The Limits of Commercialized Censorship in China

Blake Miller[*]

September 27, 2018

Abstract

Despite massive investment in China's censorship program, internet platforms in China are rife with criticisms of the government and content that seeks to organize opposition to the ruling Communist Party. Past works have attributed this "openness" to deliberate government strategy or lack of capacity. Most, however, do not consider the role of private social media companies, to whom the state delegates information controls. I suggest that the apparent incompleteness of censorship is largely a result of principal-agent problems that arise due to misaligned incentives of government principals and private media company agents. Using a custom dataset of annotated leaked documents from a social media company, Sina Weibo, I find that 16% of directives from the government are disobeyed by Sina Weibo and that disobedience is driven by Sina's concerns about censoring more strictly than competitor Tencent. I also find that the fragmentation inherent in the Chinese political system exacerbates this principal agent problem. I demonstrate this by retrieving actual censored content from large databases of hundreds of millions of Sina Weibo posts and measuring the performance of Sina Weibo's censorship employees across a range of events. This paper contributes to our understanding of media control in China by uncovering how market competition can lead media companies to push back against state directives and increase space for counterhegemonic discourse.

---

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

Introduction

Why do scathing criticisms, allegations of government corruption, and content about collective action make it past the censors in China? Past works have theorized that regime strategies or state-society conflicts are the reason for incomplete censorship. While these factors likely contribute to incomplete censorship, I suggest that incomplete censorship results in part from delegation of censorship to private companies which creates a principal-agent problem. Censorship directives are passed through a tangled network of multiple government principals and are delegated to private social media corporations. Government principals and media corporation agents are driven by competing logics: the government logic of information control and the market logic of satisfying user demand for information, respectively.

Using a unique corpus of leaked documents from a social media company, Sina Weibo, I demonstrate that these conflicting logics are the cause of much of censorship's apparent incompleteness. I find that 16% of directives from the government are disobeyed by Sina Weibo and that disobedience is driven by Sina's concerns about censoring more strictly than competitor Tencent. I also find that the fragmentation inherent in the Chinese political system exacerbates this principal agent problem. Fragmentation results in decentralization of censorship enforcement, competition between government agencies over censorship objectives, and non-uniform distribution of regulatory leverage across the many government agencies in charge of delegating censorship and sanctioning social media companies for non-compliance.

This paper contributes to our understanding of media control because it shows that market competition impacts information control outcomes, breaking with a large body of works on information control that assume market competition doesn't matter. This paper complements the work of Yang (2013) by showing that market concerns open space for contention in China, but it also emphasizes that market competition is the main driving force behind this opening. Internet businesses push back on government controls in an effort to appear "more free" than competitors, leading to expanded space for contention.

1 The Southern Weekend Incident

Each year, the liberal Chinese newspaper Southern Weekend writes a New Years editorial on a theme they would like to characterize the coming year. In January of 2013, they chose the theme "China Dream, Constitutional Dream," stressing the need for progress in the coming year on strengthening the rule of law and protecting rights enshrined in the constitution. After the editorial team submitted their final draft to the Guangzhou Propaganda Department and received no edits, they assumed everything had been approved for publication. The editors and Southern Weekend's readers were surprised when they opened their newspapers and found a fawning paean to the Party in place of the expected boundary-pushing editorial that had become the newspaper's trademark. The article boasted, "we are closer to the Chinese dream than ever before" among other platitudes.

This clumsy reworking of the editorial resulted in a strike by Southern Weekend editorial staff and sizable student protests in major Chinese cities. The strike and protests gained momentum through coordination and discussion on major Chinese social networking sites such as Sina Weibo[1], despite four strongly worded directives from the Central Propaganda Department and top leadership to cease all such discussion. This apparent incapacity to control perhaps the most threatening type of information online—information with potential to fuel student-led collective action[2]—may come as a surprise to many. The Chinese state is often represented in press and academic writing as a monolith[3], with high capacity to control information flows using the many advanced methods of censorship at its disposal. However, in many circumstances, it seems that intense pressure from the

---

[1]According to monthly active user statistics, throughout most of the log data, Sina Weibo was the second largest social network in China, behind Qzone. WeChat, China's most popular social platform is a messaging app that is similar to What'sApp, but with several social, payment, and service features tacked on. Though WeChat's monthly active users surpassed Sina Weibo in Q1 of 2012, it is not a competitor with Sina Weibo in the same way as Tencent Weibo is; both have a similar microblog platform, ostensibly inspired by Twitter. Because Tencent Weibo has been in beta for several years, Tencent does not report MAU numbers in its annual financial reports, however, most measures of active users during the period in which the logs were created put Sina Weibo comfortably in the lead.

[2]The Chinese Communist Party has experienced many student-led movements that have presented clear threats to its grip on power. Students initialized and sustained the Great Proletarian Cultural Revolution, a movement resulting in the dismantling of state and party institutions through arbitrary mass violence. Student—led movements on two separate occasions—in 1976 and in 1989—sparked mass protests which at the time seemed capable of threatening the CCP's monopoly on power.

[3]Much work on censorship assumes that either the government is a unitary actor or that the central government is the main enforcer of censorship. Formal literature, often for the sake of parsimony, defines "the government" or "the autocrat" as the singular actor and practitioner of censorship (Lorentzen, 2014; Guriev & Treisman, 2015; Gehlbach & Sonin, 2014; Chen & Xu, 2016; Egorov, Guriev, & Sonin, 2009) (King, Pan, & Roberts, 2013) draw inferences from censorship outcomes to measure a singular intention of a single government actor: to allow criticism but censor content with "collective action potential."

highest levels of the party and government fails to move social media companies like Sina Weibo to act.

2   Delegation, Fragmentation, and Agency Loss

During the Southern Weekend editorial incident, the Propaganda Department of Guangzhou[4] and the Central Propaganda Department[5] were involved in attempts to control the spread of information on the event. Despite these urgent attempts to censor all mention of this incident, a novel leaked dataset from 2011-2014 documents that popular social media company Sina Weibo willfully ignored directives to remove content related to the incident on their platform. Sina's calculus was clear. By providing more information about the strike and protests, they could attract information-seeking users away from chief competitor Tencent.[6] One log of company decisions instructs employees to "not be stricter than Tencent," and to hold off on implementing government directives until "urged to block content a second time." When told to delete users, Sina instructed employees to block users temporarily and unblock them the following day "as soon as you receive instructions."[7] The case of the Southern Weekend editorial incident illustrates how government fragmentation and delegation of censorship to private corporations can result in incomplete censorship outcomes. In this case, the Central Propaganda Department, the Guangzhou Propaganda Department, and the Beijing Municipal Government agencies directly involved in managing Sina Weibo had different objectives and were unable to adequately monitor and sanction the company. By delegating censorship to Sina Weibo, concerns over user retention and competitiveness became a factor in the company's decision to comply with

---

[4]Tuo Zhen, the head of the Guangdong Propaganda Department at the time was concerned about limiting the fallout in response to the editorial incident, especially since he was being blamed by both the public and the Central Government. See this archived analysis for more information: https://web.archive.org/web/20180730135458/http://chinamediaproject.org/2013/01/07/inside-the-southern-weekly-incident/

[5]The original directives can be found archived at the following links:
http://www.webcitation.org/71yPYfpyj, http://www.webcitation.org/71yPahJIQ,
and http://www.webcitation.org/71yPeE80B

[6]Tencent is the largest ICP company in China. It owns WeChat, Tencent Weibo, and QZone, three of China's most popular social platforms. It is Sina's most direct competitor.

[7]Full text of log from 1/5/2013: "There is a lot of related content on Tencent. After we reported [their lack of implementation] to the Network Management Office, Tencent implemented [censorship of the content]. Currently [Sina] Weibo is partially carrying out instructions to block content. First prevent retweets on content flagged [by the Network Management Office]. When urged to block content a second time, fully implement directives. With respect to banning users, for the meantime do not implement [user bans]... handle users relative to Tencent's [level of] implementation. We should not be stricter than Tencent. Today maintain user blocks, tomorrow as soon as you receive instructions release the block."

directives. By Sina's calculations, the cost of flouting government directives outweighed the benefits of increased user engagement. In this section I outline this theory in detail. In subsequent sections, I present evidence in support of this theory from censorship logs like the one mentioned above and large databases that measure censorship outcomes on the Sina Weibo platform.

2.1 Corporate Delegation Leads to Agency Loss

Private companies play a crucial role in the process of censorship, as they bear the ultimate responsibility for removing content from their platforms. Despite their central role, private companies are too often missing from models of information manipulation. In China, censorship is delegated and regulated through internet content providers (ICP) licenses. ICP licenses are necessary to operate an internet business in China. These licenses can be revoked or suspended if ICPs do not comply with government directives.

Sina Weibo and many ICPs like it have had to balance user and shareholder demand with regulatory pressures from government agencies since they were first founded. In 2014, when Sina Weibo became listed on the NASDAQ stock exchange, it made these concerns explicit. In its regulatory filing with the SEC, it included "regulation and censorship of information disseminated over the internet in China" as a major risk that could affect its share price. In the filing, Sina Weibo also noted that censorship "may adversely affect our user experience and reduce users' engagement and activities on our platform as well as adversely affect our ability to attract new users to our platform."[8] Since Sina's listing on NASDAQ the company has been fined several times by Chinese government agencies for failing to meet censorship regulations. Sina's user engagement and activity has declined in recent years, in part due to the rising popularity of WeChat, a chat-based social networking site like WhatsApp. Some analysts, however, have attributed this decline to increasing perceptions of Sina Weibo as a heavily-censored platform.[9]

I argue that "incomplete" censorship is largely a result of a clash between ICPs responding to popular demand for information and government actors responding to pressures to maintain social stability and protect their position within government. The process of delegating to private internet companies creates a principal-agent problem.

---

[8]Full article accessible here: https://web.archive.org/web/20180812134506/
[9]See: http://www.webcitation.org/72Kwc4BJ7

Government principals delegate censorship to private internet companies through directives: verbal or written instructions providing details about how these private internet companies are supposed to handle certain kinds of objectionable content. Private internet companies then decide if and how they will comply. Because there is often a misalignment of preferences between the government principal and the ICP agent, private internet companies may ignore the directive or partially implement it. Because of this disobedience, the delegating principal suffers agency loss: agents acting against their principals' interests.

Competition between private internet companies further exacerbates this problem. In China, each government principal delegates censorship to several private internet companies, meaning that there are multiple agents involved in China's system of censorship. Rundlett & Svolik (2016) have shown that when a principal has multiple agents, principals suffer agency loss. Private internet companies benefit from having compelling information on their platforms. However, compelling content may also be considered "harmful information" to the authorities. Highly motivated users who are seeking this information often hop from platform to platform when content is censored (Roberts, 2017). By censoring less, private internet companies can attract these users to their platform and away from their competitors. This competition over users results in a race to the bottom as each company strategically tries to skirt directives more than competitors. Companies seek to jointly minimize the cost of censorship (lost user retention and engagement) and the cost of non-compliance with government directives.

Private internet companies deliberately shirk, lack capacity to implement directives, or some combination of the two. Private internet companies can take advantage of hidden information about their technical capabilities and budgets to invest as little as possible in developing high-performance censorship systems. They can also shirk, as their effort censoring content is not easy to measure or observe. Private internet companies take advantage of hidden actions and hidden information to comply with directives only when doing so will benefit the company's bottom line.

2.2 Bureaucratic Fragmentation Leads to Agency Loss

Fragmentation of China's political system further complicates censorship delegation. While on paper the Chinese political system is rigidly hierarchical, in practice, contestation be-

tween state and party organs, and bureaucracies with overlapping policy domains is common throughout the system. The findings of this analysis confirm earlier suppositions about the fragmentation inherent in the Chinese political system.[10] This fragmented system, which Oksenberg & Lieberthal (1988) coined "fragmented authoritarianism" is characterized by de facto veto power of local governments when implementing central policies and the tangled lines of authority in China's vast bureaucracy. I argue that incomplete censorship is in part due to the fragmentation of China's political system. This fragmentation leads to agency loss—disobedience of, or incapacity to implement government directives—due to two major attributes of China's system of censorship: common agency and local bias.

First, common agency[11] of private internet companies gives them discretion about which agency's directives to follow. Many bureaucracies in China have the authority to delegate censorship to private internet companies. This makes monitoring and sanctioning of private internet companies difficult because each individual bureaucracy must rely on their own limited resources in order to monitor compliance after they have issued their directive. Private internet companies receive directives from multiple principals (sometimes referred to as "common agency"), so they often have discretion about which directives to follow, especially with principals do not have uniform preferences (Calvert, McCubbins, & Weingast, 1989), as is often the case in China. Because of this, it is not uncommon to observe inconsistencies in the way content is censored across private internet companies due to this discretion.[12]

Second, the power to enforce directives is locally biased, i.e. power is concentrated in the locality of an ICP's headquarters. Local agencies where media companies are headquartered have more control over what gets censored as they have access to the most proximate and effective levers of power to enforce compliance with directives. This local bias was clear in interviews conducted in Cairns (2016) that suggested that the "[Bei-

---

[10]See O'Brien (1994); Montinola, Qian, & Weingast (1995); Lieberthal (1995); Jin, Qian, & Weingast (2005); Zheng (2007); Stern & O'Brien (2012); Mertha (2009)

[11]Many government agencies have the power to issue directives to single private internet companies.

[12]For example, there is significant correlation among censored keyword lists in Chinese game chatrooms when they are created by the same parent company or developer, but very weak correlation among lists within the same Chinese provincial or city jurisdiction (Knockel, Ruan, & Crete-Nishihata, 2017). Previous work has found inconsistencies in the implementation of censorship across platforms and companies operating in China, including search engines (Villeneuve, 2008), blogging services (MacKinnon, 2009), chat apps (Crandall et al., 2013), live streaming (Crete-Nishihata et al., 2016), and mobile games (Knockel et al., 2017), which suggests companies have flexibility and discretion when interpreting and implementing censorship directives.

jing Municipal Government's] de facto political power as regulator of Sina Weibo" caused "cross-jurisdictional conflicts." When authorities outside of Beijing wanted content censored, they had to lobby Beijing-based bureaucracies. This local bias leads to agency loss because the central government and local governments outside of the ICP's jurisdiction do not have the same implements of enforcement and are easier to ignore. Because private internet companies are constantly weighing the cost of non-enforcement and expected revenue from satisfying user demand, private internet companies will pay outsized attention to local directives, despite being national platforms. Without a centralization of the power to issue and enforce directives, private internet companies can selectively ignore directives from outside of their jurisdiction.

The complexities of information control in China's fragmented system have been well-studied as they relate to traditional media. In a study of newspapers, Stockmann (2013) finds that during the reform of China's media system, fragmentation posed a challenge to state monitoring and sanctioning of commercialized newspapers. She argues that factional affiliations, rank of a media company's sponsoring agency, and geographical jurisdiction defined a "discursive space," giving media companies greater freedom to report more critically about opposing factions, lower ranked agencies, or other geographical regions. Mertha (2009) finds that policy entrepreneurs make use of the media to lobby and appeal to various fragmented interests across China's bureaucracy. Others have noted the many conflicts between central and local governments, party and state organs, propaganda departments in different localities, and media organizations and regulators (Brady, 2009; Lynch, 1999; Shambaugh, 2007).

3 Data and Methods

To test the above theoretical claims, that fragmentation and corporate delegation result in outcomes that deviate from the intentions of delegating government principals, researchers face a number of challenges. Despite the vital role private internet companies play in China's system of censorship, there has been a dearth of data on them and very little scholarship dedicated to them. Much of the available data used to study Chinese censorship consists of only content and censorship outcomes. These data do not provide any information about the interactions between individuals in government and

between private internet companies and government actors that determine what is and is not censored. Due to limitations of available data, assumptions that censorship outcomes can serve as a measure of government intent are common, despite the central role non-government actors play in the process.[13]

3.1 Log Data

To address these shortcomings, I have created a custom dataset of censorship logs—notes taken in the process of censorship at Sina Weibo, one of China's most popular social networking sites. This dataset is the first of its kind to capture the entire process of Chinese censorship, from a government directive to a private internet company, to that private internet company's decision on how or whether to comply. With the help of research assistants, I have coded these logs by content, the bureaucracy issuing directives to Sina Weibo, and whether or not Sina Weibo implemented these directives. Log data adds empirical support to the theoretical claims detailed earlier in the following ways:

First, because these logs measure disobedience, they can provide insights into the relative power of bureaucracies to delegate censorship. By examining the variance in Sina's rates of compliance with directives across different bureaucracies, we can observe how fragmentation leads to agency loss. If fragmentation results in agency loss, rates of compliance with directives should be correlated with a bureaucracy's power to monitor and sanction Sina Weibo.

Second, logs include direct and indirect information about the reasons behind non-compliance with directives. Occasionally managers explain their decisions to comply or disobey government directives. If corporate delegation leads to agency loss due to concerns about competitors and the adverse impact of censorship on user experience, Sina Weibo managers should express these concerns in the logs. Other less direct evidence of profit motivations can be observed in how thoroughly Sina censors content that is harmful to its business interests and content that is harmful to government interests. By comparing the censorship rates of Sina-related news events and censorship on behalf of a government actor, these logs can add support to claims that Sina Weibo factors market concerns into its censorship efforts. Comparing censorship rates of content that is harmful to

---

[13]State media describes the process of delegation to private actors in detail in a report here: http://www.webcitation.org/72ZsG5mc7

Sina's business interests and content that is harmful to government interests can also demonstrate that agency loss is not simply due to a lack of capacity, but rather due to deliberate and strategic shirking.

3.2 Event Data

Users seek information at different rates depending on the type of event, and more user attention is usually more threatening to the government. Because Sina is concerned with audience demand, they may increase or decrease their censorship efforts in response to audience interest and, by proxy, sensitivity. To see how censorship implementation varied by audience demand for information for sensitive events, I used data from annual "blue books," policy briefs written for government cadres, on "social opinion and emergency management" compiled by the Institute for Public Opinion Research of Shanghai Jiao Tong University and published by the Chinese Academy of Social Sciences. These blue books collate reports written by academics, government officials, and policy experts and offer suggestions for policy changes and improvements in the coming year. Using these events as a sample ensures that I am analyzing events that public opinion experts found noteworthy, and where "public opinion supervision" (舆论监督) and/or "public opinion guidance" (舆论引导) was potentially necessary. These reports include 300 events (60 events for each year from 2010-2014) ranked by an index of search volume across several services. Each year's cases are selected from the top 1200 events with the highest search volume index[14] based on their designation as a "public opinion emergency" that would be salient to opinion and thought workers, the intended audience of these "blue books."

3.3 Censorship Outcomes Data

For each of the events above, I measure how thoroughly Sina carried out censorship on its platform by searching large datasets of historical Sina Weibo posts for the exact content targeted for censorship in logs. For this analysis I use two datasets that measure censorship on the Sina Weibo platform: Free Weibo and Weiboscope. The Free Weibo dataset includes 47 million weibo posts spanning from 2009-2018 and the Weiboscope

---

[14]The search volume index is calculated as the average of news search volume, social networking platform search volume, blog search volume, Sina Weibo search volume, video search volume, and WeChat Public Accounts search volume.

dataset includes 226 million posts from only 2012.[15] Using these data I compare the instructions in the logs to actual censorship outcomes at the event level.

To compare log content to actual Weibo post content, I first had research assistants manually identify logs relevant to each of the 320 events from "blue books" by searching within a month window on either side of the event date. For each relevant log, research assistants then extracted the full content text from logs, stripping away instructions on censorship and government directives that usually go along with log text. After content from relevant logs was extracted, I searched large databases of censorship outcomes for exact or near-exact text matches to each individual log's content.

Running several searches of 273 million Weibo posts is not a trivial task. To identify exact or near-exact matches to posts mentioned in the logs, I needed to build a rudimentary search engine, building an inverted index of each of the Weibo posts in these large databases. For each log, I then scored the relevance of each post in the two large databases of Weibo Posts using the Okapi BM25 ranking algorithm. The Okapi BM25 algorithm was the gold standard for search engine ranking prior to recent advances in deep learning and was the core technology behind the Bing search engine for several years. With each post in these databases ranked according to relevance to the text extracted from the logs, research assistants then identified posts from each set of search results that exactly or almost exactly matched the post targeted in logs. Results were considered near-exact matches if they contained the entire query text with either small editions, or a a few words replaced with synonyms, homophones, or homoglyphs. Using the final dataset of retrieved matches, I estimated the proportion of posts that were actually censored by Sina Weibo for each event. The results of this procedure are visualized in Figure 1.

---

[15]Both datasets are constructed by recording a post soon after it has been created and later querying for that post at fixed intervals to see if the post has been censored or deleted.

4 Empirical Implications and Results

Figure 1: Censorship Rate of Retrieved Weibo Posts Mentioned in Logs
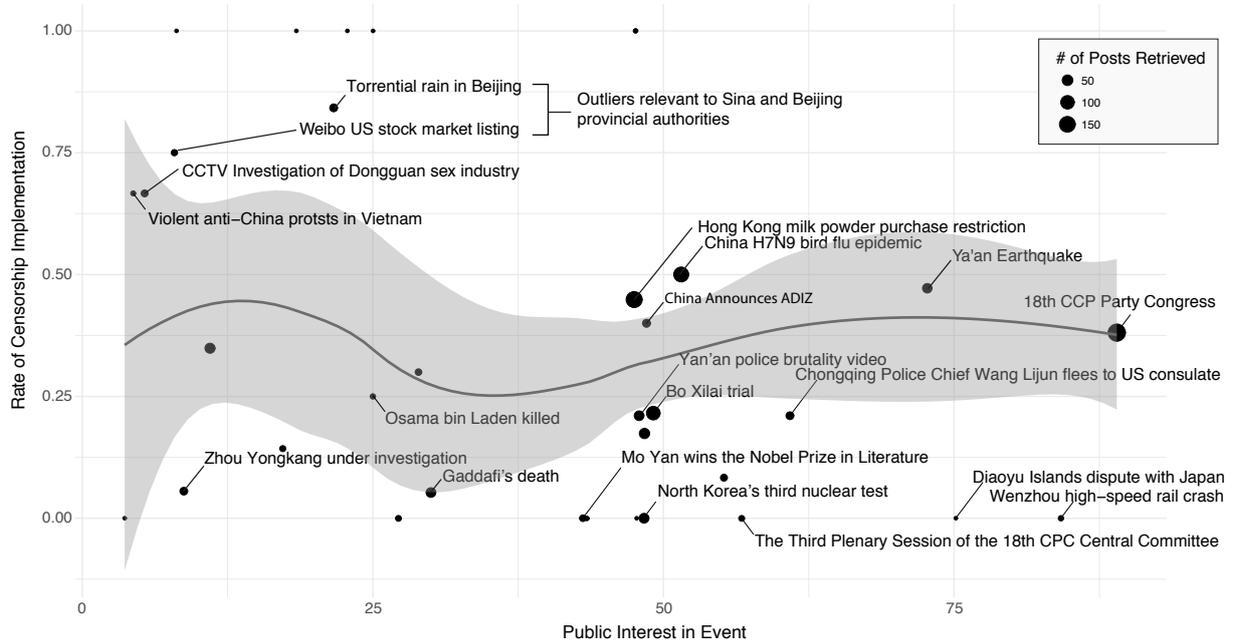
Each point on the figure represents a event where posts were retrieved and its size is proportional to how many posts were retrieved. The y-axis, "rate of censorship implementation" is the number of retrieved posts that were censored over the total number of posts retrieved. The x-axis, "public interest in the event" is the search volume index gathered from blue books. The regression line and confidence intervals are from a weighted LOESS model. The average rate of censorship implementation is .38.

4.1 Corporate Delegation Leads to Agency Loss

I argue that incomplete censorship is due in part to a clash between the market logic driving the behavior of private companies and the logic of control that guides governments. In other terms, market demand for information sometimes clashes with the informational preferences of government actors. In this section, I argue that corporate delegation leads to agency loss for three main reasons. First, market competition incentivizes Sina to prefer to censor less than its competitors, leading to a race to the bottom. Second, Sina Weibo is technically limited in its ability to comply with government directives, resulting in agency loss. Third, Sina Weibo deliberately shirks in order to minimize the cost of censorship and to minimize impact of censorship on the usability of its platform.

The logs provide many examples of Sina Weibo's concerns about censoring more than

competitor, Tencent Weibo. In one instance Sina Weibo was instructed to delete all content related to a murder case that had garnered much public interest. According to logs, Sina Weibo monitored the compliance of Tencent Weibo, alerted the Supervision Department to their non-compliance, and delayed implementation of directives until they could be assured that Tencent had also complied with directives.[16] In another example, the Network Management Office ordered Sina Weibo to remove a popular account on their service. After seeing that Tencent had not complied with removing the same user on their platform, Sina Weibo drafted a response to the Network Management Office: "Since Tencent hasn't deleted the account, we are unable to delete the account at this point in time." In several other cases, Sina waited until Tencent complied to implement directives to delete user accounts. During several sensitive events such as the annual Spring Festival Gala—a Chinese propaganda variety show and most-watched television show on the planet[17]—Sina Weibo instructed employees to avoid censoring more comments on the official television broadcaster's account than competitor Tencent. Sina also went to great lengths to limit positive news about Tencent on its platform. When news broke about a mass purge of bots on Tencent, Sina Weibo employees were instructed to prevent retweets of posts that praised Tencent's actions. Similarly, when Xi Jinping visited Tencent, Sina Weibo employees were instructed to prevent retweets of all related posts. Conservatively, these patterns in logs suggest that Sina Weibo's concerns about remaining competitive with Tencent drove many of their decisions to comply with directives. Preferences to censor slightly less than competitors, and the low level of implementation of directives is consistent with a race to the bottom.

Though we observe Sina deliberately disobeying directives in logs, in many instances, it seems that Sina's poor censorship performance is due to a lack of capacity. In Sina's IPO documents they claimed, "although we attempt to monitor the content posted by users on our platform, we are not able to effectively control or restrict content generated or placed on our platform by our users." This statement appears to be at least partially

---

[16]Log entry from December 26, 2011 reads: Supervision Department demands to eliminate related content to the Henan Shenqiu murder case that caused the death of four children and injury of one child. Currently, we told Supervision Department that there are many related posts on Tencent Weibo and asked Tan Chao to negotiate. They have not yet responded. Previously, we censored any searches of the incident but didn't eliminate everything. Currently, the parameter is to process anything attacking the party or government policy. If there are other situations, report immediately.

[17]Viewership of the New Year's Gala is around 700 million. In 2012, Guinness World Records gave the show a audience of 498.7 million and named it the, "Most Watched National Network TV Broadcast." See http://www.webcitation.org/71y5Jcwx5.

accurate. In some logs, Sina Weibo appears overwhelmed. During a major collective action incident in 2012 involving mass protest and several self-immolations in Tibet, Sina Weibo struggled to keep up with a large magnitude of takedown requests about Tibet. At one point, Sina Weibo employees sent an SOS to all department heads as they struggled to mobilize enough employees to meet increasing censorship demands.[18] Even in less dire circumstances, Sina Weibo appears to perform relatively poorly. In Figure 1, we see that Sina Weibo usually does not perfectly follow through with the decisions to comply with censorship directives. The average rate of censorship of Sina Weibo posts from the Free Weibo and Weiboscope databases that exactly match posts in the logs is .38. This low rate of censorship implementation may be for one of two reasons: unintentional low capacity, or deliberate shirking.

To test whether Sina's poor performance was exclusively due to low censorship capacity, I identified all logs that referenced news events about Sina Weibo that were potentially damaging to the company's reputation or bottom line. All posts mentioning Sina had already been labeled by research assistants, making the search easier. In total I identified three such news events in the logs: 1) discussions of Sina's IPO, 2) discussions of new research measuring the speed of Sina's censorship, and 3) Reuters interviews with former Sina Weibo employees who worked as content censors. To measure how well Sina followed through on these decisions to censor content, I retrieved exact or near-exact matches of content mentioned in the logs from the Free Weibo database using methods described in detail above. In total I retrieved 33 posts, 24 of which were censored, an implementation rate of .73. This rate of implementation is significantly higher than the Blue Book rate mentioned earlier (.38) according to a 2-sample chi-squared test of equality of proportions at the 95% confidence level. This evidence suggests that Sina Weibo has the capability to more thoroughly censor content in response to government directives, but chooses not to. I find that the low rate of actual censorship is likely at least somewhat deliberate.

These quantitative measures are consistent with Sina's general instructions about how intensely to censor content. Employees are at times instructed to deliberately shirk and to obstruct the process of censorship by delaying directives and prolonging bargaining

---

[18]Log entry from February 1, 2012 reads: Recently there have been a lot of demands to delete posts about Tibet!! There is no way we can add this many banned keywords; if we do so, the rate of posts needing investigating will increase too much. We have already added several keywords related to Tibetan independence, Communist Party [policy in Tibet], [police] killings, but the number of keywords keeps increasing... I don't know what to say!! Every department head, please disseminate! This is extremely urgent!!!

between delegating government agencies and Sina. In the days preceding the Southern Weekend editorial incident mentioned in the beginning of this paper, log documents begin with a general notice urged employees to "negotiate as much as possible" and "defer implementation of censorship requests from the Supervision Department and Internet Management Office," Sina Weibo's two chief regulators. The notice went on to instruct employees to "not process too many user posts" and to "not be too stringent." Just three months earlier, a similar general notice had suggested the opposite. Employees were urged to "tighten" their "control measures" and to "resolutely eliminate all posts relating to negative incidents, rumors about the Politburo Standing Committee and their families, collective action content, coups, power struggles related to the 18th CCP Party Congress, Ling Jihua, Central Public Security Bureau, Bo Xilai, etc." But even this strongly worded notice instructed employees to refrain from immediately implementing "especially unreasonable directives."

Finally, as is the case in any workplace, part of Sina Weibo's lack of capacity to censor content comes from agency loss due to rogue or incompetent employees. Of course, the individual who leaked this entire cache of documents to the press was a Sina Weibo employee.

4.2   Bureaucratic Fragmentation Leads to Agency Loss

Leaked censorship logs from Sina Weibo do not depict a system of censorship that is hierarchical, centralized and efficient. Rather, many bureaucracies appear to have varying degrees of authority, do not appear to coordinate their censorship directives, and appear to at times be in conflict with one another. Logs document 3 main bureaucracies in charge of censorship at Sina Weibo, and a long tail of other bureaucracies with the power to issue directives to the company (see Figure 2). To measure the relationship between fragmentation and agency loss, I calculate rates of disobedience across multiple bureaucracies, and confirm that these patterns match theoretical expectations. I then argue for two mechanisms behind this relationship. First, common agency of private internet companies gives them discretion about which agency's directives to follow. Second, the power to enforce directives is locally biased, i.e. power is concentrated in the locality of an ICP's headquarters.

Figure 2: Agencies and Individuals Influencing Censorship on Sina Weibo
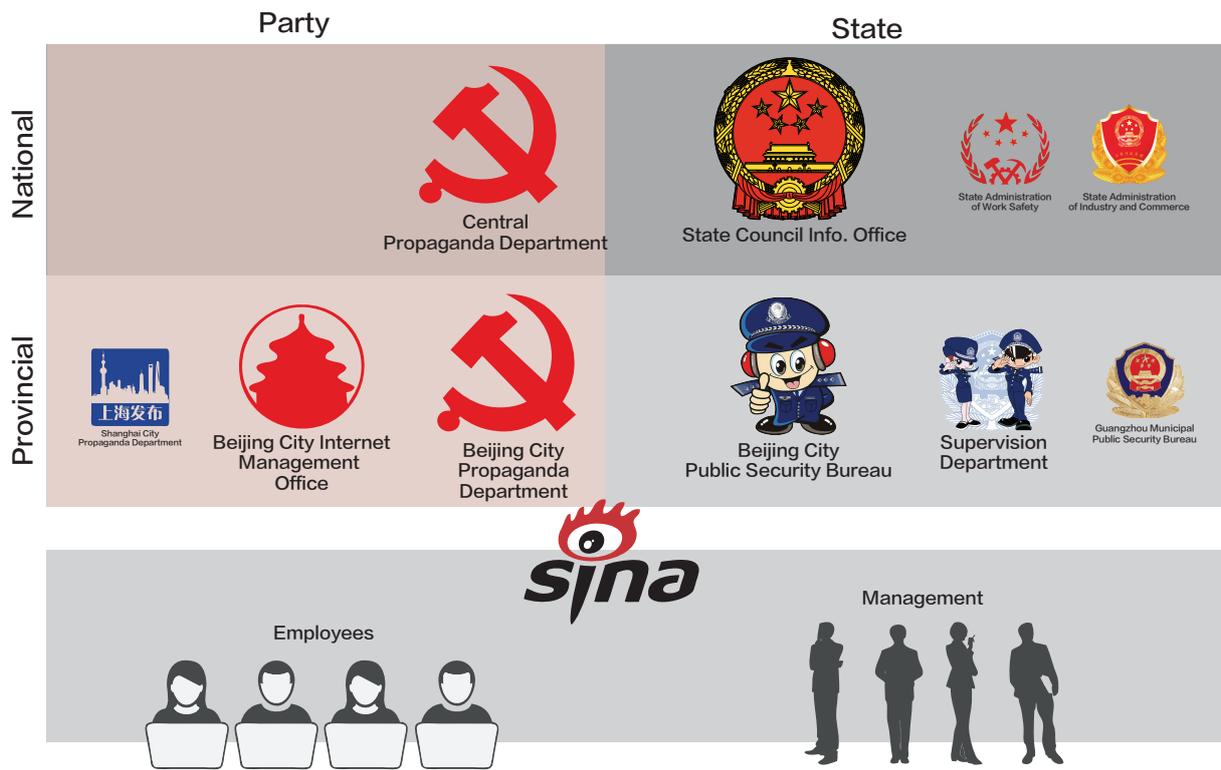
If bureaucratic fragmentation leads to agency loss, we should expect to see differences in the responsiveness of Sina Weibo to certain bureaucracies. Three bureaucracies are most commonly called: The Beijing Municipal Internet Propaganda Management Office (Internet Management Office), the Public Information and Internet Safety Supervision Department of the Beijing Public Security Bureau (Supervision Department), and the State Council Information Office (SCIO).[19] Personnel and budgets of the Internet Management Office and the Supervision Department are controlled by the Beijing Municipal Government. In a handful of cases, the State Council Information Office delegates directly to Sina Weibo, but most of the time it issues instructions to the Supervision Department and Internet Management Office to delegate to private internet companies. Though it has informal authority to delegate to these agencies, it does not have leverage over budgets and personnel. Because Beijing Municipal Government agencies are directly responsible for monitoring and sanctioning Sina, they should be obeyed at higher rates than the State Council Information Office. Beijing Municipal Government has strong incentives to keep

---

[19]The logs confirm much of the basic bureaucratic structures that monitor and direct censorship at Sina Weibo as identified in Cairns (2016).

Sina in check, as they will ultimately bear responsibility for failure to manage Sina when they are evaluated for promotion at the end of their terms. In the logs, I find that overall, 16% of all directives are disobeyed. Beijing municipal bureaucracies, the "Network Management Office" and the "Supervision Department" are disobeyed at rates of 15% and 17% respectively. By contrast, the SCIO is disobeyed 20% of the time. Though these measures of disobedience do indicate that the SCIO is disobeyed more often than Beijing bureaucracies, there are not enough SCIO directives to distinguish a statistically significant difference in the two proportions. To address this limitation, I examine the data for evidence for two theorized mechanisms linking bureaucratic fragmentation to agency loss: common agency and local bias.

If common agency of private internet companies results in agency loss, we might expect to see instances in the data where agencies delegating censorship send different directives to Sina Weibo, and where Sina Weibo chooses to implement the more lenient or less-specified directive. This is a very hard test to pass because the logs are usually not detailed enough to show discrepancies between directives from two agencies. Directives are usually not copied verbatim and are summarized, often in shorthand. Despite this, there are a handful of logs where we observe behavior consistent with common agency leading to agency loss. One such log involves large-scale protests in the city of Shifang in July of 2012 over a copper plant local residents believed was causing health problems. The Supervision Department, a Beijing municipal state public security organ directed Sina Weibo to remove all collective action content from their site. At the same time, the Internet Management Office, a Beijing municipal party propaganda organ directed Sina to remove a list of specific posts.[20] The latter order was easier to implement and was unlikely to completely shut down discussion of the event on Sina Weibo. Sina opted to ignore the first order and implement the latter order.

This evidence has its limitations. These handful of accounts confirm that common agency resulted in some agency loss, but it is impossible with these data to determine the magnitude of this agency loss. Previous work on delegation, however, has shown that in situations where there are multiple principals, agent discretion results in higher agency loss than if there was only one delegating principal (Calvert et al., 1989).

---

[20]A log from July 3, 2012 reads: "Regarding the Shifang incident, the Supervision Department is currently demanding that we eliminate any inflammatory and mobilizing content. The Internet Management Office has no clear directives but has sent over a list of individual posts to process, ordinary requests like these should be followed."

Table 1: Distribution of Bureaucracies Issuing Directives

Name	Administrative Rank	Beijing	Directive Count	Percent of Directives
Beijing Municipal Internet Management Office	Provincial-level Municipality	Yes	310	50.65
Beijing Municipal Supervision Department	Provincial-level Municipality	Yes	269	43.95
Beijing Municipal Internet Police	Provincial-level Municipality	Yes	8	1.31
Beijing Municipal Bureau of Radio and Television	Provincial-level Municipality	Yes	3	0.49
Shanghai Municipal Propaganda Department	Provincial-level Municipality	No	1	0.16
Guangzhou Municipal Supervision Department	Provincial-level Municipality	No	1	0.16
State Council Information Office	National	No	16	2.61
Ministry of Public Security of the Central People's Government	National	No	2	0.33
Central Military Commission of the People's Liberation Army	National	No	1	0.16
State Administration for Industry and Commerce of the People's Republic of China	National	No	1	0.16

If local bias results in agency loss, we can expect to observe two main things from the log data. First, it should be uncommon for non-Beijing Municipal Government agencies to appear in directives because non-Beijing bureaucracies—especially non-Beijing local governments—should have very little power to sanction Sina for non-compliance with directives. In the case of Sina, the company falls under the jurisdiction of the provincial-level municipal government of Beijing, as Sina is headquartered in Beijing. To measure the distribution of directives sources, two research assistants coded the source bureaucracy for all 8,427 unique censorship logs according to specific instructions about what constituted a bureaucratic source. I manually checked all of their labels and searched the database for any logs they may have missed. The resulting distribution of bureaucracies can be seen in Table 1. Overall, 96.6 percent of directives come from Beijing Municipal bureaucracies.

While other provincial-level bureaucracies can send directives to Sina, only two out of 611 specified directive sources in the logs are non-Beijing Municipal bureaucracies.[21]

If local bias results in agency loss, we should also see greater responsiveness to the demands of the Beijing Municipal Government than other bureaucracies, particularly the Beijing Municipal Public Security Bureau and the Beijing Municipal Propaganda Department. To test this, research assistants coded all instances in the logs where Sina Weibo employees were directed defend government Weibo accounts, monitoring and deleting comments that were offensive to that particular agency. Of the accounts Sina Weibo protected, 46% were Beijing Municipal Government accounts, 46% were national-level government accounts, and 8% were non-Beijing provincial-level government accounts. As Sina Weibo users are not overwhelmingly concentrated in Beijing, this suggests that decisions to censor content are locally biased.

Sina also appears to censor more efficiently when an event is salient to the Beijing Municipal Government. In the aftermath of a flash flood in Beijing, citizens tried to organize vigils for victims of the flood. Many logs related to this event indicate the Supervision Department and the Internet Management Office's keen interest in thoroughly removing this mobilizing content from Sina Weibo.[22] There was not much interest in this particular event according to the blue book measures (in blue books, the degree of interest in events is a proxy for their politial sensitivity to the national government), and it appears to be an outlier in Figure 1. Censorship of this event was much more efficiently implemented than other blue book events at all levels of interest in the event (see Figure 1). This is consistent with the theory that local bias due to fragmentation leads to agency loss in the process of censorship delegation.

5 Conclusion

Leaked censorship logs from Sina Weibo provide an intimate look into the conflicting informational preferences of the Chinese government and private internet companies. They depict a system of multiple principals and multiple agents and a tangled web of competing

---

[21]One log includes a directive from the Shanghai Municipal Propaganda Department and another from the Guangzhou Municipal Public Security Bureau.

[22]A log from July 28, 2012 reads: "The Supervision Department requested we ramp up handling and elimination of posts inciting and mobilizing netizens to hold vigils for victims of the Beijing torrential rain disaster. A mobilizing post was discovered tonight and reported to the Supervision Department and Internet Management Office."

informational objectives. The logs show that the outcome of censorship involves many actors and does not necessarily reflect a unified government strategy. Rather, government fragmentation and delegation to several corporate actors results in a system where the end result of censorship is generated by the aggregated and contested preferences of central leadership, subnational governments, and subnational elites passed through a final layer of distortion: media corporations. The logs document outright disobedience of directives, even in highly sensitive situations, and show how Sina Weibo strategically disobeys directives in order to gain a edge over competitor Tencent Weibo. While much of the academic literature and media depicts China's censorship apparatus as swift, centralized, focused, and sophisticated, the system is often slow, fragmented, contentious, and low-tech, making censorship orders difficult to enforce and giving social media companies a great deal of discretion over what citizens do and do not see. Delegated censorship to private companies results in significant agency loss that is then further compounded by political fragmentation.

6 Developments Since 2014

In the last few months of log data, China's information control institutions underwent significant reforms. These reforms seemed aimed at addressing the problem of fragmentation of China's system of information control. In 2014, the Cyberspace Administration of China (CAC) assumed its role as China's chief regulator of cyberspace. The CAC is a joint party and state organ that houses the SIIO (which is a continuation of the SCIO[23]), and the General Office of the Central Leading Group for Internet Security and Informatization which reports directly to the Central Committee of the Chinese Communist Party. The creation of the CAC gave party and state organs unambiguous authority over provincial and municipal bureaucracies regulating private internet companies such as Sina Weibo.[24] It is unclear from available data whether or not these reforms succeeded in reducing agency loss resulting from political fragmentation. Since these reforms, however, Sina and Tencent still appear to resist and defy regulations.

In 2015, the CAC threatened to shut down Sina Weibo due to insufficient censorship.[25]

---

[23]Technically, the SCIO was referred to as the SIIO after 2011. In late 2011 reforms, the SCIO's rank was elevated so that it directly reported to the State Council.

[24]See full report on the CAC here.

[25]See Wall Street Journal article here.

In 2017, the CAC imposed "maximum fines" on Sina Weibo, Tencent, and Baidu for "failing to fulfill their management duties and violating China's Cyber Security Law."[26] In 2018 Sina was ordered by the CAC to suspend "key portals such as its hot search site and portal on celebrities and their personal lives for a week" due to its violation of "relevant internet laws and regulations and spread illegal information."[27] In 2018 the CAC suspended popular social networking site, Zhihu for one week for "lax supervision and the spread of illegal information."[28] In late 2018 the CAC ordered the suspension of news aggregators for several weeks due to "illegal" information sharing. Sina Weibo has responded by expanding its efforts to censor content through crowdsourcing and gamification, offering iPads to the best "Weibo Supervisors," users who volunteer their efforts to help Sina clean up harmful content.[29]

These developments are interesting. While Sina appears to increase efforts to police harmful information, it does so through crowd-sourcers, in an attempt to cut costs. In recent state media, the authorities noted that despite these increased efforts, they "are not fully performing their duties."[30] Whether or not the poor performance of Sina and other internet companies like it poor performance is due to continued defiance or low capacity to censor, the CAC still seems concerned with agency loss.

7 Beyond China

Tensions and alliances between corporations and government actors, as well as government delegation to corporations, are relevant far beyond the case of censorship in China. The fraught alliance between social media companies and governments is representative of a greater trend beyond the Chinese context, in authoritarian and democratic politics alike. Corporations, especially in the realm of surveillance and data analytics are operating on the behalf of governments to spy on citizens, enforce copyright laws, censor information, and repress government opponents. These alliances represent a "broader trend of neoliberal restructuring, in which political authority and decision-making power are taken out of the public realm and transferred to private environments, often under-

---

[26]See Global Times article here: http://www.webcitation.org/71y4srDVO, and see a The Diplomat article here: http://www.webcitation.org/71y4v72U0.

[27]See http://www.webcitation.org/71y4wnS1A.

[28]See http://www.webcitation.org/71y52ogK6.

[29]See http://www.webcitation.org/71y53kuQE.

[30]See http://www.webcitation.org/71y55KoD9.

pinned by commercial and market logics" (Crouch, 2004; Hintz, 2016). Since the 2016 elections, it became increasingly clear how difficult it would be to hold social media companies accountable for their actions. Behaviors that are in the public interest, such as cleaning up bot accounts, preventing the spread of fake news, identifying and disrupting foreign influence campaigns, monitoring hate speech, preventing the spread of violence, and removing bad actors are often in conflict with fundamental profit motivations and concerns about competitiveness. It took years for Twitter to take any meaningful action on bot accounts, and they did so only under extreme public pressure, due to fears of a hit to monthly active user statistics that would reduce its stock price.[31]

As the Snowden leak revealed, the PRISM program gave the NSA authority to request that Microsoft, Facebook, Apple, and Google provide data matching keywords approved by a U.S. Foreign Intelligence Surveillance (FISA) Court ruling. At large tech companies, requests for information controls are released publicly, showing a tension between corporate and consumer preferences and states' logic of social control (Tanash, Chen, Thakur, Wallach, & Subramanian, 2015). Private companies often refuse requests out of concerns for their users' preferences, their profitability, and their reputation. An example of these conflicts is Apple's refusal to cooperate with FBI requests following the San Bernardino terrorist attack. The FBI sought to compel Apple to break their own encryption so that the FBI could obtain information on one of the suspect's iPhones. Obeying such a request, Apple said, would "threaten the security of our customers," which Apple has trumpeted as an advantage over competitor Google. In 2014, they boasted, "unlike our competitors, Apple cannot bypass your passcode and therefore cannot access [customer] data." Despite public statements about Apple's company values, their refusal is likely in part driven by concerns about profits and competitiveness. Like Sina, Apple is not keen on giving into government demands that would put it at a competitive disadvantage.

The relevance of tensions and partnerships between state and corporate actors is also indicative of trends in Chinese reform that not only includes corporate actors in the policymaking process, but uses corporations as labs to create structures and institutions that can later be subsumed into the state. For example, the very process of censorship described in this paper has already been partially subsumed into the state. Employees working in editorial functions at internet content producers (ICPs) in China can no longer

---

[31]See http://www.webcitation.org/71y56kNiL.

remain on private payrolls. This means that according to Chinese law, the employees who censor content can no longer be employees of Sina Weibo, but now must be employees of a government agency such as the Network Management Office or the Supervision Department. Whether or not this solves principal agent problems identified in the censorship system has yet to be seen, but control over the internet has been tightening at a rapid pace over the last few years.

Other such plans to subsume institutions borne out of corporations seem on the horizon. China is now relying on the infrastructure of Alibaba's Alipay to serve as the technical back-end to a national "social credit system" which merges social and financial data to give users a score that can selectively restrict access to state and commercial services, increase monitoring and policing of certain "untrustworthy" people, and even determine what jobs an individual can have. Even more interesting are the institutions that are being created within large companies such as Alibaba that mirror government institutions. Alibaba's disputes between suppliers and customers are resolved by a jury of one's peers and sentences are doled out by impartial judges; these roles are given to users on Alibaba's platform. Judges and juries decide how to split up money held in escrow for disputed transactions made in the system. This corporate system alleviates much of the strain put on rigid and ineffective legal institutions (Liu & Weingast, 2017). Despite being a competitor to the existing bureaucratic system, the government appears to tolerate its existence. It is conceivable and consistent with current trends that these corporate structures might one day inform reform of legal institutions, or be subsumed into the current legal system in China.

The trends outlined above and the relationship between Sina Weibo and government actors in China may represent a potentially transformative shift in how states and corporations interact. More nimble and adaptive corporate structures may help authoritarian governments leverage data to manage and monitor public opinion. Alternatively, corporate profit motives may prove to increase conflicts between state and society, leading to reform or instability.

References

Brady, A.-M. (2009). *Marketing dictatorship: Propaganda and thought work in contemporary china*. Rowman & Littlefield.

Cairns, C. (2016). *Fragmented authoritarianism? reforms to china's internet censorship system under xi jinping* (Unpublished doctoral dissertation). Cornell University.

Calvert, R. L., McCubbins, M. D., & Weingast, B. R. (1989). A theory of political control and agency discretion. *American journal of political science*, 588–611.

Chen, J., & Xu, Y. (2016). Why do authoritarian regimes allow citizens to voice opinions publicly? *The Journal of Politics, Forthcoming*.

Crandall, J. R., Crete-Nishihata, M., Knockel, J., McKune, S., Senft, A., Tseng, D., & Wiseman, G. (2013). Chat program censorship and surveillance in china: Tracking tom-skype and sina uc. *First Monday*, *18*(7).

Crete-Nishihata, M., Hilts, A., Knockel, J., Ng, J. Q., Ruan, L., & Wiseman, G. (2016, November). *Harmonized histories? a year of fragmented censorship across chinese live streaming applications*. Retrieved from https://netalert.me/harmonized-histories.html

Crouch, C. (2004). *Post-democracy*. Polity Cambridge.

Egorov, G., Guriev, S., & Sonin, K. (2009). Why resource-poor dictators allow freer media: A theory and evidence from panel data. *American political science Review*, *103*(04), 645–668.

Gehlbach, S., & Sonin, K. (2014). Government control of the media. *Journal of Public Economics*, *118*, 163–171.

Guriev, S. M., & Treisman, D. (2015). How modern dictators survive: Cooptation, censorship, propaganda, and repression.

Hintz, A. (2016). Restricting digital sites of dissent: commercial social media and free expression. *Critical Discourse Studies*, *13*(3), 325–340.

Jin, H., Qian, Y., & Weingast, B. R. (2005). Regional decentralization and fiscal incentives: Federalism, chinese style. *Journal of public economics*, *89*(9), 1719–1742.

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.

Knockel, J., Ruan, L., & Crete-Nishihata, M. (2017). Measuring decentralization of chinese keyword censorship via mobile games. In *7th USENIX workshop on free and open communications on the internet (FOCI 17)*. Vancouver, BC: USENIX Association. Retrieved from https://www.usenix.org/conference/foci17/workshop-program/presentation/knockel

Lieberthal, K. (1995). *Governing china: From revolution through reform*. WW Norton.

Liu, L., & Weingast, B. R. (2017). Taobao, federalism, and the emergence of law, chinese style.

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.

Lynch, D. C. (1999). *After the propaganda state: Media, politics, and" thought work" in reformed china*. Stanford University Press.

MacKinnon, R. (2009). China's censorship 2.0: How companies censor bloggers. *First Monday*, *14*(2).

Mertha, A. (2009). "fragmented authoritarianism 2.0": political pluralization in the chinese policy process. *The China Quarterly*, *200*, 995–1012.

Montinola, G., Qian, Y., & Weingast, B. R. (1995). Federalism, chinese style: the political basis for economic success in china. *World politics*, *48*(01), 50–81.

O'Brien, K. J. (1994). Implementing political reform in china's villages. *The Australian Journal of Chinese Affairs*(32), 33–59.

Oksenberg, M., & Lieberthal, K. G. (1988). *Policy making in china: Leaders, structures, and processes*. Princeton: Princeton University Press.

Roberts, M. (2017). *The censorship tax information distortion within china's great firewall.* Princeton University Press.

Rundlett, A., & Svolik, M. W. (2016). Deliver the vote! micromotives and macrobehavior in electoral fraud. *American Political Science Review*, *110*(1), 180–197.

Shambaugh, D. (2007). China's propaganda system: Institutions, processes and efficacy. *The China Journal*(57), 25–58.

Stern, R. E., & O'Brien, K. J. (2012). Politics at the boundary: Mixed signals and the chinese state. *Modern China*, *38*(2), 174–198.

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

Tanash, R. S., Chen, Z., Thakur, T., Wallach, D. S., & Subramanian, D. (2015). Known unknowns: An analysis of twitter censorship in turkey. In *Proceedings of the 14th acm workshop on privacy in the electronic society* (pp. 11–20).

Villeneuve, N. (2008). Search monitor: Toward a measure of transparency.

Yang, G. (2013). *The power of the internet in china: Citizen activism online.* Columbia University Press.

Zheng, Y. (2007). *De facto federalism in china: Reforms and dynamics of central-local relations* (Vol. 7). World Scientific.