Who Not What:

The Logic of China's Information Control Strategy

Mary Gallagher* Blake Miller†

January 13, 2019

Abstract

In this paper, we examine how the Chinese state exerts control over social media. While social media companies are responsible for censoring their own platforms, they also selectively report certain users back to the government. Our focus in this article is to understand the logic of this information control strategy - what types of users or content are reported back to the government.

We find that in the state's information control efforts, content is less important than commonly thought. It is often a question of who is posting, not what they are posting. Instead of focusing exclusively on content category, the state permits open discussion and debate on social media while controlling and managing influential social forces that may challenge the Party-State's hegemonic position. To explain our findings we rely on "ideology and organization" by Schurmann (1966), emphasizing the Party's goals of embedding itself in all social structures and limiting the ability of non-Party individuals, networks or groups to carve out separate space for leadership and social status. In the virtual public sphere that is social media, the CCP continues to apply these principles in order to co-opt, repress, and limit the reach of influential non-Party "thought leaders."

*Professor, Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: metg@umich.edu).

†Postdoctoral Fellow, Program in Quantitative Social Science, Dartmouth College, Silsby Hall, Hanover, NH 03755 (E-mail: blake.a.p.miller@dartmouth.edu).

Introduction

Neihan Duanzi was a social media portal for videos, memes, and jokes in China that developed a strong user community, similar to Reddit or 4chan. As the site gained popularity, users began to meet up and identify each other in the real world. The space that *Neihan Duanzi* users had claimed for themselves began to pose a more obvious threat to the state when its users began to communicate in secret codes of car honks at intersections. In April of 2018, *Neihan Duanzi* was shut down. What is puzzling is that the site was quite apolitical, and—if we are to believe the literature about why authoritarian governments censor content—seemingly harmless to the ruling CCP. The portal rarely featured rebukes of corrupt officials, discussions of collective action, or criticisms of the government. Why would a joke portal like this need to be shut down? *Neihan Duanzi*'s large and growing community was not a clear threat to the CCP at the time of its ban, but it had the potential to transform into one, or at least compete with the state over the public space that the CCP wishes to control. In this article, we argue that the regime attempts to maximally shrink non-official public spaces because they could threaten the CCP's authority at some point in the future. Key to shrinking this space is the careful targeting of individuals who are influential within counter-hegemonic spaces.

While the state could reassert control over these spaces by shuttering all private social media platforms, it benefits from a well-maintained and constrained private media ecosystem. The Chinese state, once at odds with free market capitalism is now using it as a source of stability. Drawing on the expertise of tech firms who continue to innovate in social networks, digital payments, and targeted advertising analytics, the government can reserve real-world and violent repression for those who are flagged by algorithms that predict future acts of state subversion and propensity to participate in collective action against the regime. These systems are already in development (Meissner, 2017; Yang, Yang, & Fei, 2017). Information control, in contrast to the pre-internet age, is enforced with a scalpel rather than a hammer.

China's repression regime is carefully constructed to encourage most citizens to be tamed via impersonal censorship algorithms. Users learn, sometimes unconsciously, to "blog about this, but not about that," "forward posts about this, but not about that."

Censorship can be ratcheted up via different tactics, suited for different kinds of users. For many, the intrusion of state security into their lives is merely virtual and often covert and unobserved. For a select few who dare to push the boundaries, online censorship may be more explicit and threatening. It may also move from the virtual realm to real world interaction – a home visit, an invitation to "drink tea", an interrogation, a detention in a black jail, or worse.

This article analyzes when delegated, sometimes automated, censorship of online content shifts back to the state's repressive apparatus. The Chinese government delegates online censorship to private internet content producers (ICPs), large companies like Sina and Tencent, that are compelled to satisfy the government's demand for information control. While much of the work of delegated censorship is run-of-the-mill deletion or hiding of sensitive content, ICPs also selectively report certain content and users back to the government. Our question here is to understand and theorize about the logic of this reporting system: who and what are targeted for handling by the state itself?

We find that "reporting up" is targeted toward influential public opinion leaders whose standing and influence may threaten the Party's hegemonic presence in China's online public sphere. Substantive topics matter less than user influence and virality in state efforts to "guide opinion." Many topics are not censored and reported up unless they achieve a certain degree of virality or influence. This includes topics that have a clear collective action threat (a political dissident, a workers' strike) (King, Pan, & Roberts, 2014). However, it also includes viral discussions that challenge the legitimacy of the regime in more indirect and oblique ways (elite defection, ridicule of top leaders).

Early work by Schurmann (1966) is instructive in explaining the logic behind a focus on "who not what." Schurmann emphasizes the importance that the CCP placed on ideology and organization in its social revolution, which he conceptualized as a process of destruction and replacement of former elites. Communist control had both ideational and spatial dimensions. Ideology replaced other values and norms of China's old society. Communist organization embedded the party in all social, economic, and political structures. The content of the ideology and the structure of the organizations have changed, of course, since Schurmann's analysis of the Maoist state. Ideology has changed: for example, class

struggle is no longer used to mobilize society. Social organizations have changed as well, for example, a dynamic private sector has developed under reform. But even in these new spaces—the private sector, the non-profit sector, the Internet—the CCP has developed tactics to infiltrate and to dominate, to crowd out other voices, and to eliminate or co-opt alternate sources of leadership and influence. The arrival of social media, in particular, has presented the CCP with the challenge of instituting CCP's ideological domination and organizational presence in the virtual public sphere. The battle for online public opinion dictates an ambitious approach that seeks to allow for online expression while also suppressing the rise of influential voices and viral discussions. Although the government is certainly worried about the Internet as a device to facilitate social movements and protests, it also worries simply when a large number of people online are talking about the same thing, what it calls "a public opinion emergency" (舆论危机).

We relate the government's desire to allow for online expression while limiting extra-Party influence to the way in which the state's management of social grievances has evolved since the violent suppression of the student movement in 1989—what is often called "responsive authoritarianism." Since 1989, the Chinese state developed new tools of targeted repression and broad responsiveness, learning from the mistakes of 1989. It has become more tolerant of protests and grievances that are narrowly expressed, localized, and economistic. At the same time, its ability to target and isolate political dissidents and to use repression to silence the regime's most vociferous opponents have improved. With the explosion of social media since 2009, however, it has become more difficult to maintain the narrow, localized and non-systemic character of grievance articulation. The viral and virtual nature of discussion online can unite similarly-aggrieved individuals across time, space, and barriers of region, class, and even language. For example, local scandals in food safety, wage arrears, or corruption can be more easily linked to systemic problems in the authoritarian political system, such as the lack of a free press, independent trade unions, or close relations between political and business elites. Social media has the potential to transform a narrow, local problems to system-challenging threats.

Maintaining responsiveness while containing the threat of the online public sphere is a delicate balancing act. In order to maintain its responsiveness, which seems to have vastly improved the government's legitimacy (Nathan, 2003; Dickson, 2016; Tang, 2016;

Zhang, 2017), the central government uses social media discussion to monitor its local agents and mitigate the severe principal-agent problems inherent in China's top-down system. However, in order to hedge against the possibility of the web acting as the "single spark that ignites a prairie fire", the government must control the virality of any online discussion and the influence of public opinion leaders online. CCP manuals to train Internet censors specifically reference this problem as the "butterfly effect (蝴蝶效应)," the need to control information before it gets out of (their) control (Wei et al., 2015; Gao & Zhang, 2011; Zou & Su, 2015; PRC State Council General Office, 2016).

This article has four parts. In the first section, we examine control dynamics over time, arguing that the regime has moved from broad-based repression to targeted repression coupled with greater responsiveness to a broader swathe of the population. The second section uses government manuals and documents on opinion guidance and internal censorship logs from a social media company to analyze when commercial censors report users and information back up to the government. This allows us to both empirically test our model and supplement our results with additional transcript evidence from government sources (Lorentzen, Fravel, & Paine, 2017). The third and final section discusses the findings.

1 The Rise of Responsive Authoritarianism and the Challenge of Virality

The Chinese Communist Party's approach to socio-political control has shifted over time. As economic reforms sped up in the 1990s, the government became far more tolerant toward protests that it saw as rooted in socio-economic grievances brought on by market reform. Small and narrow protests multiplied during this time, with protestors limiting their claims and often hiding organizations and leaders. These protests were different not only in who participated, but also in their claims, their targets, and their tactics. On the other side, the government began to overhaul the repressive apparatus of the state to improve its ability to preempt major challenges while placating reasonable grievances (Tanner, 2004).

Protesters often framed their grievances as narrow and economistic, as mainly related to economic deprivation and loss precipitated by the onslaught of dislocation brought on by rapid economic restructuring or by the externalities of rapid, often unregulated growth. These protests overwhelmingly, and perhaps even intentionally, articulated their targets as bad local actors, such as corrupt local officials, greedy enterprise bosses, or compromised local bureaucrats who failed to enforce central laws strictly when such laws interfered with growth. These lower level targets were best attacked using tactics that were wholeheartedly moralistic and normative, but also, at times, legalistic. As O'Brien and Li have described, their invocation of resistance was rooted in support for the legitimate policies and laws of the central government and their anger that local governments did not enforce them (O'brien & Li, 2006). Demands were for better rule, not for a new set of rulers.

This turn toward more give and take between state and society has been labeled bargained authoritarianism (Lee & Zhang, 2013), responsive authoritarianism (Heurlin, 2017), and contentious authoritarianism (Chen, 2012). These terms, though emphasizing different aspects of the phenomenon, hone in on the notion of protest as an accepted mode of political participation in contemporary China (Tang, 2016). State and society have both learned new tactics to make high levels of protest in an authoritarian regime sustainable, yielding neither to mass repression or a tipping point toward revolution (Heurlin, 2017; Lee & Zhang, 2013; Su & He, 2010; Cai, 2008). In this game of cat and mouse between state and society, citizens became adept at constrained escalation, normative morality plays, and hiding the degree to which protests had leaders or organization in order to avoid the most severe repression (O'Brien & Deng, 2015; Fu, 2017).

Repression has not been eliminated from the tool kit of the Chinese state during this period of responsive authoritarianism. However, repressive tactics have become more targeted, often hidden or disguised, more pre-emptive than reactive, and often with greater psychological sophistication (Ong, 2015; Chen, 2017; Cai, 2008). In a collection of articles, (O'Brien & Deng, 2015; Deng & O'Brien, 2013) examine the practices of "psychological coercion" and "relational repression" as strategies deployed by state agents who find it harder to dominate through sheer physical coercion or have lost the ability to withhold resources, especially to citizens with their own market power. An aggrieved citizen might

be compelled to give up resistance by a colleague or a relative. They note that "protest control is taking on a person-by-person quality," made possible by the vast increase in state resources dedicated to internal security and stability (O'Brien & Deng, 2015; Wang & Minzner, 2015; Ong, 2015).

China's progression toward more sophisticated repressive strategies is not unique. The use of non-physical, hidden or delegated repression has been noted in many parts of the world and as important characteristics of repression (Earl, 2003). Davenport (2007) shows how high capacity states can disable social movements without overt violence. Frantz and Taylor find that the ability to co-opt potential opposition leaders can alter an authoritarian government's need for widespread repression while enhancing their ability to target clear opponents of the regime (Frantz & Kendall-Taylor, 2014). Sophistication and "person-to-person" contact are not cheap. Observers noted in 2011 that China's internal security budget exceeded its spending on the military. Wang & Minzner (2015) detail the "rise of the security state" and its multiple tasks of bargaining, compensation, surveillance, and repression.

As many studies have shown (Tanner, 2004; Chen, 2012; Cai, 2010; O'Brien & Deng, 2015), escalation incentives are baked into this model of responsive authoritarianism and were even summarized by protesters themselves with a pithy saying: "Do nothing and get no results, do something small and get small results; do something big, and get a big result" (Tanner, 2004). With the rise of social media, escalation has become far easier with lower costs and more immediate impact. The virality of online discussion makes it more possible that narrow and limited grievances, over environmental degradation, food safety, or lapses in public safety, can be united online into encompassing demands for greater transparency, a freer media, or bottom-up accountability for political elites. Social media is an easy way to escalate and curry favor with public opinion, especially for Netizens with large followings and influence.

Responsive authoritarianism also has had clear benefits for the government. It has gained from its more lenient and accommodating stance toward protest and grievances, maintaining or even improving its legitimacy/public approval since 1989 (Dickson, 2016; Tang, 2016). Social media also provides the central government easier access to "real"

public opinion, solving the "dictator's dilemma" of not having a true sense of how people feel about a repressive regime (Wintrobe, 2000). If people feel like they can be more honest in social media, then the regime also benefits from this new data point. How can the government have its cake and eat it too? How can it sustain high levels of protest and response/bargaining within the context of massive social media use? Can responsive authoritarianism survive the onslaught of virality?

2 Empirical Analysis

This analysis draws upon two sources of data: 1) government documents and manuals about opinion guidance and 2) leaked internal documents from a social media company Sina Weibo. The former serves as transcript evidence of the government logic of control we theorize in previous sections, and the latter empirically tests our theory of information control using observational data contained in the decisions of social media companies to escalate users and content to the authorities.

A leaked dataset of internal censorship logs from Sina Weibo, a popular Twitter-like social networking site, gives us a chance to explore what kind of online content and users relayed back to the government. Although we cannot completely disambiguate between "reporting up" to facilitate repression and "reporting up" to facilitate cooptation, we know that online speech and activity is often invoked by state security agents when they engage ("drink tea") with citizens. Reports by NGOs show that online activity can be the starting off point for government repression, through "drinking tea" visits, arbitrary detention, and arrest (Tager, Glenn Bass, & Lopez, 2017). However, our analysis only covers the first step of reporting back to the government. In addition to the repressive activity reported by diehard activists, it is evident from government documents and manuals analyzed in this paper that government outreach to social media "thought leaders" involves a combination of carrots and sticks.

We find evidence the the state is most worried about controlling the virality of content, the influence of individuals, and the ability of social media discussion to mobilize and incite. Orders to report up are almost always accompanied by instructions to censor,

often via a covert form of censorship called shadow-banning[1], or out-and-out deletion. We find that this is done in order to reduce the virality of a discussion, to shape the course of a discussion by eliminating certain types of posts, and to reduce the influence of opinion leaders while giving a freer rein to the "nobodies" (for example, by allowing small users more leeway than big users).

2.1 Who, Not What: The Logic of Information Control

In this section, we present transcript evidence of our claims that the state is consistently focused on the identity and status of the user rather than the specific content of the post. The regime is interested in reducing counter-hegemonic space via targeted efforts to co-opt or repress individuals who are influential voices in the online public sphere.

This section draws on data from dozens of party- and government-produced books, documents, or manuals about "opinion guidance," "thought work," and information controls whose intended audience is government leadership cadres. The authors of these manuals and documents are high-level propaganda and public security officials as well as academics conducting research on behalf of the government.

Government logic of opinion guidance described in these manuals is consistent with our theory of "who not what." These documentary sources show that the state focuses its opinion guidance efforts on users with large followings or viral posts, what is referred to as an "opinion leader" (意见领袖). Nearly every manual on opinion guidance included at least a small section devoted to management of these "opinion leaders." Below, we will present this logic using excerpts from government manuals and documents.

Manuals suggest that the Party is concerned that it has ceded too much discourse power to voices outside the system. One manual suggests that "the mechanism for public opinion formation online has changed... [as the] means of digital dissemination has broken the [state's] traditional monopoly of discourse." This manual laments that social media has "brought about a change of position from strong to weak among elites, and from weak to strong among the grassroots." Further, the manual suggests that, "because cyberspace

[1]Shadow-banning is commonly used across social media platforms to prevent users from discovering that they have been censored. When a user's post is shadow-banned, only the original poster can see the post while it remains invisible to all others.

has no systemic barriers or binding ideological constraints, all kinds of thoughts, ideas and values have a platform. Different classes, areas, and types of media can exchange, integrate, or confront these ideas, making the public opinion environment increasingly complex" (B. Zhou, 2011, 118-122). In excerpts presented later in this section, the state appears fearful that social media's unique attributes and the power of opinion leaders as agenda setters on these platforms can quickly aggregate and organize individuals around counter-hegemonic ideas, which can undermine regime legitimacy and also lead to system-challenging collective action

To confront this expanded space for counter-hegemonic discourse, manuals stress the importance of opinion leaders because they occupy the most social real-estate in the online opinion field. Because they are so well-connected, manuals direct cadres to pay particular attention to "'Big V' users who have many fans and followers... [because they] decide the development of online opinion events or the formation of topics of discussion" (Zeng, 2015, 180-194).

Manuals not only consistently stress the importance of opinion leaders in the state's opinion guidance efforts, but they also agree about the metrics cadres should use to track and identify them. One manual suggests the importance of two key metrics: repost counts and fan/friend counts. This manual suggests that "Weibo's special model for information dissemination makes it much easier for social media surveillance workers to identify the original source of opinions and thus identify opinion leaders" (B. Zhou, 2011, 201-228). Another manual elaborates on the logic behind targeting users based on these variables, citing internal government research on Weibo:

> "[In] over 30 viral opinion incidents on Weibo between 2011 and 2012, there were only 7584 viral posts where the amount of retweets surpassed 500; 5047 of these viral posts were authored by ony 305 Weibo users. Moreover, these 5047 posts [were the root source of] 66.5% of all total posts and 80% of all retweets and comments. Though the public opinion space seems complex and untenable, in reality, it is largely controlled and led by a small number of big V users..." (Zeng, 2015, 180-194).

In this same manual, the authors then suggest that the state should balance their dual goals of legitimation and control by regularly approaching opinion leaders and not only

guiding them toward consensus, but also learning from their perspectives:

> "We must communicate with them, seek consensus, and let them make suggestions from a rational and constructive perspective, instead of publishing aggressive and inflammatory extreme speech as they please... it is necessary to maintain close contact with them, through regular or irregular meetings or online exchanges, to take the initiative to invite them to visit major public projects..." (Zeng, 2015, 180-194).

The state targets opinion leaders because they are uniquely positioned to set the agenda on social media because of their celebrity status or their online reputation that has commanded a large following (Dang, 2013, 115-123). One manual argues that opinion leaders have the power to "set a new agenda to lead the public to focus their attention elsewhere, resulting in a change in the overall direction of public opinion" (B. Zhou, 2011, 201-228). This manual discusses the case of a 2010 Shanghai high-rise fire that killed 58 people. It suggests that "changes in the agenda of discussion of this event were closely related to the actions of opinion leaders." Popular blogger Han Han, deputy editor of the magazine Caijing, Luo Changping, and Economic Observer reporter, Chou Ziming, brought to light systemic issues behind the disaster. Public outry reached a fever pitch after an opinion leader with username "Beijing Cook," organized a "public mourning event that attracted 100,000 participants in Shanghai." (B. Zhou, 2011, 201-228). This example illustrates the danger opinion leaders pose to the goveernment due to their ability to aggregate diffuse and atomistic concerns and mobilize their followers to directly challenge the CCP. Shortly after this incident, the state increased pressure on Han Han, who has since ceased his once prolific and acerbic blog writing. At the end of his blogging career, he famously remarked that "influence belongs only to those with power... they own the theater, and they can always bring down the curtain, turn off the lights, close the door, and turn the dogs loose inside" (Osnos, 2014).[2]

To prevent opinion leaders from fomenting challenges to the state, manuals stress the importance of "institutionalizing specially selected opinion leaders" as well as "discovering, cultivating, training opinion leaders, and encouraging leadership cadres to act as opinion leaders" (B. Zhou, 2011, 201-228). One manual stresses that opinion guidance should prioritize cooptation rather than repression, even in cases when users are posting

[2]See http://www.webcitation.org/74wItIZaz.

inflammatory information.

The continued importance of the revolutionary logic of "ideology and organization," detailed in Schurmann is apparent in the frequent appeals to the "united front," a concept that grew out of the revolutionary Communist Party's efforts to guide and shape the opinions and sympathies of those within non-CCP organizations such as the KMT. One such manual references principles of "united front" work in state efforts to minimize counter-hegemonic space online, suggesting that cadres should "take advantage of united front work with new media individuals" by "unifying with those outside of the party and outside of the system," "supporting those with good political values, who are familiar with online rules, and who are influential and popular 'big V' users" (T. Zhou, 2012). One manual furthers:

> "Develop an online 'united front' [...] In response to critical opinions, 'do not fight, instead look for the greatest common denominator.' Maintain the relative independence of opinion leaders, seek common ground on major issues while reserving differences on minor ones, gather and co-opt those with different points of view, deal with individuals on a case-by-case basis, support [those with the] 'correct' [opinion] and repress [those with the] 'wrong' [opinion]" (Zou & Su, 2015, 88-89).

This is echoed in another manual that stresses that not only should the state be focusing on co-opting and guiding the opinions of opinion leaders, but that it should tolerate and learn from a plurality of views:

> "When we identify ["grassroots"] opinion leaders on Weibo, we must strive to guide them in the right direction through [opinion leaders with the correct opinion]. In particular, we must pay attention to the role of network opinion leaders who are close to the identity of netizens, encourage their positive suggestions, and tolerate their radical speech." (B. Zhou, 2011, 201-228).

One other trend in the manuals is the lack of concern for content. Most manuals do not suggest clear content-based mechanisms for monitoring online speech, and when they do, it is left to the discretion of local governments. Instead, social metrics discussed above such as repost counts and fan/friend counts are key to "public opinion early warning systems." Outside of the manuals, this kind of approach is used to identify "hot topics." Public opinion emergency management reports published by the Chinese Academy of Social

Sciences identify "hot topics" using aggregate search volume on a variety of events.[3] Tracking of these metrics is done in the name of pre-emption. One manual suggests that, "using suitable technological methods, real-time monitoring of information with high repost counts can be implemented in order to contain crises at the first sign of danger" (B. Zhou, 2011, 201-228).

2.2 Evidence from Sina Weibo Censorship Logs

While government documents show the logic behind how individuals are targeted for repression and cooptation based on their behavior on social networks, we draw on leaked censorship logs to test our theoretical claims using data that measure how information control is carried out in practice.

Leaked internal logs from social media company Sina Weibo were first shared with the Committee to Protect Journalists (CPJ) by a former employee who worked in Sina Weibo's censorship division.[4] The data are logs that record government directives from various bureaucracies to remove content, report data on events, and report on individuals to the government through Beijing-based government affairs liaisons. Logs can include information about users who have been reported to higher levels, quotations or paraphrased content that the government wants removed, or requests for reports on changes in the perceived level of monitoring or general political situation at a given time. Sina managers then include decisions and instructions for the employees tasked to respond to these government directives. These logs are used to convey this information between employees working on different shifts.

Logs also give instructions to employees to "report up" or "report data" to government affairs liaisons in Beijing who cooperate with supervising Beijing propaganda and public security bureaucracies. Instructions to report data, content, and individuals serve two purposes that are difficult to disambiguate. In some cases, it is clear that reporting up

[3] Annual "blue books," policy briefs written for government cadres, on "social opinion and emergency management" compiled and published by the Chinese Academy of Social Sciences rank hot events by an index of search volume across several online services.

[4] The source of this leak is a former employee at the censorship office. We are not in direct contact with the source, but the source has been vetted by journalists at CPJ and we have communicated with the source through contacts at CPJ. The source has consented to the publication and use of these data for researchers.

of content and data serves an informational purpose that may be used to measure public opinion about a sensitive topic or monitor local agents. In others, information about users is reported to the authorities, most likely for further investigation or criminal prosecutions.

In total there are 8,427 individual logs in our data. Logs are notes related to government directives, management censorship decisions, work guidelines, employee duties, etc. Logs disseminate management decisions about how/if to implement government directives to the large team of content moderators employed by Sina. A log will usually include descriptions or excerpts of content along with instructions to content moderation employees on how to proceed with moderation of this content. While many datasets related to censorship capture only keyword-based censorship or manual content review, these data include logs about both. This dataset includes the complete set of logs from 2011-2014. For this paper, we analyze only logs from 2012. We do so in order to limit the analysis to a single year under the Hu-Wen Administration. In other research, we explore the full range of data and the changes in censorship behavior in the early Xi Jinping administration (Miller, 2018).

2.2.1 Limitations and Scope Conditions of Logs

These logs are from a single social media company, Sina Weibo. As such, there are limits to the external validity of our inferences. That being said, we have reason to believe that our inferences can generalize to other social media companies, particularly Tencent. Several logs indicate that competitor Tencent receives the same directives from the State Council Information Office, Shenzhen-based regulators under the same banner as Sina's regulators in Beijing, and more recently the Cyberspace Administration of China (CAC). Tencent, like Sina, also delays implementation or disobeys directives when doing so gives them a competitive edge (Miller, 2016). Additionally, the market capitalization of both Sina and competitor Tencent were similar at the time these logs were written, which means both had similar leverage when it came to negotiating government directives (Miller, 2016). It is reasonable, based on the data from the logs to expect that the process of reporting of users to the authorities is similar at Tencent. Together, Sina and Tencent represent the vast majority of the social media market.

The logs also indicate that tech companies in China cooperate or are influenced by each other's censorship decisions. For example, Sina Weibo cooperates with Baidu when developing keywords for blocking and content filtering. In a log from May 7, 2013, Sina instructs employees to add to a shared database of blocked keywords with Baidu in accordance with an arrangement with the company.

Even a conservative approach to inferences drawn from these logs does not diminish their significance. Sina Weibo is a very large and popular social media company in China. During the time of the logs it was the most popular microblog and was ranked in the top 3 of all domestic social media companies by monthly active user statistics. During the time of the analysis, Sina Weibo boasted over half a billion registered users.

2.2.2 Content Coding and Variable Definitions

Each log is coded according to several content variables and instruction variables. Content variables reference the topic category of content being targeted for censorship.[5] Instruction variables refer to instructions on how to censor content.[6] Logs often include conditional statements in the censorship instructions such as "delete content that is attacking the Party and government" or "report users with more than 10,000 followers." See Figure 1 for an example of a censorship log and the data points extracted from the raw text. The three content variables used in this analysis—collective action, government criticism, and local government corruption—are drawn from findings in the existing literature. King, Pan, & Roberts (2013) argue that the Chinese state tolerates government criticism while targeting threats of collective action. Dimitrov (2017) and Lorentzen (2014) argue that the government allows for social grievances to be aired in order for higher levels of government to collect accurate information about their lower level agents. Instruction variables are constructed by inductively building a keyword list by examining instruction text from all 2012 logs. The variables we use in this analysis can be seen in Table 1.

[5]These content variables are coded by two research assistants and have high intercoder reliability. Coding rules and concept diagrams can be found in the appendix along with intercoder reliability measures.

[6]A list of the keywords used can be found in the appendix.

Figure 1: Example of a Censorship Log and Variable Types

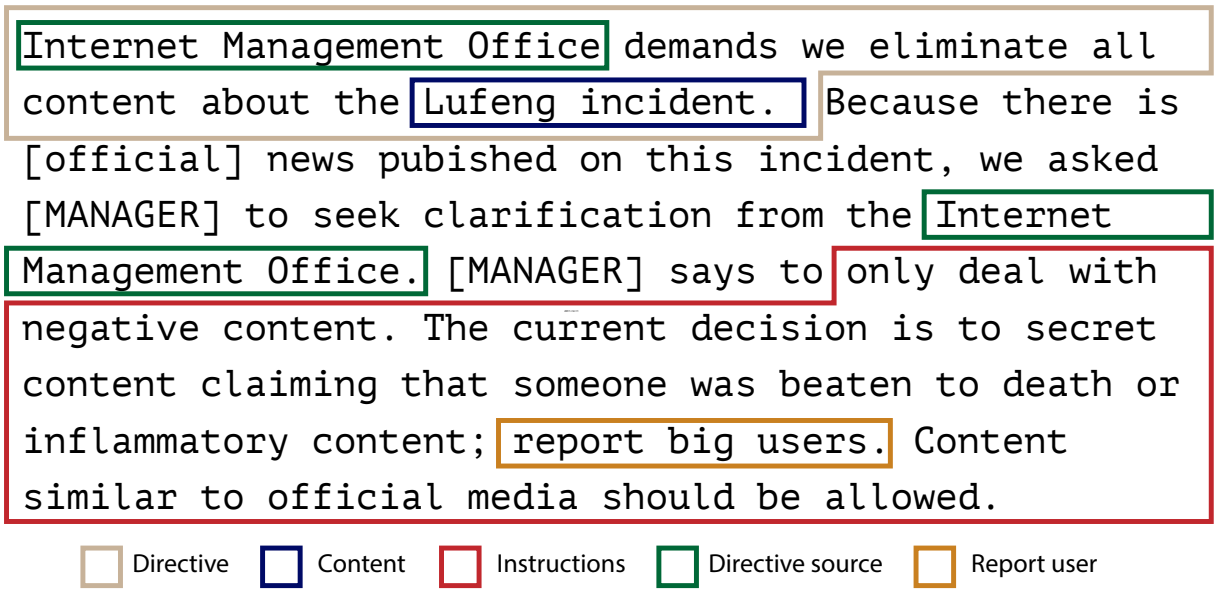


Table 1: Variable descriptions

Variable	**Description**	**Type**
User Attributes	Any differentiation or categorization of users based on attributes such as fans and following counts, big/small user, media user, foreign user, verified user, etc.	Instruction
Retweets	Any mention of the number of reposts/retweets of a log.	Instruction
Govt. Attack	Any differentiation in instructions related to an attack on government officials, policies, or institutions.	Instruction
Political Humor	Any differentiation in instructions related to mocking, satirizing, or ridiculing government officials, policies, or institutions.	Instruction
Extreme	Any differentiation in instructions for ordinary/normal vs. abnormal/extreme posts.	Instruction
Rumors	Any differentiation in instructions that relates to rumors or fake news.	Instruction
Retweets	Any differentiation in instructions based on retweet counts, or that relates to other attributes of post virality such as the number of times it is sent in a private message.	Instruction
Inciting	Any differentiation in instructions that relates to content that "incites," or "provokes"	Instruction
Govt.	The content mentions or implies a Chinese government institution, organization, or bureaucracy, a Chinese government official of any rank or position, their family members or their partners/mistresses, a Chinese government policy, or a Chinese state-owned enterprise (SOE).	Content
Local Govt. Corruption	The content mentions or implies any of the following: 1) misuse of local government office or local government funds, 2) sexual misconduct of local government officials 3) a local government official and/or his/her family financially benefiting from a government post.	Content
Collective Action	The content either 1) mentions or implies an event where a group of people took action together to achieve a common objective, or 2) mentions or implies an individual or group of individuals who are advocating on behalf of a social, religious, or ethnic group.	Content

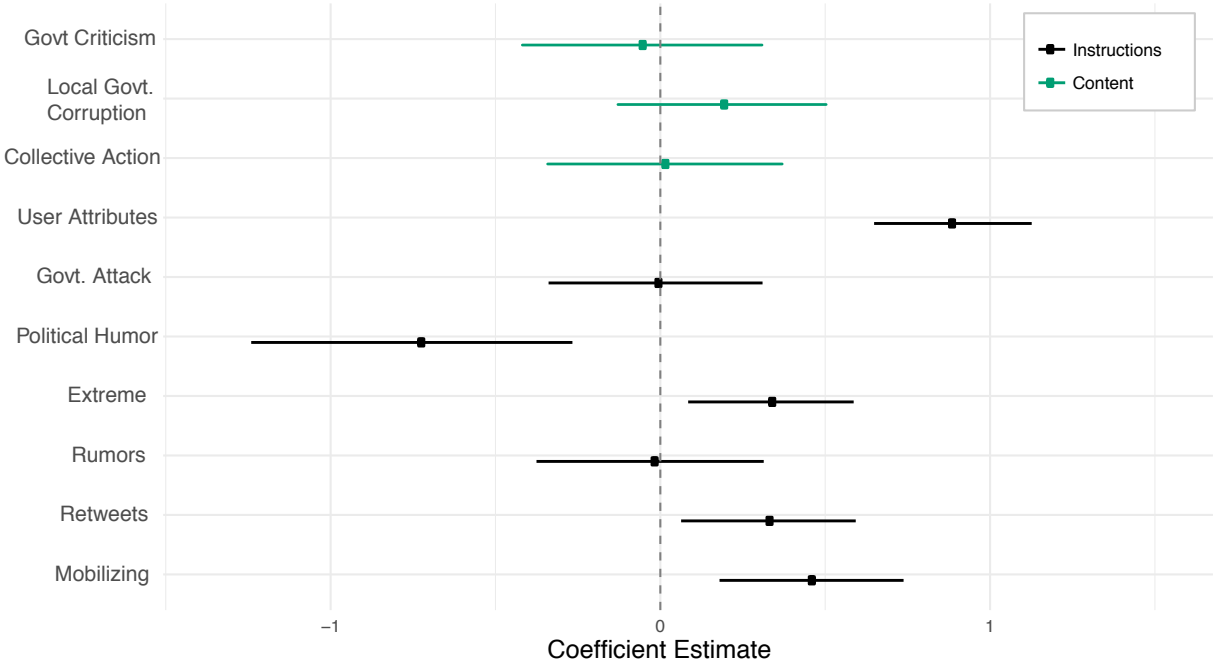
2.3 Empirical Strategy

We run a logistic regression model on log data from 2012 to measure the correlation between content and instruction variables and the reporting up of users to the authorities. We expect that content variables will not be strongly correlated with individuals being reported to the authorities. Though content does matter in the state's information control efforts, we argue that the state intervenes directly in information control efforts when a user is influential or content has gone viral. These attributes will be the strongest indicators of whether a user is reported to the authorities. The model's dependent variable is "report users," which captures whether logs explicitly instruct to report individual users to the authorities.

We illustrate these relationships using high profile cases from the logs. These cases are broadly representative of our interpretation of the dynamic between commercial censors and the state. They include one case of a viral story about the citizenship of a high profile journalist, one case of a civil society activist attempting to mobilize supporters to join his movement, and one case–the Bo Xilai incident–when the government was intensely interested in tracking public opinion regarding Bo, the former Party Secretary of Chongqing Municipality and the government's prosecution of his crimes.

We examine how the content and instruction text in logs are associated with the decision to "report users," or send information about an individual user to a Beijing-based government affairs liaison who is responsible for coordination with provincial-level bureaucracies, the State Council Information office, or the Cyberspace Administration of China. The results of our logistic regression model can be found in Figure 2. A regression table can be found in the appendix.

Figure 2: Logistic Regression Coefficient Plot



By far the biggest predictors of whether or not content or users are reported up is the presence of words that describe the influence of the user mentioned in the log (User) or how "extreme" the content is (Extreme). Mentions of retweets and how "inciting" (煽动性) content are also significantly and positively associated with whether a post is "reported up," an indication that the potential of a post to incite/provoke is strongly associated with decisions to report users up to Beijing. As we discuss below in the cases, these include posts that are clear exhortations for real world social mobilization and other posts that incite only in an abstract way.

None of the three content-based independent variables is significant. This includes content about collective action events. The censors appear to discriminate between discussions about collective action events and calls to mobilize for further collective action. For example, during an August 2012 taxi strike in the city of Hangzhou, the censors note, "Hangzhou taxis are on a collective stoppage today, protesting the rise in gas prices and the state of the roads. This news can be circulated, but any posts that incite or appeal to other cities should be secreted. Report the original poster." Similarly, in logs about the 2013 arrest of Zhang Anni, the 12-year old daughter of a Tiananmen activist who wrote a letter to Barack Obama, the censors allow for discussion of the case, but not "incitement":

"Netizens expressing support for the Zhang Anni incident in Hebei, deal with any content that is inciting on Weibo, there is no need to delete general discussion at this point."[7]

— Log entry from April 10, 2013

It's possible that the censors realize that these are hot button topics and deletion is not in their commercial interest. It's also possible that the state itself is tolerant of more open discussion because while some will sympathize with the striking taxi drivers or young activist, others will criticize them for inconveniencing customers or not being sufficiently patriotic. Han (2018) argues that the state benefits from an online atmosphere that is more open because it allows for "discourse competition" between supporters and opponents of the government. These posts reveal that careful line tread by the commercial censors. Discussion of sensitive collective action and political dissidents is tolerated up to a point. Mobilization or incitement is reported to the authorities.

The commercial censors report up information about influential "big V" users and viral, inciting news. Information is censored, but often via secreting, so that original posters might believe their posts were uncensored but just unpopular. Online information control is deployed alongside real world physical repression for the die-hard activists, as we see below in the case of Xu Zhiyong, the civil society activist.

2.4 Reporting of Users

2.4.1 Yang Lan's Citizenship

On March 7, 2012, the censors at Weibo were getting off the night shift. As they wrote up guidance and instructions for the next shift, they wrote, "Saying that Yang Lan is a US citizen has already been declared a rumor, directly delete posts of small users, report up big users."[8] A few days later, they repeated the command, "[Suggesting that] 'Yang Lan is a US citizen' should be considered fake news, if it is seen on Weibo, delete it, report big users."[9]

[7]Original Chinese text: "网友声援合肥张安妮事件，微博中处理煽动行动内容，一般讨论的暂时不必处理。"

[8]Original Chinese: 说杨澜是美国籍的已经辟谣了，小用户直接删除重要用户报一下

[9]Original Chinese: 传杨澜为美国籍为虚假消息微博见到删除重点用户报一下.

The censors were attempting to shut down a raging debate on Weibo, "Is American citizenship still valuable?" On March 6 there were more than 1.5 million posts on the topic. Some of the discussion focused on the post-Global Financial Crisis state of the US economy, but other Netizens took aim at their compatriots. "If you say China isn't good, you are a Western slave. If you say America is good, then you are an American bitch. If you say you don't want to be a Chinese, you are a wretch and a traitor... If you get a green card and stand on the corner of American streets shouting, 'I love you, China,' you are a 'patriot.'[10]

Yang Lan, a celebrity journalist and TV host, "China's Oprah Winfrey", was alleged to be this kind of person. She appeared patriotic and had a strong political background (her father was a translator for Zhou Enlai[11]), but after making a fortune in China's media market, had taken citizenship elsewhere. Yang tried to rebut the accusations. Later, she and her wealthy husband, Bruno Wu, threatened to sue.

In the Yang Lan case, the censors treat users differently based on how important and influential they are. Normal users who gossip about or criticize Yang Lan will just have their posts deleted. Opinion leaders will be deleted and reported to the government. Although it's possible that the government is worried about collective action, this seems unlikely. It is more likely that the government is concerned about its legitimacy. If someone as well-connected, wealthy, and "within-the-system" successful as Yang Lan seeks foreign citizenship, it reflects poorly on the CCP. Are the elite losing faith in the country's future? Do even the well-connected and rich have an exit option should the future take a turn for the worse? The crackdown on the Yang Lan case shuts down these questions.

2.4.2 The Arrest and Trial of Xu Zhiyong

In cases involving well-known political activists, however, the concern with influence is more directly connected to social mobilization and collective action. For example, Xu Zhiyong, a well-known political activist and liberal reformer, first became famous in 2003 as part of a group of intellectuals who challenged the state's "custody and repatriation

[10] Archived Atlantic article available here
[11] See relevant Xinhua report here

system" that detained rural migrants who did not have legal documentation to live in cities. He later opened an NGO, *Gongmeng*, (also known as the Open Constitution Initiative), which championed the rights of migrants to education in cities and investigated other scandals. He was detained in 2009 and *Gongmeng* was accused of tax evasion. At the time, Xu was working on the case of the tainted milk scandal in 2008, which had led to the deaths of several infants who consumed adulterated baby formula. In May and August of 2012, Xu Zhiyong used Weibo to ask for citizen support for *Gongmeng*, even soliciting donations online and calling for unity among Chinese citizens. Both of the posts below were instructed to be secreted (original poster can see it, but no one else), "big V" users who posted such material were reported up.

"On May 23rd, Xu Zhiyong used domestic sites (Sina Weibo and Tencent Weibo) and foreign sites (Twitter) to incite Netizens to join *Gongmeng*, to purchase a "citizen" badge. Even calling for people to sign a "citizen promise." [...] If this content is seen on Weibo, secret it. Report big users to the authorities."[12]

— Log entry from May 26, 2012

"If 10,000 citizens each give *Gongmeng* 100 yuan, we could do so many things! We earnestly request support for *Gongmeng*, please believe that we will use the money for the places that are most in need of justice. Legal aid assistance account: Xu Zhiyong, China Industrial Bank Beijing Branch, [BANK INFO] - Secret posts calling for donations to *Gongmeng*."[13]

— Log entry from October 8, 2012

In 2013, activists including Xu Zhiyong called for greater disclosure of government salaries and wealth. Xu Zhiyong championed this initiative as part of his new venture, The New Citizens Movement.[14] Demonstrations were held in several cities and led to the arrests of many people. After being under house arrest for three months, Xu Zhiyong was finally detained in July 2013, charged with "attempting to disturb public order" in August, and sentenced to four years in prison after a trial in January 2014. (He was released from prison in July 2017.)

[12]Original Chinese text: "许志永"于5月23日在境内《新浪微博》、《腾讯微博》及境外《推特》发布信息煽动网民加入"公盟"组织，购买"公民"徽章。并呼吁网民签署《公民承诺》。原文称："需要公民徽章的请给我发信[EMAIL]，我们需要再订制。其实不需要传统意义上的所谓加入，公民（公盟）是自由公民的联合，我们都是时代新公民，为公民社会共同努力！请发信至[EMAIL],常联系。这个内容微博中见到私密，重要用户发的上报负责人"

[13]Original Chinese text: "如果有一万个公民每年每人捐给公盟100元，我们能做多少事啊！恳请支持公盟，请相信我们会把钱用到最渴求正义的地方。法律援助救助账号：许志永，中国工商银行北京分行学院路支行，[BANK INFO]，支付宝兼恳请支持，望扩散号召给公盟捐钱的私密"

[14]A more detailed analysis of this movement is in Pils (2014).

During this time of renewed activism around transparency issues, the commercial censors were given many orders to censor information about Xu Zhiyong and, importantly, to report up "big users" and "special circumstances." Certain users were named in particular, including Xu's defense lawyer, Liu Weiguo, and other rights defense lawyers. On the same day of the post below, Liu was himself detained:

"Regarding the issue of several lawyers going to Beijing's Third Detention House to visit the detained Xu Zhiyong, when verifying, secret (this information). Increase the number of key words that are automatically secreted. Immediately report up any special circumstances including the lawyers, Chen Jiangang[a], Liu Weiguo[b]. These two say a lot, while verifying, pay careful attention; these words have already been verified."[15]

[a]Read more about Chen Jiangang here
[b]Read more about Liu Weiguo here

— Log entry from July 18, 2013

With well-known political activists who openly challenge the government on a number of fronts, the repressive apparatus of the state is not hidden or subtle. The commercial censors are told that the issue is of great importance and online discussion is used to monitor particular people, both virtually and eventually through physical detention and criminal charges. Many of the lawyers mentioned in these posts were later detained in the July 2015 massive crackdown on human rights lawyers across China. Many other "rights defense" lawyers have gone underground since the arrests of 2014 and 2015 (Pils, 2014).

But the Xu Zhiyong incident also reveals that the government desires to shape and guide public opinion, not just repress the actions of these few or stop collective action. During the January 2014 trial of Xu, for example, censors were instructed to carefully calibrate censorship:

"Tonight at the First Intermediate Court in Beijing, Wang Gongquan admits guilt that he and Xu Zhiyong incited people to disturb public order, he has deeply reflected and awaits trial. Related content can be discussed as long as it reflects the media reports. Anything that does not reflect the official line, secret, stop small users and small media from discussing, discussions in the big media should be verified first, posts with lots of retweets should be carefully reviewed."[16]

— Log entry from January 22, 2014

[15]Original Chinese text: "关于有几位律师和网民去北京第三看守所看望被关押的许志永一事，审核中见到都私密处理，加了部分自动私密关键词，有特殊的情况及时上涉及的律师：陈建刚刘卫国这两个人说的比较多，审核的时候多注意，已经是审核词了"

Both the Yang Lan and Xu Zhiyong cases demonstrate the targeting of influential users as the regime seeks to control what is said and heard about China's elite. The cases are different in the sense that Yang Lan's citizenship is an abstract threat to the CCP's standing while Xu Zhiyong's exhortation to fellow citizens is a concrete call for social mobilization. It's also no coincidence that these cases involve a famous lawyer and and even more famous journalist, part of China's new professionalized and cosmopolitan elite. It is critically important to the Party not only what they do or say online, but also what other people say about them.

2.4.3 Bo Xilai's Scandal

The importance of external elite voices was also apparent during the 2012-2013 power struggle, and potential coup, which pitted Chongqing Party Secretary Bo Xilai and the incoming leader, Xi Jinping. As the scandal unfolded, influential users who expressed support for Bo were reported to the authorities. Additionally, the government demanded that Sina Weibo monitor and report public opinion data at fixed intervals. In the log below, we see Sina implementing requests to report data about public opinion related to Bo Xilai. This log serves both an informational role and a repressive role:

"Handling Bo Xilai-related posts: Total posts censored, among them, support for Bo (#, %), ridicule [of Bo] (#, %), reposting foreign media (#, %), other (#, %). [Also report] whether or not there are any influential users clearly and directly supporting Bo. Also, today's Bo Xilai reporting should be calculated according to the following intervals. The reporting up schedule has been revised to the following three time slots: 10am, 6pm, and 11pm."[17]

— Log entry from April 12, 2012

"[Keep track of the] number of deleted and filtered posts about Bo Xilai and Wang Lijun, also keep track of the number of deleted and filtered comments, how many of these are articles oppose the way the central government has handled Bo Xilai, are unsatisfied with the results of the investigation of the case, how many are republished from foreign websites, and other related information. Reporting timeline has been changed as follows to these three times: 11am, 4pm, 11pm."[18]

— Log entry from April 13, 2012

[16]Original Chinese text: "王功权承认违法犯罪被取保候审】北京市一中院今晚称，王功权承认与许志永策划煽动聚众扰乱公共场所秩序的违法犯罪活动，王功权表示深刻反省；对王功权予以取保候审。相关内容目前可以按照该新闻口径讨论，凡是与新闻不一致的私密处理，小用户小媒体禁止评论，大媒体评论先审转高审。"

[17]Original Chinese text: "有关薄的处理数XX条，其中，支持薄XX条，占XX%，调侃XX条，占XX%，转发外媒XX条，占XX%，其他情况XX条，占XX%。有无影响力大的用户明显直接挺薄内容。今天薄熙来的数同时按这个时段统计这个报数时间段改为10点、18点、23点三时段"

The Bo Xilai scandal was what the Party calls a "sudden, unexpected emergency," (突发事件). As such guiding public opinion followed the common practice of pairing immediate public opinion responses with data collection and planning for more comprehensive public opinion interventions as the situation developed. More targeted interventions required information gathering which is why Sina was asked to report data to the authorities. In later stages of the crisis, more specific topics of conversation became off-limits, not because there were pre-existing procedures to target specific categories of content, but because the goal of "opinion guidance" is to reduce the visibility and influence of content propagated by counter-hegemonic voices and increase the visibility of state narratives. An example is a more specific log from the Politburo Standing Committee's inspection tour of Chongqing following the incident:

"Today [MANAGER] published content moderation instructions for the inspection tour of Chongqing by the 9 members of the Politburo Standing Committee (PSC):

1. Manage all content (including news) related to the inspection tour of Chongqing by the 9 members of the PSC.
2. Manage all rumors about government, the most important being: foreign reports and speculation, reports from foreign media where the size of discussion in the comment section is particularly large, content originating from sensitive individuals, commentary on Sina Online official accounts, and headline news.
3. For users who publish rumors, the Public Security Department will investigate and deal with the person at the time, and at the same time investigate the criminal responsibility of the person in charge of the website.
4. In mid-May, government rumors were thoroughly cleaned up and whenever one appeared, the owner of the website was required to go to a meeting at the city bureau and rectify their errors."[19]

— Log entry from April 27, 2012

Later, as Bo Xilai was tried in the Jinan Intermediate People's Court, the logs portray a more planned and carefully executed opinion guidance strategy. Many China watchers

[18]Original Chinese text: "薄和王的删除、过滤数，共删除条，过滤条，其中反对中央对薄熙来处理决定的篇；不满案件调查结果的篇；转载境外网站的篇；其他类信息篇。11点、16点、23点三时段以后时段更改如下。"

[19]Original Chinese text: "今日[MANAGER]布置的九常委视察重庆的清理要求：1：有关常委去重庆视察的内容（包括新闻，薄出事之前，所有微博账户）全部私密处理。薄被处理的新闻通稿不要动。下周一前必须完成，做好分类的数据统计。常委名字+重庆、模式、唱红、打黑。2：以后的时政谣言，从严处理，对用户必要时候上手段控制用户。3：媒体评论，北京负责审核发现，及时报备，管理不住的常委新闻，就关闭评论，天津负责清理。以上要求：监控所有人员务必执行。通知原文：1：有关常委去重庆视察的内容（包括新闻）全部处理。2：时政类谣言全部清理，主要类型为：外站报道及猜测、境内媒体的报道后面的评论及尺度比较大的一些新闻内容、敏感人士发布的内容、新浪网官方账户的评论，如头条新闻。3：发布谣言的用户，公安部门将依法查处当时人，同时追究网站负责人的刑事责任。4：5月中旬彻底清理干净时政类谣言，凡出现一条，网站负责人必须去市局开会，进行整改。"

expressed surprise when the Bo Xilai trial was covered live on Weibo, but this was done carefully and with Sina Weibo's cooperation.[20] On the opening day of the the trial, Weibo was instructed to not only target categories of off-limits content, but aid the Party in its efforts to set the agenda by heavily controlling counter-hegemonic discourse from "Big V" users and various non-official media accounts:

"... All comments discussing the matter from big media users and big V users are to be reviewed first before publication, one hour after the broadcast, lock comments. Directly prevent comments on small media accounts."[21]

— Log entry from August 22, 2013

In several related logs on the topic, content moderators were instructed to report up influential users and hide all content that was not exactly as reported by the Jinan Intermediate People's Court's Weibo account. By restricting discussion to official channels and focusing primarily on influential counter-hegemonic voices, the state was able to appear open, transparent, and legitimate in its trial of the fallen Party Secretary.

In the handling of all three, Yang Lan's citizenship, Xu Zhiyong's trial, and the intrigue over the Bo Xilai affair, commercial censors delete or secret information that mobilizes or incites public opinion, they report influential users to the authorities, and they also report up data on public opinion. Their tasks are multifaceted and to focus only on the repressive activity misses the more sophisticated goals of guiding public opinion and delivering more accurate information to leaders about what opinion leaders and the masses are "really" thinking regarding controversial topics. However, information about specific users can always be retained for future repressive activities. In an April 2011 log about a taxi and truck driver strike, the link between information and repression is clear:

"Strengthen handling of inciting and harmful information about the taxi and truck drivers' strike against the rise in gas prices, especially information that mobilizes for the June 1 strike. Accumulate related news and retain the IP addresses (of users) with information that incites or mobilizes."[22]

— Log entry from April 15, 2011

[20] See http://www.webcitation.org/74k5wmpno

[21] Original Chinese text: "所有讨论此事的大媒体评论以及大V的评论均先审后上，一小时后锁定评论，小媒体发表此事直接关闭评论。"

[22] Original Chinese text: "网监通知：进一步加强借"油价上涨"煽动出租车及货运司机罢运、罢工等的有害信息的清理工作，特别是煽动6月1日罢工的信息。相关信息自行积累，行动性、煽动性信息查IP留存。以便随时统计。昨日清理，目前未发现相关行动性信息。"

In an interview with the source of the leak, he recalls that, "when some political rumor was spread widely, or during the Tiananmen Square anniversaries, police harassed people based on the information we provided." The source also recalls that a manager once said, "hand in your big form. The police need it to arrest people" when instructing employees at Sina Weibo's censorship office in Tianjin (Wang, 2016b,a).

3 Concluding Discussion

This paper contributes to our general understanding of the governance strategies of the Chinese Party-State in the virtual public sphere of social media. Building off earlier theories that emphasize the importance of organization and ideology to Communist Party legitimation and control, we find that the government is concerned with the power of virality and influence on the Internet. Attempts to mobilize citizens collectively via online discussion, such as in the Xu Zhiyong epidsode, are threats to the government's hold on power. However, the government also attempts to limit "counter-hegemonic discourse" even when the threat of real world collective action is not imminent.

Maintaining responsiveness and more accurate monitoring of public opinion are also important facets of the government's strategy, however. In a sense, these findings underscore the information problems that confront dictators (Wintrobe, 2000). Social media discussion is desirable because it facilitates more accurate understanding of public opinion trends and social grievances. Yet, it is also too much of a good thing.

Going back to the exposition of the importance of Party leadership in Schurmann (1966), we confirm that the CCP is most vociferously interested in constraining the rise of alternative voices and influence. As the newest battleground for the hearts and minds of citizens, social media platforms are critical spaces for the CCP to inhabit and dominate. The government is intolerant of many topics once they become viral or are circulated by influential opinion leaders. The content of the post is less important than who is posting it and how many people are reposting it. The discussion over Yang Lan's citizenship, for example, seems to fall into this category. It does not immediately threaten any sort of collective action or revolt, but rather it is part of a gradual process in which the legitimacy of the Party is undermined and sabotaged by elites and influential thought leaders.

Schurmann's work is instructive in part due to its focus on the 1950s, the period immediately after the Communist Revolution when the CCP was constructing the building blocks of its rule. In the realm of today's social media, the CCP is similarly limited in its capacity to take over all the space afforded by the Internet, nor does it really wish to do so given the advantages of having more free discussion and better information about public opinion. Instead, the CCP focuses on repressing or co-opting voices that could threaten its position of leadership, to counter what Lei & Zhou (2015) call the "counterhegemonic public sphere." This is an arena in which citizens are exposed to and generate discussion often beyond the confines of the official government position. The key threat is not the degree to which a public opinion is representative, but rather how influential is. They write:

> As the propaganda official put it, in the end it is those who speak up instead of those who keep silent that influence other people and bring trouble for the Chinese government (Lei & Zhou, 2015, 455).

Despite the relevance of Schurmann's conceptualization of Party governance, it cannot fully capture the complex way in which the CCP tries to manage virtual public space in a context that is far more open, diverse, and vibrant than the Maoist period. Moreover, more information and the "appearance of freedom" have real benefits to the government, both in terms of its legitimacy and its understanding of public opinion. Repnikova (2017) finds that the authorities shape media policy by carefully scouring social media. She finds that, "alongside control, Chinese authorities scrupulously listen to and study public opinion online, engage with and respond to public grievances, and creatively mobilize the public through interactive social media" (Repnikova, 2018). Stockmann (2013) shows that allowing for greater media liberalization with continued content control can bolster people's trust in the news. Han (2018) argues that "discourse competition" between regime supporters and opponents benefits the state by revealing that Netizens are not all liberal opponents of the CCP. There are true believers, emboldened by the anonymity and the network-building available online. Finally, Roberts (2018) argues that censorship is more like a "tax" than a simple restriction on information. Most Netizens will never seek out sensitive political topics online or purchase a VPN to read the foreign media. As these leaked blogs seem to indicate, the censors do indeed allow for a great deal of

discussion online to happen while systematically targeting those few users who have the influence and the public following to cause real damage.

Appendix

Logistic Regression Results

Table 2: Logistic Regression Output

	Report Users
Image	−0.118
	(0.093)
Extreme	0.230***
	(0.067)
Attack.Govt.	−0.034
	(0.093)
User	0.469***
	(0.062)
Rumors	−0.018
	(0.083)
Retweets	0.159**
	(0.080)
Inciting	0.234***
	(0.078)
Political Humor	−0.331***
	(0.113)
Government Criticism (Content)	−0.048
	(0.093)
Local Government Corruption (Content)	0.068
	(0.079)
Collective Action (Content)	−0.041
	(0.091)
Constant	−2.048***
	(0.090)
Observations	1,412
Log Likelihood	−491.963
Akaike Inf. Crit.	1,007.926
Note:	*p<0.1; **p<0.05; ***p<0.01

Content Variables

Figure 3: Collective Action Coding Diagram

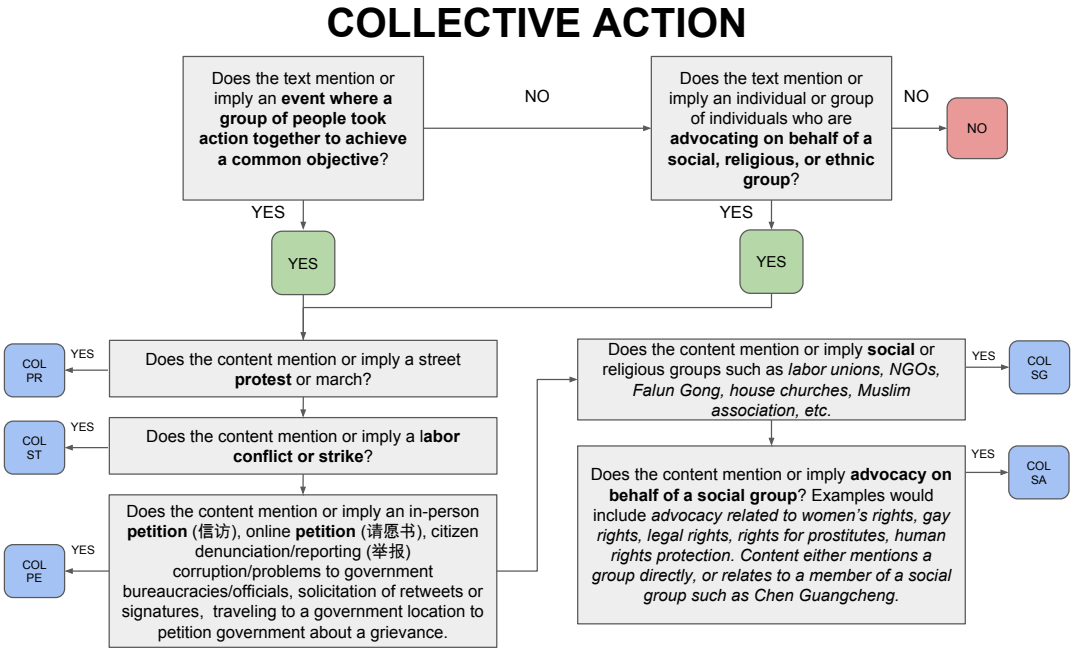


Figure 4: Corruption Coding Diagram

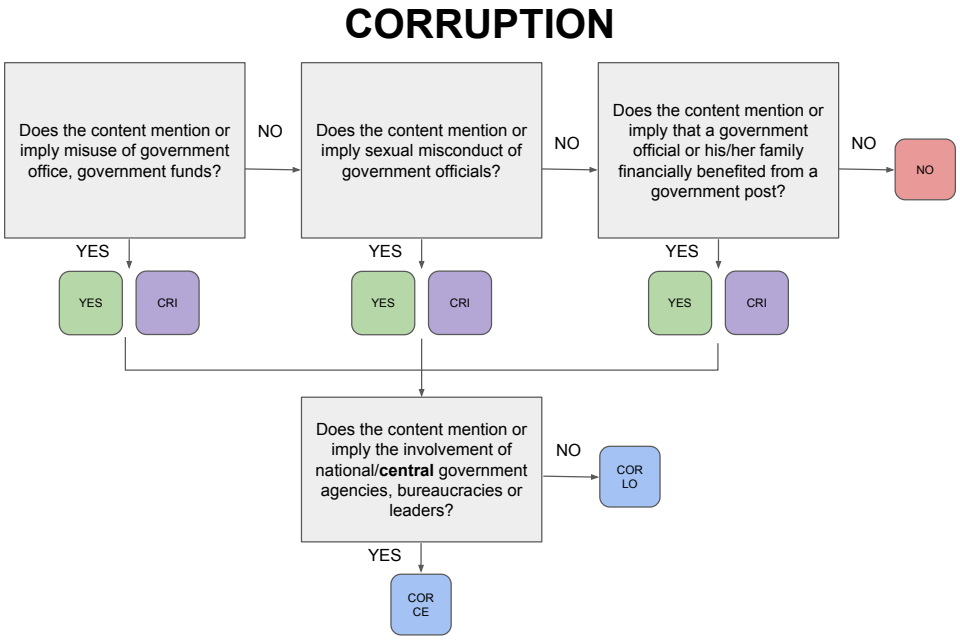
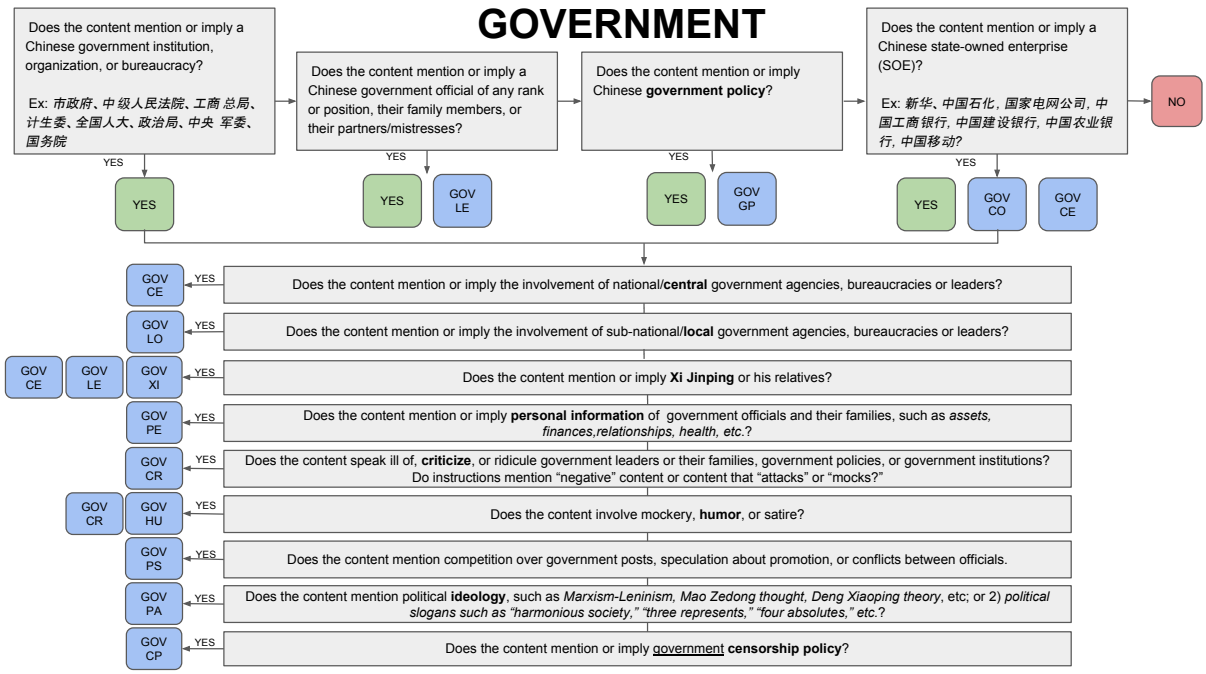


Figure 5: Government Coding Diagram



Instruction Variables

3.1 Report Users

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to a user (用户, UID):

- Report (报备, 报一下)
- Report up (上报)
- Send to Beijing (发给北京)
- Send it to the authorities (发给负责人, 提供给负责人)

3.2 Report Data

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to content (信息, 内容, 文章号, 言论, 评论):

- Report (报备, 报一下)
- Report data (报数)
- Report up (上报)
- Send to Beijing (发给北京)

- Send it to the authorities (发给负责人, 提供给负责人)

3.3 User Attributes

In the instruction text, one of the following behavioral, ideological, or personal attributes, or an attribute similar to the following, is mentioned with respect to users (用户, 账号, 号):

- Small [user] (小[用户,号])
- Big [user] (大[用户,号])
- Big V user—verified user with many fans, V—verified user, VIP (大V, V, VIP)
- Opinion leaders (意见领袖)
- Famous people (名人)
- Media [users] (媒体)
- Grassroots (草根)
- Important [users] (重要)
- Ordinary [users] (普通)
- Delete ordinary users and shadow-ban verified users (普删V私)
- [Users] with many fans (粉丝较多的)
- [Users] with around [#] fans (粉丝数[#]左右的)
- Newly-registered [accounts/users] (新注册)
- Old accounts (老账户)
- Leftist [users] (左派)
- Rightist [users] (右派)
- [Users] who post memes (段子类)
- "Reincarnated" [users]—users who continually create new accounts after deletion (转世)
- [Users] who continue to post (连续发布的用户)

3.4 Extreme

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to content (信息, 内容, 文章号, 言论, 评论):

- Extreme (过激的)

- Excessive/overboard (过分的)
- Vicious/malicious (恶毒的)
- Malicious (恶意的)
- Disgusting/gross (恶性的)
- Disgusting/abominable (恶劣的)
- Inappropriate (不行的)
- Attacking (攻击的)
- Hurling invective (谩骂的)

3.5 Attacking the Government

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to: the country, the party, the internet management system, the party, party-state, country, government, system, social system, leadership, the Central Propaganda Department, Xi Jinping (我国, 我党, 我互联网管理制度, 党, 党政, 国家, 政府, 体制, 社会制度, 领导人, 中宣部, 习近平)

- Attack _____ (攻击_____)
- Make a negative connection with _____ (负面联系_____)
- Concerning something negative about _____ (涉及_____负面的)

3.6 Rumors/Fake news

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to content (信息, 内容, 文章号, 言论, 评论):

- Fake (虚假)
- Untrue information (内容不实)
- Spreading fake news (传播虚假消息)
- Spreading rumors (造谣)
- Unconfirmed true or false (不确定真假)
- According [to rumors] (传)
- Refuting rumors (辟谣的)

3.7 Retweets

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to content (信息, 内容, 文章号, 言论, 评论):

- Retweets cannot surpass [#] (转发不能过[#])
- Retweets surpass [#] (转发超过[#]的)
- Lots of retweets (转发多的, 转发高的, 转发较高的)
- Every hour surpassing [#] retweets (每小时总量超过[#])
- Pay attention to retweets (注意转发)
- Severely restrict retweets (严禁转载)
- Negative retweets (负面转发)

3.8 Inciting

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to content (信息, 内容, 文章号, 言论, 评论):

- Inciting (煽动性)
- Mobilizing (行动性)
- Collective (群体性)
- Call on/appealing to (号召的)
- Cause a fuss, make a scene (闹事)
- Causing a panic (造成恐慌的)

3.9 Political Humor

In the instruction text, one of the following, or something closely resembling the following is mentioned with respect to: the country, the party, the internet management system, the party, party-state, country, government, system, social system, leadership, the Central Propaganda Department, Xi Jinping (我国, 我党, 我互联网管理制度, 党, 党政, 国家, 政府, 体制, 社会制度, 领导人, 中宣部, 习近平):

- Make fun of _____ (拿_____开玩笑)
- Ridicule (调侃)
- Satirize (讽刺的)

- Meme/joke (段子)
- Satirical humor(恶搞)

Intercoder Reliability

These are the results comparing the predictive power of coder 1 to coder 2's decisions using the area under the receiver operator curve (ROC AUC). ROC AUC is a better measure than raw accuracy because it is not misleading in situations where there are class imbalances. Nearly all categories are above the acceptable measure of .7 and most are above .85, which is highly reliable.

Table 3: Intercoder Reliability Measures

Category	**AUC**	**Category**	**AUC**	**Category**	**AUC**
Col. Action	0.89	Government	0.82	Reoccurring Political Event	0.85
Social Groups	0.87	Central Government	0.78	Sensitive Anniversary	0.89
Petitions	0.8	State-Owned Enterprise	0.63	Regular Political Event	0.67
Protest	0.86	Government Policy	0.72	Sexuality	0.82
Social Activism	0.85	Political Humor	0.96	Pornography	0.51
Strikes	1	Government Leadership	0.92	Sina	0.94
Commercial	0.81	Local Government	0.73	Sina Censorship	0.98
Sina's Competitors	0.51	Party Ideology	0.83	Sina Company Business	0.93
Corruption	0.93	Leadership Personal Information	0.7	Hong Kong, Taiwan, Macau	0.98
Central Govt. Corruption	0.91	Xi Jinping	0.98	Hong Kong, Macau	0.99
Local Govt. Corruption	0.84	Government Criticism	0.7	Taiwan	0.99
Crime	0.81	Censorship Policy	0.58	Ethnicity	0.98
Financial Crime	0.89	Power Struggle	0.59	Tibetan	0.94
Political Crime	0.86	Nationalism	0.79	Uighur	0.96
Violent Crime	0.84	Party History	0.67	Other Ethnic Minority	0.94
Illegal Goods/Services	0.89	Military	0.78	Entertainment	0.73
Disaster	0.91	Territorial Disputes	0.94	Terrorism	0.99
Man-made Disaster	0.9	Rumors	0.91		
Natural Disaster	0.76	Non-political Rumors	0.98		
Foreign Media	0.81	Political Rumors	0.88		

References

Cai, Y. (2008). Local governments and the suppression of popular resistance in china. *The China Quarterly*, *193*, 24–42.

Cai, Y. (2010). *Collective resistance in china: Why popular protests succeed or fail*. Stanford University Press. Retrieved from http://www.sup.org/books/title/?id=16868

Chen, X. (2012). *Social protest and contentious authoritarianism in china*. Cambridge University Press.

Chen, X. (2017). Origins of informal coercion in china. *Politics & Society*, *45*(1), 67-89. Retrieved from http://dx.doi.org/10.1177/0032329216681489 doi: 10.1177/0032329216681489

Dang, S. (2013). *Wangluo yulun hudie xiaoying yanjiu: Cong "weineirong" dao yulun fengbao*. People's University Press.

Davenport, C. (2007). State repression and political order. *Annu. Rev. Polit. Sci.*, *10*, 1–23.

Deng, Y., & O'Brien, K. J. (2013). Relational repression in china: using social ties to demobilize protesters. *The China Quarterly*, *215*, 533–552.

Dickson, B. (2016). *The dictator's dilemma: The chinese communist party's strategy for survival*. Oxford University Press.

Dimitrov, M. K. (2017). The political logic of media control in china. *Problems of Post-Communism*, *64*(3-4), 121–127.

Earl, J. (2003). Tanks, tear gas, and taxes: Toward a theory of movement repression. *Sociological Theory*, *21*(1), 44–68.

Frantz, E., & Kendall-Taylor, A. (2014). A dictator's toolkit: Understanding how co-optation affects repression in autocracies. *Journal of Peace Research*, *51*(3), 332–346.

Fu, D. (2017). Disguised collective action in china. *Comparative Political Studies*, *50*(4), 499–527.

Gao, H., & Zhang, Z. (2011). *Wangluo yuqing yu shehui wending.* Xinhua Press.

Han, R. (2018). *Contesting cyberspace in china: Online expression and authoritarian resilience.* Columbia University Press.

Heurlin, C. (2017). *Responsive authoritarianism in china.* Cambridge University Press.

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in china allows government criticism but silences collective expression. *American Political Science Review*, *107*(02), 326–343.

King, G., Pan, J., & Roberts, M. E. (2014). Reverse-engineering censorship in china: Randomized experimentation and participant observation. *Science*, *345*(6199), 1251722.

Lee, C. K., & Zhang, Y. (2013). The power of instability: unraveling the microfoundations of bargained authoritarianism in china. *American Journal of Sociology*, *118*(6), 1475–1508.

Lei, Y.-W., & Zhou, D. X. (2015). Contesting legality in authoritarian contexts: Food safety, rule of law and china's networked public sphere. *Law & Society Review*, *49*(3), 557–593.

Lorentzen, P. (2014). China's strategic censorship. *American Journal of Political Science*, *58*(2), 402–414. http://www.peterlorentzen.com/research_files/Lorentzen_censor_2014.pdf.

Lorentzen, P., Fravel, M. T., & Paine, J. (2017). Qualitative investigation of theoretical models: the value of process tracing. *Journal of Theoretical Politics*, *29*(3), 467-491. Retrieved from https://doi.org/10.1177/0951629816664420 doi: 10.1177/0951629816664420

Meissner, M. (2017, May). *China's social credit system.* Retrieved from https://www.merics.org/fileadmin/user_upload/downloads/China-Monitor/merics_ChinaMonitor_39_englisch_Web.pdf

Miller, B. (2016). *The limits of commercialized censorship in china.* (Unpublished paper draft)

Miller, B. (2018). *Reassessing the targets of china's online censorship apparatus.* (Unpublished paper draft)

Nathan, A. J. (2003). Authoritarian resilience. *Journal of Democracy*, *14*(1), 6–17.

O'Brien, K. J., & Deng, Y. (2015). Repression backfires: tactical radicalization and protest spectacle in rural china. *Journal of Contemporary China*, *24*(93), 457–470.

O'brien, K. J., & Li, L. (2006). *Rightful resistance in rural china.* Cambridge University Press Cambridge.

Ong, L. H. (2015). 'thugs-for-hire': State coercion and everyday repression in china.

Osnos, E. (2014). *Age of ambition: Chasing fortune, truth, and faith in the new china.* Macmillan.

Pils, E. (2014). *China's human rights lawyers: advocacy and resistance.* Routledge.

PRC State Council General Office. (2016). Guowuyuan bangongting guanyu zai zhengwu gongkai gongzuo zhong jin yibu zuohao zhengwu yuqing huiying de tongzhi. Retrieved from https://goo.gl/983hh8

Repnikova, M. (2017). *Media politics in china: Improvising power under authoritarianism.* Cambridge University Press.

Repnikova, M. (2018, November). *China's 'responsive' authoritarianism.* Retrieved from https://www.washingtonpost.com/news/theworldpost/wp/2018/11/27/china-authoritarian/?noredirect=on&utm_term=.749732649430

Roberts, M. E. (2018). *Censored: Distraction and diversion inside chinas great firewall.* Princeton University Press.

Schurmann, F. (1966). *Ideology and organization in communist china.* Univ of California Press.

Stockmann, D. (2013). *Media commercialization and authoritarian rule in china.* Cambridge University Press.

Su, Y., & He, X. (2010). Street as courtroom: state accommodation of labor protest in south china. *Law & Society Review*, *44*(1), 157–184.

Tager, J., Glenn Bass, K., & Lopez, S. (2017). *Forbidden feeds: Government controls on social media in china.* Pen America. Retrieved from https://pen.org/wp-content/uploads/2018/06/PEN-America_Forbidden-Feeds-report-6.6.18.pdf/

Tang, W. (2016). *Populist authoritarianism: Chinese political culture and regime sustainability.* Oxford University Press.

Tanner, M. S. (2004). China rethinks unrest. *The Washington Quarterly*, *27*(3), 137–156.

Wang, Y. (2016a, March). *The business of censorship: Documents show how weibo filters sensitive news in china.* Retrieved from https://web.archive.org/web/20161003182432/https://cpj.org/blog/2016/03/the-business-of-censorship-documents-show-how-weib.php

Wang, Y. (2016b, March). *Read and delete: How weibo's censors tackle dissent and free speech.* Retrieved from https://web.archive.org/web/20170515022534/https://cpj.org/blog/2016/03/read-and-delete-how-weibos-censors-tackle-dissent-.php

Wang, Y., & Minzner, C. (2015). The rise of the chinese security state. *The China Quarterly*, *222*, 339–359.

Wei, Y., Tian, X., Wang, H., Nie, G., Du, J., Feng, W., ... Deng, Y. (2015). *Gongan jiguan yuqing fenxi yu yulun yindao.* China Legal Publishing House.

Wintrobe, R. (2000). *The political economy of dictatorship.* Cambridge University Press.

Yang, Y., Yang, Y., & Fei, S. (2017, July). *China seeks glimpse of citizens' future with crime-predicting ai.* Retrieved from https://www.ft.com/content/5ec7093c-6e06-11e7-b9c7-15af748b60d0

Zeng, S. (2015). *Wangluo yuqing yingdui jiqiao.* Guangdong People's Publishing House.

Zhang, Z. (2017). *Reza hasmath and jennifer yj hsu, eds. ngo governance and management in china.* Springer.

Zhou, B. (2011). *Weibo wenzheng yu yuqing yingdui.* People's Publishing House.

Zhou, T. (2012). *Buduan kuochong shehui yulun de gongtong xinxi jichu.* Study Times.

Zou, C. L., Hongqiang, & Su, G. (2015). *Lingdao ganbu: Wangluo yuqing gongzuo zhinan work guide for public opinion.* People's Daily Press.