The University of Chicago

Carrot and Stick:
A Model of Censorship and Co-optation

Sizhen Fang

May 2022

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts Program in the Committee on International Relations

Faculty Advisor: Scott Gehlbach
Preceptor: Manuel Cabal

Abstract

Authoritarian leaders use a variety of tools in their rulings of citizens, including censorship and co-optation. In using these tools, the leader contracts an “authoritarian bargain” with the citizens where it provides citizens with better public goods in exchange for more centralized power. This paper applies the global coordination game to show leader’s optimal censorship strategy on her power centralization and public goods provision. An authoritarian leader with intention to attack the existing institutions to centralize her political power needs a high censorship rate to hide her personalistic intention, while at the same time, benefits from a low censorship rate on public goods provision since it serves as a useful channel to gauge public opinions. The paper collects and analyzes novel text data from Chinese social media using machine learning and statistical methods to provide empirical support of the theory. Results show that criticisms on the leader's personalistic behaviors are 42.5% more likely to be censored than criticisms on the government's public goods supply.

Acknowledgements

This project would not have been possible without the support of many people. Many thanks to my adviser, Scott Gehlbach, whose expertise was invaluable in formulating the research questions and methodology. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level. I would also like to thank my preceptor, Manuel Cabal, for his valuable guidance throughout my studies. Also thanks to Burcu Pinar Alakoc, Henry Cotton, Apoorva Malarvannan, Maya Krishnan, Michael J. Reese, Linnea Turco, everyone from my first preceptor group and second-year thesis workshop, and especially Matthias Staisch for generous advice and suggestions.

Thanks to the Committee of International Relations for awarding me generous fellowship, providing me with the financial means to complete this project. And finally, thanks to my friends, Grace Cai, Yanqing Shen and Celia Yinzhi Xu, for always offering emotional support and love.

Table of Contents

Abstract 2
Acknowledgements 3
List of Figures 5
List of Tables 6
Chapter 1. Introduction 7
Chapter 2. Literature Review 11
Chapter 3. Theories of Authoritarian Censorship 13
Chapter 4. Formal Model 17
4.1 Environment 17
4.2 Equilibrium Characterization 20
4.3 Equilibrium Analysis 21
4.4 Implication 22
Chapter 5. Censorship in China 25
Chapter 6. Empirics 27
6.1 Background 27
6.2 Data 31
6.3 Analysis 32
6.4 Robustness Check 37
Chapter 7. Conclusion 38
References 40
Appendices 43
Appendix I. List of Keywords 43
Appendix II. Summary Statistics 44
Appendix III. Classification: Identifying Relevant Posts 45
Appendix IV. Propaganda 46
Appendix V. Example of Posts 49

List of Figures

Figure 1: Time Series Plot 33
Figure 2: Time Series Plot with Propaganda 47

List of Tables

Table 1: Citizen's Payoff 19
Table 2: Incumbent's Payoff 20
Table 3: Optimal Public Goods 23
Table 4: Optimal Public Goods with Pro-illiberal Leaders 23
Table 5: Regression Results 36
Table 6: Bootstrap 37
Table 7: Summary Statistics 44
Table 8: Constitutional Amendment Classification 45
Table 9: Retirement Reform Classification 45
Table 10: Air Pollution Classification 45

Chapter 1. Introduction

The media control of authoritarian regimes has received a great deal of attention from political scientists from the past decades. The conventional wisdom of authoritarianism believes that the critical function of censorship is to suppress criticisms and it is often assumed that less democratic government will do their best to "stifle independent criticism" (Geddes and Zaller 1989). Indeed, totalitarianism is partly defined by unchallenged state control over the mass media (Brzezinski 1965), and the loss of control over the media has been linked to the collapse of authoritarian regimes (Lawson 2002). As a result, any negative reporting, or criticisms by the media under an authoritarian government is often considered to be the result of weakness, indicating the regime's failure to achieve the goal of absolute media control (Levitsky and Way 2010).

However, more recent studies have recognized that, paradoxically, some independent, negative reporting can in face help the authoritarian rulers. Lorentzen (2014) shows that optimally permitting investigative reporting on lower-level officialdom can help regimes maintain power by improving governance (P. Lorentzen 2014). Also, tolerating regular small-scale protests can be an effective information-gathering tool that enhances the regime's stability (P. L. Lorentzen 2013). Empirically, King et al. (2013) show that China's censorship program curtails collective actions but allows criticisms to circulate, regardless of its content (King, Pan, and Roberts 2013). Malesky (2019) further point out that a significant portion of sampled posts in King et al.'s analysis coincide with state-led consultation campaigns to solicit public input on policy proposals (Gueorguiev and Malesky 2019).

The logic of why authoritarian leaders tolerate negative comment is simple and intuitive: From the government point of view, negative comments on social media make the leaders' poor governance common knowledge but also serves as a useful information channel for the leader to gauge public opinions. Leaders learn from the critical posts citizens' opinions and expectations about specific policies. Furthermore, complete cleanup of sensitive content can impair the regime's ability to learn from bottom-up information and address social problems before they become threatening.

In this account, the essential reason of why authoritarian regime would not relentlessly censor or even ban social media is that censorship is a double-edged sword. A high level of censorship will decrease the probability of the leaders' poor performance being known to the citizens, which could threaten their power in place. However, high censorship also makes it more difficult for the leader to learn the preferences of citizens through which the leader could improve their public goods provision to mollify the citizens. Therefore, once collective actions are prevented, social media can be an excellent way to gauge public opinion about specific policies and. Loosening up the restrictions on public expression can be a useful government tool for leaders to learn how to satisfy and mollify the citizens (King, Pan, and Roberts 2013).

My contribution in this paper is to extend this account by arguing that, in the theoretical optimal, the authoritarian leader adopts selective censorship on criticisms. The leader has a high tolerance for critiques relates to their supply of public goods but has a low tolerance for comments that threatens its centralization of power. For instance, despite allowing a free expression on social media regarding various public policies, ranging from the retirement reform to the air pollution control, President Xi Jinping's speech at the National Propaganda and Ideology Work Conference in August 2013 also revealed the Party's hostility to dissidents

online. Xi stressed the need to suppress controversies and "mistaken viewpoints" on every social media platform (Garrick and Bennett 2016). Moreover, Xi empathized on creating a "positive" online environment and filtering out 'negative' speech, much of which criticized the Party or its leaders (Cairns 2016).

In this paper, I develop a formal model in which an authoritarian leader chooses a level of censorship on her authoritarian deinstitutionalization and a level of censorship on her provision of public goods. The key to my argument is the "authoritarian bargain" between the leader and the citizens. In authoritarian regimes, the citizens relinquish political rights for economic security (Desai, Olofsgård, and Yousef 2009). A key feature in the model is that not only do citizens face uncertainty about the intention of the leader, but the leader also faces uncertainty about the citizens' types. Most importantly, the model shows that when the leader can persuade the citizens to accept the status quo authoritarian ruling and can afford to supply citizens' ideal level of public goods, lowering censorship rate on public goods can help the leader improve their public provision and stabilize their ruling.

In the second part of this paper, I provide empirical evidence using novel data collected from one of the biggest Chinese social media, Sina Weibo, to support the applicability of the model. My empirical application shows that the Chinese censorship program adopts a much higher censorship rate on issues regarding the power centralization of the party and its top leaders (for e.g., the constitutional amendment in 2018) but and lower censorship rate on issues regarding the party's public goods provision.

The paper proceeds as follows. In the next section, I review recent literature on censorship and authoritarianism. In the model section, I present a simple model that formalizes

the theoretical logic. Following the formal model, I present empirical evidence I draw from the Chinese social media to support for my argument. A final section concludes.

Chapter 2. Literature Review

Deciphering the intention of authoritarian leaders is always of interest to political scientist scholars. While earlier western scholars adopt methodologies such as Kreminology or Sovietology (Breslauer 1992; Yahuda 1972), contemporary scholars rely more on government statistics, public opinion surveys, and interviews (Guo 2009; Kung and Chen 2011; Pan 2019).

With the boom of the Internet, social media platforms provide a great scope to decipher the intention of authoritarian leaders. Scholars collect first-hand data from social media platforms to study how censorship is conducted and how the Internet may increase the space for public discourse. King et al. (2013) argue that the primary motive for censorship is to quash potential for collective action but not squelch criticism of the government (King, Pan, and Roberts 2013). Gueorguiev and Malesky (2019) further point out that a significant portion of sampled posts in King et al.'s analysis coincide with state-led consultation campaigns to solicit public input on policy proposals (Gueorguiev and Malesky 2019). Moreover, Qin et al. (2017) show many posts on highly sensitive topics, such as protests and strikes, circulate on social media (Qin, Strömberg, and Wu 2017), and these posts help predict and spread collective actions (Qin, Strömberg, and Wu 2019). Following the previous studies, this project will offer novel and recent Chinese social media data to identify and explain the Chinese government's censorship strategy, specifically its selective censorship on criticisms.

The formal model in this project builds on two strands of literature on media censorship. The first strand focuses on the benefits of censorship for leaders. Censorship as an information manipulation tool blocks the bad news that can potentially hurt the leader's reputation and reduces political turnover (Besley and Prat 2006). It can also help the leader mobilize citizens to

further her political goals which may not be of citizens' best interests (Gehlbach and Sonin 2014).The second strand stresses the opportunity cost of censorship and explains why the government wants information circulation. Critiques, protests, and unpromising elections, which conventional wisdom considers only harmful for authoritarian leaders, serve as a valuable outside information source. Lorentzen (2014) shows that optimally permitting investigative reporting on lower-level officialdom can help regimes maintain power by improving governance (P. Lorentzen 2014). Also, tolerating regular small-scale protests can be an effective information-gathering tool that enhances the regime's stability (P. L. Lorentzen 2013). Similarly, authoritarian rulers gain information that reduces the risk of their violent removal from office through elections (Cox 2009). Furthermore, Chen and Xu show that when public opinions are divided, allowing public communication has a discouraging effect on collective actions (Chen and Xu 2017).

Chapter 3. Theories of Authoritarian Censorship

It is widely agreed that the ultimate purpose of censorship is to maintain the leaders' status quo governance. Previous scholars believe that leaders should prevent all negative comments that can potentially threaten their regime from spreading and circulating among the public to achieve this goal (Geddes and Zaller 1989). Past studies identify two types of negative comments that can jeopardize the stability of the authoritarian regime. First is *state critique*, which contains dissidents and attacks or finds faults with the state policies or leaders. Leaders may find these critiques threatening because critiques make their poor performance a common knowledge, deepening citizens' dissatisfactory with their governance. The second type of comments leaders fear is the *collective action potential* comments, in which people join to express themselves collectively and have the potential to spur collective actions. Collective actions, such as protests and strikes, are often thought to be authoritarian regimes' death knell. Small, local riots may grow and spread into national-wide revolutions, which eventually overthrow the regime.

As past studies predict, authoritarian leaders prefer not to have any information that can potentially threaten their power in place to circulate among citizens. Hence, Leaders do not welcome unauthorized critiques which signal their incompetence to the public, nor do they welcome collective actions that challenge their power in place. Therefore, the leadership's optimal choice is to censor all adverse comments, including those that criticize the state and spur collective actions.

The existing theory explains the Chinese government's deletion of social media posts that have collective action potentials. Leaders firmly regulate collective action expressions because

they fear the mobilizing nature of social media. However, the existing theory fails to offer an adequate explanation for why authoritarian leaders would allow for any critiques. On the contrary, previous scholars believe that authoritarian leaders will censor all critiques, as they censor collective actions, since critical comments make leader's incompetence known to the public, damaging leaders' reputation and threatening their power in place.

The underlying assumption of the existing theory is that the circulation of critical posts brings only threats to authoritarian leaders but no benefits. In other words, authoritarian leaders gain nothing from tolerating critiques but risk revealing their misconduct to the public. However, King et al. (2013) 's findings imply that leaders may gain from allowing for critical comments on social media. The benefits of critical comments are twofold. They serve as a valuable information source for authoritarian leaders to gauge public opinions. Leaders learn from the critical posts citizens' opinions and expectations about specific policies. Allowing citizens to express opinions about current policies, especially critiques, can help authoritarian leaders maintain power by improving their governance. Furthermore, complete cleanup of sensitive content can impair the regime's ability to learn from bottom-up information and address social problems before they become threatening.

In this sense, the empirical pattern discovered by King et al. (2013) may be the theoretically optimal strategy for an authoritarian regime to use social media to maximize its political survival. Once collective actions are prevented, social media can be an excellent way to gauge public opinion about specific policies and. Therefore, loosening up the restrictions on public expression can be a useful government tool for

leaders to learn how to satisfy and mollify the citizens (King, Pan, and Roberts 2013).

At first glance, King et al. (2013) 's conclusion that the Chinese government would allow any criticisms seems to contradict the Chinese government's recent behaviors. For example, President Xi Jinping's speech at the National Propaganda and Ideology Work Conference in August 2013 revealed the Party's hostility to dissidents online. Xi stressed the need to suppress controversies and "mistaken viewpoints" on every social media platform (Garrick and Bennett 2016). Moreover, Xi empathized on creating a "positive" online environment and filtering out 'negative' speech, much of which criticized the Party or its leaders. Following this conference, a nationwide Internet Cleaning-up Campaign extensively deleted posts containing views deviating from those of the Party (Cairns 2016).

Xi's speech and policy seem to conflict with the conclusion that the Chinese government welcomes any criticisms, as King et al. (2013) state. The Chinese government cannot let all dissidents circulate online while also deletes "negative" and "deviated" opinions. Thus, a reasonable conjecture is that the Chinese censorship program adopts a selective censorship strategy.

To further disentangle Chinese leaders' convoluted attitudes towards critiques, I hypothesize that the Chinese censorship program has a high tolerance for critiques relates to its supply of public goods but has a low tolerance for critiques on issues related to the leader’s personalistic behaviors.

The main drawback in King et al.'s theories is that they assume all critiques weigh the same to the authoritarian leaders. In other words, they fail to recognize different types of criticism may associate with varying dimensions of legitimacy, which leads to different concerns for the censors. On the other hand, the amendment I propose allows authoritarian leaders to have different sensitivity levels to critiques on various issues. When citizens complain about

governments' failures to deliver decent public service, leaders can quickly mollify them by improving their supply of public goods, as King et al. suggest. However, when citizens attack the centralized power of the Chinese Communist Party and its top leaders, implementing such political reforms to mollify citizens is way too costly for authoritarian leaders since they risk overthrow to improve governance - a Gorbachev option.

The more profound logic behind this is the "authoritarian bargain" between the leader and the citizens (Desai, Olofsgård, and Yousef 2009). The authoritarian leader provides the citizens with welfare and public goods in exchange for more centralized political power. Thus, the leader needs a low censorship rate on public goods-related critiques to learn the citizens' preferences. Simultaneously, the leader needs a high censorship rate on politically sensitive matters to centralize power. Therefore, the authoritarian leaders likely have a high tolerance for critiques related to their supply of public goods but have a low tolerance for critiques on politically sensitive issues such as the top leaders' power centralizing behaviors

Chapter 4. Formal Model

In this section, I present a baseline model to illustrate the mechanism of information manipulation.

4.1 Environment

Players and Preferences. The model has three players. An incumbent of an institutionalized authoritarian regime (she) and two citizens (he) (Citizen $i = A, B$). The leader can be of two types: the pro-liberal type and the pro-illiberal type (ω_$\{G\} \in \{0, 1\}$). The pro-illiberal incumbent non-strategically chooses to backslide in the game's first move, and the pro-liberal incumbent never backslides.

Citizen i's type is denoted as $\omega_i \in \{0,1\}$, with 0 and 1 representing a satisfied type and a dissatisfied type, respectively. The satisfied type has a partisanship preference favoring the incumbent. The dissatisfied type, on the contrary, intrinsically opposes the incumbent. Furthermore, I assume that both citizens have a desire to protect the status quo institutions and thus are against the incumbent's personalistic behavior.

Information. The incumbent's type as well as citizens' types are all private information. In the beginning of the game, Nature assigns the type of the incumbent and the citizens. For simplicity, I assume that both citizens share a common prior that an incumbent is pro-liberal is one-half:

$$Pr(\omega_G = 0) = \frac{1}{2}.$$

The incumbent and the citizens also share a common prior that a citizen is satisfied with a probability one-half:

$$Pr(\omega_i = 0) = \frac{1}{2},\ i = A, B$$

Furthermore, the two citizens' types are i.i.d.

Signaling and Censorship. Despite the types being private information, citizens receive an outside signal (for example: media, whistle blowers or travelling abroad) $s \in \{b, \emptyset\}$ indicating the true type of the incumbent. If the incumbent is pro-illiberal, then the citizens receive $s = b$. Otherwise, the citizens receive $s = \emptyset$. Citizens can also send a signal in the public information space indicating their types. This signal is observable by the other fellow citizen and the incumbent.

The incumbent can block the signals from the outside source and from citizens. Specifically, the incumbent can choose different censorship levels for blocking the outside signals and the public communication between citizens.

Let $c_1 \in \{0,1\}$ be the incumbent's censorship policy targeting the outside signaling and $c_2 \in \{0,1\}$ be the censorship policy targeting the public communication among citizens. When $c_1 = 0$, the outside signal is fully revealed. When $c_1 = 1$, the outside signal is blocked, and the citizens receive the null signal. Likewise, when $c_2 = 0$, both citizens are allowed to freely send signals indicating his type. When $c_2 = 1$, no citizen's voice will be heard by the incumbent and the fellow citizen.

Public Goods Provision. Another tool in the incumbent's toolkit is to provide public goods to citizens. Public goods are modeled as material transfers equally distributed to each citizen. In this game, the incumbent can promise to offer each citizen a public good $r \in (0, \bar{r})$ if neither of the citizens participate in the revolution ($\bar{r}$ is the limit of public goods budget of the incumbent

whose value I will clarify in later sections). If neither citizen revolt, then the public goods r will be realized as promised.

Collective Action and payoffs. Citizens simultaneously choose to keep the status quo or to rebel. If both citizens rebel, the status quo is abandoned, and the incumbent ousted. If only one citizen rebels, the incumbent is ousted with probability $\lambda\left(0<\lambda<\frac{1}{2}\right)$. There is a fixed cost $k>0$ for a citizen to participate in a revolt.

In making their decisions to rebel or not, the citizens balance their partisanship preferences, their desire to prevent the institutions being attacked, and the public goods provision. Their payoff from the status quo is normalized to 0. For a satisfied citizen, successfully ousting the incumbent induces him a payoff $v_0<0$. A dissatisfied citizen, on the other hand, received a payoff of $v_1>0$ by ousting the incumbent. Furthermore, both types of citizens receive a payoff of $d>0$ if successfully ousting an incumbent who backslides.

Citizen i's payoff is summarized as below:

Table 1: Citizen's Payoff

	Revolt(j)	Not Revolt(j)
Revolt(i)	$\theta_i(\omega_i,\omega_G)-k$	$\lambda\theta_i(\omega_i,\omega_G)-k$
Not Revolt(j)	$\lambda\theta_i(\omega_i,\omega_G)$	r

Here, $\theta_i(\omega_i,\omega_G)=v_i+d\cdot\omega_G$.

The incumbent's payoff depends on the citizens' collective actions, her choice of public provision and her decision to backslide, as summarized below:

Table 2: Incumbent's Payoff

A = (0,0)	A = (0,1) or (1,0)	A = (1,1)
$-2r$	$-\lambda\rho$	$-\rho$

Here, $\rho > 0$ is the cost that the incumbent bears if there is a full-scale revolt (i.e., if both citizens participate), and $\lambda\rho$ is the repression cost if there is a partial revolt (if only one citizen revolt).

Timeline. The timing of actions is as follows:

1. Nature chooses the type of two citizen: satisfied or dissatisfied, and the type of the incumbent: pro-liberal or pro-illiberal.
2. Citizens receives an outside signal revealing the leader's type, which is blocked by a probability c_1.
3. Citizens signals each other about their types (also observable by the leader), which is blocked by a probability c_2.
4. The incumbent promise to offer public goods for both citizens if and only if neither citizen revolts.
5. Citizens choose their action.

4.2 Equilibrium Characterization

I focus on the equilibrium where the citizens always truthfully reveal their type. Such equilibrium satisfies the following requirements:

1. The incumbent's censorship strategy on the outside signal, c_1, maximizes its expected payoff given citizens' equilibrium strategies, beliefs, and the government equilibrium public good provision.
2. The incumbent's censorship strategy on public communication, c_2, maximizes its expected payoff given citizens' equilibrium strategies, beliefs, and the incumbent's equilibrium public good provision.
3. When public communication is (not) allowed, the incumbent's choice of public good provision maximizes its expected payoff given citizens' equilibrium strategies and beliefs.
4. Each citizen in the collective-action stage maximizes his own welfare, given the equilibrium belief about the other one's equilibrium strategy.

4.3 Equilibrium Analysis

Collective Actions. Since multiple Nash Equilibrium exists, I focus on the Nash Equilibrium that is the Pareto-optimal. If two Nash Equilibriums are both Pareto-optimal, then I focus on the Nash Equilibrium that both citizens choose to remain the status quo. In any equilibrium when citizens truthfully reveal their types if public communication is allowed, a citizen *i*'s action is uniquely defined such that, when the incumbent did not attack the existing institutions, a satisfied citizen never protests, and a dissatisfied citizen protest if and only if the other citizen abstains from the protest and enough public goods is offered. When the incumbent attacks the institutions, then both types of citizens will only be willing to keep the status quo if enough public goods are offered and the other citizen is also willing to abstain from the protest. A crucial point then is to calculate how much public goods is sufficient to mollify the citizens in each case.

How much public goods are enough? When public communication is allowed, there are six cases based on the citizens' types and the incumbent's action of personalistic.

First, if the incumbent does not attack the institution, a satisfied citizen will never protest. Hence, if the incumbent faces with two satisfied citizens and such an incumbent does not attack the institutions, she does not need to provide any public goods for citizens to remain the status quo. Denote the minimal amount of public goods the incumbent has to offer to each citizen for both citizens to remain the status quo as $r^*\left(\left(\omega_i, \omega_j\right), \omega_G\right)$, $r^*\left((0,0), 0\right) = 0$. Further, when one citizen is satisfied, and the other citizen is dissatisfied, the dissatisfied citizen will abstain from the protest if $r \geq \lambda v_1 - k$, thus $r^*\left((0,1), 0\right) = \lambda v_1 - k$. Finally, when both citizens are dissatisfied, they will abstain from the protest if $r \geq v_1 - k$.

Now consider the case when the incumbent has attacked the institution. If two citizens are both satisfied, then each of them needs a public good $r^*\left((0,0), 1\right) = v_1 - d + v_0 - k$ to abstain from revolting. When one citizen is satisfied, and the other citizen is dissatisfied, the dissatisfied citizen will abstain from the protest if $r \geq \lambda(d + v_1) - k$. Finally, when both citizens are dissatisfied, they will abstain from the protest if $r \geq d + v_1 - k$. Table 1 summarizes the optimal public goods the leader offers in each case.

Table 3: Optimal Public Goods

ω_G	c_2	$\omega_A = \omega_B = 0$	$\omega_A \neq \omega_B$	$\omega_A = \omega_B = 1$
0	0	0	$\lambda v_1 - k$	$v_1 - k$
	1	$\frac{1}{2}(1+\lambda)v_1 - k$		
1	0	$d + v_0 - k$	$\lambda(d + v_1) - k$	$d + v_1 - k$
	1	$\frac{1}{2}(1+\lambda)(v_1 + d) - k$		

When does the incumbent block outside signals? When citizens receive the outside signal $s = b$, both infer that the leader is pro-illiberal. In contrast, when $s = \emptyset$, by Bayes' rule, both citizens share the posterior belief that

$$Pr(\omega_G | s = \emptyset) = \frac{1}{1 + c_1}.$$

Similarly, the optimal public goods the pro-illiberal leader offers in each case, given the option of blocking the outside signal are summarized in Table 3.

Table 4: Optimal Public Goods with Pro-illiberal Leaders

c_1	c_2	$\omega_A = \omega_B = 0$	$\omega_A \neq \omega_B$	$\omega_A = \omega_B = 1$
0	0	$d + v_0 - k$	$\lambda(\mathrm{d} + \mathrm{v}_1) - \mathrm{k}$	$d + v_1 - k$
	1	$\frac{1}{2}(1+\lambda)(\mathrm{v}_1 + \mathrm{d}) - \mathrm{k}$		
1	0	$\frac{1}{2}[(1+\lambda)v_0 + d] - k$	$\frac{1}{2}[(1+\lambda)\mathrm{v}_1 + \mathrm{d}] - \mathrm{k}$	$v_1 + \frac{d}{2} - k$
	1	$\frac{1}{4}(3+\lambda)v_1 + \frac{d}{2} - k$		

From Table 4, we can also see that whether public communication is allowed or not, and given any combination of the citizen's types, the optimal amount of public goods the incumbent needs to offer the mollify the citizens are always lesser if the incumbent blocks the outside signal. Hence, when a pro-illiberal incumbent intends to backslide, the optimal choice is always to block the outside signal and hide her type from the citizens.

Incentive to allow public communication. When and why does the incumbent wish to allow public communication? From the last section, we learned that the pro-illiberal incumbent always benefits from censoring the outside signal, so when citizens face a pro-illiberal incumbent, they always receive $s = \emptyset$. When the incumbent is pro-liberal, the outside signal is also null since the pro-liberal incumbent never intends to backslide. In other words, the citizens always receive $s = \emptyset$ with regardless of the incumbent's type.

Both types of incumbents thus choose to offer the same amount of public goods given their belief on the citizen's types. The expected payoff for the incumbent if public communication is disallowed is:

$$U_G(c_2 = 1) = -\frac{1}{2}(3 + \lambda)v_1 - d + 2k$$

and if the incumbent allows for public communication, her expected payoff is:

$$U_G(c_2 = 0) = -\frac{1}{2}(2 + \lambda)v_1 - \frac{1}{4}(1 + \lambda)v_0 - d + 2k$$

Note that $U_G(c_2 = 0) > U_G(c_2 = 1)$ for all λ, v_0, v_1 and k. In other words, it is optimal for the incumbent to allow for public communication, regardless of their types.

Chapter 5. Censorship in China

China has sought to monitor and control how its citizens use the Internet for as long as Internet existed in China. Its tactics frequently include strict media controls through monitoring systems and firewalls, the closure of publications or websites, and the imprisonment of dissident journalists, bloggers, and activists (Xu and Albert 2017). In 2000, China launched its first censorship and surveillance program, the Golden Shield Project, a massive censorship and surveillance mechanism aimed at restricting content, identifying and locating individuals, and providing instant access to personal records (The Great Firewall of China 2011). Since Chinese President Xi Jinping came to power in 2013, censorship of all forms of media has tightened. In February 2016, Xi announced a new media policy for party and state news outlines: “All the work by the party’s media must reflect the party’s will, safeguard the party’s authority, and safeguard the party’s unity.” Xi emphasizes that state media must align with the party leadership’s “thought, politics, and actions. (Press 2016)”

Today's China has roughly 983 million active internet users (Thomala Lai Lin 2022). As the number of Chinese internet users increases, so does the scope of Chinese censorship. In 2016, Freedom House ranked China last for the second year in a row out of sixty-five countries that account for 88 percent of the world's internet users. In its 2016 worldwide press freedom index, the France-based watchdog group Reporters Without Borders ranked China 176th out of 180 countries (Countries and Territories 2016). More than a dozen government agencies review and enforce laws governing information flow within, into, and out of China.

The Communist Party's Central Propaganda Department (CPD) is the most powerful monitoring body, coordinating with the General Administration of Press and Publication and the State Administration of Radio, Film, and Television to ensure that content promotes party

doctrine (Xu and Albert 2017). In addition, the Chinese government employs many people to monitor and censor the Chinese media. Over two million employees review internet posts using keyword searches and compiling reports for "decision-makers." Both the government and private companies employ these "public opinion analysts" to constantly monitor China's internet (Jia 2019).

Chapter 6. Empirics

The model offers a framework to think about authoritarian leaders' censorship choices on personalistic behaviors and their public goods provisions. The key idea is that, since the leader provides more public goods to co-opt the citizens in exchange for more tolerance to backslide, leader with the intention to backslide adopts a high censorship rate on their personalistic intention and a low censorship rate on citizen's expression on public goods provisions, as far as the leader can provide enough public goods to mollify the citizens.

In the empirical session, I use evidence from China to show that the implication from the model is consistent with real-world data. China is one of the biggest authoritarian regimes in the world, and it has censorship over all media capable of reaching a wide audience. The Chinese communist party led by President Xi Jinping fits the model as an authoritarian regime with a leader who has personalistic intention and has strong economic capacity to provide its citizens with ample public goods. As predicted by the model, we expect to see a high censorship rate regarding policy issues related to the leader's personalistic behavior, and a low censorship rate on issues regarding the leader's provision of public goods.

6.1 Background

I examine three policy issues that have triggered great massive attention and discussion among the Chinese citizens. One of these events involves the Chinese leader's behavior of deinstitutionalization, and the other two concern the leader's provision of public goods.

Constitutional Amendment. Many consider the fifth amendment to the People's Republic of China's constitution, passed in 2018, to be one of the most significant authoritarian personalistic

events in the country's history. One of the most significant changes is the abolition of term restrictions for both the President and Vice President, allowing Chinese President Xi Jinping to serve indefinitely. Other changes include the inclusion of Xi Jinping Thought in the Constitution's Preamble, making Xi the "first Chinese leader ever to have his theories enshrined in the constitution during his own lifetime. (Ken 2018)"

The amendment removed Deng Xiaoping's earlier restriction on key leadership positions, which he imposed in 1982. Prior to 1982, key leadership positions were not subject to term limits. All but the chair of the Central Military Commission was given a two-term limit (10 years total) to reflect Deng's determination to lay a long-term institutional foundation for domestic stability and modernization, and to prevent the country from reverting to the days of "Mao Worship".

Rumors regarding President Xi's plan to violate Deng's decade-rules and attempt to stay in office after his second terms started circulating in international media a few years before Xi's constitutional revision proposal (Huang 2017). Despite this, there was no news in Chinese media. On February 25th, 2018, Xi surprised everyone by announcing in a dryly written proposal a constitutional amendment. The motion was passed by the 13th National People's Congress with almost unanimous votes on March 11th, 2018, with 2,958 votes in favor, two against, and three abstentions (France-Presse 2018).

Retirement Reform. The current statutory retirement age in China is 60 years for male workers and 55 years for female. The Chinese government has been preparing to raise the retiring age for both men and women to 65 years old.

The Chinese Academy of Labor and Social Security announced in 2008 that when "the conditions are ripe," the "relevant department" would postpone the retirement ages. The retirement age will be raised “small step by small step” commencing in 2010 for women and 2015 for males, with one year being added every three years. By 2030, the retirement age will have increased to 65 years old (Song 2008).

On March 12th, 2018, with the retirement reform plans still in the works, the Vice Minister of Human Resources and Social Security announced that retirement reform is an “irresistible trend”, and confirming that the plan will be implemented in a phased manner (Dong and Zhu 2018).

Many people believe that the objective of the retirement reform is to reduce the Chinese government's financial burden. The Chinese society is more likely to face the challenge of population ageing because of the One Child Policy's persistently low birth rate and rising life expectancy. As a result, delaying retirement age can save the government roughly 5 billion dollars in pension funds (Shen 2020).

Despite decades of public consultation, the concept continues to elicit strong resistance from Chinese individuals. Supporters of this reform initiative say that raising the retirement age is reasonable due to increased life expectancy. Opponents, on the other hand, argue that this strategy serves the national interest at the expense of individual Chinese citizens' interests: ordinary people will have to work and pay for social security for another five to ten years before they can retire and receive their pension.

Air Pollution in Beijing. Air pollution, or the smog, has been one of the biggest public health concerns for Chinese citizens in the recent decade. Beijing suffers from some of the worst air

pollution in the country due to its enormous industrial clusters and high population densities. As early as in 1990, the word “Beijing Cough”, a slang for a persistent throat tickle experienced by travelers to Beijing have spread among the foreign community and locals. Back then, however, the Chinese had not identified the smog in Beijing as the cause of the “Beijing Cough”.

In 2012, when the air quality in Beijing was unprecedentedly bad, the U.S. Embassy in Beijing started to post daily statistics on PM2.5, PM10.0 and Air Quality Index (AQI) on Sina Weibo. PM2.5, which was known very little by Chinese people back then, are tiny particles in the air that causes the smog and when inhaled, can post great health risks.

The posts by the U.S. Embassy in Beijing triggered strong protest by the Chinese government. The Chinese government accused the U.S. Embassy of “intervening in its internal affairs” and urged the U.S. Embassy to stop releasing data on PM2.5. The tension between the Chinese government and the U.S. Embassy nevertheless attracted great attention of the Chinese citizens and spiked heated discussion online. The rising public concerns forced the Chinese government to take actions. On March 4, 2014, the Chinese premier, Li Keqiang, promised at the National People’s Congress that he will “resolutely declare war against pollution (Greenstone 2018).”

The results of China’s anti-air pollution campaign are controversial. On the one hand, research show that anti-air pollution campaign made huge progress by reducing the concentrations of fine particulates in the air by 32 percent and raising the life expectancy of the residents by 2.4 (Greenstone 2018). On the other hand, the air quality still does not meet the expectation of Chinese citizens. On March 13th, 2018, when the 13th National People's Congress was held in Beijing, the levels of PM2.5 in Beijing reached 230 micrograms per square meter, 23 times higher than WHO's recommended daily level(北京空气质量指数月统计历史数据 2021).

Even though it was not a topic that the government actively solicited comments during the 13th National People's Congress, the air quality in Beijing and the anti-pollution campaign discussed intensively on the Internet.

6.2 Data

I use censored and uncensored social media post from Sina Weibo, often referred to as the Chinese equivalence to Twitter. Launched by Sina Corporation on 14 August 2009, Sina Weibo is a leading social media platform in China, with over 445 million monthly active users as of in 2018.

The censored Weibo posts are gathered from FreeWeibo, a website that makes available content Sina Weibo that has been censored and deleted by Chinese authorities under the Great Firewall. Its home page is constantly updated to show the latest most-censored Weibo content.

FreeWeibo makes public censored content from Weibo collected by WeiboScope, a Chinese social media data collection project founded in January 2011 based on Hong Kong University. Using a combination of computer programs on Linux and scripts that query the Sina Weibo API, WeiboScope obtained a random sample of Sina Weibo user accounts and their latest posts (Fu and Chau 2013).

In searching for relevant posts, I design a set of keywords targeting at each event. The selection of keywords is optimized by computer-assisted methods (King, Lam, and Roberts 2017). I first identify a list of keywords that are relevant to each event, and then collect posts containing each keyword. Among the posts collected, I select a random sample of 1,000 posts, label their relevance and fit a machine learning text classifier to predict the relevance of all collected posts. Finally, I select a list of keywords among all words in the relevant posts that are

contained in the most posts. For each event, I iterate the steps two to three times to mine for the optimal list of keywords.

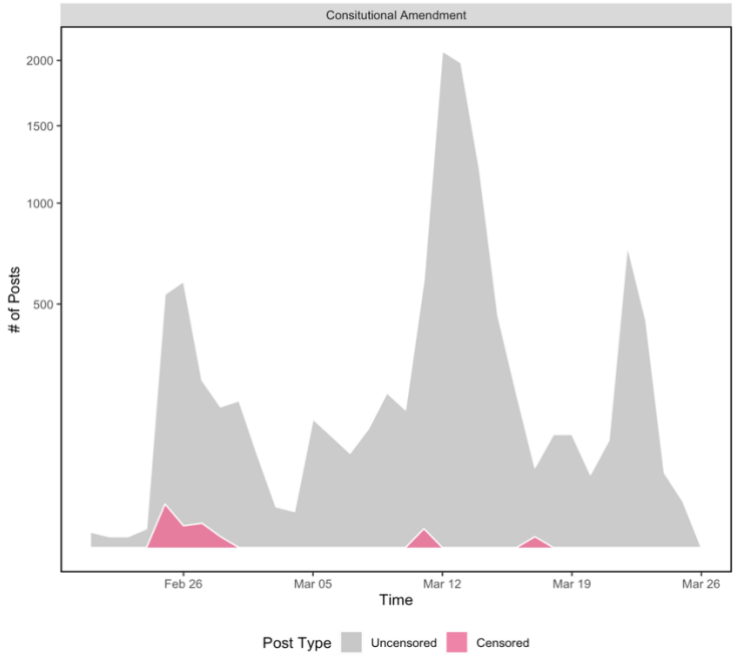
From the optimized keywords, I collected from Sina Weibo and FreeWeibo 56,216 posts on constitutional amendment, 9,642 posts on retirement reform and 8,139 posts on Beijing air pollution from February 21st, 2018 to March 26th, 2018. I then use machine learning to identify posts that are truly relevant. After testing various widely used classification algorithms, I choose among them the model with best performance (random forest classifier) to identify the relevance of the posts from a 1000 posts sample randomly drawn from posts on each event. Then I used the fitted model from the training sample to predict the relevance of all posts on each event. The analysis sample contains 10,832 posts on constitutional amendment, 1,159 posts on retirement reform, and 1,832 posts on Beijing air pollution from February 21st, 2018 to March 26th, 2018.

6.3 Analysis

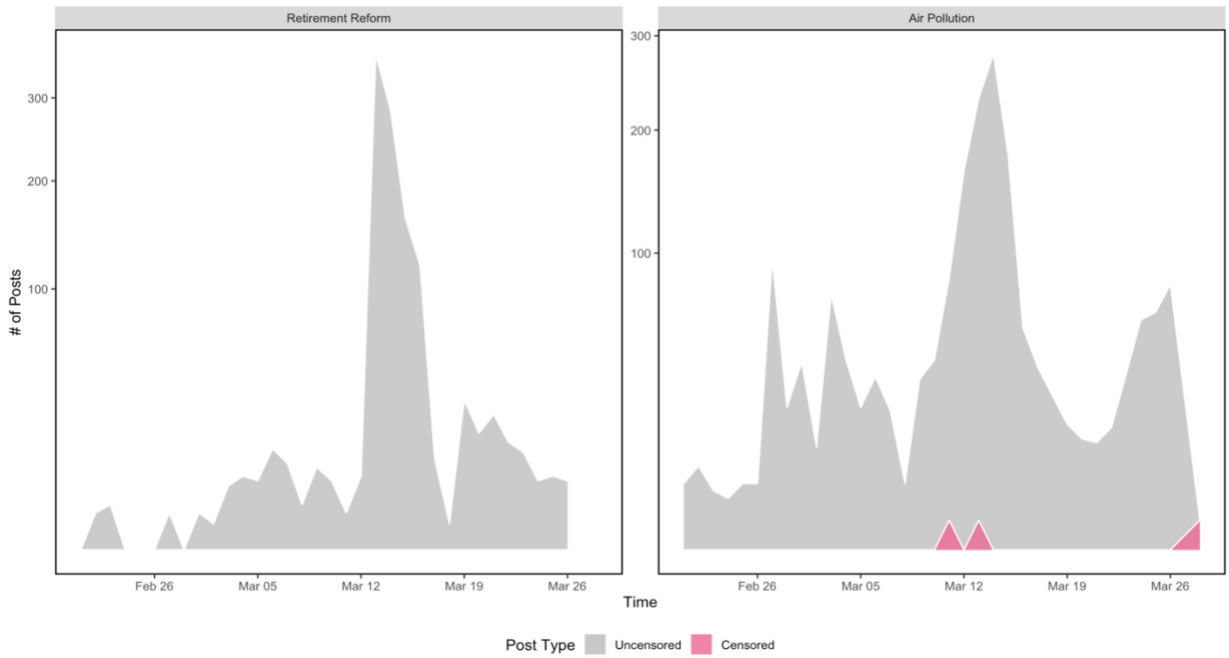
In this section, I present a time series plot of censored and uncensored posts. I then show a few examples of censored and uncensored posts, and finally present results from regression testing the relationship of topic of the post and the probability of it being censored.

Figure 1: Time Series Plot

(a)



(b)



Notes: Figure shows a stacked-area time series of uncensored and censored posts. Panel (a) plots the posts on the constitutional amendment, and panel (b) plots posts on the government's public goods provision. The left subplot of panel (b) shows the posts on the retirement reform, and the right subplot on the air pollution. Areas in different colors represent the number of posts in each category.

Figure 1 shows the time series area plot of posts. The subplot on the left shows the data on the retirement reform, and the subplot on the right shows the data on the constitutional amendment. The x-axis of the plots is date, from February 21, 2018, to March 26, 2018. The y-axis is the number of posts, quadratically scaled. Each observation in the plots shows the number of posts in each day. Different posts are represented by areas of different color shades. The grey area shows the number of posts that are uncensored public discussion, i.e., posts that are uncensored and non-propaganda. The pink area shows the number of censored posts.

We notice there are volume bursts around the days that the event occurred. For instance, there are two volume bursts in the postings of constitutional amendment: one around February 25th, 2018, when the proposal for constitutional amendment is released; the other around March 11th, 2018, when the constitutional amendment is proved in the 13th National People's Congress. Similarly, there is a volume burst on retirement reforms around March 12th, 2018, when the vice minister of the Ministry of Human Resources and Social Security announced that ministry was researching on the details of the retirement reform. There is also a volume burst on air pollution around March 12th, 2018, when the city of Beijing was hit by a serious smog.

We see from Figure 1 that the censorship rates on the constitutional amendment is higher than that of the retirement reform and the air pollution issue. There is a significant amount of public discussion on the retirement reform and the air pollution circulating on the internet, and few of them are censored. On contrary, public discussion on the constitutional amendment is mostly censored.

To test if the policy issue that the post mentions is associated with its censorship rate, I run a regression with censorship as the dependent variable, and the type of policy that the post mentions as one of the dependent variables. A possible confounder in this case is the sentiment

of the posts, i.e., whether the post supports or criticizes the policy issue. It is possible that the post is censored not because it mentions a sensitive policy issue, but because it criticizes rather than supports the government policy.

A possible confounder in this case is the sentiment of the posts, i.e., whether the post supports or criticizes the policy issue. It is possible that the post is censored not because it mentions a sensitive policy issue, but because it criticizes rather than supports the government policy. To control for the sentiments, I label the sentiment of each post in the analysis sample[1]. Sentiment is coded as -1, 0, or 1, where -1 means that this post opposes or criticizes the government policy, 0 means the post is neutral, and 1 means the post supports the policy.

I then fit the data with the following linear models:

$$c_i = \beta_0 + \beta_1 p_i$$

$$c_i = \beta_0 + \beta_1 p_i + \beta_2 s_i$$

$$c_i = \beta_0 + \beta_1 p_i + \beta_2 s_i + \beta_3 p_i \times s_i,$$

where c_i is a binary variable indicating whether the post i is censored or not, p_i is a binary variable indicating whether the post i mentions a personalistic issue or a public goods issue, and s_i indicates the whether the post i criticizes or supports the government.

[1] If the size of the sample of large, training a machine learning model can be extremely helpful in predicting the sentiments of the posts. However, since the analysis sample only contains 1,729 posts, I manually labelled them.

Table 5: Regression Results

	Dependent variable:		
	(1)	(2)	(3)
Personalistic Post	0.0136***	0.155***	0.425***
	(0.009)	(0.010)	(0.023)
Neutral Post		-0.013*	-0.004
		(0.007)	(0.007)
Positive Post		-0.045***	0.002
		(0.009)	(0.010)
Personalistic x Neutral			-0.276***
			(0.028)
Personalistic x Positive			-0.369***
			(0.027)
Observations	1,792	1,792	1,792
R^2	0.106	0.118	0.200
Adjusted R^2	0.105	0.116	0.198
Residual Std. Error	0.127 (df = 1790)	0.126 (df = 1788)	0.120 (df = 1786)
F Statistics	211.736*** (df = 1; 1790)	79.374*** (df = 3; 1788)	89.219*** (df = 5; 1786))

*p<0.1; **p<0.05; ***p<0.01

Note: Table shows the regression results. Column (1) shows the result from model 1 (equation 6). Column (2) from model 2 (equation 7), and column (3) from model 3 (equation 8).

We see from the results in Table 5 that, in model 1 when the event type is the only independent variable, it is significantly associated with the probability that the post is censored. Particularly, the estimate for event type is 0.136 with a standard error of 0.009 and is significant at p = 0.001. When the post discusses a personalistic issue, the probability of it being censored is 13.6% higher than if the post mentions a public good provision issue.

Controlling in for sentiment (model 2), the event type is still significant at the 0.001 level. The estimate for event type is 0.155 with a standard error of 0.010, suggesting that when the post discusses a personalistic issue, the probability of it being censored is 15.5% higher than if the post mentions a public good provision issue. Furthermore, positive are significantly less likely to be censored than a negative post, regardless of its theme.

Model 3 adds the interaction term of policy type and sentiment. Negative comments on the personalistic issue are significantly more likely to be censored than positive comments on personalistic. If a post is a negative comment, then the probability of it being censored is 42.5%

higher if it discusses a personalistic issue than a public good provision issue. On the contrary, if a post is positive, then the probability of it being censored is 5.6% higher if it discusses a personalistic issue than a public good provision issue.

6.4 Robustness Check

As a robustness check, I resample the data with replacement 1,000 times to estimate the coefficient and standard error in model3 above. I resample the data 1,000 times, runs the regression model 3 and saves the coefficients and standard errors from the model, and compute the means of the coefficients and the standard errors.

Table 6: Bootstrap

	Dependent variable:
Personalistic Post	0.428
	(0.023)
Neutral Post	-0.0037
	(0.0065)
Positive Post	0.002
	(0.0097)
Personalistic x Neutral	-0.277
	(0.028)
Personalistic x Positive	-0.373
	(0.027)
Observations	1,792
R^2	0.200
Adjusted R^2	0.198
Residual Std. Error	0.120 (df = 1786)
F Statistics	89.219*** (df = 5; 1786))

Note: Table shows the bootstrap estimates of coefficients and standard errors in model 3.

This table shows the bootstrap estimates of coefficients and standard errors in model. The first column shows the coefficients, and the second column shows the standard errors. Overall, the statistics are very close, which proves the stability of our results.

Chapter 7. Conclusion

The purpose of this paper is to contribute theoretically and empirically to the vast literature on authoritarian censorship. By extending the global coordination game, the article explains how an autocratic leader chooses between censorship and co-optation. In my framework, the incumbent commits to a level of censorship on the outside signal, revealing her intention to backslide, a level of censorship that allows or disallows public communication among citizens. The incumbent learns the optimal amount of public goods sufficient to mollify the citizens. Censoring public communication has an opportunity cost since it deprives the incumbent of information about the citizen's types to decide how much public goods to offer.

As the model demonstrated, that the government benefits from a high censorship level on signals revealing her personalistic intentions and a low censorship level on public communication that helps her decide the amount of public goods to offer, as far as the government is economically strong enough to provide citizens with the amount of public goods they desire.

The empirical evidence studies the censorship strategy of the Chinese government. It shows that while the Chinese leaders can be very welcoming of critiques on their public goods policies, they do not allow public discussion on policies through which the leader seeks to attack established institutions and backslide. Furthermore, the results suggest a positive relationship between censorship and propaganda. On topics where the censorship rate is high, we observe more propaganda.

This paper offers a new angle to think about leadership's censorship strategy. Leaders could be more sophisticated in choosing censorship levels than we may have expected. Rather than simple-mindedly deciding to allow or censor all negative comments, the leader strategically

chooses what materials to censor to maximize her political survival and economic interests. Moreover, the paper sheds light on how the leadership combines various tools in her ruling.

References

Besley, Timothy, and Andrea Prat. 2006. “Handcuffs for the Grabbing Hand? Media Capture and Government Accountability.” *American Economic Review* 96(3): 720–36.

Breslauer, George. 1992. “In Defense of Sovietology.” *Post-Soviet Affairs* 8(3): 197–238.

Brzezinski, Carl J. Friedrich, Zbigniew K. 1965. *Totalitarian Dictatorship and Autocracy*.

Cairns, Christopher. 2016. “Fragmented Authoritarianism? Reforms to China’s Internet Censorship System under Xi Jinping.” : 36.

Chen, Jidong, and Yiqing Xu. 2017. “Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?” *The Journal of Politics* 79(3): 792–803.

“Countries and Territories.” *Freedom House*. https://freedomhouse.org/countries/freedom-world/scores (May 6, 2022).

Cox, Gary. 2009. “Authoritarian Elections and Leadership Succession, 1975-2004.”

Desai, Raj, Anders Olofsgård, and Tarik Yousef. 2009. “The Logic of Authoritarian Bargains.” *Economics and Politics* 21: 93–125.

Dong, Qing, and Chuange Zhu. 2018. “延迟退休最新回应：这些人或将工作到 65 岁 部分职工可自主选择--2018 全国两会--人民网.” *People*. http://industry.people.com.cn/n1/2018/0314/c413883-29867345.html (November 28, 2021).

France-Presse, Agence. 2018. “Xi Jinping Can Be President For Life As China Parliament Ends Term Limits.” *NDTV.com*. https://www.ndtv.com/world-news/chinas-parliament-abolishes-presidential-term-limits-1822398 (November 28, 2021).

Fu, King-wa, and Michael Chau. 2013. “Reality Check for the Chinese Microblog Space: A Random Sampling Approach.” *PLOS ONE* 8(3): e58356.

Garrick, John, and Yan Chang Bennett. 2016. *China’s Socialist Rule of Law Reforms Under Xi Jinping*. Routledge.

Geddes, Barbara, and John Zaller. 1989. “Sources of Popular Support for Authoritarian Regimes.” *American Journal of Political Science* 33(2): 319–47.

Gehlbach, Scott, and Konstantin Sonin. 2014. *Government Control of the Media*. Rochester, NY: Social Science Research Network. SSRN Scholarly Paper. https://papers.ssrn.com/abstract=1315882 (December 5, 2020).

Greenstone, Michael. 2018. “Four Years After Declaring War on Pollution, China Is Winning.” *The New York Times*. https://www.nytimes.com/2018/03/12/upshot/china-pollution-environment-longer-lives.html (December 13, 2021).

Gueorguiev, Dimitar D., and Edmund J. Malesky. 2019. “Consultation and Selective Censorship in China.” *The Journal of Politics* 81(4): 1539–45.

Guo, Gang. 2009. “China’s Local Political Budget Cycles.” *American Journal of Political Science* 53(3): 621–32.

Huang, Zheping. 2017. “All the Signs That China’s Xi Jinping Is Planning on a Third Term.” *Quartz*. https://qz.com/1030850/all-the-signs-that-chinas-xi-jinping-is-planning-on-a-third-term/ (November 28, 2021).

Jia, Lianrui. 2019. “What Public and Whose Opinion? A Study of Chinese Online Public Opinion Analysis.” *Communication and the Public* 4(1): 21–34.

Ken, Suzuki. 2018. “China’s New ‘Xi Jinping Constitution’: The Road to Totalitarianism.” *nippon.com*. https://www.nippon.com/en/in-depth/a05803/ (November 28, 2021).

King, Gary, Patrick Lam, and Margaret Roberts. 2017. “Computer-Assisted Keyword and Document Set Discovery from Unstructured Text.” *American Journal of Political Science* 61(4): 971–88.

King, Gary, Jennifer Pan, and Margaret Roberts. 2013. “How Censorship in China Allows Government Criticism but Silences Collective Expression.” *American Political Science Review* 107(2 (May)): 1–18.

Kung, James Kai-Sing, and Shuo Chen. 2011. “The Tragedy of the Nomenklatura: Career Incentives and Political Radicalism during China’s Great Leap Famine.” *American Political Science Review* 105(1): 27–45.

Lawson, Chappell. 2002. *Building the Fourth Estate: Democratization and the Rise of a Free Press in Mexico*. Berkeley: University of California Press.

Levitsky, Steven, and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes after the Cold War*. Cambridge University Press.

Lorentzen, Peter. 2014. “China’s Strategic Censorship.” *American Journal of Political Science* 58(2): 402–14.

Lorentzen, Peter L. 2013. “Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime.” *Quarterly Journal of Political Science* 8(2): 127–58.

Pan, Jennifer. 2019. “How Chinese Officials Use the Internet to Construct Their Public Image.” *Political Science Research and Methods* 7(2): 197–213.

Press, Associated. 2016. “Xi Jinping Asks for ‘absolute Loyalty’ from Chinese State Media.” *The Guardian*. https://www.theguardian.com/world/2016/feb/19/xi-jinping-tours-chinas-top-state-media-outlets-to-boost-loyalty (May 6, 2022).

Qin, Bei, David Strömberg, and Yanhui Wu. 2017. “Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda.” *Journal of Economic Perspectives* 31(1): 117–40.

———. 2019. “Social Media, Information Networks, and Protests in China.” : 49.

Shen, Yan. 2020. “分析：中国渐进式延迟退休真的来了？ 如何落地恐怕并不容易.” *Reuters*. https://www.reuters.com/article/china-postpone-retirement-idCNKBS28307R (November 28, 2021).

Song, Fangcan. 2008. “‘65 岁退休’说引发网络热议 网友疑拄拐杖上班.” *Chinanews*. https://www.chinanews.com.cn/gn/news/2008/11-12/1445802.shtml (November 28, 2021).

“The Great Firewall of China: Background Torfox.” 2011. https://cs.stanford.edu/people/eroberts/cs181/projects/2010-11/FreedomOfInformationChina/the-great-firewall-of-china-background/index.html (May 6, 2022).

Thomala Lai Lin. 2022. “China: Social Media Penetration Rate 2021.” *Statista*. https://www.statista.com/statistics/234991/penetration-rate-of-social-media-in-china/ (May 6, 2022).

Xu, Beina, and Eleanor Albert. 2017. “Media Censorship in China | Council on Foreign Relations.” https://www.cfr.org/backgrounder/media-censorship-china (May 6, 2022).

Yahuda, Michael. 1972. “Kremlinology and the Chinese Strategic Debate, 1965–66.” *The China Quarterly* 49: 32–75.

“北京空气质量指数月统计历史数据.” 2021. *Aqistudy*. https://www.aqistudy.cn/historydata/monthdata.php?city=%E5%8C%97%E4%BA%AC (December 13, 2021).

Appendices

Appendix I. List of Keywords

The following keywords are used to search for and scrape posts:

Constitutional Amendment:

修宪(Constitutional Amendment), 宪法(Constitution), 修改(Amendment), 终身(life-long), x法(X law), 连任(reelection), 任期(term limit), 称帝(come to the throne), 袁世凯(Yuan Shikai)

Retirement Reform:

退休年龄(retirement age), 退休金(retirement fund), 养老金(pension), 延迟退休(retirement delay)

Beijing Air Pollution: 北京雾霾(Beijing smog), 两会雾霾(Two Sessions smog), 帝都雾霾(Capital smog), 北京空气(Beijing air), 两会空气(Two Sessions air), 帝都空气(Capital air)

Appendix II. Summary Statistics

Table 7: Summary Statistics

Event	% Censored	% Propaganda
Constitutional Amendment	14.15%	98%
Retirement Reform	0%	36%
Air Pollution	0.25%	16%

Appendix III. Classification: Identifying Relevant Posts

Table 8: Constitutional Amendment Classification

Model	**Accuracy**
Naive Bayes	79.75%
SVM	85%
Logistic Regression	86.75%
KNN	84%
Random Forest	**90.5%**
XG Boost	83.25%

Table 9: Retirement Reform Classification

Model	**Accuracy**
Naive Bayes	83.5%
SVM	89.5%
Logistic Regression	90.25%
KNN	86%
Random Forest	**93.25%**
XG Boost	86.75%

Table 10: Air Pollution Classification

Model	**Accuracy**
Naive Bayes	74.75%
SVM	79.75%
Logistic Regression	**81.25%**
KNN	75.75%
Random Forest	81%
XG Boost	77.5%

Appendix IV. Propaganda

Social media is not only a space for citizens to express their opinions on policy issues, but also a tool that the government uses intensively for propaganda purposes. In Sina Weibo, the Chinese government spread propaganda through government owned or affiliated institutions, including state media and party affiliated organizations. For example, on March 11th, 2018, the day when constitutional amendment was passed by the 13th National People's Congress, Sina Weibo was flooded with government propaganda promoting the righteousness of the policy. Here's an example of a propaganda post published by People's Daily, one of the biggest state media, and re-posted by thousands of other users:

宪法是治国安邦的总章程，是党和人民意志的集中体现。在保持宪法连续性、稳定性、权威性的基础上，推动宪法与时俱进、完善发展，这是我国法治实践的一条基本规律。

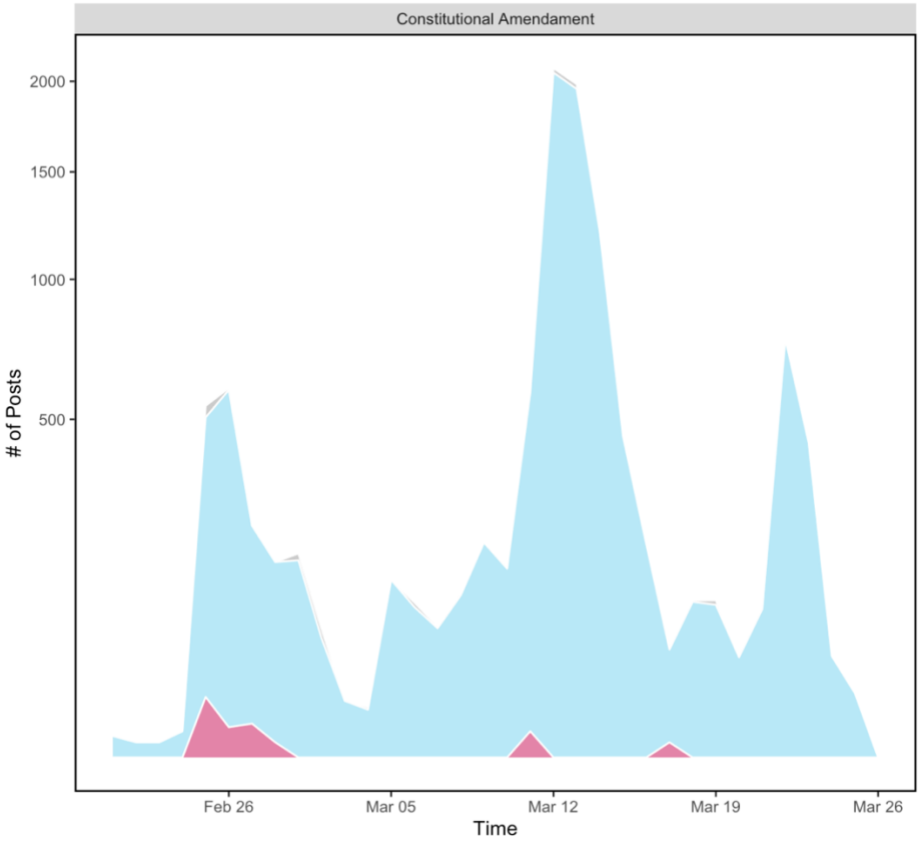
"The constitutions are the primary agenda for governing the country and is the concentrated reflection the will of the Party and its people. In addition to ensure the consistency, stability, and authority of the constitutions, facilitating the constitutions to developed with time is a basic principle of the legal practice in our country."

Since our interest is to examine the censorship on public discussion, we need to distinguish between public discussion and government propaganda. The method to identify propaganda posts is like that identifies relevant posts. A potential issue with the identification of propaganda posts is that, on highly sensitive policy issues, the proportion of propaganda posts can be as high as over 95 percent among all relevant posts. Applying machine learning classifier on these highly skewed sample often yield a less ideal result especially when our goal is to identify the minority. To solve this issue, I first exclude from the analysis sample all posts

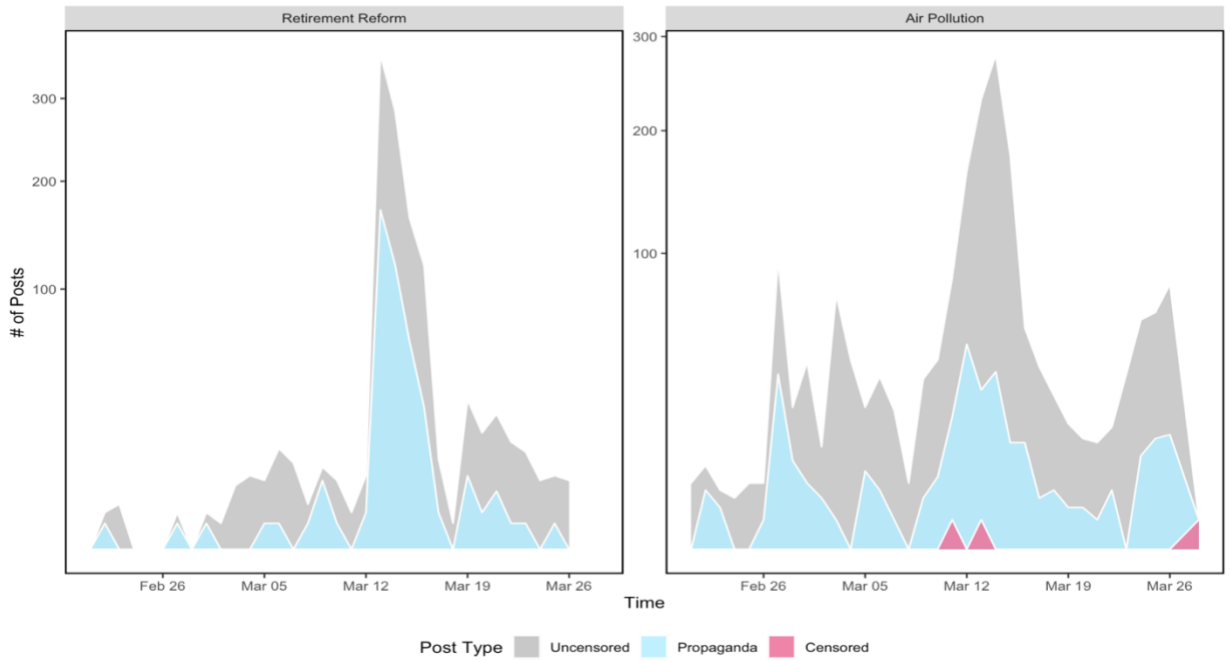
published by government affiliated users and mark them as propaganda. The remaining sample published by individual users is more balanced of propaganda and non-propaganda posts. From there, I select a random sample of 1,000 posts and label posts as propaganda if they are direct quotes and re-posts from states news and propaganda. Finally, I apply machine learning classifier to training sample use the fitted model to predict propaganda posts.

Figure 2: Time Series Plot with Propaganda

(a)



(b)



Note: Figure shows a stacked-area time series of uncensored public discussion, propaganda, and censored posts. Panel (a) plots the posts on personalistic (the constitutional amendment), and panel (b) plots posts on the government's public goods provision. The left subplot of panel (b) shows the posts on the retirement reform, and the right subplot on the air pollution. Areas in different colors represent the number of posts in each category.

Figure 2 shows the time series area plot of posts, including propaganda posts. The blue area shows the number of propaganda posts. The relationship between propaganda and censorship confirms with the recent findings on positive correlation between censorship and propaganda (e.g., Gehlbach et al. 2021).

Appendix V. Example of Posts

Here I provide some examples of censored and uncensored public discussion posts in the sample to show what is going on in the Chinese social media.

First, here is a censored post that strongly opposes the constitutional amendment:

我坚决反对对宪法第十四条的修改。这次对任期的修改是 du cai 统治的体现。

推翻帝制百余年，改革开放四十年，不能再有这样的体制。

"I firmly object the amendment on Article 14 of the constitution. The change in term limit is a signal of dictatorship. A hundred years after we abandon the monarchy, and 40 years after the economic reform, we cannot have this kind of political system again."

Here is another censored post which, on the contrary, strongly supports the constitutional amendment and compares President Xi Jinping with other strongman leaders in the world:

看到一群人在意淫，瞎掰活。不懂吧，看看世界潮流。普京从出任总理迄今已经 19 年，仍在职。默克尔从首次成为总理到如今执政，已经 13 年。日本的安倍晋三 2006 年首次出任首相到如今执掌日本首相，迄今 12 年。新加坡的李显龙任职已经 14 年。世界各国首脑，只要水平尚可获得人民认可的，其任期就突破 10 年。

"I see many people talking nonsense. You guys don't understand. Look at the trend worldwide. Putin has been president for 19 years, and still in office. Merkel has been president for 13 years. Shinzo Abe has been in office for 12 years since he first became the Prime Minister of Japan. Lee

Hsien Loong has been the Prime Minister of Singapore for 14 years. Heads of government over the world, as far as they are capable and recognized by their people, will be in office for more than 10 years."

On the other hand, there are many criticisms of the public goods provision circulating on the internet. Here is an example of uncensored post that opposes the retirement reform:

延迟退休极为不人道 现在中国的工厂 2 班 12 小时工作超越人体极限生命早已燃烧殆尽 已不在是人 是工具比牲畜差不多

"Raising the retirement age is extremely inhuman. Working in Chinese firms with two shifts of 12 hours is well above the limit of human beings. Lives are already burn out. We are not human anymore, but rather like tools and animals"

Another uncensored post opposing the retirement reform points to the privileges of high-ranking government officials as the root of the problem:

取消国家省部级高级干部的待遇终身制对于医疗改革非常重要 如果不是他们花费 70%的资源和费用 老百姓也就不需要延迟退休了 财政部也就不需要搞延迟退休哦 因为养老金缺钱啊 被他们花掉了 这对底层公民不公平的

"It is very important to terminate the life-long benefits of the provincial and ministerial senior officials. If they did not use 70% of the resources and funds, then the common people won't need to retire later, and the Ministry of Finance won't need to propose a raise in retirement ages. Because the pension fund is short, and they (the senior officials) spent it. This is unfair for the citizens."