Censoring the Intellectual Public Space in China: What Topics Are Not Allowed and Who Gets Blacklisted?

Xiaojun Yan and La Li

Censorship is one of the main forms of political coercion deployed by modern states to control and regulate public expression. In this article, we examine the political censorship of China's intellectual public space, which has long been underexplored. We apply unsupervised machine learning to examine the database of a leading intellectual portal website, which serves as an archive of both published and censored intellectual writings between 2000 and 2020 and includes over 740 million Chinese characters. We identify a strategic censorship mechanism that consists of thematic and persona censorship elements. Thematic censorship involves the state filtering out writing that competes with the official policy narrative, historiography, and values. Persona censorship involves the complete muting of individual intellectuals who have previously made derogatory attacks on the supreme leaders of the Communist Party, which represents a symbolic act of open defiance.

Censorship is one of the main forms of political coercion deployed by modern states to control and regulate public expression. Over the past two decades, China has developed the largest state censorship operation in the information age. The logic and *modus operandi* behind China's vast network of online surveillance and control have been extensively examined. Studies have indicated that, as an instrument of political coercion, censorship is used selectively and strategically by the state in China to quash undesirable political expression (e.g., King, Pan, and Roberts, 2013; Lorentzen, 2014; Roberts, 2018; Han and Shao, 2022). However, the criteria used by state censors in strategic censorship remain a matter of

debate. Various conflicting theories about which topics are not allowed and who gets blacklisted by the Chinese state have been proposed.

For example, King, Pan, and Roberts (2013; 2014) argue that the collective action potential of a social media post is the decisive factor that leads to its deletion by the state, while online criticism against the state, its leaders, or its policies without such potential is often tolerated. Gueorguiev and Malesky (2019) argue that if online criticism on social media platforms is solicited by the state regarding specific topics during officially designated consultation periods, it may be tolerated, but not if it is unsolicited and outside of the state-permitted time windows. Tai and Fu (2020, 18) note that social media messages with higher "specificity" — the extent to which they involve specific terms — along with those that signal internal or external conflicts, are more likely to be censored, to prevent such discussions from becoming "focal points" that can encourage readers to "think toward undesirable directions". Gallagher and Miller (2021) suggest that state censors often target online public opinion leaders with greater socio-political influence, who are more likely to prompt viral discussions on topics that may challenge the hegemony of the state. Esberg (2020) focuses on the historical case of Chile and argues that the preferences of key constituencies of the state — particularly their moral values — may have influenced censorship decisions.[1]

Unlike other studies, we identify two factors that are essential for discerning these variations in the criteria of

---

Xiaojun Yan 🔟 *is an Associate Professor of politics and public administration at the University of Hong Kong (xyan@hku.hk). He obtained an AM and a PhD in political science from Harvard University. He is a comparative political scientist with special expertise in the politics of China.*

La Li 🔟 *is a PhD Candidate in Department of Politics and Public Administration, University of Hong Kong (larry95@hku.hk). His research interest focuses on China's propaganda system.*

strategic state censorship: the scope of the censorship decision and the context in which political expression takes place. We argue that neither factor has received sufficient scholarly attention to date. First, the *scope* of a censorship decision may affect the criteria, as the daily removal of content is likely to be based on substantially different criteria from that applied when preventing an author from making any public expression, which is a much rarer event. Although account blocking and content deletion have been rightly distinguished in the literature on China's censorship practice (e.g., King, Pan, and Roberts, 2014; Tai and Fu, 2020), why and how censorship decisions with different scopes are made and implemented has not been fully examined so far. State censors selectively delete certain topics and tolerate others, but their decisions to completely mute and erase a specific author from the public domain, regardless of what they actually write, reflects an "offense" allegedly committed by the victim that is implicitly fundamental and non-negotiable in the eyes of political authorities. Distinguishing the rationales behind censorship decisions aimed at removing specific content and those that blacklist individual authors is necessary.

Second, the *context* of the political expression matters. The literature has sufficiently implied that context is highly consequential for political expression, as the specific spatial-temporal structure in which political expression takes place may mitigate or magnify the power of even the same expressive activity (Chang and Manion, 2021; Han and Shao, 2022). Insulting or spreading rumors about a sovereign in the context of a masquerade is fundamentally different from making the same criticisms through openly lèse-majesté remarks in *Forum Romanum.* Thus, the censorship criteria applied by state censors may need to be differentiated according to the context of political expression. In the information age, social media platforms and intellectual portal websites provide different venues for such expression. Social media platforms are communicative, interactive, and often anonymous networks through which opinions can be delivered in few words and often address immediate concerns, while intellectual portal websites publish much longer articles and include distinct author identities, effectively serving as a basis for broadcasting intellectual ideas. Through their writings, non-anonymous intellectuals who are "thought leaders" equipped with "agenda-setting power" (Gallagher and Miller, 2021, 1019) exercise their moral leadership on a public platform. Their open defiance, as illustrated through making derogatory remarks about heads of state under their real names, is much more politically and symbolically significant than the spreading of rumors about leaders by unnamed Internet users on one of the many social media platforms. Thus, state censors may want to apply a different set of censorship standards based on the specifics of the context or platform.

In this research, we apply unsupervised machine learning to examine an unprecedented backend database leaked from a leading Chinese intellectual portal website ("the website" hereafter), which contains a comprehensive collection of public intellectual writings from the past two decades.[2] Since its founding, the website has tasked itself with collecting and republishing a thorough collection of Chinese intellectual writings that have been published elsewhere; it has thus accumulated a relatively complete digital archive of such writings. With the leaked data from the website, we construct a database containing the full text of every article collected and published on the website between January 1, 2000 and August 1, 2020. This consists of both publicly viewable and censored articles (made unavailable to the public). The database contains about 740 million Chinese characters in 144,280 articles written by 28,494 authors. Among these, 5,406 articles by 769 authors have been censored by order of the state regulators. The corpus of the censored texts contains more than 23 million Chinese characters and, to the best of our knowledge, is the largest of its kind, and thus provides a rare opportunity to examine state censorship of intellectual political expression in China.

Our research demonstrates that state censorship in the Chinese intellectual public space consists of two elements: the selective deletion of articles based on their content, referred to as thematic censorship, and the complete blacklisting of some public intellectuals, or persona censorship. We find that antithetical narratives concerning basic national policies (*jiben guoce*), official historiography, or values advocated by the state are more likely to undergo thematic censorship, while a previous record of making derogatory attacks on the supreme leaders of the Communist Party appears to be the main predictor of personal censorship. We also find that factors such as the topic discussed, the influence of the author, whether they had participated in major national resistance movements, overseas work or study experience, and belonging to the "political establishment" have little or no effect on the Chinese state's decision to completely silence an individual author.

Theoretical Contribution

Through this research, we make three theoretical contributions. First, we deepen the scholarly understanding of state censorship by distinguishing two understudied censorship mechanisms in a relatively under-explored field. Previous studies have focused on investigative media, primarily covering localized incidents, or blog posts and social media, which represent more of a popular discourse and general mood and often provide information about incidents that may not have appeared in the regular media. We examine the elite public intellectual discourse, which plays an important but different role, shaping the views of

both citizens and elites on where the country should go; this research thus extends the study of censorship beyond social media to the landscape of long-form articles and sensitive public debates. Differing from findings that even vitriolic criticisms against the top Chinese leaders would not be censored so long as they do not possess mobilization potential (King, Pan, and Roberts, 2014)[3], we find that public intellectuals on the website who make personal attacks against the supreme leaders of the Chinese state suffered the most severe penalty possible — being completely erased from the public domain. In addition, we do not find evidence that influential opinion leaders (Gallagher and Miller, 2021) or articles with more specificity (Tai and Fu, 2020) are more likely to be censored — practices that have been convincingly revealed and demonstrated by prominent research into the Chinese social media universe over the past decade. Instead, intellectual writings against official policy lines, approved historical narratives, or state-advocated moral values fall victim to state censorship. Our findings suggest that in the intellectual public space, a quite different set of censorship criteria is used by Chinese censors.

Second, we enrich the theory of strategic repression and targeted coercion of modern states. Rightfully, studies have highlighted that when states deploy coercive power to realize their political goals, they often do so with strategic precision and adaptability, so as to reduce the potential cost and amplify the deterring effect incurred by such undertakings (e.g., Greitens, 2016; Xu, 2021; Pop-Eleches and Way, 2023). However, relatively less has been said about how states tailor the use of their coercive capacity to different contingencies. Through an empirical analysis of China's state censorship system, we discover two critical yet long-overlooked factors that shape the state's strategic deployment of coercive power: the context in which state coercion takes place and the scope, or intensity, of such undertakings. Taking state censorship as an example, we demonstrate that the very standard applied by state regulators when they make censorship decisions varies substantially according to the venue (social media vs. intellectual portal sites) and scope (content deletion vs. author ban) of such decisions.

Third, we also contribute to the theory of authoritarianism by discerning the priority of authoritarian state concerns over different kinds of political threats with convincing empirical evidence. Censorship criteria often credibly expose the intention of the state, particularly the state's perception of political threats (King, Pan, and Roberts, 2013). By comparing the standards behind the Chinese state's undertakings of censorship at different levels of intensity, we empirically demonstrate how authoritarian rulers perceive and rank political threats of different nature — at least in the intellectual public space. In this research, we reveal that there are two mechanisms of censorship: "thematic censorship" and

"persona censorship." In *thematic censorship*, only the specific content that challenges the official discourse of the state is deleted. *Persona censorship* involves the complete ban of a particular intellectual who has openly ridiculed the top leadership of the state in the public domain. The different levels in the severity of censorship of antithetical discourses versus that of discourses against lèse-majesté effectively show that state authorities perceive the latter as a far more grave threat and more severe political trespassing. The personas of authoritarian leadership, both past and current, are still at the central position of the symbolic authority of an authoritarian regime — the open violation of which is to be firmly nipped in the bud.

Censorship In The Intellectual Public Space

State censorship and the persecution of intellectuals is a global phenomenon with a long history. In the Roman Empire, scholars and philosophers who violated the majesty of the sovereign would be exiled and silenced, their works burned to ashes (Cramer, 1945). The history of states silencing intellectuals and banning their writings extends from Tudor and Stuart England (Cressy, 2005) to Ancien Régime France (Kelly, 1981), from the revolutionary regimes of Cuba (Black, 1989) and Mexico (Camp, 1981) to the theocracy of Iran (Kurzman, 2001), and from the underdeveloped Zimbabwe (Ngoshi, 2021) and Eritrea (Schmidt, 2010) to the more prosperous Singapore (Tan, 2016). A recent atrocity is the tragedy of Jamal Khashoggi, a public intellectual and commentator who was cruelly murdered for criticizing the Crown Prince of Saudi Arabia in his published writings (Martinez, 2018).

Intellectuals speak to society and offer moral leadership through public writing, which is an important form of political expression. Modern states face the dilemma of increasingly being dependent on the practical knowledge of intellectuals, who may also be major critics of how the state operates, thus calling into question the legitimacy of the social order and its political structure (Lipset and Dobson, 1972). No ruler in modern times can risk completely closing down the intellectual public space, but they also cannot afford to take a laissez faire attitude toward their national intelligentsia. Permitting public writing about certain topics in certain degree of scope "will always be preferable to complete censorship" (Lorentzen, 2014, 403).

For modern states, censorship must be tailored to the specific context. As Gallagher and Miller (2021, 1012) note, "the state enforces information control and repression with a scalpel rather than a hammer". Indiscriminate censorship is likely to backfire and induce a range of negative consequences that undermine the state. For instance, such censorship may attract even more attention to the prohibited content (Hobbs and Roberts, 2018), harm the credibility of the state's disclosed information

(Gläßel and Paula, 2020), further mobilize societal resistance (Pan and Siegel, 2020), or block crucial information channels that allow rulers to learn about underlying grievances in the population (Egorov, Guriev, and Sonin, 2009; Dimitrov, 2017). State censors are also found to possess various instruments to implement strategic and adaptive censorship, including the total blocking of specific information sources (MacKinnon, 2008), selectively deleting writings and messages deemed to be offensive to the state (Stockmann, 2013), distracting audiences' attention from the prohibited content by deploying state-sponsored "trolls" (Han, 2015; King, Pan, and Roberts, 2017), adding friction to the public's access to undesirable information (Roberts, 2018; Sanovich, Stukal, and Tucker, 2018), or undertaking behind-the-scenes censorship by outsourcing some of the operations to the private sector (Zhao, 2000; Sun and Zhao, 2021; Ruan et al., 2021). The state may also alter its censorship strategy to signal to other countries its change of approach (Weiss, 2014; Cairns and Carlson, 2016) or the venting of social frustration (Hassid, 2012).

Given its distinctive nature as a political expression venue, the intellectual public space is critical in a state's censorship strategy. Unlike popular public spaces, in which participation in collective deliberation and action is mostly anonymous, the intellectual public space involves members of the intelligentsia exerting political influence through either discourse, which shapes the ideological and moral landscape of a nation, or through iconic symbols of overt defiance (Finkel, 2007). Its connective structure is a radiating network in which individual intellectuals are the major nodes of influence. Intellectual writing is politically significant, as it can help to disseminate alternative discourses that may conflict with the official rhetoric (Davies, 2007; Zarycki, 2009), lead to the development of a coherent dissident group (Flam, 1999), or cultivate the next generation of anti-regime youth (Wasserstrom, 1991).

This distinction between popular and intellectual public spaces had led censorship mechanisms to adapt to the specific pathways of influence, sources of power, forms of content, and connective structures embedded in specific venues for political expression. In the following, we assess the content of a leading intellectual portal site in China and apply unsupervised machine learning to examine the censorship regime that the Party-state of China imposes on the nation's intellectual public space. We can then identify the criteria that the state uses to determine which topics cannot be discussed and who gets blacklisted.

Data

The database we analyzed is leaked from the backend database of one of China's leading portal websites for conceptual critiques, op-eds, and current affairs commentaries (Yan and Li, 2023). The website serves as a *de facto* archive of intellectual work in the social sciences and humanities and of serious discussions of current affairs and state policies. The website reprints content from other online intellectual platforms and strives to republish a comprehensive collection of Chinese intellectual writings. Our dataset can thus be best understood as a collection, archive, or digital library of China's public intellectual writings between 2000 and 2020.[4]

Given its influence, the website is watched closely by state censors. Three Party-state agencies (and their local branches), the Central Propaganda Department of the CCP, the Office of the Central Cyberspace Affairs Commission of the State Council (*zhongyang wangxinban*), and the Internet police of the Ministry of Public Security (*wangjian*), have the authority to censor any item published on the website that is deemed inappropriate. When a censorship order is issued, the agency demands swift deletion, and failure to do so on the part of the managerial team of the website may result in penalties in the form of fines or a temporary shutdown of the website. Overall, the observation and censorship mechanism in place for the website is carefully applied and is always operational.[5] However, although the censored articles disappear from public view, they are nevertheless stored in the backend database of the website. This enables us to discern the censored from the uncensored (and thus publicly viewable) articles.

We construct a database containing all articles that have ever appeared on the website. Research highlights the difficulty of obtaining reliable data "about both what was banned and what was permitted" (Esberg, 2020, 825), particularly over a long period (King, Pan, and Roberts, 2013). Our dataset addresses this through a clearly labeled set of published and censored articles. This offers a unique opportunity to study the Chinese Party-state censorship mechanism deployed in the online intellectual public space over a continuous 20-year period, and thus almost from its inception.[6] The database contains the main texts, author names, numbers of clicks, and publication dates of all 144,280 articles. Among these, 138,874 items written by 28,290 authors survived state censorship and were publicly viewable on August 1, 2020, while 5,406 articles written by 769 authors were published but later deleted following the instructions of the state censors. The overall censorship rate is thus 3.89%. Table 1 provides a summary of the database.

Three important caveats about the scope of this research should be mentioned. First, we are aware that self-censorship is a pervasive phenomenon. Contributors to both social media websites and intellectual portal sites self-censor their work to varying degrees. Website managers and editors also exercise censorship during the selection

Table 1
Summary of the website Database

	Uncensored Authors	Partially Censored Authors	Completed Muted Authors
Authors	27,725	565	204
(Active Authors[Note 2]*)*	(5,966)	(425)	(35)
Published Articles	111,798	27,076	0
(By Actives Authors)	(90,548)	(26,166)	(0)
Deleted Articles	0	1,939	3,467
(By Active Authors)	(0)	(1,454)	(3,303)

Note 1: This table summarizes the censorship status of authors in the website database. The Uncensored Authors category consists of authors whose articles were all accessible on the day we collected the data (August 1, 2020). The Partially Censored Authors category consists of authors for whom some of their articles were accessible, but some were not accessible to the public. The Completely Silenced Authors category consists of authors who had all of their articles deleted. Deleted articles unavailable to the public are permanently stored in the backend database.
Note 2: The category Active Authors consists of authors who have three or more articles published on the website.

process based on their understanding, best knowledge of, or even guesswork about the state censorship criteria. In other words, topics that are frequently censored are those that authors misjudged to be within-bounds but are then proved not to be. In this research, we focus on state censorship (i.e., the state's proactive attempts to regulate, control, and shape public expression in the intellectual public space) and regard self-censorship as a constant.

Second, due to the nature of the website and the data, we primarily examine *post hoc* censorship (implemented after an article is being published) rather than *ex ante* censorship (implemented before an article is published). As the data source collects and reprints articles from all over the Internet, the articles being gathered and published by the website may have already survived one or more rounds of censorship elsewhere — particularly the automatic keyword filtering censorship system, customarily called the "Great Firewall." This means that we may have underestimated the censorship rate. This concern is nonetheless alleviated by the fact that intellectual writings are normally long and sophisticated texts. Scholars have long argued that text censorship relies more on the hand-censoring of state censors and less on automatic keyword filtering (King, Pan, and Roberts, 2017). This allows a time window for the website to collect the articles in question and leave a record of their censorship in the database.
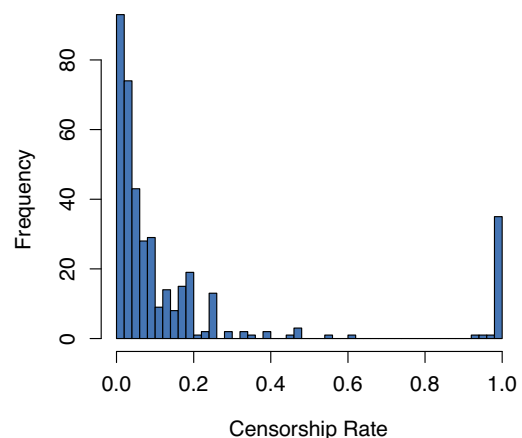
Third, as the data do not contain precise records of the times when deletion instructions were issued, we cannot conduct a strict time series analysis of the dynamics of censorship over small timescales. However, as confirmed in the literature on state censorship, instructions to delete an article are typically issued within 24 hours of said article's first appearance on the website (King, Pan, and Roberts, 2014). Thus, for the purpose of this research, we can safely assume that the censorship time is roughly the same as the publication time. We conduct an explorative analysis of the time variation in the censorship of articles

regarding China's One-Child Policy based on this assumption.

Two Types of Censorship Instructions

Two types of instructions are issued by the state censors. In one type of order, the censoring agency specifies the title of the article in question and requests its immediate removal. In the other type of censoring order, the censoring agency demands that all articles written by a particular author be deleted, regardless of topic. In the latter scenario, the censoring agency also demands that the author in question be banned from future publication on the website. The distribution of censorship rate by author indeed shows a clear bimodal pattern (see Figure 1), which identifies one group of completely silenced authors (*censorship rate* = 1.0)

Figure 1
Binomial Distribution of Censorship Rate by Author

Notes: This figure shows the censorship rate of active authors on the website. The censorship rate demonstrates a binomial distribution.

5

and another group of authors who only have some of their writings censored.

Put simply, in China's intellectual public space, some articles are censored because of their content and other articles are censored because of their author. This raises two interesting questions. First, what criteria does the state use to determine what content should be deleted? Second, what are the reasons for blacklisting some authors?

What Topics Are Not Allowed?

We use topic modelling to identify the relative frequency of the censoring of different topics. The basic assumption is that if only some of a particular author's publications are censored, the censorship decision is based on the content of the deleted articles and not on the identity of the author. We call this group of authors the "partially censored" authors. To find out which topics are more likely to be censored by the state, we compare the content of the censored articles written by the partially censored authors with the content of all of the publicly viewable articles on the website as of August 1, 2020. We estimate the contributions of each topic identified from this corpus to the eventual censorship decision. This process, described below, produces a list of topics ranked by censorship magnitude.

We construct a corpus of the 1,939 deleted articles written by the partially censored authors and the 138,874 articles that survived state censorship. The articles published by completely silenced intellectuals are excluded from this subsample, as they may not have been censored because of the content of the articles.

The voluminous size and high dimensions of this corpus pose challenges to conventional text mining methods. On average, each article in our dataset contains 5,132 Chinese characters. This is much longer than the length of social media posts, which are the data used in most research on China's censorship mechanisms (as a reference, each post on *Weibo* — the Chinese version of *Twitter* — allows a maximum of 144 Chinese characters).

To reduce the dimension of the textual data, we first deploy an advanced graph-based ranking algorithm, *TextRank*, for text pre-processing. *TextRank* creates abstracts for each article using between 50 and 200 keywords extracted from the text.[7] The number of keywords is proportional to the length of the article — the longer the article, the more keywords are selected. This method reduces the noise created by long texts with minimal sacrifice of meaning and interpretability (Milhalcea and Tarau, 2004).

We then use the Latent Dirichlet Allocation (LDA) topic modeling approach developed by Blei, Ng, and Jordan (2003) to transfer each article into a vector of 160 topics; thus, each article could be seen as a distribution over 160 topics.[8] To identify the most frequently and least frequently censored topics, we use a logistic regression model where the dependent variable is whether article x is censored, and the independent variable is the topic

distribution of article x.[9] This way, we are able to identify each topic's contribution to the state's decision to censor (or not censor) an article, namely the "censorship magnitude" of each topic. When a topic has a greater censorship magnitude, articles discussing this topic are more likely to be censored.
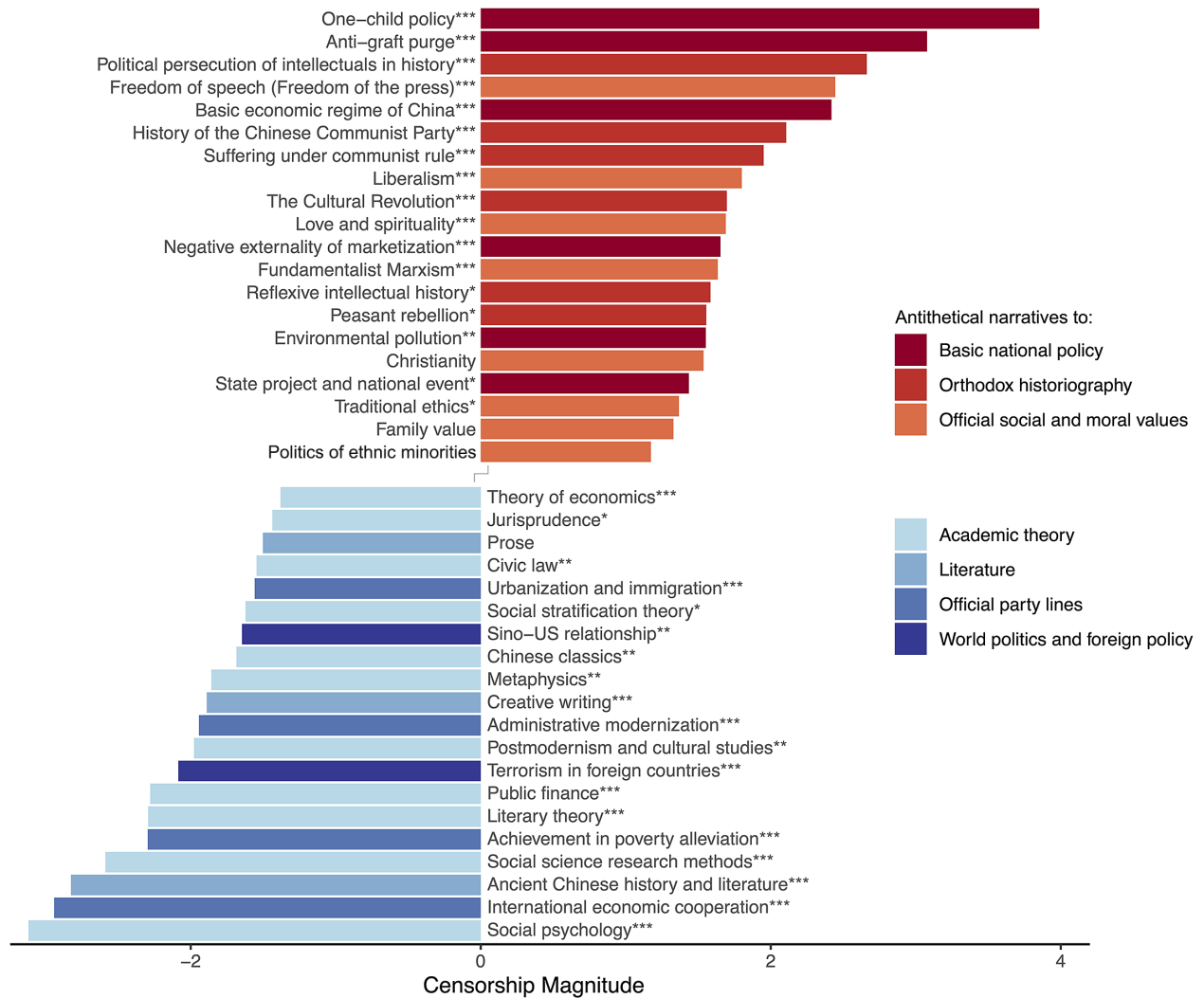
Thematic Censorship: Prohibited Topics

Figure 2 shows the topics that have the most and least contribution to the censorship of an article (for details about the topics, see Appendix I). Three findings are particularly salient. First, for intellectual writings, the state censors are most likely to block discussions about alternative policies that challenge the basic national policies, such as those questioning the scientific validity of the One-Child Policy, arguing against the post-2012 anti-corruption campaign under the presidency of Xi Jinping (*e.g., accusing the campaign as a disguised political purge*), commenting on the grave inequality under China's socialist market economy, disclosing the social costs incurred by the state's environmental policies, or protesting decisions about important state projects and national events (*such as the 2008 Beijing Olympics*).

The reversal of the One-Child Policy affords a valuable opportunity to test our findings. Since its establishment in the early 1980s, the One-Child Policy has gained increased significance in China's official discourses and was framed as a "basic national policy." This policy was strictly enforced by the Chinese state from the 1980s until its swift about-face in the early 2010s (Mattingly, 2020), throughout which time the policy remained the most censored topic in our dataset. When a basic national policy sees a quick turnaround, public discussion and intellectual writing still need time to adjust their direction (Yan and Li, 2023). Given this discursive inertia, we expect state censors to be busier censoring antithetical writings in the intellectual public space that are no longer compatible with the new policy direction of the state. In other words, we expect greater volume and higher frequency of censorship undertakings around critical moments of policy turnaround.

Our dataset shows exactly this pattern. Three critical points mark the course of the turnaround of China's One-Child Policy. On each of these three occasions, we observe a significant peak in the censorship rate of the topic of the One-Child Policy. In 2008, the Chinese state openly hinted the possibility of "reconsidering" the One-Child Policy. In 2013, the gradual relaxation of the policy started, and a "Conditional Two-Child Policy" was put in place. In 2018, the state bureaucracy in charge of the enforcement of the One-Child Policy — the National Commission on Family Planning — was merged with the National Health Commission (Alpermann and Zhan, 2019). Figure 3 shows the total number of publications on the One-Child Policy and the percentage that have

Figure 2
Topics with Highest and Lowest Censorship Magnitude

Notes: This Figure illustrates the 20 topics with the highest censorship magnitude (in red) and the 20 topics with the lowest censorship magnitude (in blue) of the 160 topics identified by the LDA topic model. The asterisks after the topic label indicate the statistical significance of the topic's censorship magnitude. The color of each bar indicates the general topic category. More details of the topics can be found in Appendix I.
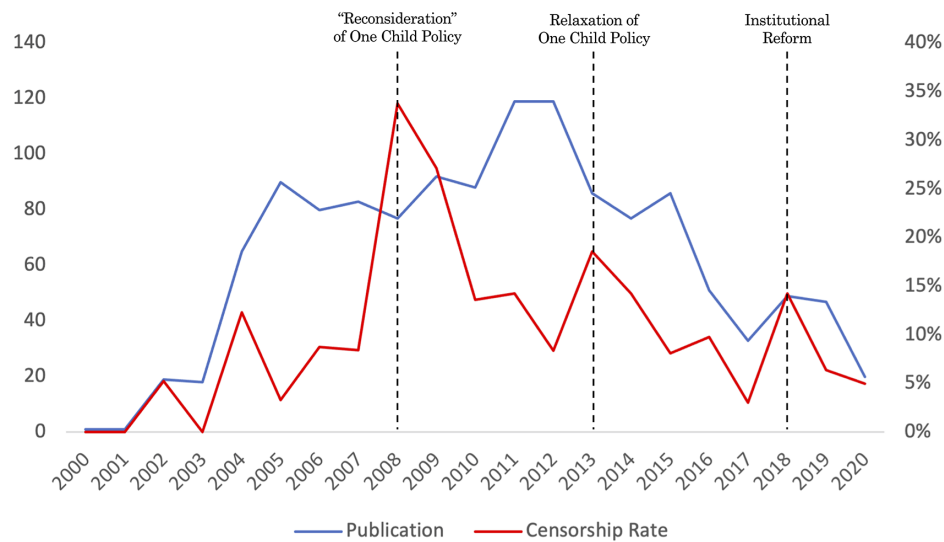*$p < .1$, **$p < .05$, ***$p < .01$

been censored (censorship rate) for each year over the past 20 years. The intensity of state censorship on this topic peaks in 2008, 2013, and 2018, which reflects the state's attempt to ensure a smooth and controlled policy change for each critical juncture in the process, curbing uncontrolled discussion about this important policy about-face.

Second, intellectual writings on China's contemporary history, particularly history after the founding of the CCP in 1921, invite intense attention from censors. Censored topics include competing historiographical interpretations of the Chinese communist revolution (*such as discussions about the internal factional struggles among the communist leaders and red army generals during the 1930s*), unflattering stories about the Communist Party's past (*such as discussions about the cultivation of opium in the communist-controlled areas during the Sino-Japanese War*), and often nuanced accounts of personal experiences of political persecution and suffering under communist rule by members of the intelligentsia. It shows that the Party-state is on high alert to prevent the emergence of any unofficial, competing, or even antithetical historiography.

Third, our results also show that articles related to social and moral values — particularly those that are regarded by the CCP as fundamentally alien and even harmful to the officially sanctioned morality — are more likely to be censored. These topics include the usual suspects such as

7

Figure 3
Articles on the One-Child Policy

Notes: This figure shows the number of publications (blue line) and the censorship rate (red line) of articles about the One-Child Policy. It shows that the censorship rate peaks in 2008, when the Party-state declared a "reconsideration" of this basic national policy. The censorship intensity has gradually decreased since 2008, as the policy was slowly dismantled. The three peaks in censorship in 2008, 2013, and 2018 may indicate the state's desire to ensure a smooth and controlled policy overhaul.

Western liberalism or Christianity, but also counterintuitive items such as the traditional ethic of filial piety, which is considered harmful to the norms of the state-endorsed socialist legality (*e.g., one censored article discusses the view that family members should not report each other's crimes*), and Marxist fundamentalism, which is negatively critical about China's market reforms (*e.g., one censored article described China today as a crony capitalist country according to orthodox Marxist standards*).

In contrast, as shown in Figure 2, metaphysical discussions of academic theories of economics, sociology, and cultural studies are less likely to be censored, as are articles on the negative side of domestic policies in foreign countries (*e.g., terrorism in Western countries*), or issues in China's foreign policy such as the Sino–US relationship. Of course, intellectual writings on topics that are in line with the official political rhetoric (*such as achievements in poverty alleviation, national development, or administrative modernization*) have particularly low censorship magnitude. Details of each of the 160 topics are provided in Appendix I, which presents the key words (in English and Chinese) under each topic, the censorship magnitude for each topic, and the yearly topic prominence and censorship rate overall.

Who Gets Blacklisted?

Authors Are Not Blacklisted Because of Topics
Now we turn to *persona censorship*. Occasionally, the state issues an order to blacklist a particular author, demanding the complete erasure of all of her writings regardless of content or topic. One might assume that the likelihood of a person being blacklisted is positively related to the frequency of her writing on topics that are most likely to be censored; that is, blacklisting may be just an extreme variant of thematic censorship. However, a comparison of the writings of the partially censored authors with those of the completely blacklisted authors refutes this theory. Our comparison of various dimensions consistently shows that the writings of the blacklisted intellectuals are more similar to the corpus of uncensored and publicly viewable articles on the website than to the corpus of the deleted articles in the partially censored authors subsample. In other words, our empirical findings strongly suggest that decisions to completely blacklist a particular scholar have highly different motivations than content-based thematic censorship decisions.[10]

First, a comparison of the average number of censored articles per a partially censored author and per a completely silenced intellectual suggests that there are two different censorship mechanisms at work. On average, the 425 partially censored authors have 3.4 censored articles per person, whereas the 35 completely blacklisted authors have 91.9 censored articles per person. This remarkable disparity supports the hypothesis that different censorship mechanisms are applied to the two groups of authors.

A further comparison shows that the topics of the deleted articles written by the blacklisted authors are more similar to the topics in the publicly viewable articles

written by the partially censored authors (*cosine similarity* = 0.227) that survived state censorship than to their deleted articles (*cosine similarity* = 0.291, *t*(3302) = –87.734, *p-value* = 2.2e-16**). We further verify whether an author is completely muted due to the last few articles she published before the ban. We found that the last n articles (n ∈ 2,3,5) published by blacklisted authors are robustly more similar to the survived pool of articles than the censored pool of articles (see Figure A3 in Appendix C for more details). This confirms that it is unlikely that the state's blacklisting of an author is triggered by the content of articles she last published prior to the ban.

In fact, our findings show that, even when the blacklisted authors write on the least censored topics, their articles are still completely censored. When partially censored authors write on the 20 topics that are least likely to be censored (see Figure 2, they have a low censorship rate of 0.49%, whereas the censorship rate of the completely blacklisted author remains 100%. It is worth noting that the frequency of publication on the least censored topics does not significantly vary between the two groups of authors — over the past two decades, the completely blacklisted authors published an average of 8.4 articles per person on the 20 topics that are least likely to be censored, whereas the average of the partially censored authors is 10.06 articles (*t*(58.082) = –0.751, *p-value* = 0.456). This shows that the completely blacklisted authors on the website are no less likely to publish on topics welcomed by the Party-state than their peers; yet, not even their articles on "benign" topics escape complete erasure.[11]

Possible Contributing Factors for Blacklisting

Then, why are some authors blacklisted by the state at all? We use feature selection models to evaluate six possible factors that are discussed in the literature. Appendix D introduces the data collection processes for these variables. For detailed descriptive statistics of the variables, see Appendix E.

Article Topic: Intellectuals who write more about politically taboo topics may be more likely to be silenced or exiled by the state, lest they spread subversive ideas and antithetical discourses (Finkel, 2007). We use the percentage of each author's articles that focus on any of the 20 topics that are most likely to be censored (*TopCensored20*) to measure this factor.

Public Influence: Intellectuals who are socially or politically influential may be more likely to be blacklisted, as they could facilitate the organization of or even instigate the mobilization of potential social movements (Coser, 1997; Gallagher and Miller, 2021; Pan and Siegel, 2020; Ngoshi, 2021). We measure the public influence of an author with three variables. First, we use the ratio of the number of followers to the number of posts on an author's real-name Weibo account to measure her influence on

social media (*WeiboInf*). Second, we use the number of articles mentioning a particular intellectual's name in China Digital Times, a prominent overseas opposition media platform, to estimate the author's political influence within dissident networks (*CDTcount*). Third, we use the number of articles mentioning a particular intellectual's name in the People's Daily (*renmin ribao*), the official mouthpiece of the Party-state, to measure the author's political influence within the Party-state establishment (*PPDcount*).[12]

Opposition Movement: An intellectual's participation in national opposition movements can be seen as a credible sign of her anti-regime tendency; thus, intellectuals with a record of participation in national opposition movements are more likely to be blacklisted by the state (Gasster, 1969; Flam, 1999). In the past decades, the two most prominent national opposition movements in China have been the June Fourth Movement in 1989 and the "Charter 08" Movement in 2008. To measure this factor, we first determine whether the author was on the state's Most Wanted List issued after the June Fourth movement or on a list of persons who were prohibited from entering China because of participation in the June Fourth Movement (*JuneFourth*). Then, we determine whether the author was a signatory of "Charter 08", an anti-state manifesto drafted by a group of intellectuals led by Liu Xiaobo (*Charter08*).[13]

Overseas Experience: Intellectuals who have overseas experiences may be more likely to be blacklisted for the following reasons: (a) they may spread subversive Western ideologies and discourses (Zweig and Yang, 2014); or (b) their international connections and fame make suppression against them politically costly. In this case, a complete blacklisting could be an economical choice for the state (Camp, 1985). We measure this factor on two dimensions. First, we record whether an intellectual is non-Chinese or is currently sojourning overseas and thus is beyond the jurisdiction of the Chinese Party-state (*Foreign*). Second, we collect the information about each author's degree and work experience and check if they have obtained any higher degree from overseas institutions of higher learning (*OverseasDegree*) or if they had full-time work experiences in a foreign country (overseas military or diplomatic postings for the People's Republic of China (PRC) is excluded) (*OverseasWork*).

Political Status: Intellectuals in China are customarily categorized as "establishment," and "non-establishment," or "independent". Generally, "establishment intellectuals" work in state-funded institutions and thus tend to be more closely controlled and monitored by the regime than their more "independent" counterparts, who are less connected to state institutions (Hua, 1994). Thus, intellectuals who are more embedded in the state establishment may be less likely to be blacklisted. To measure an intellectual's

involvement in the establishment, we use the following criteria: (a) work experience in state-funded institutions (*EstExp*); (b) a leadership role in state-funded institutions (*EstLeader*); or (c) is/was a deputy or member of the People's Congress or the Political Consultative Conference — the two legislative organs of the PRC — at any level (*Sessions*).

Lèse-Majesté: Intellectuals tend to be harshly penalized for making pejorative remarks or personal attacks on the supreme leaders of the state. This is because using abusive language to attack a recognized supreme leader — incumbent or retired — in published writings, constitutes a damaging symbolic act of defiance (Kelly, 1981; Black, 1989; Streckfuss, 1995). The state, presumably, would use all available means to prevent this kind of open and symbolic defiance from happening and diffusing to a larger sphere. We check whether an author has made abusive or pejorative remarks or personal attacks toward any of the Party-state's supreme leaders (*LeaderAtk*). We define supreme leaders as those whose ideological concepts are included in the Constitution of the CCP: Mao Zedong, Deng Xiaoping, Jiang Zemin, Hu Jintao, and Xi Jinping.

Deciphering Persona Censorship

To weigh different contributing factors, we first fit a logistic regression model to gain a preliminary understanding of the correlations between various factors and the response variable, namely whether an intellectual is blacklisted by the state. We also control for the intellectual's age (*BirthYear*), academic discipline (*Discipline*), and type of institutional affiliation (*AffiType*).[14] Then, to weigh the contribution of each factor against the state's eventual decision to blacklist a particular intellectual, we deploy two feature selection models from the least absolute shrinkage and selection operator (LASSO) family: the adaptive LASSO and the group LASSO. For a more detailed discussion of the methods used, see Appendix F.

The results of the logistic regression show that making personal attacks against the supreme leaders of the Chinese Party-state or using abusive or pejorative language when writing about them is the most, if not the only, important motivation for a complete ban by state censors. The odds ratio denotes that, other things being equal, attacking the supreme leaders in a publication makes an intellectual 30.18 times more likely to be blacklisted. Other variables, such as the topic of an article and an author's participation in national opposition movements, public and social influence, overseas study and work experience, or political status in relation to the Party-state establishment, have little or no effect on the state's decision to completely silence an author in the intellectual public space. This result holds when we control for gender, academic discipline, and age (see Table 2).
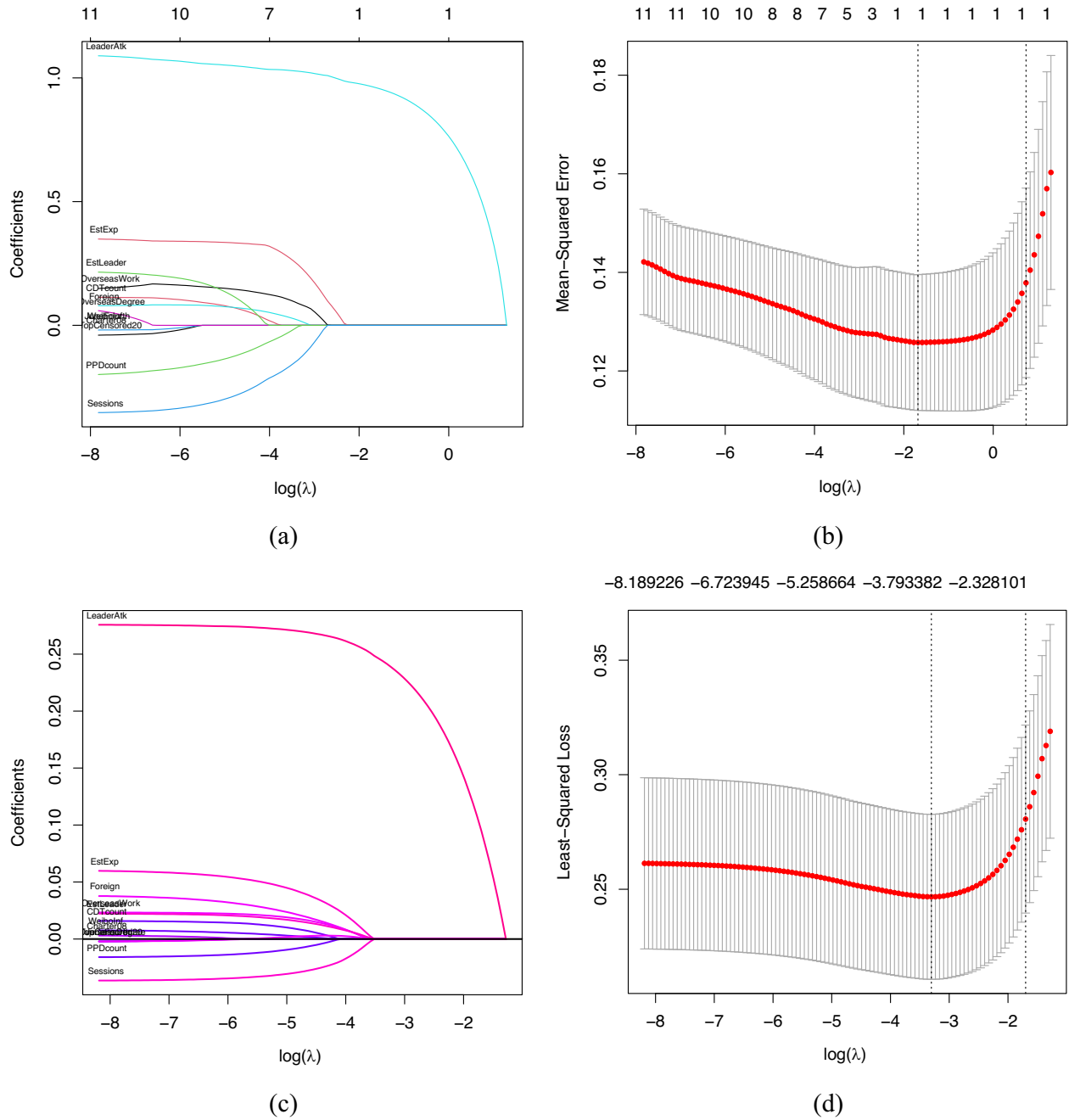
Table 2
Deciphering Persona Censorship with Logistic Regression

	Blacklisted	
	(1)	**(2)**
TopCensored20	−0.221	−0.234
	(1.034)	(1.041)
WeiboInf	0.00003	0.00002
	(0.0001)	(0.0001)
CDTcount	0.002	0.002
	(0.004)	(0.004)
PPDcount	−0.006	−0.006
	(0.008)	(0.008)
JuneFourth	−0.046	−0.077
	(1.517)	1.505
Charter08	−0.061	−0.021
	(0.607)	(0.621)
Foreign	0.209	0.222
	(0.745)	(0.754)
OverseasDegree	0.212	0.192
	(0.644)	(0.647)
OverseasWork	0.356	0.326
	(0.731)	(0.739)
EstExp	0.891	0.814
	(0.600)	(0.616)
EstLeader	0.453	0.430
	(0.519)	(0.520)
Sessions	−1.555	−1.544
	(1.220)	(1.212)
LeaderAtk	3.407**	3.327**
	(0.551)	(0.568)
Discipline		0.032
		(0.212)
BirthYear		−0.0001
		(0.0001)
AffiType		−0.006
		(0.022)
Constant	−4.255**	−4.039**
	(0.698)	(0.970)
Observations	399	399
Log Likelihood	−83.925	−83.682
Akaike Inf. Crit.	195.849	201.364

Notes: The robust standard errors are in parentheses. Model (1) is the baseline model. Model (2) includes controls for the intellectual's gender, discipline, and year of birth.
* p<.05, **p<.01 (two-tailed test).

Figure 4 shows the results of the adaptive LASSO and the group LASSO regressions, which penalize the variables (or group of variables) with relatively less predictive power and thus select the stronger predictors. As the penalization weight λ increases, the relatively unimportant predictors decrease toward zero, and the variables with more predictive power will thus be revealed.[15] The results of both LASSO models show that at optimized values of λ (i.e., λ_{min} and λ_{1se}), *LeaderAtk* (i.e., whether an intellectual has attacked national supreme leaders) becomes the only variable with predictive power to anticipate whether an intellectual will be blacklisted by state censors.[16]

Figure 4
Relative Importance of Variables for Predicting Blacklisting Using LASSO Models

(a)

(b)

(c)

(d)

Notes: This figure shows the results of the adaptive LASSO regression and the group LASSO regression. Panel (a) shows the importance of each variable (represented by each colored line). As λ — the penalization weight — increases, the contributions of all of the variables tend to decrease. The black line at the top of the chart is *LeaderAtk*, illustrating its prominence. Panel (b) shows the choice of λ upon cross-validation. The left dashed line is λ_{min}, which is the minimum mean cross-validation error, and the right dashed line is λ_{1se}, which is the most regularized model, with the cross-validation error held to within one standard error of the minimum. Panel (c) and (d) shows the trace plot of the group LASSO model and the choice of λ thereof.

Can authors return from persona censorship? In other words, once an author is blacklisted, are they forever silenced? In this research, we define a blacklisted author as one who has at least three articles being published, which have all been ordered to be deleted by the state. If an author was once blacklisted and then permitted to publish again, we may observe the following: (1) the deletion of all (at least three) of the author's publications

Table 3
Deciphering Persona Censorship with LASSO Models

	Blacklisted		
	(1)		**(3)**
(Intercept)	−2.839 (-2.467)		−0.825 (-0.825)
TopCensored20	. (.)	A	. (.)
WeiboInf	. (.)	B	. (.)
CDTcount	. (.)	B	. (.)
PPDcount	. (.)	B	. (.)
JuneFourth	. (.)	C	. (.)
Charter08	. (.)	C	. (.)
Foreign	. (.)	D	. (.)
OverseasDegree	. (.)	D	. (.)
OverseasWork	. (.)	D	. (.)
EstExp	. (.)	E	. (.)
EstLeader	. (.)	E	. (.)
Sessions	. (.)	E	. (.)
LeaderAtk	0.963 (0.501)	F	0.242 (0.095)

Notes: This table shows the output of the two LASSO-based models. Model (1) is the adaptive LASSO model; Model (2) is the group LASSO model. The table reports the coefficients of each variable λ_{min} and λ_{1se} (in parentheses). A dot signifies that the variable is eliminated under the given λ because of insignificance. The letter indexes in Model (2) indicate the grouping of the variables.

from before time point t; (2) the author's resumption of publication at time point t + 1, with at least one article uncensored; and (3) an interval between time points t and t + 1 that is significantly longer than the mean interval of the author's publication before time point t. We identify only one author whose publication record meets all three conditions.[17] The author is an expert in the philosophy of art, with no public record of making pejorative comments about the Party-state's past and present supreme leaders. It is likely that the censorship of his three articles before 2009 was the result of individual censorship instructions (i.e., content deletion) rather than author blacklisting. In general, the existing records to date show that authors muted by the state rarely recover from being blacklisted.

Robustness Checks

We use several more traditional and transparent methods to check the robustness of our main results. First, Table 4 presents a 2 × 2 crosstab between *LeaderAtk* and *Blacklisted*. It shows that authors who have issued personal attacks against the supreme leaders are 12.69 times more likely to be blacklisted than those who have made no such attacks. Second, the main results reported include the simple logistic regression results with all 13 variables and three groups of control variables. Here, we also conduct logistic regressions with each of the 13 variables (with normalization) one by one as robustness checks (see Table A2 of Appendix G). The results confirm that *LeaderAtk* is the variable with the greatest contribution to the state's eventual decision regarding whether an article is to be censored.

Figure A6, Figure A7, and Table A3 in Appendix G show the results of the basic LASSO regression (L1 regularization) and ridge regression (L2 regularization). These robustness tests support our main results.

To cross-check the results from the LASSO models, we also use two tree-based models, the random forest model and Boruta algorithm (Figure A8 and Figure A9 in Appendix G). The results of the tree models prove that making personal attacks on supreme leaders is the most robust predictor of the state's decision to blacklist an intellectual.

A potential concern is that the six factors that we consider as motivations for blacklisting a member of the intelligentsia are interrelated. In other words, there may be a high level of multicollinearity between these variables, and thus they might collectively describe a particular type of intellectual that is highly susceptible to state silencing. The correlation matrix presented in Figure A5 of Appendix E shows this is not the case. The main predictor of interest, *LeaderAtk*, demonstrates weak correlations with other predictors, eliminating multicollinearity issues. Other predictors also show weak intercorrelations in general, with the only exception being the three

Table 4
Relations Between *LeaderAtk* and *Blacklisted*

	LeaderAtk = 1	*LeaderAtk* = 0
Blacklisted = 1	22	13
Blacklisted = 0	25	339
Blacklisted Ratio	46.81%	3.69%

Notes: This table shows the relationship between *LeaderAtk* and *Blacklisted*. It shows that if an author has a record of attacking the supreme leader of the Party state, her possibility of being blacklisted is 46.81%. However, the possibility of being blacklisted for an author who has no record of Lèse-Majesté is only 3.69%.

variables measuring overseas experience (i.e., *Foreign*, *OverseasDegree*, and *OverseasWork*), which are understandably tightly clustered. Perhaps somewhat surprisingly, the two variables measuring social movement participation (i.e., *JuneFourth* and *Charter08*) are almost uncorrelated. This may be due to the two-decade-long time span between the two movements, as well as the fact that the prominent participants of the June Fourth Movement had already been penalized by the time of Charter 08 and faded away from the public sphere: it is highly unlikely that the participants of the two movements were the same group of people.

Another potential concern lies in our measurement of the intellectuals' participation in national opposition movements. In the main results, we measure this factor through the actual records of the intellectuals' direct participation in the June Fourth Movement and the "Charter 08" Movement. However, the authors publishing on the website may at some point in their lives have joined, encouraged, or memorialized the opposition movements and thus have been indirectly involved in these movements. Therefore, in the robustness check, we substitute *CDT64* and *CDT08* for *JuneFourth* and *Charter08.* These two new variables indicate whether an intellectual has written about the two prominent national opposition movements on the dissident platform China Digital Times. Our main results hold (see Table A4 and Table A5 in Appendix G).

For the same reason, we also substitute the political status of each intellectual's publication venues for the political status of the individual's institutional affiliation and career attributes (*EstExp*, *EstLeader*, *Sessions*). The assumption is that publishing in the CCP's mouthpieces could serve as a certificate of political trustworthiness and thus shield the intellectual from blacklisting. We create two variables: (1) *Organ* indicates whether an intellectual has published in an official newspaper or journal of the Central Committee of the CCP, namely the *People's Daily* and *Qiushi Magazine*, and (2) *OfficialPress* indicates whether an intellectual has published in the "People's Press" at any level; these are the official publishing houses directly run by the Communist Party committees at the central and provincial levels. Our main results hold (see Table A6 and Table A7 in Appendix G).

We also substitute *TopCensored20* and *TopCensored10* for *TopCensored30*, tightening and loosening, respectively, the definition of "politically sensitive topics" (see Table A8 and Table A9 in Appendix G). Our main results hold in all of these tests.

Is There a Mechanism of Spike Suppression?

An alternative explanation is that there may be a "spike suppression" mechanism at work when the state makes censorship decisions. For instance, King, Pan, and Roberts (2013) argued that state censors pay more attention to a certain topic when there is a "spike" of interest in it on social media platforms so as to avoid the risk of that topic becoming focal. Lorentzen (2014) also contended that investigative reports are censored more stringently when there are more bad stories to tell, compared with quieter periods. Our data of intellectual public writings do not support either argument. In Appendix H, we show that the prominence of a topic (measured by the proportion of a given topic in all published articles on the website in a given year) is negatively related to the censorship rate of that topic (see Table A10).

We also find that the prominence of an article (measured by the proportion of each topic in the article weighted by the prominence of each topic in a given year) is not significantly related to whether it is censored (*coef* = –1.77, *p-value* = 0.737). It may also be argued that the state is more likely to take down writings of more prolific authors to contain their social influence (Gallagher and Miller, 2021). Our data do not support this hypothesis. In Appendix H, we demonstrate that the relationship between the prolificacy of an author and the possibility of said author's articles being censored is actually negative (see Table A11 of Appendix H). Furthermore, an author's prolificacy is not a valid predictor of the likelihood of said author being blacklisted by state censors (see Table A12 of Appendix H). Combining these findings, we find the "spike suppression" hypothesis to be invalid. The Chinese state does not particularly target prominent topics, prominent articles, or prolific authors, at least in the intellectual public space we study.

Concluding Remarks

In 21st century states, censors must patrol both popular and intellectual public spaces. Most studies of state censorship in China focus on the popular public space, which is mainly made up of the many social media platforms that have developed since the early 2000s. Studies of censorship regarding these platforms indicate that state censors tend to block "the spread of common knowledge about collective action events (and not grievances)" (King, Pan, and Roberts, 2017, 497), delete unsolicited criticisms of the state (Gueorguiev and Malesky, 2019), remove messages that are specific or that signal conflict (Tai and Fu, 2020), or "repress and limit the reach of influential non-Party 'thought leaders'" (Gallagher and Miller, 2021). However, none of these standards appear to be applicable to the Chinese state's censorship of the intellectual public space. Political expression in this context has considerable significance but has received limited scholarly attention so far. We draw on the database of a leading Chinese intellectual website and find that two types of state censorship operate in parallel in this intellectual public space, each following a different rationale and set of standards.

Thematic censorship is deployed to block writings that oppose the official narrative of national policies, orthodox historiography, and officially endorsed values and moral norms. Persona censorship is used to completely silence a small group of intellectuals who dare to make pejorative remarks about the incumbent and past supreme leaders of the state, which represents a symbolic gesture of open defiance against the state's authority. These two elements are combined in a mechanism that represents China's censorship apparatus in the intellectual public space.

Our findings differ from those of other studies regarding the criteria used for censorship. We argue that this difference is due to our identification of the above-mentioned factors, both of which have theoretical and practical significance. Our extensive dataset of 144,280 published and censored intellectual public writings over a continuous 20-year period includes a total of 740 million Chinese characters, and enables us to conduct a comprehensive investigation of China's censorship system in the intellectual public space. First, we argue that other studies do not distinguish censorship decisions in terms of scope, and thus regard the state's selective censorship of content on a day-to-day basis as the same as the complete and non-negotiable banning and silencing of individual authors on rarer occasions. We, however, find that the standards used for selectively labeling certain content as inappropriate are fundamentally different from those that determine which authors should be completely banned from future public expression. Antithetical narratives at odds with the basic national policies, official historiography, and the moral values advocated by the Party-state are forbidden topics, but pejorative personal attacks directed at supreme leaders result in an author being completely banned from the public sphere. As the criteria of censorship often reflects the intent and goals of the government (King, Pan, and Roberts, 2013), our findings indicate that rulers may be more threatened by the direct and open ridicule of the supreme leader as an individual, as this is a sign of public and symbolic defiance and thus requires a comprehensive ban of any output by the intellectual who has committed such an "offense."

The context in which political expression takes place can also help us understand state censorship. Unlike other research that focuses on social media, we consider state censorship of the intellectual public space, which has a unique political significance. Unlike the netizens who wield the combative power of the masses through collective expression in the popular public sphere, the intelligentsia are politically significant for the entire nation in terms of their discursive agency and moral leadership. They provide the discursive agency to construct a potentially parallel consensus that advocates new norms, discourses, and narratives that compete with the official ones. Intellectuals have a significant role to play in the forming

of public opinion, as they are those "who speak in the name of the social whole" and are also "mandated …to tell the group what the group thinks" (Bourdieu, 2020, 45). As the Chinese philosopher Huang Zongxi (1610–1695) wrote, "ultimately right and wrong are to be determined by scholar-philosophers in the schools, for they are the custodians of the Truth" (deBary, 1957, 197). If this normative and moral authority is not restrained, intellectuals can lead public opinion into a parallel social consensus that is at odds with that upheld by the state; they may then call for alternative systems of symbolism, power, and authority. Thus, intellectuals have the power to legitimize or delegitimize the state at a fundamental level without calling for immediate collective action. The collective action potential standard used in the censorship of social media platforms may not be equally applicable in the state's undertaking to control the intellectual public space.

Intellectuals may also have a political impact through their moral leadership. These "moral counter elites" (Reddaway and Glinski, 2001, 140) make symbolic gestures of overt defiance to the state authority. Those who personally engage in public displays of disdain, contempt, and defiance against the state or ruling elite may even voluntarily submit to the subsequent state violence, and thus become symbols of dissidence (Flam, 1999). Thus, unlike social networking platforms that mainly have "flat" structures, the intellectual public space radiates out from the demonstrative persona and charisma of its most symbolic members. These intellectuals do not seek to connect but to display and thus take a position of moral leadership. Their symbolic gestures of defiance may also have a "broken window effect," and they may shatter the power foundation of the state. The unique channels through which the intelligentsia exert political power suggest that the state must apply a different set of criteria when censoring topics or authors in the intellectual public space.

This research contributes to the general literature on authoritarianism by revealing how an authoritarian state assesses threats and selectively applies coercive power to a sector of society. Studies show that authoritarian ruling elites tend to be selective, tactical, and discreet in their use of state coercive power (Greitens, 2016; Gerschewski, 2013; Xu, 2021; Pop-Eleches and Way, 2023). Such rulers rationally restrict their use of coercion to occasions when the perceived threat to the regime is greatest, thus reducing the potential for a backlash. However, it is unclear how authoritarian states identify, classify, and evaluate threats. In this study, we demonstrate that the authoritarian state recognizes the moral leadership of those intellectuals who dare to openly defy it by publishing lèse-majesté remarks of ruling elites who are on the pinnacle of power. Thus, our findings reveal the "personal" aspect of modern authoritarian regimes (Geddes, Wright, and Frantz, 2018; Shirk, 2018), which has long been downplayed or ignored by the predominantly institutionalist

literature on authoritarian regimes. To quote Anthony Giddens ([1987], 304), "a key aspect of totalitarianism, without which the rest would not be possible, or at least would not be unified into a cohesive system of rule, is the presence of the leader figure". The authoritarian state's persona censorship mechanism, which completely silences a small group of authors, reminds us of the penalties traditionally (and still current in a few countries today) imposed on offenses of lèse-majesté; these are "purely discursive crimes" based on national security, but they "do not physically threaten the state but erode the state's construction of what it contends is a sacred national identity" (Streckfuss, [1995], 448). This mechanism represents the state's protection of its own core symbolic authority.

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/S1537592723002815.

Acknowledgments

Notes

1 Likewise, Balsekar ([2014]) finds that in a context of electoral democracy (India), censoring an allegedly offensive film may be a political strategy pursued by otherwise resource-poor candidates to appeal to minority constituencies who are expected to be offended by the film in question.

2 The proposal of this research has been reviewed and approved by the Institutional Review Board (IRB) before any data processing or fieldwork. Research was undertaken according to the approved procedures and protocols.

3 It is worth noting that Pan and Siegel's ([2020], 113-4) subsequent study in Saudi Arabia provides a different argument, suggesting that "express[ing] dissatisfaction with or criticizing the Saudi monarchy including specific royal family members, members of the religious establishment such as state-sanctioned clerics, or religious doctrine associated with the monarchy," to the extent that they "challenge the legitimacy of the religious monarchy," and in fact "likely represents the most intolerable form of online expression for the Saudi regime". Although Pan and Siegel ([2020], 114) regard lèse-majesté as a subset of "criticism," they do distinguish the criticism of the supreme leader from that of specific policies, and identify the latter form of criticism as "less problematic for the regime as it challenges its policies but not its underlying legitimacy".

4 As the website focuses on public intellectual writings, it does not include academic research articles published in peer-reviewed journals.

5 Censorship orders received by the website always demand complete deletion of the articles in question. State censors never demand the editing/ rewriting/redaction of only some of an article's content. The situation may vary for other publication venues.

6 Hereafter, for simplicity, "published" articles refer to those that were publicly viewable on the website on August 1, 2020, and "censored" articles refer to those that had been published on the website but were later deleted according to state censorship orders.

7 Existing research shows that normally around 10 keywords are representative enough for articles of the length of an academic paper (Zhou, Yang, et al., [2020]; Zhou, Shi, et al., [2022]). In this research, we extract around 50 to 200 keywords for each article to make a generous representation that assails the topic modeling.

8 Topic modeling is preferred in textual analysis because it provides both depth (by putting words into contexts) and width (by abstracting features using the whole text) (Grimmer and Stewart, [2013]; Ying, Montgomery, and Stewart, [2021]). By doing so, according to Mueller and Rauh ([2018], 358), we "let the data speak without losing interpretability of the results". For the choice of the number of topics $K = 160$, see Appendix A.

9 For details of the logistic regression model, see Appendix B.

10 All our study on blacklisting are based on the 35 "active" blacklisted authors, who have at least three attempts of publication. The non-active authors — those who published one or two articles and all were censored — have too few publications that makes it impossible to decide whether the articles were censored one by one (and thus happened to be all censored), or the author was muted (and thus all her publications were erased wholesale). For this reason, we only focus on the active blacklisted authors, who on average publishes 91.9 articles, and all were censored.

11 Similarly, the blacklisted authors are not significantly more likely to publish on the 20 most censored topics ($t(39.349) = 1.239$, p-value = 0.223).

12 Some intellectuals have common Chinese names, and thus the counts of articles written by them may not be accurate. To address this problem, we treat the number of articles written by someone with a common name as missing data. We then use multivariate imputation by chained equations (MICE) methods to

generate multiple predictions for each missing value based on the observed data. See Buuren et al., 2006.

13 We use publicly available information to compile the list. For the "Charter 08" Movement, a full list of signatories is available online (China Digital Times, 2019). For the June Fourth Movement, we combine three name lists: (a) the major "conspirators" of the movement identified by a report of the Beijing Municipal Government published in 1989 (Chen, 1989); (b) an online document leaked from Guangdong Province listing the prominent participants of the June Fourth Movement who should be denied entry to China (Zeng, 1995); and (c) a list of the 21 most wanted student and labor-union leaders during the June Fourth Movement published by the Ministry of Public Security of China (People's Daily, 1989a,b).

14 For the measurement of each control variable, see Appendix D.

15 The LASSO models uses cross-validation to identify the optimized λ, which resamples the dataset to use different portions of the data as training and testing sets to train the model, and iterate for 10 times. This way, the choice of the λ, a hyperparameter, is unlikely to be influenced by a few outlying incidents.

16 For the numeric results of the LASSO regressions, please refer to Table 3.

17 This author had 3 publications before year 2009 — all being censored, and 14 uncensored articles after year 2015.

References

Alpermann, Björn and Shaohua Zhan. 2019. "Population Planning after the One-Child Policy: Shifting Modes of Political Steering in China". *Journal of Contemporary China* 28 (117): 348–66.

Balsekar, Ameya .2014. "Seeking Offense: Censorship as Strategy in Indian Party Politics". *Comparative Politics* 46 (2): 191–208.

Black, Georgina Dopico. 1989. "The Limits of Expression: Intellectual Freedom in Postrevolutionary Cuba". *Cuban Studies* (19): 107–42.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation". *Journal of Machine Learning Research* 3: 993–1022.

Bourdieu, Pierre. 2020. *On the State: Lectures at the Collège de France 1989-1992*, ed. Patrick Champagne et al. Cambridge, UK and Medford, MA: Polity Press.

Buuren, Stef van et al. 2006. "Fully Conditional Specification in Multivariate Imputation". *Journal of Statistical Computation and Simulation* 76 (12): 1049–64.

Cairns, Christopher and Allen Carlson. 2016. "Real-world Islands in a Social Media Sea: Nationalism and Censorship on Weibo during the 2012 Diaoyu/

Senkaku Crisis". *The China Quarterly* 225 (Mar): 23–49.

Camp, Roderic A. 1981. "Intellectuals: Agents of Change in Mexico?" *Journal of Interamerican Studies and World Affairs* 23 (3): 297–320.

Camp, Roderic A. 1985. *Intellectuals and the State in Twentieth-Century Mexico*. Austin, TX: The University of Texas Press.

Chang, Charles and Melanie Manion. 2021. "Political self-censorship in authoritarian states: The spatial-temporal dimension of trouble". *Comparative Political Studies* 54 (8): 1362–92.

Chen, Xitong. 1989. 关于制止动乱和平息反革命暴乱的情况报告 (Report on the Situation of Stopping Unrest and Quelling Counterrevolutionary Riots). url: https://www.gov.cn/gongbao/shuju/1989/gwyb198911.pdf (visited on 09/28/2021).

China Digital Times. 2019. Charter 08 Signatories (Total of 5 Batches). url: https://chinadigitaltimes.net/space/%E3%80%8A%E9%9B%B6%E5%85%AB%E5%AE%AA%E7%AB%A0%E3%80%8B%E7%AD%BE%E5%90%8D%E4%BA%BA%E5%91%98%E5%90%8D%E5%8D%95%E6%B1%87%E6%80%BB%EF%BC%88%E5%85%B1%E4%BA%94%E6%89%B9%EF%BC%89 (visited on 09/24/2021).

Coser, Lewis A. 1997. *Men of Ideas: A Sociologist's View*. New York, NY: Free Press.

Cramer, Frederick H. 1945. "Bookburning and Censorship in Ancient Rome: A Chapter from the History of Freedom of Speech". *Journal of the History of Ideas* 6 (2): 157–96.

Cressy, David. 2005. "Book Burning in Tudor and Stuart England". *The Sixteenth Century Journal* 36 (2): 359–74.

Davies, Gloria. 2007. "Habermas in China: theory as catalyst". *The China Journal* 57: 61–85.

deBary, William Theodore. 1957. "Chinese Despotism and the Confucian Ideal". *Chinese Thought and Institutions*, ed. John King Fairbank, 163–203. Chicago, IL: The University of Chicago Press.

Dimitrov, Martin K. 2017. "The Political Logic of Media Control in China". *Problems of Post-Communism* 64 (3-4): 121–27.

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

Esberg, Jane. 2020. "Censorship as Reward: Evidence from Pop Culture Censorship in Chile". *American Political Science Review* 114 (3): 821–36.

Finkel, Stuart. 2007. *On the Ideological Front: The Russian Intelligentsia and the Making of the Soviet Public Sphere*. New Haven, CT: Yale University Press.

Flam, Helena. 1999. "Dissenting Intellectuals and Plain Dissenters: The Cases of Poland and East Germany". *Intellectuals and Politics in Central Europe*, ed. András Bozóki, 19–41. Budapest, HU: Central European University Press.

Gallagher, Mary and Blake Miller. 2021. "Who Not What: The Logic of China's Information Control Strategy". *The China Quarterly* 248 (1): 1011–36.

Gasster, Michael. 1969. *Chinese Intellectuals and the Revolution of 1911: The Birth of Modern Chinese Radicalism*. Seattle, WA: The University of Washington Press.

Geddes, Barbara, Joseph Wright, and Erica Frantz. 2018. *How Dictatorships Work: Power, Personalization, and Collapse*. Cambridge, UK and New York, NY: Cambridge University Press.

Gerschewski, Johannes. 2013. "The three pillars of stability: legitimation, repression, and co-optation in autocratic regimes". *Democratization* 20 (1): 13–38.

Giddens, Anthony. 1987. *The Nation State and Violence*. Berkeley, CA and Los Angeles: University of California Press.

Gläßel, Christian and Katrin Paula. 2020. "Sometimes Less is More: Censorship, News Falsification, and Disapproval in 1989 East Germany". *American Journal of Political Science* 64 (3): 682–98.

Greitens, Sheena Chestnut. 2016. *Dictators and their Secret Police: Coercive Institutions and State Violence*. New York, NY: Cambridge University Press.

Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts". *Political Analysis* 21 (3): 267–97.

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

Han, Rongbin. 2015. "Defending the Authoritarian Regime Online: China's 'Voluntary Fifty-cent Army'". *The China Quarterly* 224 (Dec): 1006–25.

Han, Rongbin and Li Shao. 2022. "Scaling Authoritarian Information Control: How China Adjusts the Level of Online Censorship". *Political Research Quarterly* 75 (4): 1345–59.

Hassid, Jonathan. 2012. "Safety Valve or Pressure Cooker? Blogs in Chinese Political Life". *Journal of Communication* 62 (2): 212–30.

Hobbs, William R. and Margaret E. Roberts. 2018. "How Sudden Censorship Can Increase Access to Information". *American Political Science Review* 112 (3): 621–36.

Hua, Shiping. 1994. "One Servant, Two Masters: The Dilemma of Chinese Establishment Intellectuals". *Modern China* 20 (1): 92–121.

Kelly, George Armstrong. 1981. "From Lèse-Majesté to Lèse-Nation: Treason in Eighteenth-Century France". *Journal of the History of Ideas* 42 (2): 269–86.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression". *American Political Science Review* 107 (2): 326–43.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2014. "Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation". *Science* 345 (6199): 891.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2017. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument". *American Political Science Review* 111 (3): 484–501.

Kurzman, Charles. 2001. "Critics within: Islamic Scholars' Protests against the Islamic State in Iran". *International Journal of Politics, Culture, and Society* 15 (2): 341–59.

Lipset, Seymour Martin and Richard B. Dobson. 1972. "The Intellectual as Critic and Rebel: With Special Reference to the United States and the Soviet Union". *Daedalus* 101 (3): 137–98.

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

MacKinnon, Rebecca. 2008. "Flatter world and thicker walls? Blogs, censorship and civic discourse in China." *Public Choice* 134 (1-2): 31–46.

Martinez, Gabriela. 2018. Why did Saudi Arabia want to silence Jamal Khashoggi? url: https://www.pbs.org/newshour/world/why-did-saudi-arabia-want-to-silence-jamal-khashoggi (visited on 11/16/2021).

Mattingly, Daniel C. 2020. "Responsive or Repressive? How Frontline Bureaucrats Enforce the One Child Policy in China". *Comparative Politics* 52 (2): 269–88.

Milhalcea, Rada and Paul Tarau. 2004. "Textrank: Bringing Order into Text". *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, Barcelona. Association for Computational Linguistics*. Stroudsberg, Pennsylvania, 404–11.

Mueller, Hannes and Christopher Rauh. 2018. "Reading Between the Lines: Prediction of Political Violence Using Newspaper Text". *American Political Science Review* 112 (2): 358–75.

Ngoshi, Hazel Tafadzwa. 2021. "Repression, Literary Dissent and the Paradox of Censorship in Zimbabwe". *Journal of Southern African Studies* 47 (5): 799–815.

Pan, Jennifer and Alexandra A. Siegel. 2020. "How Saudi Crackdowns Fail to Silence Online Dissent". *American Political Science Review* 114 (1): 109–25.

People's Daily. June 1989a. "公安部转发北京市公安局通缉令, 通缉'高自联'在逃的头头和骨干分子 (Ministry of Public Security Forwards Wanted Notice from the Beijing Public Security Bureau

for the Fugitive Leaders and Key Elements of the Beijing Students' Autonomous Federation". *People's Daily*.

People's Daily. June 1989b. "各地依法取缔'高自联'等组织,一批煽动、组织动乱和暴乱的头目被捕 (In Accordance with the Law, Various Localities Ban Organizations Such as the Beijing Students' Autonomous Federation, Leaders Who Organized Unrest and Riots Arrested)". *People's Daily*.

Pop-Eleches, Grigore and Lucan A. Way. 2023. "Censorship and the Impact of Repression on Dissent". *American Journal of Political Science* 67 (2): 456–71.

Reddaway, Peter and Dmitri Glinski. 2001. *The Tragedy of Russia's Reforms: Market Bolshevism against Democracy*. Washington, D.C.: United States Institute of Peace Press.

Roberts, Margaret E. 2018. *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton, NJ: Princeton University Press.

Ruan, Lotus et al. 2021. "The Intermingling of State and Private Companies: Analysing Censorship of the 19th National Communist Party Congress on WeChat". *The China Quarterly* 246: 497–526.

Sanovich, Sergey, Denis Stukal, and Joshua A. Tucker. 2018. "Turning the Virtual Tables: Government Strategies for Addressing Online Opposition with an Application to Russia". *Comparative Politics* 50 (3): 435–82.

Schmidt, Peter. 2010. "Postcolonial Silencing, Intellectuals, and the State: Views from Eritrea". *African Affairs* 109 (435): 293–313.

Shirk, Susan L. 2018. "The Return to Personalistic Rule". *Journal of Democracy* 29 (2): 22–36.

Stockmann, Daniela. 2013. *Media Commercialization and Authoritarian Rule in China*. New York, NY: Cambridge University Press.

Streckfuss, David. 1995. "Kings in the Age of Nations: The Paradox of Lèse-Majesté as Political Crime in Thailand". *Comparative Studies in Society and History* 37 (3): 445–75.

Sun, Taiyi and Quansheng Zhao. 2021. "Delegated Censorship: The Dynamic, Layered, and Multistage Information Control Regime in China". *Politics and Society*. 1–31.

Tai, Yun and King-wa Fu. 2020. "Specificity, Conflict, and Focal Point: A Systematic Investigation into Social Media Censorship in China". *Journal of Communication* 70 (6): 842–67.

Tan, Kenneth Paul. 2016. "Choosing What to Remember in Neoliberal Singapore: The Singapore Story, State Censorship and State-Sponsored Nostalgia". *Asian Studies Review* 40 (2): 231–49.

Wasserstrom, Jeffrey N. 1991. *Student Protests in Twentieth-Century China: The View from Shanghai*. Stanford, CA: Stanford University Press.

Weiss, Jessica Chen. 2014. *Powerful Patriots: Nationalist Protest in China's Foreign Relations*. Oxford, UK: Oxford University Press.

Xu, Xu. 2021. "To Repress or to Co-opt? Authoritarian Control in the Age of Digital Surveillance". *American Journal of Political Science* 65 (2): 309–25.

Yan, Xiaojun and La Li. 2023. "Greasing the Wheels of Policy Reversal: Discursive Engineering and Public Opinion Management During the Relaxation of China's Family Planning Policy". *Governance*, Forthcoming.

Yan, Xiaojun and La Li. 2023. "Replication Data for Censoring the Intellectual Public Space in China: What Topics Are Not Allowed and Who Gets Blacklisted?" *Harvard Dataverse*, https://doi.org/10.7910/DVN/AXBEYC.

Ying, Luwei, Jacob M. Montgomery, and Brandon M. Stewart. 2021. "Topics, Concepts, and Measurement: A Crowdsourced Procedure for Validating Topics as Measures". *Political Analysis* doi:10.1017/pan.2021.33, 1–20.

Zarycki, Tomasz. 2009. "The Power of the Intelligentsia: The Rywin Affair and the Challenge of Applying the Concept of Cultural Capital to Analyze Poland's Elites". *Theory and Society* 38 (6): 613–48.

Zeng, Weiyan. 1995. "北京之春"获广东边防局文件,中共黑名单重点限制49人入境 (Beijing Spring Receives Documents from Guangdong Border Authorities, 49 Individuals Blacklisted by the Chinese Communist Party and Prohibited from Entering China). url: https://www.64memo.com/b5/6661.htm (visited on 09/28/2021).

Zhao, Yuezhi. 2000. "From commercialization to conglomeration: The transformation of the Chinese press within the orbit of the party state". *Journal of Communication* 50 (2): 3–26.

Zhou, Ning, Wenqian Shi, et al. 2022. "TextRank Keyword Extraction Algorithm Using Word Vector Clustering Based on Rough Data-Deduction". *Computational Intelligence and Neuroscience*: 1–19.

Zhou, Qingyu, Nan Yang, et al. 2020. "A Joint Sentence Scoring and Selection Framework for Neural Extractive Document Summarization". *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 28: 359–63.

Zweig, David and Feng Yang. 2014. "Overseas Students, Returnees, and the Diffusion of International Norms into Post-Mao China". *International Studies Review* 16 (2): 252–63.