters problem. The estimates of β are significant across all specifications. However, the implied marginal effects on event probabilities are smaller.

A.3   Monte Carlo simulations

We run a set of Monte Carlo simulations to assess the Nickel bias in the coefficient estimates 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 but for simplicity without Weibo penetration and controls. We then generate data using the estimated parameters, adjusted so that $\delta_t + \delta_i \geq 0$ and estimate the model on this data. We repeat this procedure 100 times.

Figure A4 in the Appendix plots the distribution of t-statistics of coefficients α, β and γ in 2 against the standard normal density.

Figure A4 shows the distribution of β-estimates from Monte Carlo simulations. The graphs to the left show results from estimations without interaction-fixed effects (Equation 2). The bias is very small, the difference between the true and the mean estimated β is in the third value digit for both protests and strikes. The graphs to the right are based on regressions that include interaction-fixed effects, corresponding to the results in Table 4. These graphs show a clear, albeit small, bias. The blue line shows the coefficients from the estimations without interaction-fixed effects (Equation 1) while the red line shows the mean coefficients from the simulated data. The green line shows the interactions-fixed effects estimate in the original (not simulated) data.17 The slightly larger estimates from the interaction-fixed effects

model are explained by the Nickell bias. The bias-corrected estimates from the interaction-fixed effects model would be almost identical to those from the model without these effects.

Table 1: Protests by cause 2006-2013

```
                    |                         size
            cause   |    <100      100s     1,000s    10,000s    unknown  |     Total
--------------------+-------------------------------------------------------+---------
government policy or |     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 healt |     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 relatio |      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

industry	<100	100s	size 1,000s	>=10,000s	unknown	Total
manufacturing	49	255	158	2	42	506
taxi	24	54	13	2	17	110
transportation	131	77	13	0	22	243
services	48	38	2	0	8	96
education	29	23	5	0	6	63
construction	6	12	4	0	1	23
retail	8	5	2	0	3	18
mining	0	6	2	0	1	9
other	1	4	0	0	0	5
unknown	0	0	1	0	0	1
Total	296	474	200	4	100	1,074

Table 3. Event spread across locations

	I	II	III	IV
	Protest	Protest	Strike	Strike
y_{it-1}	0.015**	0.015**	0.028***	0.027***
	(0.007)	(0.007)	(0.007)	(0.007)
$\sum d_{ij} y_{jt-1}$	-0.030	-0.030	0.168***	0.160***
	(0.037)	(0.036)	(0.059)	(0.053)
$\sum f_{ijt} y_{jt-1}$	0.167***	0.163***	0.103***	0.093***
	(0.043)	(0.042)	(0.029)	(0.027)
Weibo posts	0.005***	0.005***	0.007**	0.005**
	(0.002)	(0.002)	(0.003)	(0.002)
Controls	No	Yes	No	Yes
Observations	668,626	668,626	665,630	665,630
R-squared	0.017	0.017	0.023	0.023
QPtest	0.02	0.23	0.13	0.18

Results from a linear regression on an event dummy variable. The unit of observation is prefecture by day. The regression includes prefecture and day fixed effects. Controls are ln($\sum f_{ijt}$+1), population, GDP, tertiary share, industrial share, and the number of cell phone users and landline users. Standard errors are two-way clustered by date and location. The QP statistic reports the p-value of the test for serial correlation in fixed-effects model of Born and Breitung (2016).

Table 4. Event spread across locations, allowing time-constant heterogenous spread

	I	II	III	IV
	Protest	Protest	Strike	Strike
$\sum f_{ijt} y_{jt-1}$	0.214***	0.210***	0.089**	0.081**
	(0.060)	(0.059)	(0.041)	(0.039)
Weibo posts	0.004**	0.003*	0.005	0.005
	(0.002)	(0.002)	(0.004)	(0.003)
Observations	668,626	668,626	665,630	665,630
R-squared	0.223	0.223	0.229	0.229
QPtest	0.00	0.00	0.00	0.00
Controls	No	Yes	No	Yes

Results from a linear regression on an event dummy variable. The unit of observation is prefecture by day. The regression includes prefecture and day fixed effects. Controls are ln($\sum f_{ijt}$+1), population, GDP, tertiary share, industrial share, and the number of cell phone users and landline users. Standard errors are two-way clustered by date and location. The QP statistic reports the p-value of the test for serial correlation in fixed-effects model of Born and Breitung (2016).

Table 5. Event spread, within and across categories

	I	II
	Protest	*Strike*
Within		
Number events 1-2 days prior, cumulative forwards weighted	0.0617***	0.0413***
	(0.0128)	(0.0121)
Number events 1-2 days prior, distance weighted	-0.0085	0.1144***
	(0.0106)	(0.0303)
Across		
Number events 1-2 days prior, cumulative forwards weighted	0.0083***	0.0058**
	(0.0028)	(0.0023)
Number events 1-2 days prior, distance weighted categories	-0.0016	0.0054
	(0.0026)	(0.0033)
Observations	8,692,138	6,656,300
R-squared	0.0078	0.0144
Controls	Yes	Yes
Category	Cause	Industry
Mean within	0.0013	0.0014
Mean across	0.0152	0.0128

Results from a linear regression on an event dummy variable. The unit of observation is prefecture by day. The regression includes prefecture and day fixed effects. Controls are ln($\sum f_{ijt}$+1), population, GDP, tertiary share, industrial share, and the number of cell phone users and landline users. Standard errors are two-way clustered by date and location. The QP statistic reports the p-value of the test for serial correlation in fixed-effects model of Born and Breitung (2016).

Table 6: Effects on incidence: dependent variable event dummy

	I	II	III	IV	V	VI
		Strike			Protest	
Weibo posts	0.163***	0.131***	0.134***	0.199***	0.157***	0.110***
	(0.041)	(0.042)	(0.044)	(0.043)	(0.041)	(0.029)
Weibo posts, inland			-0.012			0.229***
			(0.050)			(0.045)
Observations	22,142	21,938	21,938	22,083	21,622	21,622
R-squared	0.169	0.173	0.173	0.144	0.152	0.158
Controls	No	Yes	Yes	No	Yes	Yes

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

Figure 1. Strike count by prefecture 2007-2013 and protest count by prefecture 2006-2013

Strike count (top coded at 50)

Protest count (top coded at 50)

Weibo posts (log10)
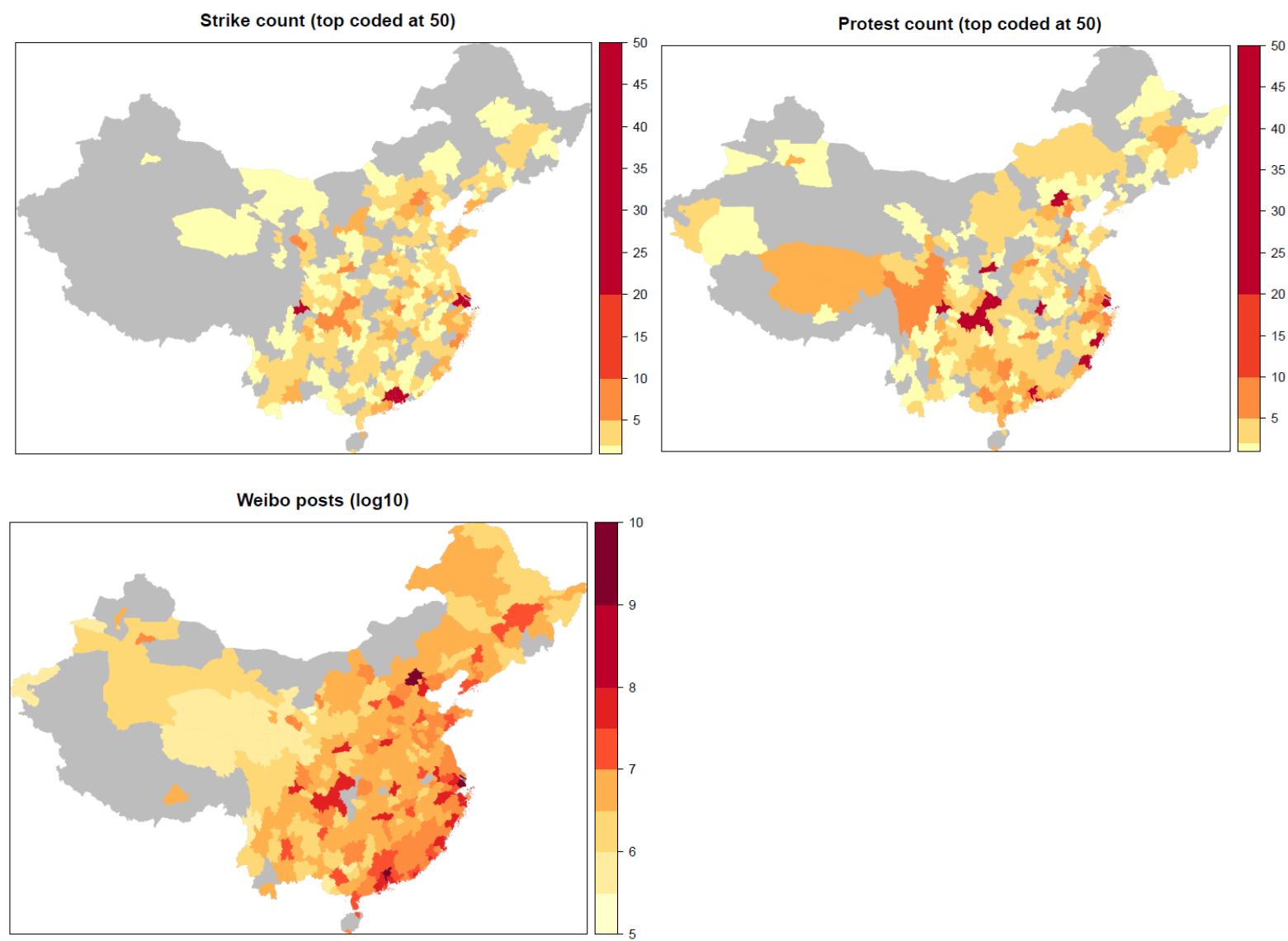
Figure 2. Number of events and Weibo posts per capita per month
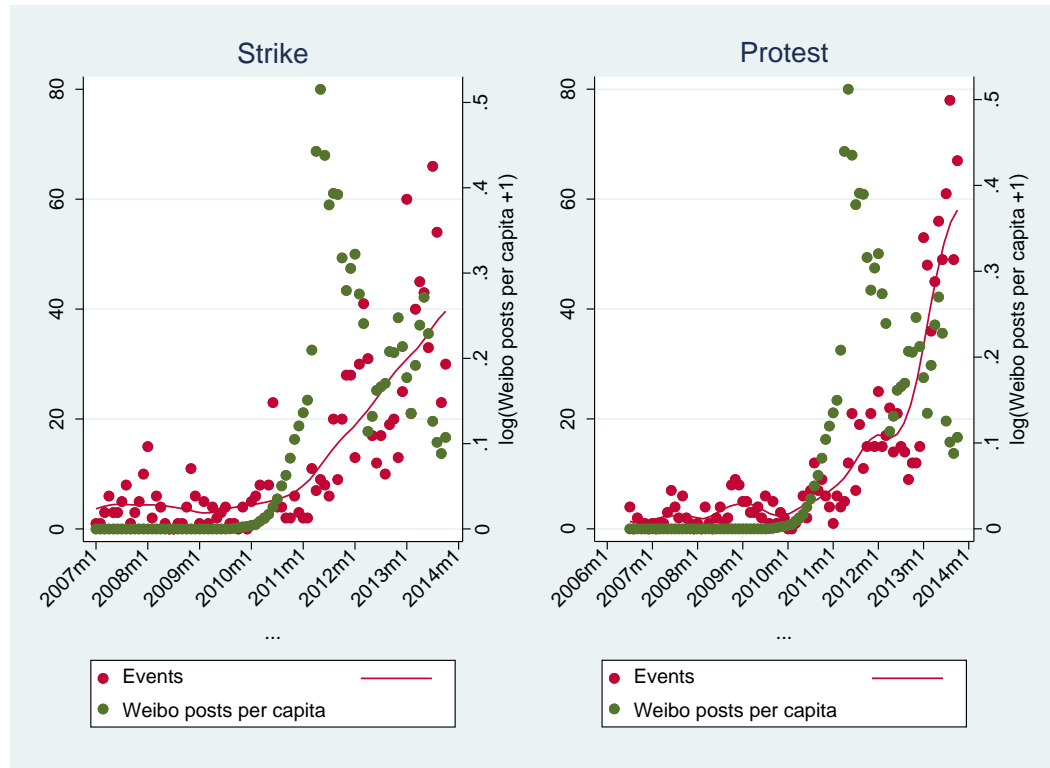
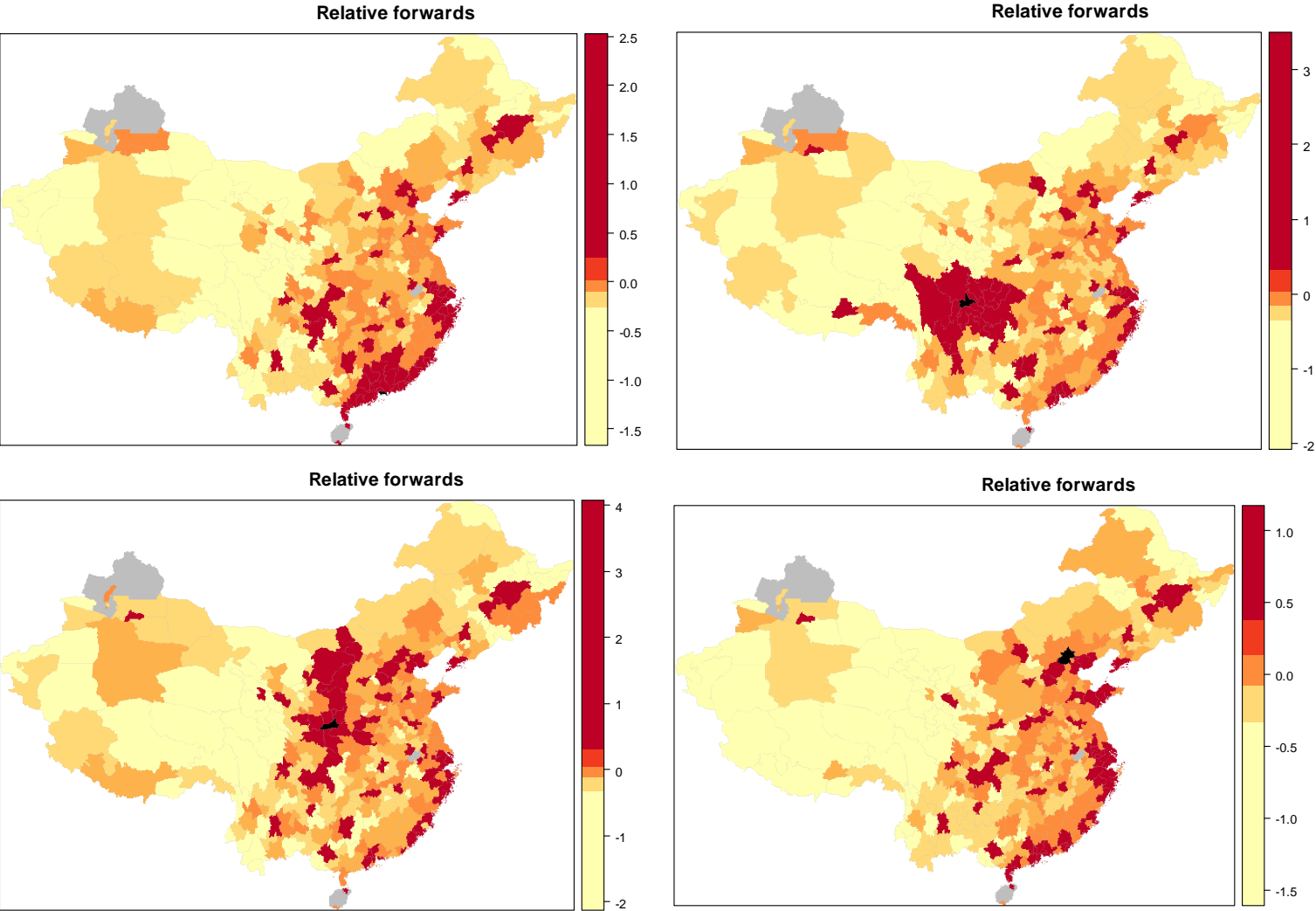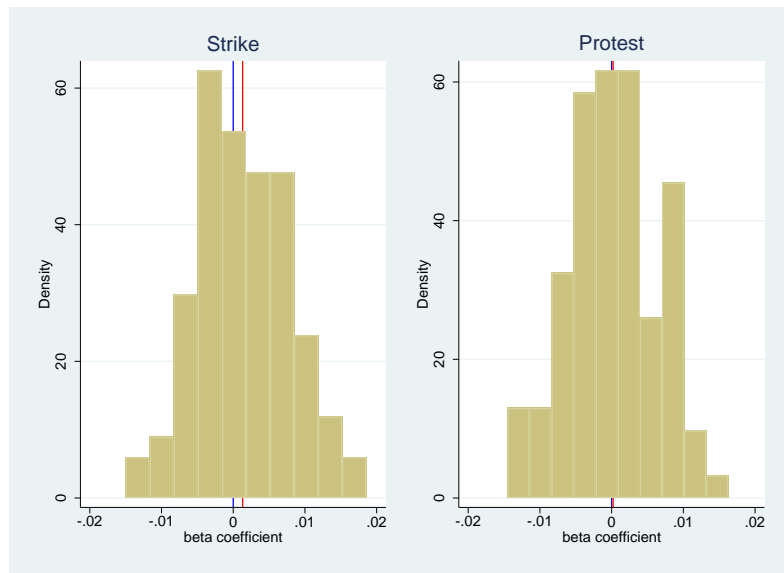Figure 3. Forward connections from prefectures with many strikes

Figure 4. Monte Carlo simulations with observability driven by Weibo and no network spread effect

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

Figure 5. Time-varying coefficients and constant forwarding matrix
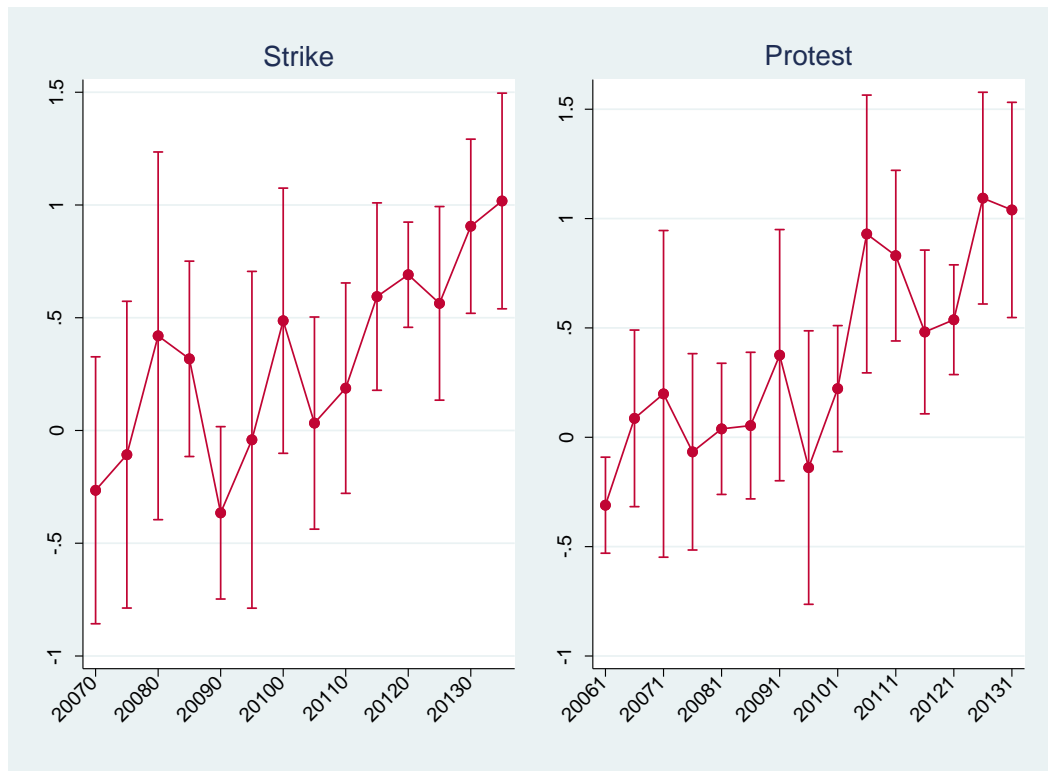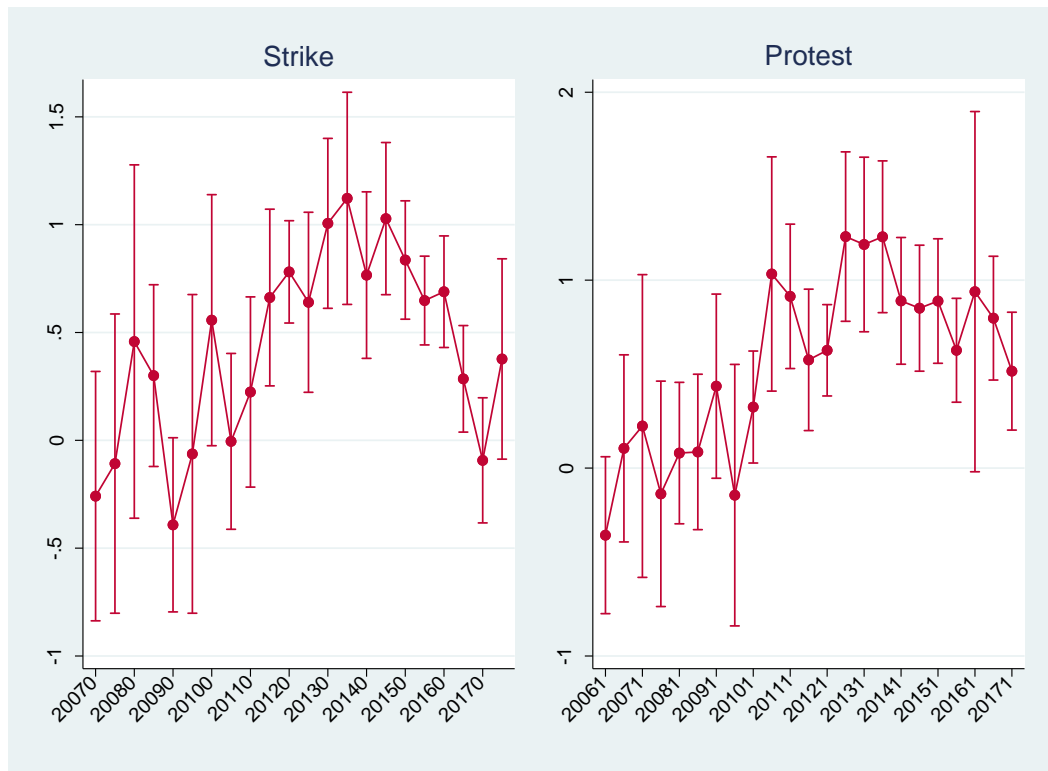
Figure 6. Time-varying coefficients and constant forwarding matrix post 2013
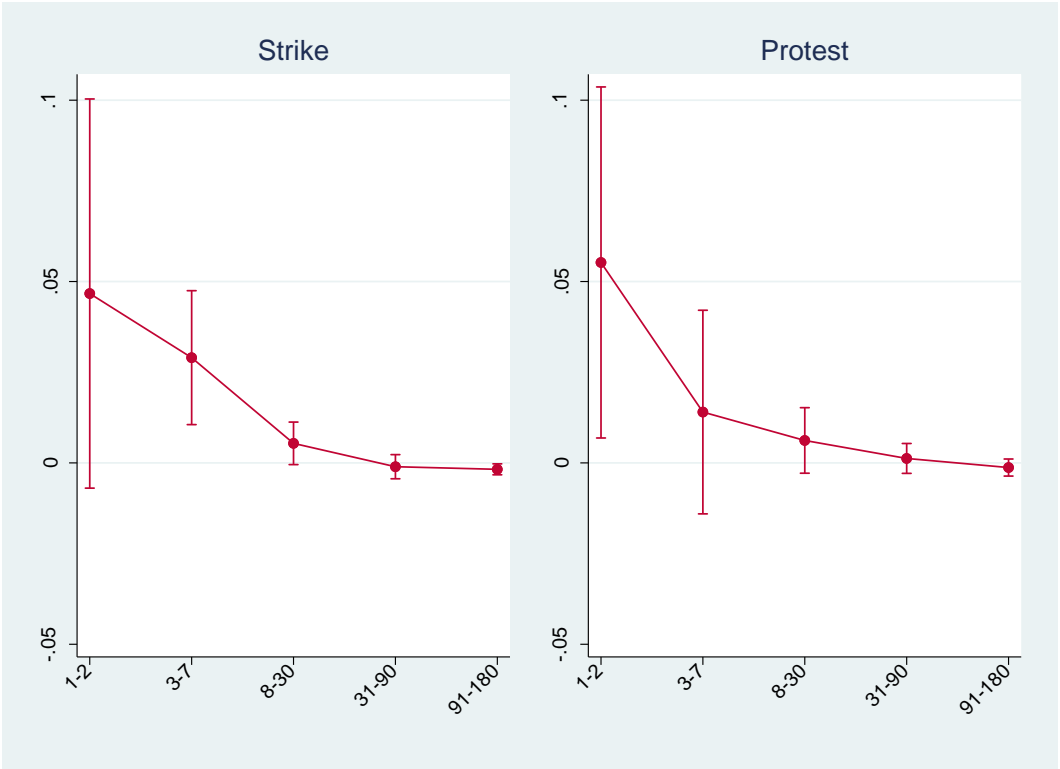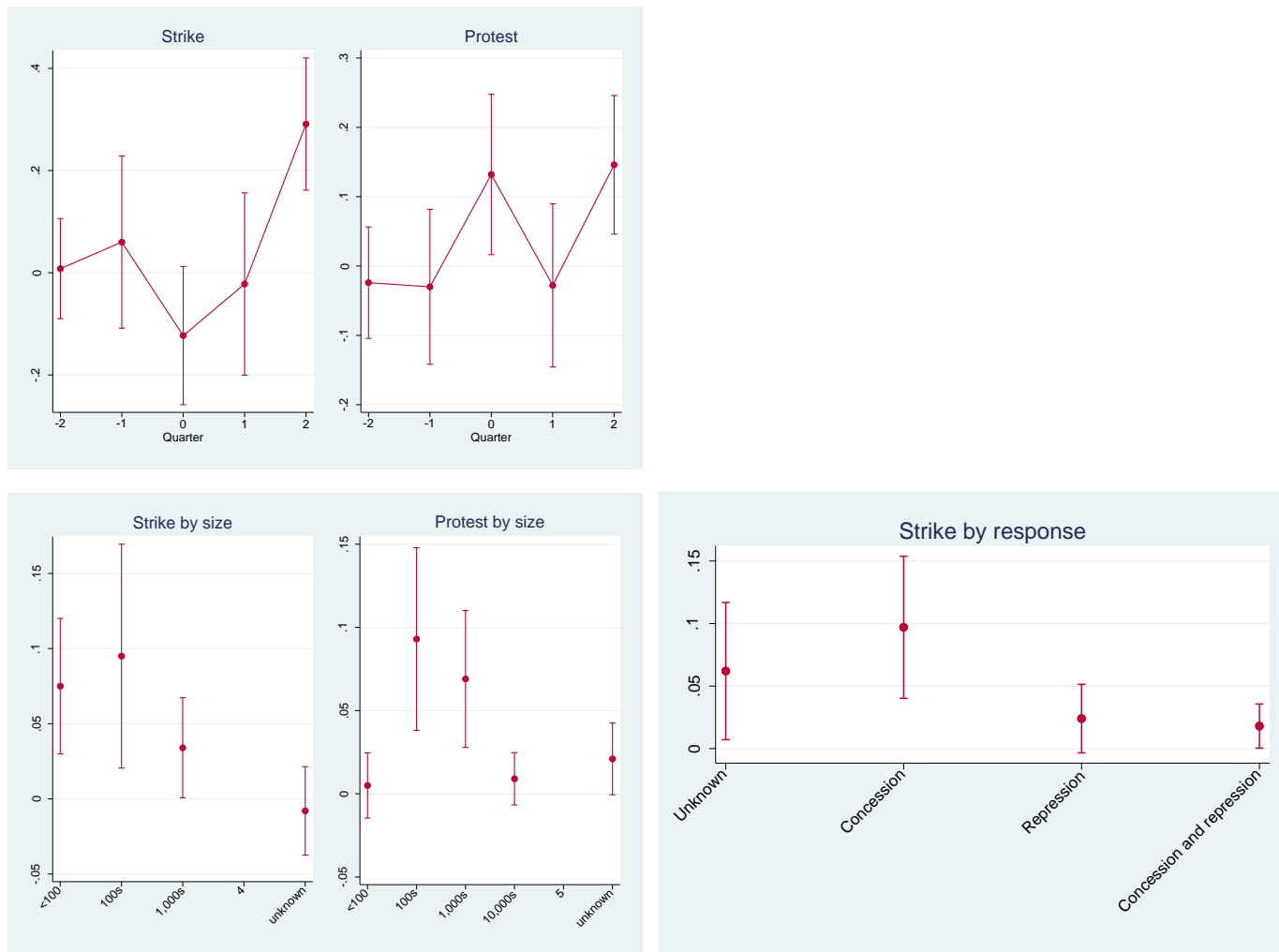
Figure 7. Effect duration

Figure 8. Dynamic and heterogenous effects on incidence (beta0)

The upper panel shows the dynamic effects. The lower panel shows heterogenous effects by size and response.

Table A1. Event spread across locations – probit model

VARIABLES	I	II	II	IV		V	VI	VII	VIII
			Strike					Protest	
y_{it-1}	0.327***	0.328***	0.192***	0.192***		0.176**	0.162**	0.104**	0.096**
	(0.066)	(0.066)	(0.035)	(0.035)		(0.079)	(0.077)	(0.048)	(0.047)
$h(\sum d_{ij} y_{jt-1})$	2.683***	2.693***	0.663***	0.665***		1.983	2.306*	0.420	0.493
	(0.912)	(0.927)	(0.224)	(0.227)		(1.448)	(1.386)	(0.330)	(0.319)
$h(\sum f_{ijt} y_{jt-1})$	0.763***	0.769***	0.238***	0.239***		0.597**	0.471*	0.203**	0.164**
	(0.275)	(0.279)	(0.084)	(0.085)		(0.272)	(0.251)	(0.083)	(0.078)
Weibo posts	0.082***	0.076**	0.081**	0.074**		0.209***	0.214***	0.207***	0.211***
	(0.031)	(0.038)	(0.032)	(0.038)		(0.047)	(0.060)	(0.047)	(0.060)
Observations	519,149	519,149	519,149	519,149	561,129	561,129	561,129	561,129	519,149
Controls	No	Yes	No	Yes		No	Yes	No	Yes
h function	X	X	ln(5x+1)	ln(5x+1)		X	X	ln(5x+1)	ln(5x+1)

Results from probit regressions on event dummy variable. All regressions contain prefecture fixed effects and quadratic time trends. Unit of observation: prefecture by day. Standard errors clustered by prefecture in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table A3. Event spread across locations

|	I	II	III	IV
	Protest	*Protest*	*Strike*	*Strike*
y_{it-1}	0.009**	0.009**	0.016***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)
$h(\sum d_{ij}y_{jt-1})$	-0.008	-0.007	0.036***	0.034***
	(0.008)	(0.008)	(0.013)	(0.012)
$h(\sum f_{ijt}y_{jt-1})$	0.055***	0.053***	0.033***	0.028***
	(0.013)	(0.012)	(0.008)	(0.008)
Weibo posts	0.006***	0.005***	0.008***	0.006***
	(0.002)	(0.002)	(0.003)	(0.002)
Controls	No	Yes	No	Yes
Observations	668,626	668,626	665,630	665,630
R-squared	0.017	0.017	0.022	0.023
QPtest	0.06	0.26	0.16	0.20

Results from a linear regression on an event dummy variable. The unit of observation is prefecture by day. The function h(x)=ln(5x+1). The regression includes prefecture and day fixed effects. Controls are ln($\sum f_{ijt}$+1), population, GDP, tertiary share, industrial share, and the number of cell phone users and landline 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 location: *** p<0.01, ** p<0.05, * p<0.1.

TableA A4. Eventspread across locations

	I	II	III	IV
$h(\sum f_{ijt}y_{jt-1})$	0.049***	0.045***	0.028***	0.023***
	(0.013)	(0.013)	(0.009)	(0.008)
Weibo posts	0.006***	0.004**	0.007**	0.005**
	(0.002)	(0.002)	(0.003)	(0.002)
Observations	668,626	668,626	665,630	665,630
R-squared	0.222	0.223	0.229	0.229
QPtest	0.00	0.00	0.00	0.00
Controls	No	Yes	No	Yes

Results from a linear regression on an event dummy variable. The unit of observation is prefecture by day. The function h(x)=ln(5x+1). The regression includes prefecture and day fixed effects. Controls are ln($\sum f_{ijt}$+1), population, GDP, tertiary share, industrial share, and the number of cell phone users and landline 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 location: *** p<0.01, ** p<0.05, * p<0.1.

Figure A1. Distribution of collective action events across prefectuers
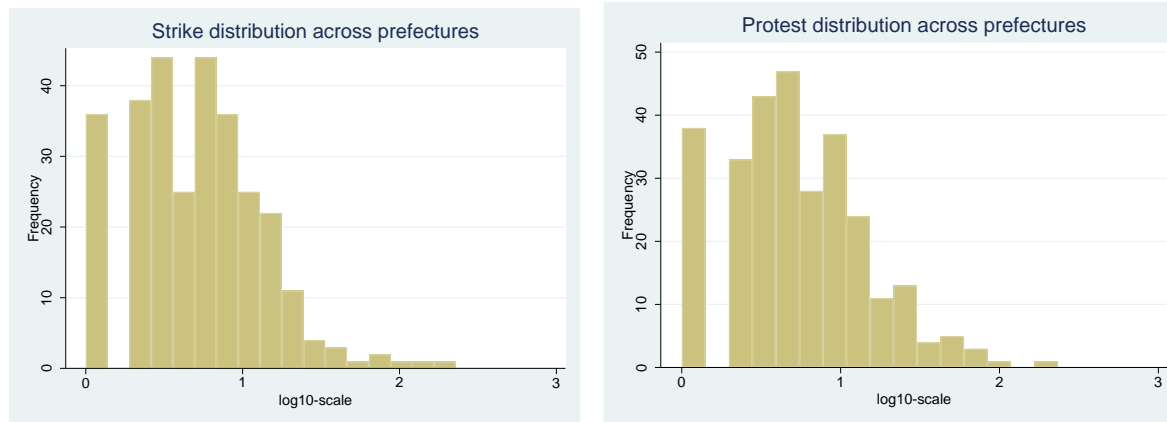
Dofile event_pm_graphs -> event_weibo.gph

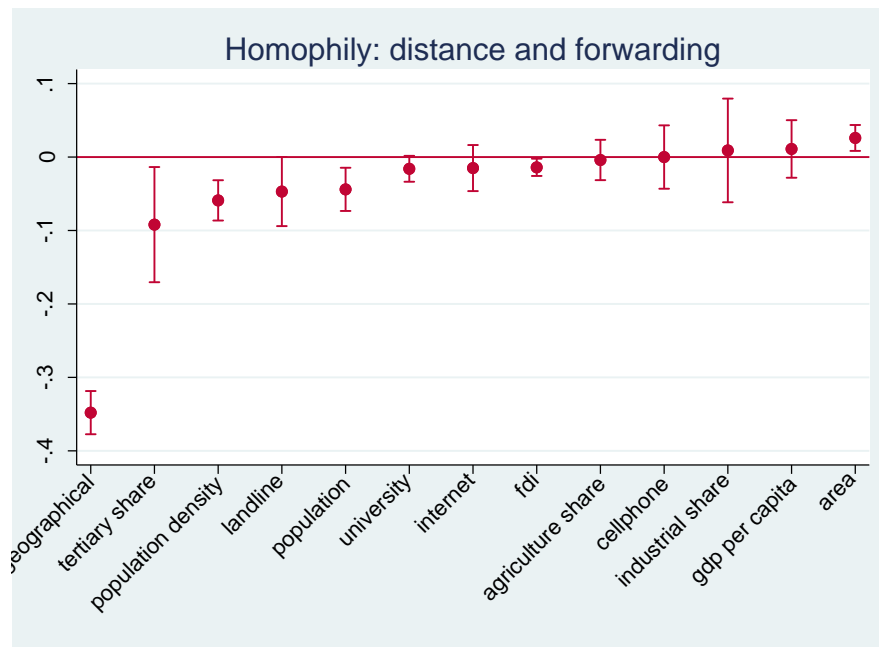Figure A2: Predictors of relative forwarding between cities
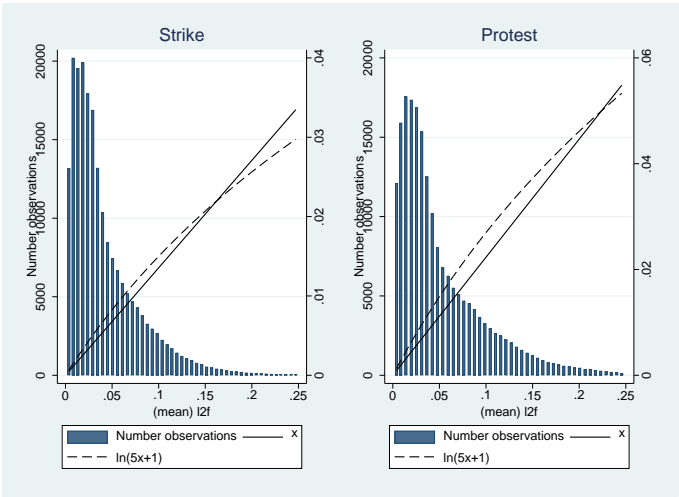
Figure A3. Lines + density.

Figure A4. Monte Carlo Simulations: distribution of t-statistic of estimated parameter = true parameter
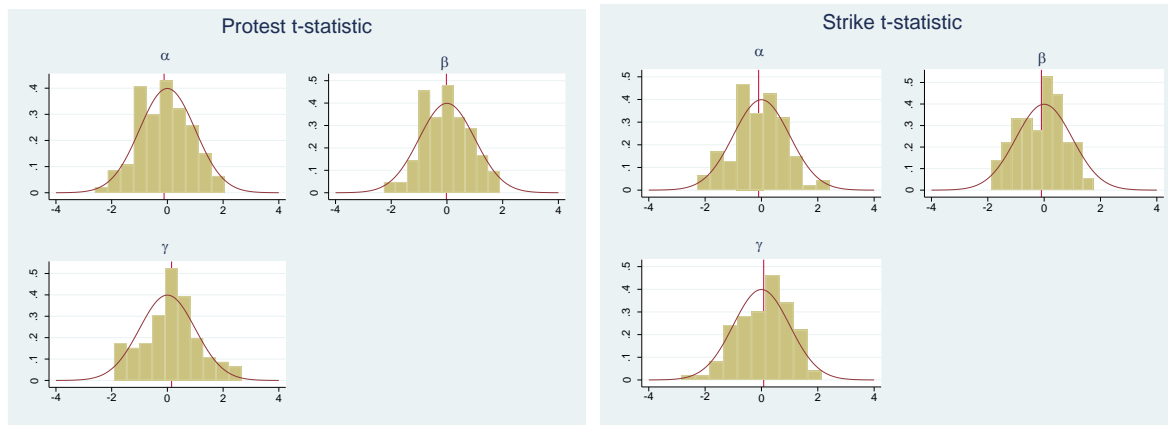

Figure A5. Monte Carlo Simulations: distribution of coefficients around DGP parameter values.