Public Participation Increases Support for Censorship in Authoritarian Regimes

Tony Zirui Yang*

July 22, 2022

Abstract

Contrary to the conventional top-down view of government censorship, ordinary citizens in authoritarian regimes frequently participate in censorship by reporting online content. This study theorizes a bottom-up perspective of censorship in authoritarian regimes and analyzes its prevalence and consequences on public opinion toward censorship in the case of China. I argue that public participation increases support for censorship by diminishing the government's responsibility and strengthening citizens' perceived empowerment. Using an original survey in China, I show that more than 50% of the respondents self-report having flagged content online and such participation is positively correlated with support for government censorship. I further conducted an experiment embedded in custom-engineered, simulated social media pages. Consistent with my theory, respondents that are encouraged to report simulated posts display significantly higher support for government censorship. My study highlights the role of ordinary citizens in facilitating authoritarian control and the normalization of repressive policies such as censorship.

Keywords: Censorship, Authoritarian Regime, China, Public Participation, Public Opinion

*PhD Candidate, Department of Political Science, Washington University, St Louis, Campus Box 1063, One Brookings Drive, MO 63130, Email: yangzirui@wustl.edu

Introduction

Conventional wisdom of government censorship in authoritarian regimes regards it as a top-down tool of the authoritarian government to silence criticism and prevent collective actions that destabilize the regime (Gueorguiev and Malesky 2019; King, Pan and Roberts 2013; Miller 2018; Pop-Eleches and Way 2021; Roberts 2018; Shadmehr and Bernhardt 2015). In light of such understanding, existing research has extensively studied how unpopular government censorship can cause backlash (Nabi 2014; Pan and Siegel 2020; Roberts 2018, 2020), and how ordinary citizens circumvent and resist unpopular censorship activities (Chang et al. 2022; Chen and Yang 2019; Gläßel and Paula 2020; Han 2018; Hobbs and Roberts 2018; Roberts 2020).

Yet, contrary to conventional views, citizens in authoritarian regimes frequently participate in censorship by reporting online content. In Turkey, for example, Twitter users weaponize the report function against political opponents in a systematic manner (Tufekci 2017). In China, ordinary citizens made over 163 million censorship requests in 2020 alone, according to government statistics (Xinhua 2021). Moreover, contrary to the conventional depiction of government censorship as unpopular, surveys across the world demonstrate significant levels of popular support for government censorship, especially in authoritarian regimes like China (Dickson 2016; Wang and Mark 2015), Russia (Nisbet, Kamenchuk and Dal 2017), and Middle Eastern monarchies (Martin, Martins and Wood 2016; Wike and Simmons 2015). *Why are many citizens in authoritarian regimes supportive of government censorship? Is it possible that participating in censorship increases citizens' support for it?*

This study takes a new step in this research field by exploring the consequences of a novel, bottom-up perspective of censorship. I theorize that when ordinary users actively participate in the censorship process by flagging online content that they disapprove of, their support for censorship increases. Specifically, public participation diminishes the government's responsibility for censorship. When it is common for users to participate in censorship, citizens will be less likely to think that the government is responsible for censorship events. Even

if citizens do not like certain censorship events, they might think that it is due to other users' reports, and are therefore less likely to blame the government. In addition, public participation increases citizens' perceived empowerment by censorship. When users can report posts that they disapprove of, they are less likely to regard themselves as victims of censorship. Instead, they are more likely to think censorship benefits them by limiting the information they do not like. As a result, censorship is less likely to be viewed as top-down political suppression imposed by the government upon ordinary users, and public opinion toward government censorship becomes more positive.

Using an original online survey in China, I provide the first empirical description of public participation in the censorship process in authoritarian regimes. I find that more than half of the respondents self-report having flagged online content to be removed from the platforms. More than a quarter have reported online content at least once every few months. In particular, half of the "participating respondents" have the experience of reporting political content, including political news, opinions, and rumors. Such public participation in censorship is prevalent across several demographic groups and the political spectrum. In addition, I also find a significant correlation between participation and support for censorship.

To causally test the theory, I conducted an original online experiment in which I construct a custom-engineered, simulated social media environment. The simulated social media page is not interactive, meaning each respondent still completes the study independently. During the experiment on the simulated social media page, I give respondents in the treatment group an "encouragement" to participate in censorship by reporting the social media posts they encounter. I then use an instrumental variable analysis to estimate the local average treatment effect (LATE) of such experimentally induced censorship participation. Consistent with my theory, participation in censorship significantly increases individuals' support for government censorship. I also find suggestive evidence that merely providing respondents with the option to report online content can increase support for censorship.

This study contributes to two streams of literature on authoritarian politics: government

censorship and public participation. First, my study implies that censorship in authoritarian regimes should be perceived as a symbiotic relationship between the government and citizens. While the extant literature examines censorship in authoritarian regimes solely from a top-down angle (Gueorguiev and Malesky 2019; King, Pan and Roberts 2013; Lorentzen 2014; Miller 2018; Roberts 2018; Shadmehr and Bernhardt 2015), my study provides a novel theory to explain how this relationship could also be viewed through a bottom-up perspective. I show that ordinary users widely participate in the censorship process and such participation, in turn, shapes their opinion toward government censorship. More importantly, the bottom-up perspective helps to reconcile the empirical puzzle for why repressive apparatus such as censorship continues to garner widespread popular support in authoritarian regimes (Dickson 2016; Mitchell and Walker 2021; Wang and Mark 2015; Wike and Simmons 2015).

Second, I extend the literature on public participation in authoritarian regimes by showing a novel and perhaps insidious form of public participation. A wealth of literature has illustrated the causes and consequences of public participation in quasi-democratic institutions (Distelhorst and Hou 2017; Gandhi 2008; He and Warren 2011; Manion 2015; Truex 2016, 2017), and contentious social movements (Fu and Distelhorst 2018; Huang, Boranbay-Akan and Huang 2019). My study shows that public participation is also important in the implementation of repressive policies like censorship. Encouraging citizens' participation in the censorship process consolidates public support for authoritarian regimes' repressive apparatus. This semblance of public participation ironically suppresses individual rights to free speech and contributes to the durability of authoritarian regimes in the Internet era.

Participatory Censorship: A Bottom-Up Perspective

Although censorship in authoritarian regimes involves multiple levels of governments and non-governmental agencies (Han 2018; Lorentzen 2014; Tai and Fu 2020), it has always been understood as a top-down process imposed by the state and social media platforms upon

ordinary citizens. Roberts (2018) defines censorship as "the restriction of the public expression of or public access to information *by authority* [emphasis added] when the information is thought to have the capacity to undermine the authority by making it accountable to the public." Han (2018) also regards censorship as "tools used *by the state* [emphasis added] to limit the boundaries of online expression." Even though recent research has found that the purposes of government censorship extend beyond silencing political opposition to rewarding regime supporters (Esberg 2020), there is nevertheless a consensus that censorship is a top-down political tool of authoritarian regimes. However, this traditional top-down view of censorship overlooks the important role ordinary users play in censorship. In this study, I explore a novel, bottom-up perspective of censorship where ordinary citizens participate in censorship by reporting online content that they disapprove of.

Many authoritarian regimes have long histories of public participation in repressive policies and political campaigns. During the Cultural Revolution of China, for instance, ordinary citizens reported their friends, colleagues, and even families, voluntarily or involuntarily, to the communist government as "counter-revolutionaries," which often led to brutal state repression of the reported persons (Dikötter 2016; Thurston 1984; Yang 2021). Similar reporting behaviors were also common in other dictatorships (Gregory 2009).

The emergence of new communication technologies has not diminished the prevalence and significance of public participation in authoritarian repression. Today, almost every social media platform has the function to report online content. Despite legitimate reasons to allow users to flag inappropriate and socially harmful content, governments and citizens can and do take advantage of such a feature to achieve their political goals. For example, Tufekci (2017) documents that Turkish users on Twitter report political opponents as "spam" in an organized manner. Such mass reporting of opponents is often successful in getting accounts suspended temporarily or even permanently (Tufekci 2017). In China, participation in censorship is especially prevalent. Official statistics claim that the government received over 163 million censorship requests in 2020 alone. Public figures ranging from outspoken

dissidents like Fang Fang, to entertainment stars like Zhang Zhehan, to even government propagandist Hu Xijin, were all targets of these reports and many of their accounts were subsequently closed down (BBC News 2020; Reuters 2021). Some observers even refer to the rising "report culture" online as "China's digital cultural revolution," highlighting the disturbing parallels to the Mao era (BBC News 2020; Costigan and Xin 2020).

The Chinese government welcomes such public input and even encourages citizens to report online content. Since the establishment of the Cyberspace Administration of China (CAC) in 2014, encouraging public participation in censorship has been one of its main objectives. The CAC has an entire department, namely *jubao zhongxin*, dedicated to soliciting and handling censorship requests from ordinary users. It even ran an official propaganda campaign in Shanghai in 2020 to promote public participation in censorship (CNS 2020).

Why do authoritarian governments allow and even encourage such public participation? Theories of political participation in authoritarian regimes highlight the challenges of authoritarian governments in gathering various types of valuable information, such as public policy preference, popular grievance, and potential social unrest (Truex 2016). Due to the lack of democratic institutions, authoritarian governments rely on alternative quasi-democratic institutions (Distelhorst and Hou 2017; He and Warren 2011; Manion 2015; Truex 2016, 2017), and contentious social movements (Fu and Distelhorst 2018; Huang, Boranbay-Akan and Huang 2019), to gauge public opinion and social unrest.

In the context of censorship, authoritarian governments need to identify messages that are threatening to the regime (Gueorguiev and Malesky 2019; King, Pan and Roberts 2013). Yet, even in China, one of the most sophisticated censorship regimes, it is still difficult for the government to completely control the Internet (Roberts 2018). To effectively censor from a large amount of information, the Chinese government allegedly hires millions of censors (King, Pan and Roberts 2017), uses automated keyword filtering (Han 2018; Ng 2015), and uses a "friction" strategy to make all sensitive content harder to get (Roberts 2018).

Encouraging ordinary users to participate in the censorship process is another strategy

that authoritarian governments can and do employ to alleviate the information gathering problem on the Internet (Tufekci 2017). Instead of monitoring every corner of the entire Internet, users' reports provide valuable signals for the government and social media firms to conduct censorship. Sina Weibo, for example, establishes an algorithm (*quanzhong*) that takes users' reports into account when determining the publicity and censorship of every account and every post. By mobilizing millions of ordinary users to participate in the censorship process and establishing an algorithm that automatically filters widely reported posts, the cost of monitoring the Internet is significantly reduced. As a result, the dynamic of censorship in China is no longer solely top-down control but a mixture of top-down control and bottom-up participation.

Participation and Public Support for Censorship

How does public participation in censorship affect public opinion toward it? Existing research has shown that government censorship of political opposition is likely to cause public backlash against the regime (Pan and Siegel 2020; Roberts 2020). However, such a backlash effect is based on two assumptions: citizens believe (1) that the government is responsible for the censorship event, and (2) that censorship negatively affects their welfare. I argue that public participation in the censorship process diminishes the government's responsibility. In addition, citizens are more likely to feel empowered by participating in censorship and are less likely to view themselves as victims of censorship. As a result, public support for censorship is more likely to increase.

Diminishing Government's Responsibility

Authoritarian governments are cautious not to take the blame for suppressing political opposition. The Chinese government, in particular, regularly outsources repression to non-state actors to avoid the impression that it is directly repressing political dissent (Ong 2022). In

the case of censorship, existing studies have identified two main "scapegoats" for the Chinese government. First, the Chinese government strategically mobilizes its pro-regime base to fabricate millions of posts and counter online critics while avoiding direct censorship of political opposition (Chen and Xu 2017; King, Pan and Roberts 2017; Miller 2018). Such a strategy diverts the focus of regime critics from the government to their pro-regime counterparts. Second, the Chinese government also uses social media platforms as another "scapegoat," openly criticizing them when backlash against online censorship occurs (Miller 2018). As a result, the blame for censorship is diverted from the government to social media companies.

Public participation in the censorship process further contributes to the diffusion of the responsibility of the government. First, public participation gives the government a passive image in censorship events. It creates the perception that the government is not the initiator of censorship and is merely responding to public demand. As such, even if citizens do not like certain censorship events, they might think that it is due to other users' reports, and therefore less likely to blame the government. Second, public participation also makes censorship activities seem more popular. Traditionally, censorship is often perceived as government suppression of free speech against the will of the people. When the public is actively participating in the censorship process, it creates the perception that censorship is not just the will of the government but also the will of many ordinary users. This perception of popularity and public acceptance itself can reduce citizens' inclinations to speak out against the regime and its censorship policy (Huang and Cruz 2021).

Increasing Citizens' Perceived Empowerment

In addition to diffusing the responsibility, authoritarian governments also frequently engage in justification of their repression to increase citizens' perceived empowerment by repression. The goal of such justification is to persuade the public that they are not the victims, but rather the beneficiaries of repressive policies (Esberg 2018; Williamson and Malik 2019). The Chinese government also tries to shift public perception of censorship in this regard. For

example, Yang (2022) finds that the Chinese government expands the range of censorship targets beyond politically threatening content to other seemingly harmless non-political content such that the public no longer regards censorship as political repression but rather a normal government policy that benefits ordinary users.

Because participation in censorship allows ordinary users to report content that they disapprove of, it reduces citizens' perception that they are victims of censorship by increasing their perceived control over the censorship outcome. Because ordinary users can report online content that affects subsequent censorship activities, it delegates some, albeit small, amount of censorship power to ordinary users and therefore increases their perception of control over the content being censored. Existing studies in social and political psychology have found that individuals with a low perception of control are more susceptible to conspiracies, less politically engaged, and more likely to attribute influence or blame to political enemies (Nyhan and Zeitzoff 2018; Smith 2021; Sullivan, Landau and Rothschild 2010; Whitson and Galinsky 2008). When censorship is solely a top-down process imposed upon ordinary users, they are more likely to have cynical views of the government and its censorship policy, perceive themselves as victims of censorship events, and blame the government for such negative consequences. On the contrary, because of the increased perception of control over censorship outcomes, individuals are more likely to view government censorship activities as enforcing their censorship preferences. Hence, they are more likely to feel empowered by censorship and view themselves as beneficiaries of censorship rather than victims.

The changes in the perceived government responsibility and citizen empowerment, in turn, likely affect how ordinary citizens view censorship. From the perspective of ordinary users, their participation in the censorship process redefines the government's role as an arbitrator of conflicting interests on the Internet rather than a manipulator of public opinion. They are less likely to view censorship as top-down political suppression imposed upon them, and they are more likely to treat censorship as a tool they can use to their advantage: to suppress political opponents. As a result, ordinary citizens are more likely to support the

government's censorship policy.

Hypothesis: As individuals participate more in the censorship process, they should display greater levels of support toward government censorship.

Study 1: Online Survey

Although censorship participation can be observed in various types of autocracies (Tufekci 2017), I focus on China, one of the most sophisticated censorship regimes, to illustrate the features and consequences of participatory censorship. To gauge the prevalence of public participation in censorship among the Chinese public and its correlation with support for censorship, I conducted an original online survey in December 2021.[1] The survey recruited 1,124 respondents via a Chinese online survey platform. After they agreed to participate, the respondents were re-directed to an American-based website, Qualtrics, where they completed the survey anonymously. I employ a quota sampling strategy to recruit respondents with a wide range of socioeconomic backgrounds. However, like many other online surveys in China, the sample will inevitably be younger and better-educated than the general Internet population (Huang 2018; Huang and Yeh 2019; Pan and Xu 2020). To address this concern, I weight the survey sample using the iterative proportional fitting algorithm such that the weighted sample will be almost identical to the Chinese Internet census in terms of gender, rural/urban location, region, and age distributions.[2]

Measurement

Participation in Censorship

To measure *Participation in Censorship*, I directly ask the respondents if they have reported online content and speeches before. Respondents are given five choices: never, once

[1]This study was approved by the Institutional Review Board (IRB) at the researcher's home institution and was pre-registered on Open Science Framework.

[2]For more information on the survey sample and weighting, see appendix.

or twice only, once per few months, once per month, and multiple times per month. The responses are coded on a five-point scale. As a robustness check, I also re-code the variable as a binary variable indicating whether the respondent has reported before. Although surveys in authoritarian regimes often face social desirability bias problems, it is unlikely to be a concern here. As discussed in previous sections, censorship participation is a prevalent behavior in China and the Chinese government holds a favorable view toward such public participation. Hence, it is unlikely that respondents are fearful of reporting their previous behaviors.

For those respondents who self-report having participated before, I further query those respondents with participation experiences about the specific types of content they requested to censor by providing a list of content for them to choose from. Specifically, I am interested in whether they have requested censorship of political content, including political news, political commentary, political opinions, political rumors, and foreign media coverage of China. I also include other non-political content such as entertainment, advertisement, vulgar language, and pornography in the list.

Outcome Variables

The main dependent variable of the analysis is *Support for Government Censorship*. I ask the respondents if they agree that the government should actively control the Internet and remove content that it deems inappropriate. The variable is measured on a five-point scale. I further measure their *Support for Censorship of Political Content* and *Support for Censorship of Non-Political Content* by asking whether the government should control online discussion of government policies and party leadership, as well as entertainment stars and popular culture.

To test the mechanisms proposed in the theory, I measure their *Perceived Empowerment* by asking the respondents whether ordinary people are the victims or the beneficiaries of the current censorship activities. Second, I measure respondents' *Perceived Government Responsibility* by asking the respondents who they think should be responsible for most of

the deleted content: the netizens, the government, or the platforms?

Control Variables

I include two different sets of control variables: demographic covariates and predisposition covariates. Demographic covariates include *education*, *age*, *gender*, and *urban/rural location*. Predisposition covariates include *party membership*, *political interests*, *political ideology*, *economic ideology*. The first two covariates are commonly used in surveys in China (Huang 2018). The third and fourth questions measuring ideology are directly borrowed from Pan and Xu (2020).

Results: Prevalence of Public Participation

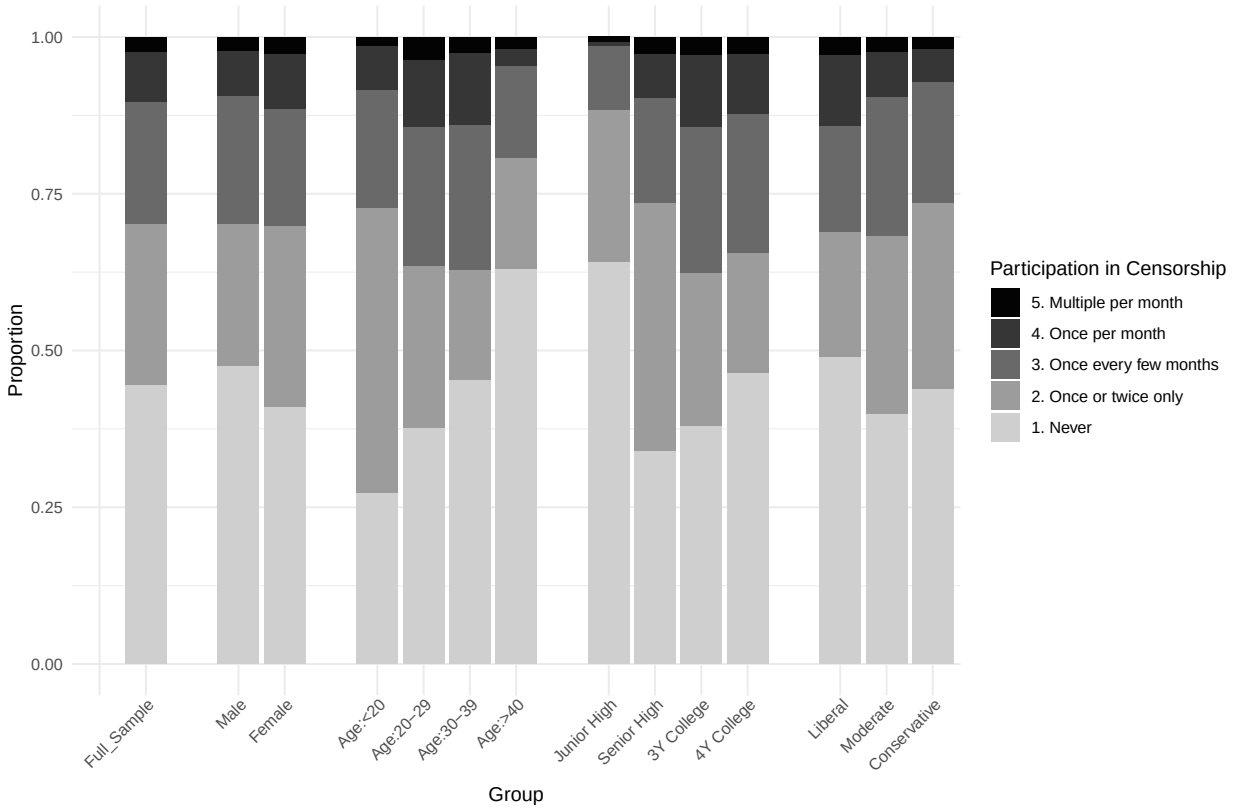
How prevalent is public participation in censorship among the Chinese public? Figure 1 presents the distribution of self-report participation in censorship using the weighted sample. As shown in the first bar on the left, more than 50% of the weighted sample self-report having previously participated in censorship. More than 25% of the weighted sample have participated at least once every few months. These results demonstrate that public participation in censorship is prevalent among Chinese internet users. Such behavior is especially ubiquitous among the younger generation. Almost three-quarters of respondents under 20-year-old self-report having such censorship experience, and around two-thirds of respondents in their 20s report similar experiences. The younger generation, they are regularly exposed to online discussions about the use of reporting as a strategy to censor other opposing opinions (Luo and Li forthcoming). As such, participation in censorship is both normal and a useful tool for these young people on the Internet.

Moreover, the participation rate does not vary significantly by political ideology.[3] To be clear, there is not a clear political cleavage in the Chinese society Pan and Xu (2020), but existing studies have found groups that self-identify as more liberal and critical of the

[3]I measure political ideology by asking whether they agree or disagree that: From a long-term perspective, multi-party democracy is suitable for China. The question is borrowed from Pan and Xu (2020). Those who agree with the statement are categorized as liberal, whereas those who disagree are conservative.

regime compared to other groups that are more supportive of the regime. These ideologically diverging groups both interact with each other as well as the state censorship apparatus (Han 2018). The results here show that even politically liberal citizens choose to weaponize the "report" function for political gains. This implies that even though many individuals might object to the authoritarian system as a whole, they do not resist participating in the repressive apparatus so long as it serves their immediate interests.

Figure 1: Distribution of Self-Report Participation in Censorship



Note: All observations are weighted by gender, rural/urban location, region, and age group. Unweighted sample shows a slightly higher proportion of respondents self-reporting participation in censorship.

What specific content did these "participating respondents" report? Most "participating respondents" (around 90%) have flagged inappropriate content online including pornography, advertisement, and vulgar language. A significant proportion has also reported political content. In general, around 50% of the "participating respondents," or 25% of all respon-

dents, self-report having participated in censorship of political content. Similar to the general trend described above, younger generations and the better-educated are significantly more likely to participate in the censorship of political content. Both liberals and conservatives have engaged in political censorship with a slightly higher participation rate among liberals. About one-third of the 'participating respondents" have reported entertainment and cultural content. In sum, the descriptive analysis demonstrates significant levels of public participation in censorship overall and across different demographic groups and different categories of online content.

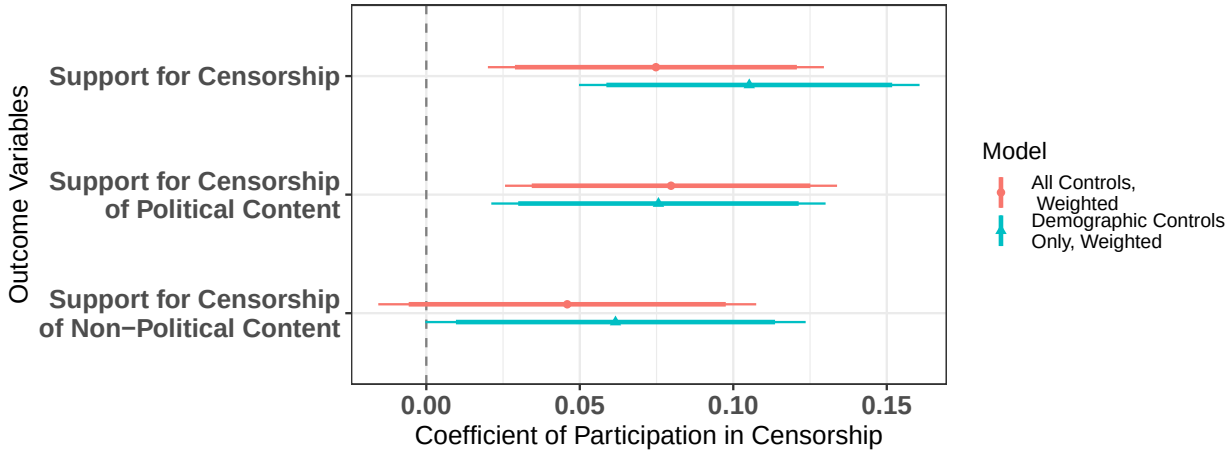
Participation and Censorship Support

Do individuals with higher levels of participation in censorship hold more favorable views toward government censorship? Figure 2 reports the results of regressing *Support for Censorship* on *Participation in Censorship* using OLS models. Consistent with the hypothesis, individuals' participation in the censorship process positively and significantly correlates with individuals' support for government control of the Internet. This result is robust to models with control variables and sample weighting. All else equal, respondents that have a higher participation level by one-unit express stronger support for government censorship by 0.074 units, which is equivalent to 7.35% of a standard deviation of the dependent variable. In terms of respondents who participate in censorship multiple times per month (around 3% of the sample), they score higher on the censorship support scale than those who have never participated in censorship (around 43% of the sample) by 0.296 units, a difference that is comparable to existing survey experiments in China (Huang 2018).[4]

The effects of censorship participation on *Support for Censorship of Political Content* and *Support for Censorship of Non-Political Content* are also positive. However, it loses statistical significance for censorship of non-political content. This is perhaps due to the

[4]In the Online Appendix, I replicate the analyses with binary measures of participation in censorship. The effect on support for censorship is still positive but only significant at the 0.1 level when controlled for all covariates.

Figure 2: Correlation between Participation in Censorship and Support for Censorship



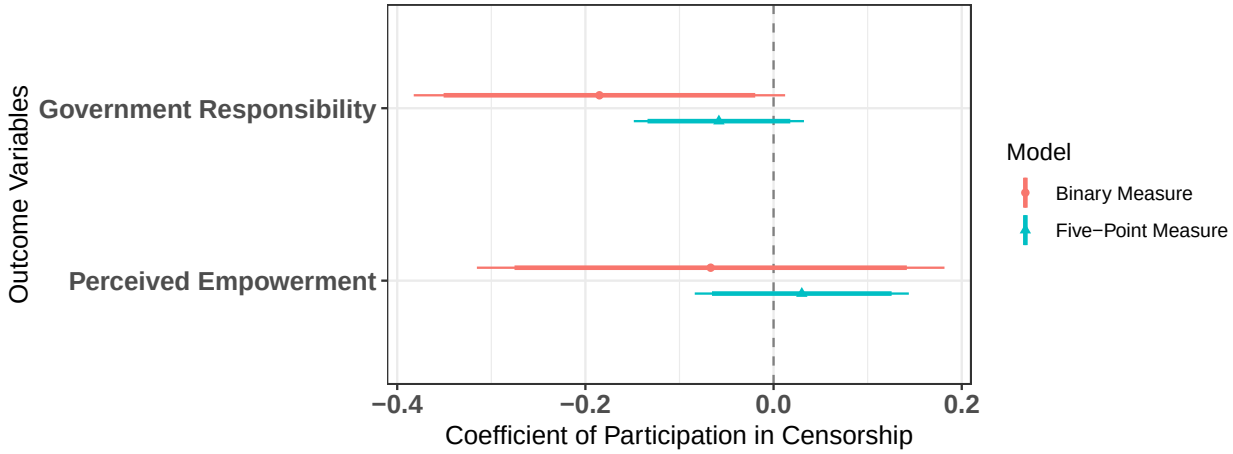
Note: All outcome variables are measured on a five-point Likert scale. Participation in censorship is measured on a five-point scale: never participated, once or twice only, once per few months, once per month, and multiple times per month. Bars indicates 90% and 95% confidence intervals.

lower proportion of respondents that have reported entertainment and cultural content. As mentioned in the previous section, only one-third of the "participating respondents" have reported entertainment compared with one-half who have reported political content.

I then test the mechanisms proposed in the theory that public participation increases support for censorship by diminishing the government's responsibility and strengthening citizens' perceived empowerment I perform OLS regressions to test the correlations between *Perceived Government Responsibility*, *Perceived Empowerment* and *Participation in Censorship*. As shown in Figure 3, consistent with the theory, respondents who participated in censorship are less likely to believe that the government is responsible for censorship events. However, the effect is only significant at the 0.1 level when using binary measures of participation and not significant when using five-point measures. The data provide less support for the perceived empowerment mechanism with both models showing a null effect. One possible explanation for the insignificant results is the lack of variation in responses. The variances of both mechanism outcomes are lower than censorship support questions and responses are concentrated on the middle choices. This might suggest that respondents lack a

clear understanding of these questions.

Figure 3: Correlation between Participation in Censorship and Mechanism Outcome Variables



Note: All models include demographic and predisposition control variables, as well as weights. Binary measure models use binary measure of participation in censorship, whereas five-point measure models use five-point measure of participation in censorship. Bars indicates 90% and 95% confidence intervals.

Taken together, the results from the observational analyses support my claim that public participation in censorship is common among the Chinese public. About one-half of the respondents self-report participated in censorship. Moreover, such participation is positively correlated with support for censorship agreeing with my main hypothesis, although the data provide limited support for the mechanisms proposed in theory. These findings highlight the relevance of the bottom-up perspective of censorship in authoritarian regimes. If one-half of the online population is actively engaged in reporting online content, it is important to analyze its consequences in order to fully understand how citizens in authoritarian regimes view censorship and repressive government policies in general.

However, the analyses above do not rule out the possibility of a reversed causal arrow: individuals who support censorship in the first place are more likely to participate in it. To provide stronger support for the effect of public participation on support for censorship, I will need to randomly assign participation in censorship and causally test the hypotheses.

Study 2: Survey Experiment

Building on the first study, I conducted an original online survey experiment on a custom-engineered, simulated social media page.[5] The simulated social media environment is not interactive, meaning each participant still completed the survey independently. The goal of the experiment is to test the effect of participation in censorship on public support for government censorship. Because it is difficult to directly manipulate censorship behavior, I use two different strategies to test my theory. First, I use an instrumental variable approach, i.e., intention-to-treat design. I give respondents in the treatment group an encouragement to participate in censorship and measure the local average treatment effect (LATE) on support for censorship among compliers (Aronow and Carnegie 2013; Marbach and Hangartner 2020).

Second, I test the effect of *providing the opportunity to participate in censorship* on support for censorship. Admittedly, providing participation opportunities is not the same as actual participation. Nonetheless, my theory about how participation in censorship leads to higher support for censorship applies to providing the opportunity to participate as well. Such an analysis is a useful complement to the instrumental variable approach and provides a fuller picture of how public participation leads to censorship support.

Using such a design has several advantages. First, it avoids ethical concerns associated with a similar field experiment. The current political climate in China is hostile toward political field experiments and might put both participants and researchers at higher risk of authoritarian repression. Survey research reduces such risk significantly. In addition, encouraging respondents to participate in censorship in the real world will also be normatively undesirable and might further contribute to the reporting culture on the Chinese Internet. Conducting the experiment in a simulated setting limits the potential negative impact of the research. Finally, reporting behaviors are not publicly observable in the real world and, as confirmed in my experiment, a substantial proportion of respondents object to the idea of

[5]This study was approved by the Institutional Review Board (IRB) at the researcher's home institution and was pre-registered on Open Science Framework.

reporting and refuse to participate in censorship even after encouragement. An experiment embedded in a simulated environment enables me to measure whether respondents participated, estimate the treatment effects among compliers, and identify those who refuse to participate regardless of treatment.

Procedure

The survey experiment was conducted in June 2022 in mainland China. Similar to the online survey in study 1, I recruited 4,008 respondents from a Chinese online survey platform and then directed them to an American-based website, Qualtrics, where they completed the survey anonymously. As was the case with Study 1, the sample covers a wide range of socioeconomic backgrounds but is younger and better-educated than the general Internet population.

The experiment consists of three parts. First, I measure pre-treatment covariates. Second, I randomly assign respondents to one of the three groups: a control group, a treatment group, and a blank control group. All respondents read the same 10 social media posts on a heatedly debated current event, Xuzhou chained woman incident, on my simulated social media page. All posts are adapted from real Weibo posts with modified user names and avatars. Respondents are fully informed about how these posts are created from real posts, and therefore no deception is used. Finally, I measure respondents' support for censorship.

On the simulated social media page, I build multiple buttons that the respondents can click under each post. In both the control group and the treatment group, these buttons are: "Like," "Share," "Comment," and "Report." To manipulate respondents' participation in censorship in the simulated social media environment, I give the respondents in the treatment group an "encouragement" to use the "Report" button. Specifically, respondents in the treatment group were shown the following paragraph:

> *We are especially interested in what posts you want to report. Please choose at least two posts that you think should be removed by the Internet regulator, and*

press the Report button to let us know.

In the blank control group, I only include three buttons that the respondents can click: "Like," "Share," and "Comment." In other words, respondents in the blank control group do not have the opportunity to report. Naturally, they do not receive the "encouragement" to report as well. In the instrumental variable analyses, I only include the treatment and control groups but not the blank control group. Table 1 summarizes the treatment assignment.

Table 1: Treatment Summary

Groups	Blank Control	Control	Treatment
Buttons Under Simulated Posts	Like, Share, Comment	Like, Share, Comment, Report	Like, Share, Comment, Report
Encouragement Message	No	No	Yes

Notes: The instrumental variable analysis only includes the control and treatment groups. To test the effect of providing participation opportunities, I only include the blank control and the control groups, but not the treatment group.

Selection of Topic

The Xuzhou "chained-woman" incident is a recent event that was heatedly debated on Sina Weibo. In January 2022, a video of a trafficked woman held in chains in a hut in Fengxian County, Xuzhou City, Jiangsu Province for years went viral. The local government tried to cover up the story by issuing contradictory statements. This caused widespread public anger on the Internet. The government and officials were heavily criticized for both causing such tragedy and trying to cover it up. However, the incident coincided with the 2022 Beijing Winter Olympics. Such timing prompted many patriotic regime supporters to argue that this is a conspiracy to defame China. They urged the public not to give further attention to this incident.

In the experiment, I selected ten posts that discuss the event. Among the ten posts, five

of them could be seen as pro-government or nationalistic, while the other five could be seen as anti-government or pro-individual rights. The order of the posts is randomized.

Measurement

To measure respondents' participation in censorship, I use a binary variable to indicate whether they click any of the "Report" buttons. Similar to the observational survey, my main dependent variable of the analysis is *Support for Government Censorship*. I also measure their *Support for Censorship of Political Content* and *Support for Censorship of Non-political Content*. To test the mechanisms, I include similar survey items in study 1: *Perceived Empowerment* and *Participation in Censorship*.

I use ten covariates to check the balance between the treatment and control groups. They are also included in the regression analyses. Among the ten covariates, four are demographic variables, including *education*, *age*, *gender*, and *region*. These demographic covariates are widely used in experiments across different contexts. The remaining six covariates are predisposition covariates including *party membership*, *nationalism*, *political interests*, *ideology*, *social media usage*, *foreign connection*. All covariates are balanced across the three experimental groups.

Instrumental Variable Analysis

The main identification strategy is an instrumental variable analysis. I use the encouragement treatment as an instrument to test the local average treatment effect (LATE) of participation in censorship (clicking the report button) on support for censorship. Formally:

$$\begin{aligned} \text{Clicking the Report Button}_i &= \alpha + \gamma \cdot \text{Encouragement Treatment} + \lambda Z_i + \epsilon_i \\ Y_i &= \xi + \beta \cdot \text{Clicking the } \widehat{\text{Report}} \text{ Button}_i + \delta Z_i + \mu_i \end{aligned}$$

where Y_i is the outcome measure; Z_i is a vector of pre-treatment covariates; and β is the LATE (or the average treatment effect among compliers). This analysis only includes the treatment and control groups, but not the blank control group.

In the control group, 43% of the respondents clicked the "Report" buttons on the simulated social media page, whereas 64% of the respondents in the treatment group clicked. Hence, there are roughly 20% of the respondents clicked the "Report" buttons due to the encouragement treatment. The instrumental variable analysis estimates the causal effects of censorship participation induced by the encouragement treatment.

Table 2 reports the results from instrumental variable analyses and the LATE of participating in censorship in the simulated social media environment. Consider, first, column 2. After controlling for pre-treatment covariates, participation induced by the encouragement treatment significantly increases respondents' general support for government censorship ($\beta = 0.402$, $p = 0.017$). This again provides direct support for the main hypothesis and more importantly, it alleviates the concerns in study 1 that the causal arrow might be reversed. The magnitude of the treatment effect on support for censorship is equivalent to 46.3% of a standard deviation. It is a substantial increase considering the baseline support for censorship is already high in the control group (3.59 out of 5).

Censorship participation induced by the encouragement treatment also increases specific support for censorship of political content ($\beta = 0.305$, $p = 0.072$) and non-political content ($\beta = 0.319$, $p = 0.081$), although both coefficients are only significant at the 0.1 level. The lower level of significance is partly consistent with the observational analyses in study 1. This might suggest that while in general, the results are consistent with the theoretical expectations, it is more difficult to move individuals' opinions on specific censored content.

To check the robustness of the treatment effect, I use an alternative measurement of participation in censorship and re-run the instrumental variable analyses. Instead of a binary variable indicating whether the respondents clicked any of the "Report" buttons, I use the number of times the respondents clicked a "Report" button. Table 3 reports the LATE

Table 2: Local Average Treatment Effect (ATE among Compliers) of Participating in Censorship on Support for Censorship

	Support for Censorship		Support for Censorship of Political Content		Support for Censorship of Non-Political Content	
	(1)	(2)	(3)	(4)	(5)	(6)
Report Click	0.417*	0.402*	0.330†	0.305†	0.299†	0.319†
	(0.169)	(0.168)	(0.169)	(0.170)	(0.180)	(0.183)
Constant	3.379**	1.977**	3.392**	2.097**	3.579**	2.216**
	(0.092)	(0.193)	(0.092)	(0.195)	(0.098)	(0.211)
Covariates		✓		✓		✓
N	2,647	2,493	2,653	2,499	2,645	2,492

Notes: Dependent variables are indicated in column headings. Standard errors in parentheses. Report click is a binary variable indicating whether the respondents have clicked any of the "Report" buttons on the simulated social media page. All individual survey items were measured on a five-point scale.
†p < .1; *p < .05; **p < .01

of censorship participation. Consistent with the main analyses, additional clicking of the "Report" buttons induced by the encouragement treatment significantly increases support for censorship.

I then test the mechanisms by running the same instrument variable models to identify the LATE of the encouragement treatment on *Perceived Government Responsibility* and *Perceived Empowerment*. As Figure 4 demonstrates, although both estimates are in the correct direction as expected by the theory, both of them are statistically insignificant. This echoes the findings in study 1 that participation in censorship does not significantly increase perceived empowerment as measured by the survey item.

Profiling Compliers

Although the results from the instrumental variable analyses show positive and significant treatment effects of censorship participation on support for government censorship, they only capture the effects of the treatment among a subgroup of the sample: the compliers. Hence, it is important to compare the background attributes of compliers and non-compliers.

Table 3: Robustness Check of the LATE (ATE among Compliers) of Censorship Participation on Support for Censorship

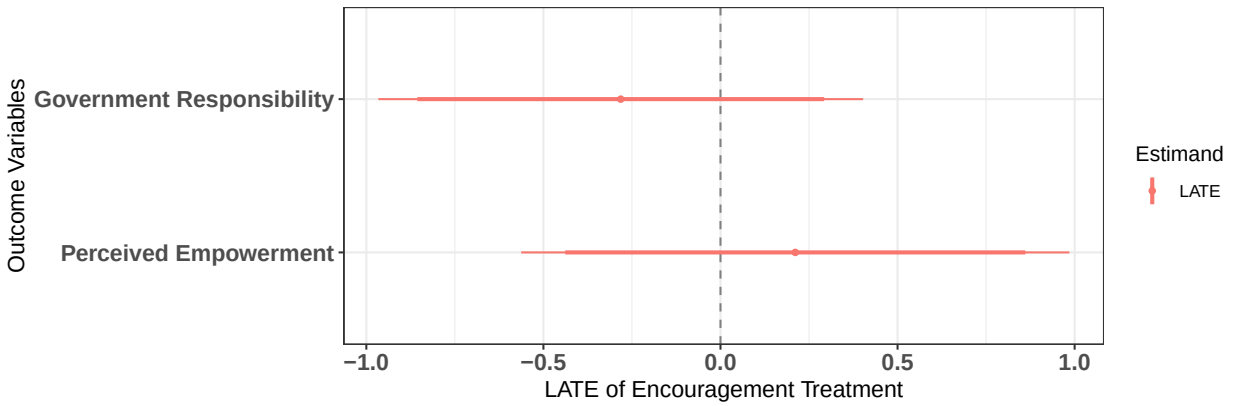
	Support for Censorship		Support for Censorship of Political Content		Support for Censorship of Non-Political Content	
	(1)	(2)	(3)	(4)	(5)	(6)
Report Click #	0.143*	0.138*	0.113†	0.104†	0.102	0.109†
	(0.059)	(0.059)	(0.059)	(0.059)	(0.062)	(0.064)
Constant	3.385**	1.934**	3.398**	2.066**	3.584**	2.183**
	(0.092)	(0.211)	(0.091)	(0.212)	(0.096)	(0.229)
Covariates		✓		✓		✓
N	2,647	2,493	2,653	2,499	2,645	2,492

Notes: Dependent variables are indicated in column headings. Standard errors in parentheses.
Report click number is the number of the "Report" buttons that the respondents clicked on the simulated social media page. All individual survey items were measured on a five-point scale.
LATE = Local Average Treatment Effect
†p < .1; *p < .05; **p < .01

Following Marbach and Hangartner (2020), I plot out the characteristics of compliers, always-takers, and never-takers, assuming there are no defiers. As shown in the top-left panel of Figure 5, compliers only account for around 20% of the full sample, whereas never-takers and always-takers consist of about 36% and 43% respectively.

Although such a low proportion of compliers might be a problem for the generalizability of the treatment effects, a comparison of pre-treatment covariates among compliers, never-takers, and always-takers suggests that it might be less concerning. As Figure 5 illustrates, always-takers are more likely to be young, social media savvy, and non-party members, whereas never-takers are more likely to be old, nationalists with few foreign connections and social media exposure. Hence, those never-takers are likely to be more pro-regime and they are not necessarily the subgroup that the regime tries to win over by promoting participation in censorship. Compliers, in comparison, are young, better-educated, more economically liberal, and less nationalistic people with more foreign connections. Hence, because of the similarity in background attributes between compliers and always-takers, it is reasonable to assume that the treatment effect of participation in censorship is likely to

Figure 4: LATE (ATE among Compliers) of Censorship Participation on Mechanism Outcomes Variables



Note: All models include pre-treatment covariates. Bars indicates 90% and 95% confidence intervals. LATE = Local Average Treatment Effect

apply to always-takers. More importantly, compliers and always-takers are more likely to be on the Internet and therefore more likely to be the target population of my theory on public participation in censorship.

Intention-To-Treat Analysis

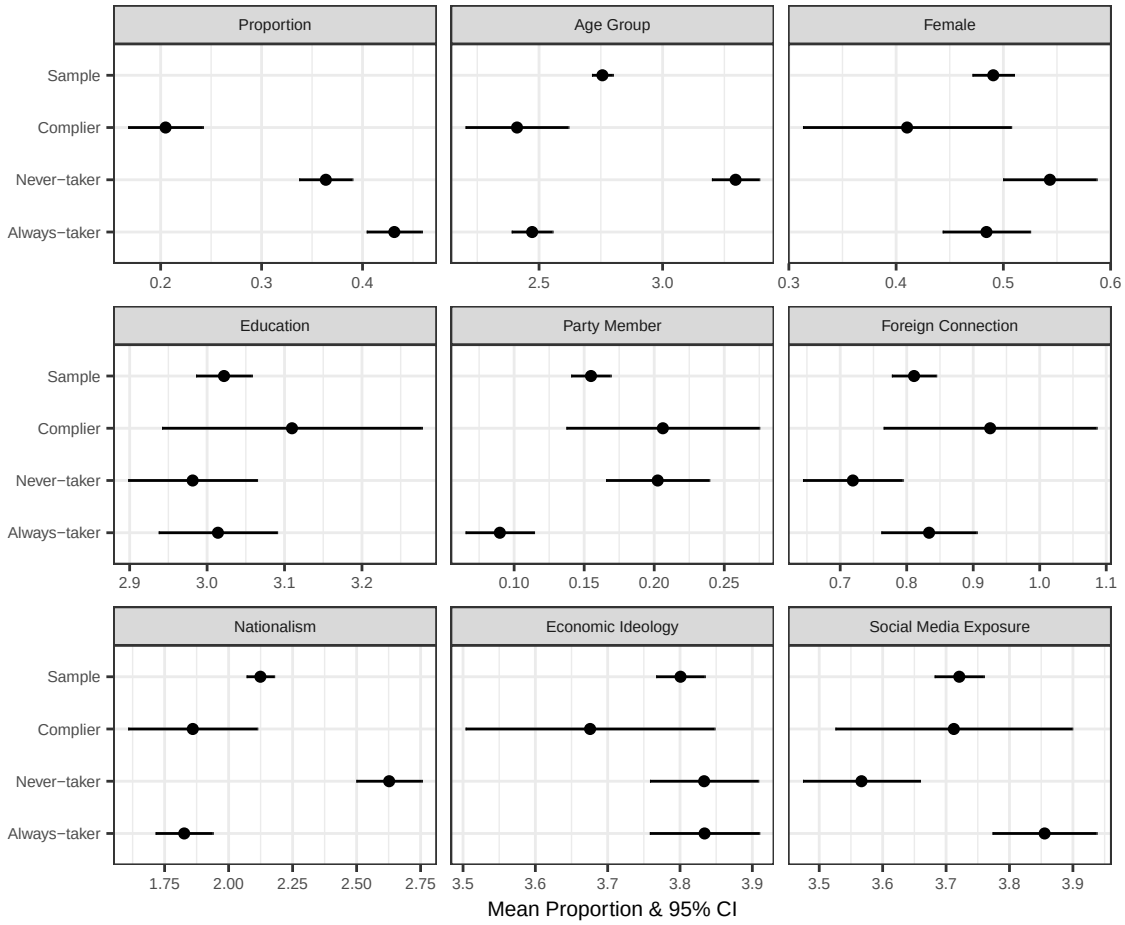
Because of the low proportion of compliers, I also conduct an intention-to-treat analysis in which respondents are analyzed according to the group they were assigned to rather than whether they actually clicked the "Report" button. In other words, I use simple t-tests and OLS regressions to estimate the effect of treatment assignment. Formally:

$$Y_i = \alpha + \beta \cdot \text{Encouragement Treatment} + \gamma Z_i + \epsilon_i$$

where Y_i is the outcome measure; Z_i is a vector of pre-treatment covariates; and β is the average treatment effect of the encouragement treatment.

Figure 6 shows the results from the intention-to-treat analysis. Consistent with the instrumental variable analysis, the encouragement treatment has a positive and significant effect on general support for government censorship ($\beta = 0.086$, $p = 0.011$) and specific

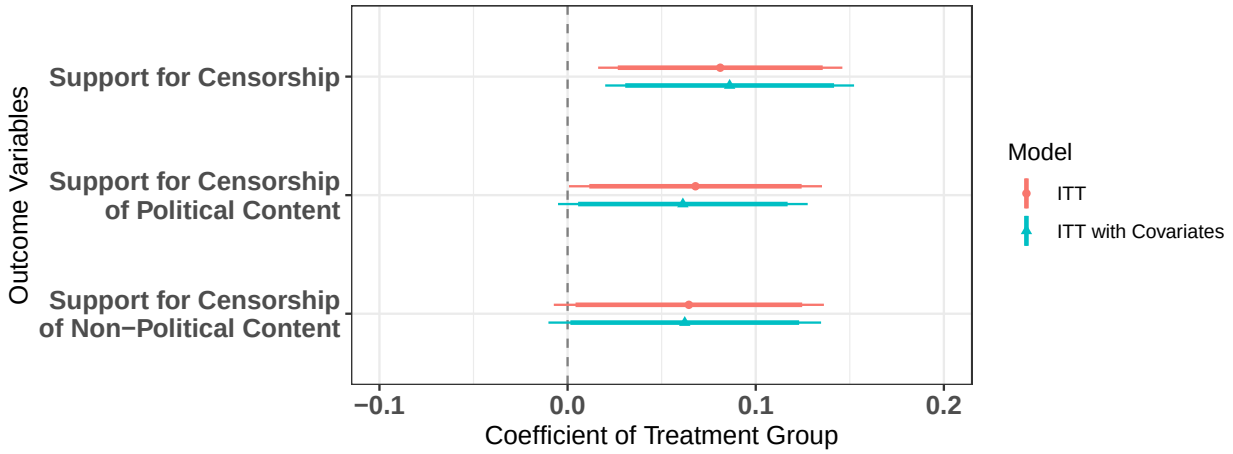
Figure 5: Profiling Compliers, Never-Takers, and Always-Takers



Note: The first panel indicates the estimated proportion of compliers, never-takers, and always-takers. The remaining eight panels demonstrate the estimated group means of the full sample, compliers, never-takers, and always-takers across eight different pre-treatment covariates. CI = Confidence Interval

support for censorship of political content ($\beta = 0.068$, $p = 0.048$). The treatment effects on support for censorship of non-political content ($\beta = 0.062$, $p = 0.092$) is only significant at the 0.1 level. In contrast to the LATE, however, the magnitude of the intention-to-treat effect is much smaller, only equivalent to 9.9% of a standard deviation of the dependent variable. This is understandable since the proportion of compliers is low. In sum, both the instrumental variable analysis and the intention-to-treat analysis show a significant treatment effect on support for censorship.

Figure 6: The Intention-To-Treat Effects of the Encouragement Treatment on Support for Censorship



Note: All outcome variables are measured on a five-point Likert scale. The analyses only include the control group and the treatment group. Bars indicates 90% and 95% confidence intervals. ITT = Intention-to-Treat

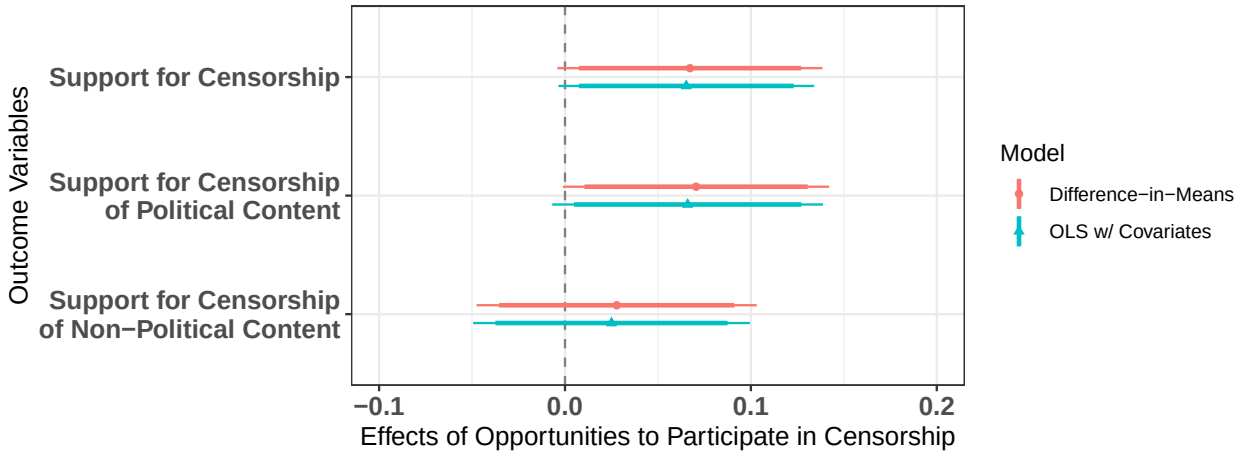
The Effects of Providing Participation Opportunities

The analyses above demonstrate that participation in censorship induced by the encouragement treatment significantly increases support for censorship. To complement the intention-to-treat design, I further test the effect of providing the opportunities to report online content on support for censorship. Theoretically, providing such opportunities might also diminish perceived government responsibility for censorship and increase citizens' perceived empowerment by the censorship apparatus. Hence, I should expect providing participation opportunities to have a positive effect on individuals' support for censorship.

In this analysis, I compare the blank control group to the control group using both t-tests and OLS regressions with pre-treatment covariates. The only difference between the two groups, and therefore the "treatment," is that the control group has an additional "Report" button under each post, whereas the blank control group does not. As shown in Figure 7, providing the opportunity to participate in censorship through the "Report" buttons increases individuals' support for censorship in general ($\beta = 0.067$, $p = 0.065$) and censorship of political content in particular ($\beta = 0.066$, $p = 0.076$). However, both relationships

are only significant at the 0.1 level. The effect on support for censorship of non-political content is positive but insignificant. Hence, the results provide suggestive evidence that merely providing the opportunity to flag online content can increase individuals' support for censorship. Yet, its effectiveness might not be as large and significant as actual participation in censorship.

Figure 7: The Effect of Providing Participation Opportunities on Support for Censorship



Note: All outcome variables are measured on a five-point Likert scale. The analyses only include the blank control group and the control group. Bars indicates 90% and 95% confidence intervals. OLS = Ordinary Least Squares

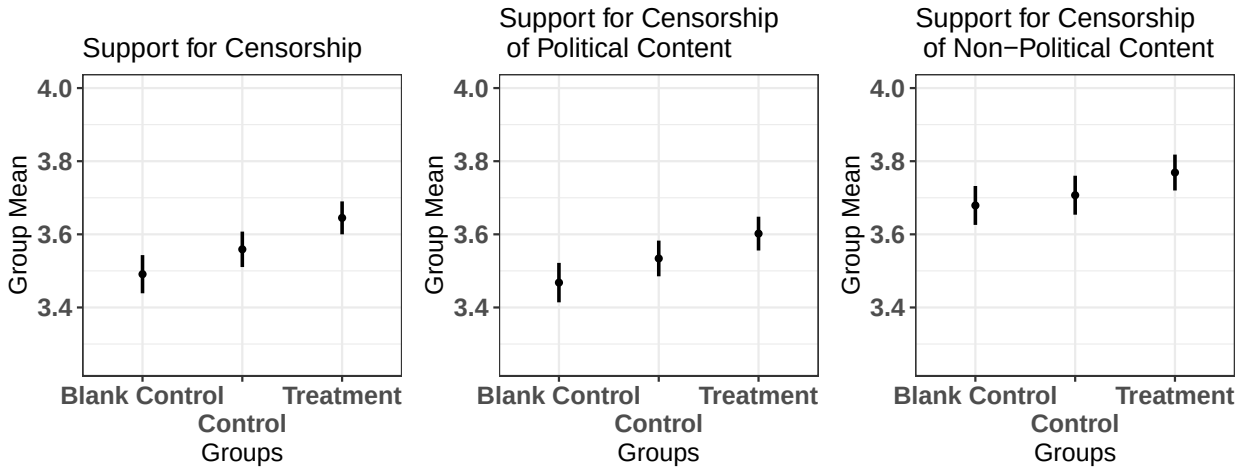
Discussion

To sum up the findings in the experiment, I find a consistent increase in support for censorship first from the blank control group to the control group, and then from the control group to the treatment group (see Figure 8).[6] This indicates that both experimentally induced participation in censorship and providing the opportunities to participate increase support for censorship to various degrees. The instrumental variable analysis further demonstrates that the treatment effects among compliers are large and significant. The LATE on support for censorship is equivalent to almost one half of a standard deviation of the dependent

[6]The effects on support for censorship of non-political content is smaller and less significant. It might be due to the topic I chose in the experiment is a political one.

variable.

Figure 8: Group Means of the Dependent Variables



Note: All three items are measured on a five-point Likert scale.

The implications of the experimental results are three-fold. First, it echoes the findings in the observational survey that public participation in the censorship process is indeed a prevalent phenomenon in authoritarian China. Always-takers account for more than 40% of the full survey sample, meaning a substantial proportion of the respondents voluntarily participate in the censorship process even without experimental treatment. Putting compliers and always-takers together, a simple encouragement message can lead to almost two-thirds of the respondents participating in censorship.

Second, the results reaffirm the positive relationship between widespread participation in the censorship process and public support for government censorship. It indicates that the button-up perspective of censorship has significant consequences on public opinion. A more comprehensive approach to authoritarian censorship that combines top-down control and bottom-up participation is needed to better understand the dynamic of censorship and authoritarian control. Moreover, the results suggest that while the government might benefit from merely providing participation access to the public, it is the encouragement to participate that eventually leads to a significant increase in censorship support.

Third, the profiling compliers analysis suggests that the effect is primarily driven by young, well-educated, pro-market individuals with many foreign connections. This group of individuals are traditionally been understood as more open to Western values and less susceptible to authoritarian controls (Huang 2015). This suggests that public participation in repressive policies like censorship might be an effective mechanism through which the authoritarian regime generates support from these young, well-educated, pro-market, and global-minded citizens, who are traditionally less inclined to support the regime. By encouraging censorship participation among these subgroups, the regime successfully sustains popular support across generations.

However, as mentioned above, there might be concerns about the generalizability of the survey experiment. First, participation behaviors might be affected by the selected topic and posts. Indeed, a regression analysis of the reporting behaviors suggests that while party members are in general more likely to participate in censorship, they are less likely to click the "Report" buttons in this specific experiment. In other words, the content selected in the experiment might affect the composition of compliers, always-takers, and never-takers. Thus, future research should test the same arguments and hypotheses in different scenarios to provide stronger support and improve generalizability. Second, as shown in the profiling compliers analysis, the proportion of compliers is relatively small. Only 20% of the respondents complied with the experimental design. Future studies should explore more creative design that expands the proportion of compliers to better estimate the average treatment effect among the entire sample.

Alternative Explanations

An important limitation of the study is that I do not find strong support for either of the proposed mechanisms. Although in the observational survey, self-report participation in censorship negatively correlates with belief that the government is responsible for censorship,

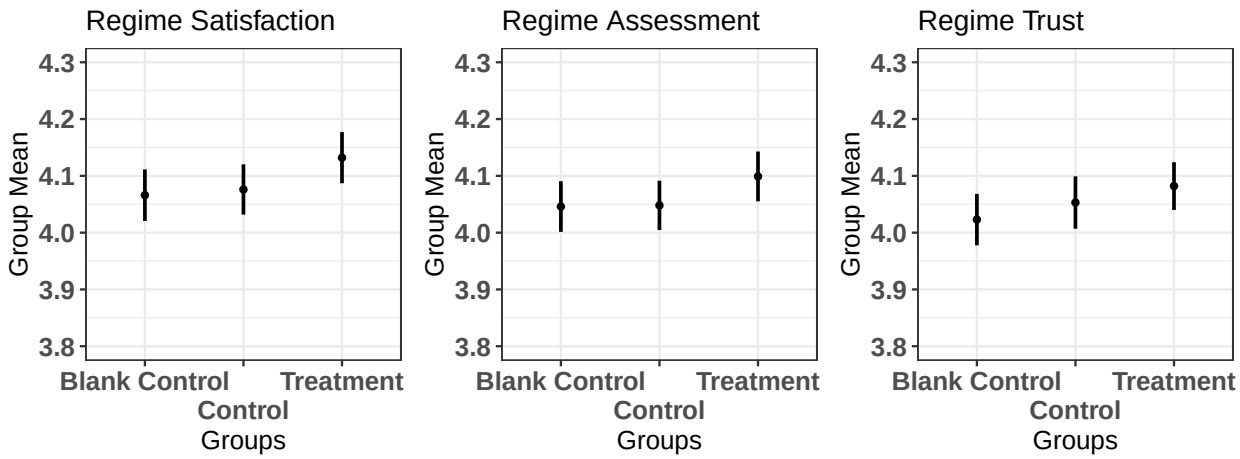
most other mechanism tests yield null results. One possible alternative explanation is cognitive dissonance (Festinger 1957). It claims that encountering inconsistencies between one's beliefs and attitudes causes psychological discomfort, which motivates individuals to reduce the inconsistencies. While an individual might initially oppose censorship, participating in the censorship process creates cognitive dissonance and subsequently forces the individual to change their attitude toward government censorship.

Another possible alternative explanation is the system justification theory (Jost 2020). It states that people are motivated to defend, bolster, and justify the social, economic, and political institutions and arrangements on which they depend (Jost 2020). Such a system justification process often happens at a non-conscious level. This implies potential psychological needs for individuals in authoritarian regimes to justify behaviors and experiences they encountered in authoritarian regimes. Hence, when individuals are constantly engaged in reporting online content, they will unconsciously justify such behavior and the censorship apparatus.

Although I cannot directly test these alternative explanations, one observable implication from both the cognitive dissonance theory and the system justification theory is that participation in censorship should have a stronger effect on increasing individuals' support for the censorship apparatus than increasing support for the authoritarian regime more broadly. Such a hypothesis is partly supported by the results in study 2. As shown in Figure 9, there is no observable difference between the control group and the blank control group across three regime support survey items.[7] Respondents in the treatment group express higher support for the regime, but the difference is only significant at the 0.1 level for *Regime Satisfaction*, and not significant for *Regime Assessment* and *Regime Trust*. In short, the results from study 2 do suggest a stronger effect on censorship support than regime support. However, directly testing the cognitive dissonance theory and the system justification theory remains a task for future research.

[7]For more information about the regime support survey items, see the Online Appendix.

Figure 9: Group Means of Regime Support Survey Items



Note: All three items are measured on a five-point Likert scale.

Conclusion

Censorship in authoritarian regimes has long been understood as top-down suppression of political opposition. To my knowledge, this study provides the first theorization, empirical description, and quantitative analysis of bottom-up public participation in censorship in authoritarian regimes. Using two original online surveys in China, one observational and one experimental, I demonstrate that public participation in censorship is both prevalent and significant in shaping public opinion toward the censorship apparatus.

The findings in this study highlight the discrepancy between the common understanding of repressive authoritarian apparatus, such as censorship, in the Western world and how ordinary citizens in authoritarian regimes perceive and interact with these repressive apparatus. For many citizens in China, censorship and other repressive institutions have been normalized as part of the rules of political life. Therefore, instead of fighting against the rules, they take advantage of it and use censorship to suppress opposing views. Such a mentality has significant downstream political implications as citizens no longer view the regime as the oppressor, but rather as a powerful arbitrator deciding which content should be removed, and who they should win over in their internal fights against fellow citizens.

Beyond autocracies, censorship has become an important social issue in many democracies including the United States, and public participation in censorship has also become more prevalent. Twitter, for example, introduced a community-based bottom-up content moderation project called "Birdwatch." Although democracies might care about different policy implications, such as electoral integrity, compared with their authoritarian counterparts, it is still important to analyze the consequences of these content moderation projects involving public participation, because the balancing act of fighting misinformation and preserving freedom of speech is difficult yet critical for sustaining a healthy democracy.

References

Aronow, Peter M and Allison Carnegie. 2013. "Beyond LATE: Estimation of the Average Treatment Effect with an Instrumental Variable." *Political Analysis* 21(4):492–506.

BBC News. 2020. "China where Fang Fang is From: the Report Culture is Spreading along with the Rise of Nationalism." *BBC News Chinese Version* .
URL: *https://www.bbc.com/zhongwen/simp/chinese-news-52560990*

Chang, Keng-Chi, William R. Hobbs, Margaret E. Roberts and Zachary C. Steinert-Threlkeld. 2022. "COVID-19 Increased Censorship Circumvention and Access to Sensitive Topics in China." *Proceedings of the National Academy of Sciences* 119(4).

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

Chen, Yuyu and David Y Yang. 2019. "The Impact of Media Censorship: 1984 or Brave New World?" *American Economic Review* 109(6):2294–2332.

CNS. 2020. "2020 Shanghai shi wangluo jubao xuanchuan yue zhengshi qidong [2020 Shanghai Internet Reporting Promotion Month Officially Launched]." *China News Agency* .
URL: *https://www.chinanews.com/sh/2020/09-02/9280670.shtml*

Costigan, Johanna M. and Xu Xin. 2020. "China's Digital Cultural Revolution." *The Diplomat* .
URL: *https://thediplomat.com/2020/04/chinas-digital-cultural-revolution/*

Dickson, Bruce. 2016. *The Dictator's Dilemma: The Chinese Communist Party's Strategy for Survival*. Oxford University Press.

Dikötter, Frank. 2016. *The Cultural Revolution: A People's History, 1962—1976*. Bloomsbury Publishing USA.

Distelhorst, Greg and Yue Hou. 2017. "Constituency Service Under Nondemocratic Rule: Evidence from China." *The Journal of Politics* 79(3):1024–1040.

Esberg, Jane. 2018. "The Audience of Repression: Killings and Disappearances in Pinochet's Chile." *Available at SSRN 3246120* .

Esberg, Jane. 2020. "Censorship as Reward: Evidence from Pop Culture Censorship in Chile." *American Political Science Review* 114(3):821–836.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Vol. 2 Stanford university press.

Fu, Diana and Greg Distelhorst. 2018. "Grassroots Participation and Repression Under Hu Jintao and Xi Jinping." *The China Journal* 79(1):100–122.

Gandhi, Jennifer. 2008. *Political Institutions Under Dictatorship*. Cambridge University 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):682–698.

Gregory, Paul R. 2009. *Terror by Quota*. Yale University Press.

Gueorguiev, Dimitar D and Edmund J Malesky. 2019. "Consultation and Selective Censorship in China." *The Journal of Politics* 81(4):1539–1545.

Han, Rongbin. 2018. *Contesting Cyberspace in China: Online Expression and Authoritarian Resilience*. Columbia University Press.

He, Baogang and Mark E Warren. 2011. "Authoritarian Deliberation: The Deliberative Turn in Chinese Political Development." *Perspectives on Politics* 9(2):269–289.

Hobbs, William R and Margaret E Roberts. 2018. "How Sudden Censorship Can Increase Access to Information." *American Political Science Review* 112(3):621–636.

Huang, Haifeng. 2015. "International Knowledge and Domestic Evaluations in a Changing Society: The Case of China." *American Political Science Review* 109(3):613–634.

Huang, Haifeng. 2018. "The Pathology of Hard Propaganda." *The Journal of Politics* 80(3):1034–1038.

Huang, Haifeng and Nicholas Cruz. 2021. "Propaganda, Presumed Influence, and Collective Protest." *Political Behavior* pp. 1–24.

Huang, Haifeng, Serra Boranbay-Akan and Ling Huang. 2019. "Media, protest diffusion, and authoritarian resilience." *Political Science Research and Methods* 7(1):23–42.

Huang, Haifeng and Yao-Yuan Yeh. 2019. "Information from Abroad: Foreign Media, Selective Exposure and Political Support in China." *British Journal of Political Science* 49(2):611–636.

Jost, John T. 2020. *A Theory of System Justification*. Harvard University Press.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107:326–343.

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.

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

Luo, Zhifan and Muyang Li. forthcoming. "Participatory Censorship: How Online Fandom Community Facilitates Authoritarian Rule." *New Media & Society* .

Manion, Melanie. 2015. *Information for Autocrats: Representation in Chinese Local Congresses*. Cambridge University Press.

Marbach, Moritz and Dominik Hangartner. 2020. "Profiling Compliers and Noncompliers for Instrumental-Variable Analysis." *Political Analysis* 28(3):435–444.

Martin, Justin D, Ralph J Martins and Robb Wood. 2016. "Desire for Cultural Preservation as a Predictor of Support for Entertainment Media Censorship in Saudi Arabia, Qatar, and the United Arab Emirates." *International Journal of Communication* 10:23.

Miller, Blake. 2018. Delegated Dictatorship: Examining the State and Market Forces behind Information Control in China PhD thesis.

Mitchell, Amy and Mason Walker. 2021. "More Americans Now Say Government Should Take Steps to Restrict False Information Online than in 2018." *Pew Research Center* 18.

Nabi, Zubair. 2014. "Resistance Censorship is Futile." *First Monday* .

Ng, Jason Q. 2015. "Politics, Rumors, and Ambiguity: Tracking Censorship on WeChat's

Public Accounts Platform." *Munk School of Global Affairs* .

Nisbet, Erik C, Olga Kamenchuk and Aysenur Dal. 2017. "A Psychological Firewall? Risk Perceptions and Public Support for Online Censorship in Russia." *Social Science Quarterly* 98(3):958–975.

Nyhan, Brendan and Thomas Zeitzoff. 2018. "Fighting the past: Perceptions of control, historical misperceptions, and corrective information in the Israeli-Palestinian conflict." *Political Psychology* 39(3):611–631.

Ong, Lynette H. 2022. *Outsourcing Repression: Everyday State Power in Contemporary China*. Oxford University Press.

Pan, Jennifer and Alexandra A. Siegel. 2020. "How Saudi Crackdowns Fail to Silence Online Dissent." *American Political Science Review* 114:109–125.

Pan, Jennifer and Yiqing Xu. 2020. "Gauging Preference Stability and Ideological Constraint under Authoritarian Rule." *Working Paper, Available at SSRN* .

Pop-Eleches, Grigore and Lucan A Way. 2021. "Censorship and the Impact of Repression on Dissent." *American Journal of Political Science* .

Reuters. 2021. "China boycotts actor after photos of visit to Japan's Yasukuni Shrine." *Reuters* .

URL: *https://www.reuters.com/world/china/china-boycotts-actor-after-photos-visit-japans-yasukuni-shrine-2021-08-15/*

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

Roberts, Margaret E. 2020. "Resilience to Online Censorship." *Annual Review of Political Science* 23:22.1–22.19.

Shadmehr, Mehdi and Dan Bernhardt. 2015. "State Censorship." *American Economic Journal: Microeconomics* 7(2):280–307.

Smith, Brianna A. 2021. "It's All Under Control: Threat, Perceived Control, and Political Engagement." *Political Psychology* .

Sullivan, Daniel, Mark J Landau and Zachary K Rothschild. 2010. "An existential function of enemyship: Evidence that people attribute influence to personal and political enemies to compensate for threats to control." *Journal of personality and social psychology* 98(3):434.

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):842–867.

Thurston, Anne F. 1984. "Victims of China's Cultural Revolution: The Invisible Wounds: Part I." *Pacific Affairs* 57(4):599–620.

Truex, Rory. 2016. *Making Autocracy Work: Representation and Responsiveness in Modern China*. Cambridge University Press.

Truex, Rory. 2017. "Consultative authoritarianism and its limits." *Comparative political studies* 50(3):329–361.

Tufekci, Zeynep. 2017. *Twitter and Tear Gas*. Yale University Press.

Wang, Dakuo and Gloria Mark. 2015. "Internet Censorship in China: Examining User Awareness and Attitudes." *ACM Transactions on Computer-Human Interaction (TOCHI)* 22(6):1–22.

Whitson, Jennifer A and Adam D Galinsky. 2008. "Lacking control increases illusory pattern perception." *Science* 322(5898):115–117.

Wike, Richard and Katie Simmons. 2015. "Global Support for Principle of Free Expression, But Opposition to Some Forms of Speech." *Pew Research Center* 18.

Williamson, Scott and Mashail Malik. 2019. "Contesting narratives of repression: Experimental evidence from Sisi's Egypt." *Journal of Peace Research* .

Xinhua. 2021. "2020 nian quanguo shouli wangluo weifa he buliang xinxi jubao 1.63 yi jian [In 2020, 163 million reports of illegal and inappropriate online content nationwide.]." *Xinhua News Agency* .

URL: *http://www.xinhuanet.com/politics/2021-01/19/c_1126996991.htm*

Yang, Jisheng. 2021. *The World Turned Upside Down: A History of the Chinese Cultural Revolution*. Farrar, Straus and Giroux.

Yang, Tony Zirui. 2022. "Normalization of Censorship: Evidence from China." *Working Paper, Available at SSRN* .