Attraction or Distraction? Impacts of Pro-regime Social Media Comments on Chinese Netizens

Stan Hok-Wui Wong1 · Jiachen Liang1

Accepted: 6 August 2021 / Published online: 4 September 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Despite heavy Internet regulations, government critics and political satires are not completely absent in the cyberspace of most authoritarian regimes. Some argue that these regimes deliberately tolerate somewhat critical online comments as a way to monitor mass sentiments. To counterbalance critics’ influences, they often mobilize and amplify pro-regime voices. We empirically examine whether such pro-regime voices succeed in changing public opinions in favor of the authorities. Based on two online surveys and an embedded survey experiment that we implemented in China, we find that when given a choice, our Chinese respondents self-select to expose themselves to comments that deviate from the official discourses. In addition, exposure to diverse comments undermines individuals’ policy support. The findings call into question the effectiveness of the “soft propaganda” that authoritarian regimes orchestrate in cyberspace.

Keywords Soft propaganda · Selective exposure · Online comments · Policy approval · Regime legitimacy · China

Introduction

Authoritarian governments have many ways to control the media. In addition to censorship, they can force media outlets to broadcast nothing but praise of political leaders or news that glorifies the regime. However, such official propaganda, or in other words, “hard propaganda,” is not an effective tool to win people’s hearts and minds because its one-sidedness may look excessive and unreal, thereby weakening its appeal (Huang, 2018). For this reason, regimes also utilize “soft propaganda” to guide public opinions (Huang, 2015, p. 419). Instead of completely censoring unfavorable information, these governments allow the coexistence of both critical and
pro-regime voices in the online sphere, in which citizens can view or even discuss some controversial issues (Egorov et al., 2009; Qin et al., 2017). The authorities then neutralize regime critics by recruiting undercover commentators (Gunitsky, 2015; King et al., 2017), mobilizing regime supporters to leave comments (Han, 2015a, 2015b), and amplifying the voices of the patriots (Fang & Repnikova, 2018). This approach is potentially more influential because, compared to one-sided propaganda, lively online discussions look more authentic and trustworthy.

China, as one of the states that have tight control of the Internet, has actively adopted soft propaganda to guide public opinions. Recent studies provide suggestive evidence for its success. For instance, some observe a rising trend of nationalism on the Chinese Internet (Hyun et al., 2014; Jiang, 2012). Others point out the proliferation of “little pinks” and “volunteered 50-cent members” (ziganwu)—netizens who volunteer themselves to defend the regime (Fang & Repnikova, 2018; King et al., 2017). There are also “wolf warriors,” netizens who manage to bypass the “Great Firewall,” so that they can visit banned foreign social media and rebut foreigners who have issues with China (Teixeira, 2020).

While these examples seem to attest to the Chinese government’s success in shaping public opinions, empirical evidence remains woefully inadequate. In this article, we try to fill this gap by exploring two questions: (1) to what extent do these pro-regime comments appeal to ordinary Chinese netizens?; and (2) would the presence of pro-regime discourses make citizens more supportive of the regime?

To answer the first question, we use an online survey to track Chinese social media users’ self-selected comment exposure over time. To answer the second question, we make use of an online experiment to examine if netizens’ policy support will change in the absence of online comment exposure. Online comments in this paper are comments taken from the comment section of Sina Weibo,¹ which is one of the most popular and most regulated social media platforms in China (Vuori & Paltemaa, 2015).

The empirical results show that when given an option, the survey respondents tend to self-select to expose themselves to critical comments rather than pro-regime comments over time. We also find that compared to those who do not have a chance to view comments, those who can view comments—either critical or pro-regime—are significantly less supportive of the underlying policies.

The findings advance our understanding of online propaganda and information control in authoritarian regimes. We argue that although the Chinese authorities have gone to great lengths to shape public opinion through cajoling and coercion, ordinary Chinese may not passively and mechanically accept soft propaganda, especially when they can access critical comments. Nor do they seem to be mindless information recipients who would self-select to expose themselves to state-sponsored information. The highly regulated media environment seems unable to quell their demand for alternative information. In addition, the presence of pro-regime comments seems unable to neutralize the effects of critical comments by increasing public support of the regime. This could be why the Chinese government does not

¹ In this article, we use the terms “Sina Weibo” and “Weibo” interchangeably.
relinquish the use of hard censorship to control the media, despite the seeming proliferation of soft propaganda.

**Literature Review**

Social media play a vital role in political communication. Many authoritarian regimes impose tight controls over the Internet for this reason (Pan & Siegel, 2020). Although some scholars recognize social media as a liberalization tool that facilitates democratization (Diamond, 2010; Lynch, 2011), a growing body of literature argues that non-democracies are able to employ social media to reinforce state ideology and strengthen authoritarianism through information controls and regulations (Creemers, 2017).

In China, there are two recognized approaches to manipulating online information (Chen & Xu, 2017). The first approach is online censorship, which refers to the practice of online content deletion that aims to prevent netizens from accessing information deemed unfavorable to the regime (King et al., 2013). The second approach, which is less conspicuous, is to co-opt social media as an instrument to reinforce regime legitimacy (Gunitsky, 2015). Studies find that the Chinese government has been recruiting commentators to post online comments that promote the interest of the party-state (Han, 2015a, 2015b). Those comments, also known as 50-cent comments or *shuijun* comments, which seem to be authentic voices from ordinary netizens, are carefully crafted by the authorities for propaganda purposes (Bremmer, 2010; Deibert & Rohozinski, 2010; King et al., 2017; Miller, 2016). Such “soft propaganda” is expected to subtly influence public opinions and persuade citizens to believe in the state ideology (Greitens, 2013; Huang, 2015; Hung, 2010).

In addition to the above, the state also actively incorporates regime supporters’ discourses into its propaganda campaigns. Recent studies show the emergence of “little pink,” a term describing ordinary citizens who like to express their fervent support of the regime (Fang & Repnikova, 2018). By amplifying their voices and filtering out the voices of dissent, the state can turn patriotic netizens into “thought work collaborators” (Hung, 2010; Repnikova & Fang, 2018, p. 763).

Existing studies primarily focus on identifying pro-regime comments generated through astroturfing (King et al., 2017; Miller, 2016); seldom do they study the effects of fabricated and filtered comments on netizens. In this study, we focus on two outcome variables of interest: (1) individuals’ selective exposure to different online comments and (2) the effects of online discussions on individuals’ political attitudes toward regime legitimacy.

**Theories and Hypotheses**

**Online Comments as an Echo Chamber**

A surge in nationalism, together with tight ideology controls, has polarized China’s cyberspace, where clashes between those who firmly support the regime and those
who dare to dissent from the government are commonplace (Chase & Mulvenon, 2002; Pan & Xu, 2018; Wu, 2013; Yang, 2014; Yang & Tang, 2018; Zhao, 2016). The growing nationalistic sentiment is partially due to a prolonged nationwide patriotic education (Hyun & Kim, 2015; Kennedy, 2009; Zhao, 1998). Fervent supporters of the party-state ideology are ubiquitous in China’s cyberspace and their voices are particularly noticeable when international tensions exist between China and other countries like Japan and the United States (Jiang, 2012; Qiu, 2006). Zhou (2005) shows that even well-educated and informed individuals in China may hold strong anti-Western and nationalistic values.

The echo chamber effects theory suggests that people favor information conforming with their pre-existing beliefs and tend to avoid ideologically uncongenial perspectives (Colleoni et al., 2014). The basic premise of the theory is that people’s cognitive resources are limited. When consuming information in a high-choice media environment, individuals do not have much time and patience to browse arguments from all sides, so they simply seek sources that reinforce, instead of challenging, their opinions (Arceneaux et al., 2012; Garrett, 2009). If this applies to Chinese social media, we shall expect to observe a selective consumption of information among Chinese social media users: regime supporters prefer information that defends the party-state, while others like to view more critical messages. Thus, we hypothesize that:

**Hypothesis 1:** Given the constraints of time or cognitive resources, individuals will self-select to expose themselves to online information with similar ideological contents.

**Online Comments as an Alternative Source of Information**

Confirming one’s belief is possibly not the only reason why people consume online comments. From an evolutionary perspective, human beings have developed sophisticated sensory organs to process external information to enhance survival chances. Excluding information that is at odds with our beliefs is not conducive to self-preservation. On the contrary, exposure to diverse information helps us assess external risks and formulate appropriate responses. In the context of authoritarian regimes, history is replete with examples where the authorities were unreliable information providers, especially in times of crises. After the explosion of a nuclear reactor at Chernobyl in 1986, the initial reaction of the Soviet authorities was to cover up the event. Workers at the nuclear plant were prohibited from sharing the news about the explosion, while telephone communications with the affected city were suspended. Similarly, during the early phase of the COVID-19 outbreak, the public security bureau arrested Li Wenliang, the whistle-blowing Wuhan-based medical doctor, on the charge of making false comments and spreading rumors (Ludwig, 2020; Zhang, 2020).

To avoid the risk of relying on official information, citizens in authoritarian regimes learn to share and obtain unofficial information (Nip & Fu, 2016; Tong & Lei, 2013). Zhu et al. (2013) show that grapevine news in China can significantly
change popular perceptions. Chen and Yang (2019) show that once Chinese people are incentivized to access uncensored information, their demand for such information will increase. Hobbs and Roberts (2018) also provide empirical evidence that sudden blocking of information improves people’s motivation to search for blocked contents. These findings provide support to Geddes’s and Zaller’s (1989) argument that many citizens in authoritarian regimes are able to decipher reality from propaganda.

The demand for unofficial information may affect the consumption of online comments in authoritarian regimes. Pro-regime comments tend to echo the official discourse. As such, they provide little new information to recipients who look for diverse information. On the contrary, user-generated contents that deviate from the official discourse may offer comment viewers informed perspectives into an issue at hand that would otherwise be unavailable from the official media. For this reason, we posit our second hypothesis:

**Hypothesis 2:** Given the constraints of time or cognitive resources, individuals will self-select to expose themselves to online information that deviates from the official discourse.

**Effects of Online Comments on Regime Legitimacy**

The reason why the authorities permit somewhat critical contents to coexist with regime-supportive contents in China’s cyberspace is the belief that they can guide public opinions or even drum up regime support with “soft propaganda.” In practice, this may not necessarily be the case. Studies show that online deliberations help expose individuals to different arguments and media framings. As such, individuals can have an opportunity to learn competing perspectives or ideas of which they were previously unaware. In other words, exposure to alternative information has the potential to stimulate critical thinking, thereby influencing individuals’ political attitudes (Tang & Huhe, 2014).

This exposure effect is also relevant to the Chinese context. Western-style democracy is absent in China. Chinese citizens have no direct way to select their political leaders beyond the village level. Accepting the political status quo, rather than challenging it, seems to be a sensible choice for them. Note, however, that the acceptance is not solely driven by scruples about the Leviathan. Historically, Chinese people tended to view the government as the head of a big family, whose decisions should be supported and followed (Shi, 2001). Rapid economic development in the reform era has further strengthened the political legitimacy of the ruling party. Empirical works also show that Chinese people have a high degree of trust in the central government even under the authoritarian control of the party-state (Chen, 2017; Li, 2016; Wang, 2005; Yang & Tang, 2010). Li (2004) observes that Chinese citizens generally believe that official policies, especially those implemented by the central government, bear beneficent intent. In other words, without arousing their awareness of diverse viewpoints, Chinese citizens in all likelihood are favorably disposed to the regime. Hence, we posit that:
Hypothesis 3: Compared with individuals who lack access to online comments, individuals who are granted such access tend to show less policy support.

Note that Hypothesis 3 is related to policy support rather than regime support. The reason is that measuring regime support is problematic. In addition to official prohibition, it is also unlikely that survey respondents give truthful responses for fear that their discontent with the regime will cause them trouble (Krumpal, 2013). For this reason, we need an indirect measure to address this issue (Fisher, 1993). One way to examine regime support is to measure citizens’ support of the state’s proposed policies. In China, where the ruling party dictates the policy-making process, policies generally reflect the interests of the party-state. As a result, policy support may be viewed as an indicator of regime support (Chen et al., 1997; Lü, 2014; Tang & Huhe, 2014).

Research Design

Comments on Sina Weibo

Although outright disparagement of the regime or national leaders is nowhere seen in China’s cyberspace, political satires, innuendoes, and pungent comments on policies are ubiquitous. On many social media sites, one can see the co-existence of pro-regime comments and comments that are more critical. While in theory netizens can expose themselves to both types of comments, the government has made efforts to highlight regime-supportive comments in the comment section as a way to guide public opinions (King et al., 2017). Weibo is a case in point. If a post involves topics about the government and government officials, social policies, or political events, the comment section of that post is likely to show “recommended comments by the account holder.” These recommended comments are predominantly pro-regime. This is a way for the state to endorse and promote regime-supportive discourses from the grassroots (Repnikova & Fang, 2018).

Figure 1 displays an example of a comment section of a post (for contents of the post, see Fig. 3), which shows a local government’s decision to establish a rigorous disease reporting mechanism after the outbreak of coronavirus in Wuhan as a way to prevent any cover-up of the disease. The Hubei government has been accused of hiding the truth about the outbreak at an early stage, and this post reports the government’s response in late January. As shown in Fig. 1, those comments are rather supportive of the government. For instance, the third comment from the top praises the government for its confidence and transparency, while the fourth commentator emphasizes his or her trust in the state. The last two comments are cheerleading slogans.

The users of Weibo can still find critical netizens’ comments, but in a different section. Similar to the retweet button on Twitter, the repost button of Weibo allows people to share a post on their personal Weibo page. They may also comment on the shared post. Weibo users can find out what people who share a post would say about it by clicking the “repost” (zhuanfa) button. Figure 2 provides an example.
Fig. 1  Comment section of the *Sina Weibo* post

Fig. 2  Repost section of the *Sina Weibo* post
As is seen, “reposted comments” are decidedly more critical and satirical than the comment section (pinglun) displayed in Fig. 1. Most of them mock the government’s slow responses to the outbreak and question the government’s inaction at an early stage (Fig. 3).

In other words, the design of Weibo provides netizens a choice to view different types of comments. If they click the comment section, they expose themselves to those recommended pro-regime comments, be they produced by astroturbers or by the “little pinks”; if they click the repost section, they will see more critical and satirical opinions expressed by other netizens.

**Online Survey Designs**

We conducted two separate online surveys in November 2018 (Study 1) and February 2021 (Study 2), respectively. Online surveys implemented with online panels or outsourcing platforms have become increasingly popular in social science research (Clifford et al., 2015). While the generalizability of the findings based on online samples is a potential concern, Li et al. (2017) show that online samples in China
provide a useful tool for research that makes inferences about the country’s netizens. Because we are primarily interested in studying Chinese netizens, whether our online samples are representative of the overall Chinese population is less of a concern here. That said, we took steps to maximize the generalizability of our findings. In addition to using quota sampling to recruit respondents representative of the Chinese population, we conducted two separate surveys to ensure that the findings are robust to a particular online sample recruited from a particular online platform. We discuss the details of each study in what follows.

**Study 1**

We issued the survey to 523 respondents through a well-established survey platform (51diaocha.com) in mainland China targeting Chinese Internet users (i.e., Weibo). The sampling quota is based on the demographic distribution recorded in the 2010 population census of China based on five factors: gender, age, educational level, income and province. To avoid repeated responses, we only accepted responses with a unique IP address and with a completion time of no less than one minute (Goodman et al., 2013).

We model our survey upon the comment choices available to Weibo users. All respondents were given five actual Weibo posts regarding five somewhat controversial social policies (For exact wording, see Online Appendix 1-Study 1). In addition to the news script itself, they were provided with an opportunity to view one type of comment related to that piece of news. In particular, we provide the respondents with three options: (1) *recommended comments* (精选评论); (2) *netizens’ comments* (网友评论); and (3) *skip comments* (忽略). The first two options are analogous to Weibo’s comment and repost sections as shown in Figs. 1 and 2. We use the term “*recommended comments*” to make this option reminiscent of the respondents’ Weibo experience. This term, together with its political connotations, should be familiar to Chinese social media users, as it can be found on various social media platforms (i.e., Weibo and Wechat). In order to hint at unorthodox voices, we use the term “*netizens’ comments.*” To ensure that our respondents can detect the nuance between *recommended comments* and *netizens’ comments*, we include in the first policy news (i.e., the example news) one sample comment next to the *recommended* and *netizens’ comments* buttons, respectively (see Fig. 4). Respondents will not see sample comments again in the remaining policy news.

If a respondent clicks the *recommended comments* button, they will see five pro-government or pro-policy comments. These comments either praised the policy or highlighted its importance. If a respondent clicks the *netizens’ comments* button, they will see five critical comments that question the feasibility and effectiveness of the policy. These comments were curated from actual comments, with some minor editing, related to the posts (for exact wordings, see Online Appendix 1).

In Weibo, users can choose to ignore comments about a post by clicking neither the *comment* nor the *repost* button. For this reason, we also provide the *skip comments* option to our survey respondents. The reason for allowing the respondents to view only one type of comment at a time is that online information seldom comes effortlessly and without cost. As discussed, people face time constraints and have
limited cognitive resources. It is impossible for people to browse all kinds of information online. As such, they have to constantly decide the kind of information to which they want to expose themselves. We force them to choose in order to learn their preference for information.

Study 2

The basic design of Study 2 is similar to Study 1. We conducted Study 2 for two reasons. The first is to minimize potential house-effects of survey companies: we commissioned a different survey company to ensure that the results are not peculiar to the online panel of a specific company. The second reason is that Study 1 is not robust to some known potential biases such as question order effects. We designed Study 2 to minimize these potential biases. We detail the improvements below.

Unlike Study 1, we randomized all policy news in Study 2 to ensure that the results are not driven by a particular sequence of news being shown to respondents (for exact wordings of all policies, see Online Appendix 1). In addition, we also randomly assigned our respondents into two groups with different order of comment choice options. They saw either recommended comments placed on top of netizens’ comments and skip comments or netizens’ comments placed on top of recommended comments and skip comments. Done this way, we were able to minimize the order-effects of comment choice options. We also finetuned the wording of the options. For “recommended comments,” we rephrased this option to “comments recommended by the account holder” (博主精选评论), while we changed “netizens comments” to “comments reposted by netizens” (网友转发评论). By adding “account holder” and “reposted” to these options, we are able to bring these choices closer to the actual Weibo setting.

To further minimize the potential bias pertaining to a particular news example, we prepared three pieces of example news and drew only one random piece at time

---

**Fig. 4** The example policy news. Notes: All respondents read this example policy news in Study 1 before viewing other policy news. As shown in this figure, we added a sample comment after the recommended comments and netizens’ comments buttons to give them an idea of the type of comments (i.e., pro-government or critical) they expect to see. In subsequent pieces of policy news, we did not provide any sample comment. This policy news is about a new rule concerning the purchase of multiple train tickets.

2 For consistency, we will use recommended comments and netizens’ comments from now on.
Estimations and Results

Estimations of Comment Choice Over Time

Our variable of interest is $N_t$, the sequence of their exposure to five (Study 1) or six (Study 2) pieces of policy news. We control for the respondents’ initial comment choice at time $t = 1$, as their initial choice is likely correlated with their attitude toward the government.

Our outcome of interest is respondents’ comment choice. In particular, each respondent read five (for study 1) or six policies (for study 2). After reading each policy news, they were allowed to choose between recommended comments, netizens’ comments, or skip comments. Because the outcome is categorical, we analyze the data using a multinomial logistic regression model (Gelman & Hill, 2006). When using multinomial logistic regressions, researchers need to specify an outcome as a baseline category for comparison. In our case, we choose netizens’ comments as the baseline group because we are primarily interested in finding one’s preference for pro-regime comments conditional on one’s decision to view a comment.4

---

3 Statistical profiles of online survey respondents in Study 1, Study 2, and Studies 1 and 2 combined are reported in Online Appendix 3. Replication data and codes are available online at https://doi.org/10.7910/DVN/CMFPFR.

4 If we were interested in finding the attractiveness of pro-regime comments vis-à-vis not viewing comments at all, then it would have made sense to choose skip comments as the baseline group. However, in
Given that we have three outcomes, the multinomial logistic regression requires us to run two multinomial logit equations, contrasting the baseline outcome with each alternative outcome. Because we are primarily interested in finding whether recommended or netizens’ comments appeal to the respondents, the multinomial logit equation of interest is as follows:

\[
\Pr \left( Y_{it} = \frac{\text{Recommended comments}}{\text{Netizens’ comments}} \right) = e^{\alpha_i + \beta N_t + \gamma_1 A_i + \gamma_2 B_i + X_i \omega + \varepsilon_{it}}
\]

where \( Y_{it} \) denotes the probability of individual \( i \) choosing recommended comments at time \( t \) over choosing the baseline category netizens’ comments, \( \alpha_i \) is individual \( i \)’s intercept, \( \beta \) is the coefficient on the sequence of policy news \( N_t \), and \( \gamma_1 \) and \( \gamma_2 \) are coefficients on respondent \( i \)’s initial comment choice of recommended comments \( A_i \) and netizens’ comments \( B_i \). \( \omega \) is a vector of coefficients on a list of covariates \( X_i \) and \( \varepsilon_{it} \) is the error term clustered at the individual level and assumed to be independent and identically distributed.

The multinomial logistic regressions compare the choice of recommended comments vis-à-vis netizens’ comments. It would be instructive to pit recommended comments against non-recommended comments (i.e., netizens’ comments and skip comments). For this reason, we also converted the outcome variables into a binary choice (recommended v. non-recommended) and re-ran the specifications using random-effects logistic regressions.

Table 2 presents the regression results of different regression specifications. For multinomial logistic regressions, we show only the results of choosing recommended comments over the baseline group, netizens’ comments. After controlling for respondents’ initial choices of recommended comments and netizens’ comments, the coefficients on time variable \( N_t \) are negative across all specifications. The results are consistent in Study 1, Study 2, and Studies 1 and 2 combined. We also find similar results in both the multinomial logistic model and the random-effects logistic model. In Specifications 3 and 4, we control for eight measures of individuals’ political attitudes, including their political efficacy, their resistance against the West. The coefficients on \( N_t \) remain the same. Thus, the results suggest that when considering respondents’ comment choice variation over time, the probability of respondents choosing recommended comments over netizens’ comments declines significantly. Rather, they are more likely to choose netizens’ comments as they continue to view more policies.

Footnote 4 (continued)

expectation, it is almost certain that no one would read comments indefinitely, which implies that everyone will eventually choose skip comments. The coefficient on the variable of interest \( N_t \) is likely negative.

6 Both \( A_i \) and \( B_i \) are dummy variables. \( A_i \) (\( B_i \)) takes a value of “1” if respondent \( i \) chooses recommended comments (netizens’ comments) at time \( t = 1 \) and “0” otherwise.

7 The results of choosing skip comments over netizens’ comments are presented in Online Appendix 4.
Table 2  Respondents’ choice of comments over time

Dep. var. choice of recommended comments

| Study 2 | Study 2 | Study 1 | Study 1 and Study 2 |
|---------|---------|---------|--------------------|
|         | Multinomial | Random-effects | Multinomial | Random-effects | Multinomial | Random-effects |
|         | (1)       | (2)       | (3)       | (4)       | (5)       | (6)       | (7)       | (8)       |

|         |         |         |         |         |         |         |         |         |
|         | $N_i$   |         |         |         |         |         |         |         |
|         | $-0.065^{**}$ | $-0.115^{**}$ | $-0.066^{**}$ | $-0.115^{**}$ | $-0.082^{**}$ | $-0.087^*$ | $-0.077^{**}$ | $-0.112^{**}$ |
|         | (0.021) | (0.024) | (0.021) | (0.024) | (0.031) | (0.035) | (0.017) | (0.020) |
|         |         |         |         |         |         |         |         |         |
|         | Recommended comments as initial choice |         |         |         |         |         |         |         |
|         | $1.092^{**}$ | $3.891^{**}$ | $1.139^{**}$ | $3.902^{**}$ | $1.197^{**}$ | $2.958^{**}$ | $1.171^{**}$ | $3.585^{**}$ |
|         | (0.310) | (0.320) | (0.319) | (0.319) | (0.262) | (0.282) | (0.197) | (0.216) |
|         |         |         |         |         |         |         |         |         |
|         | Netizens’ comments as initial choice |         |         |         |         |         |         |         |
|         | $-0.655^*$ | $1.802^{**}$ | $-0.600$ | $1.859^{**}$ | $-0.138$ | $1.407^{**}$ | $-0.436^*$ | $1.646^{**}$ |
|         | (0.319) | (0.324) | (0.329) | (0.323) | (0.271) | (0.280) | (0.205) | (0.217) |
|         |         |         |         |         |         |         |         |         |
|         | Male |         |         |         |         |         |         |         |
|         | 0.122 | 0.068 | 0.113 | 0.077 | 0.110 | 0.160 | 0.133 | 0.147 |
|         | (0.108) | (0.137) | (0.108) | (0.137) | (0.118) | (0.139) | (0.079) | (0.099) |
|         |         |         |         |         |         |         |         |         |
|         | Age |         |         |         |         |         |         |         |
|         | 0.009 | 0.007 | 0.011 | 0.009 | 0.016 | 0.005 | 0.013^{**} | 0.007 |
|         | (0.005) | (0.007) | (0.006) | (0.007) | (0.009) | (0.010) | (0.005) | (0.006) |
|         |         |         |         |         |         |         |         |         |
|         | CCP member |         |         |         |         |         |         |         |
|         | $-0.121$ | $-0.123$ | $-0.157$ | $-0.182$ | $0.080$ | $0.140$ | $-0.085$ | $-0.068$ |
|         | (0.131) | (0.165) | (0.136) | (0.169) | (0.185) | (0.214) | (0.104) | (0.132) |
|         |         |         |         |         |         |         |         |         |
|         | Education level |         |         |         |         |         |         |         |
|         | 0.097 | 0.026 | 0.104 | 0.032 | 0.018 | 0.020 | 0.049 | $-0.009$ |
|         | (0.062) | (0.078) | (0.062) | (0.077) | (0.068) | (0.077) | (0.043) | (0.053) |
|         |         |         |         |         |         |         |         |         |
|         | Income level |         |         |         |         |         |         |         |
|         | $-0.012$ | $-0.008$ | $-0.006$ | $-0.002$ | $-0.077$ | $-0.048$ | $-0.036$ | $-0.028$ |
|         | (0.035) | (0.046) | (0.036) | (0.046) | (0.047) | (0.055) | (0.028) | (0.035) |
|         |         |         |         |         |         |         |         |         |
|         | Political interest |         |         |         |         |         |         |         |
|         | 0.024 | 0.013 | 0.053 | 0.020 | 0.311^{**} | 0.429^{**} | 0.090 | 0.149 |
|         | (0.101) | (0.130) | (0.103) | (0.132) | (0.099) | (0.114) | (0.069) | (0.087) |
|         |         |         |         |         |         |         |         |         |
|         | Frequency of social media news consumption |         |         |         |         |         |         |         |
|         | $-0.241^*$ | $-0.132$ | $-0.241^*$ | $-0.143$ | $-0.164$ | $-0.101$ | $-0.208^*$ | $-0.137$ |
|         | (0.122) | (0.154) | (0.123) | (0.155) | (0.134) | (0.153) | (0.089) | (0.111) |
|         |         |         |         |         |         |         |         |         |
|         | Frequency of social media use |         |         |         |         |         |         |         |
|         | 0.093 | 0.097 | 0.093 | 0.084 | 0.285^* | 0.178 | 0.124 | 0.060 |
|         | (0.112) | (0.136) | (0.113) | (0.135) | (0.123) | (0.141) | (0.081) | (0.100) |
Table 2 (continued)

| Dep. var. | choice of recommended comments | Study 2 | Study 2 | Study 1 | Study 1 and Study 2 |
|-----------|---------------------------------|---------|---------|---------|-------------------|
|           | Multinomial Random-effects      | Multinomial Random-effects | Multinomial Random-effects | Multinomial Random-effects |
|           | (1) (2)                         | (3) (4)                         | (5) (6)                         | (7) (8)                         |
| Most recent social media use | $-0.006$ 0.046 | $-0.002$ 0.055 | 0.012 0.083 | 0.011 0.084 |
|           | $(0.062)$ $(0.072)$            | $(0.061)$ $(0.072)$            | $(0.067)$ $(0.077)$            | $(0.045)$ $(0.054)$            |
| Frequency of social media discussion | $-0.021$ 0.066 | $-0.018$ 0.048 | $-0.026$ 0.072 | 0.005 0.087 |
|           | $(0.094)$ $(0.119)$            | $(0.097)$ $(0.120)$            | $(0.098)$ $(0.117)$            | $(0.067)$ $(0.084)$            |
| Internal political efficacy | 0.135 0.079 |          |          |          |
|           | $(0.071)$ $(0.091)$            |          |          |          |
| External political efficacy |          | $-0.062$ $-0.072$ |          |          |
|           |          | $(0.069)$ $(0.086)$            |          |          |
| Resistance against the West | 0.005 $-0.001$ |          |          |          |
|           | $(0.079)$ $(0.097)$            |          |          |          |
| Collectivism | 0.108 0.189 |          |          |          |
|           | $(0.080)$ $(0.099)$            |          |          |          |
| National pride | $-0.018$ 0.050 |          |          |          |
|           | $(0.134)$ $(0.164)$            |          |          |          |
| Information transparency | $-0.128$ $-0.126$ |          |          |          |
|           | $(0.067)$ $(0.084)$            |          |          |          |
| Life satisfaction | $-0.048$ 0.042 |          |          |          |
|           | $(0.098)$ $(0.120)$            |          |          |          |
| Society satisfaction | 0.104 0.103 |          |          |          |
|           | $(0.102)$ $(0.131)$            |          |          |          |
Table 2 (continued)

Dep. var. choice of recommended comments

| Study 2 | Study 2 | Study 1 | Study 1 and Study 2 |
|---------|---------|---------|---------------------|
|         | Multinomial | Random-effects | Multinomial | Random-effects | Multinomial | Random-effects | Multinomial | Random-effects |
| (1)     | (2)       | (3)      | (4)            | (5)           | (6)          | (7)          | (8)          |
| Number of obs | 3912 | 3906 | 3912 | 3906 | 2615 | 2615 | 6527 | 6521 |
| Log pseudolikelihood | $-3029.198$ | $-2115.744$ | $-3003.382$ | $-2110.626$ | $-1865.099$ | $-1416.006$ | $-5045.113$ | $-3570.704$ |
| Pseudo $R^2$ | 0.246 | 0.253 | 0.259 | 0.259 | 0.259 | 0.230 | 0.230 | 0.230 |
| Wald Chi$^2$ | 345.910 | 354.930 | 354.930 | 354.930 | 293.26 | 293.26 | 599.260 | 599.260 |

Notes: The dependent variable is respondents’ comment choice of recommended comments at time $t=1, 2, \ldots, 5$ or 6. For multinomial logistic regressions, the baseline group is netizens’ comments. For random-effects regressions, the baseline group is non-recommended comments (i.e., both netizens’ comments and skip comments). Independent variables include policy number $N_t$ and respondents’ initial comment choice of recommended comments and netizens’ comments at time 1 ($t=1$). Internal political efficacy, External political efficacy, Resistance against the West, Collectivism, National pride, and Information transparency are measured on a 4-point scale where 1 indicates “Disagree strongly” and 4 “Agree strongly”; Life satisfaction and Society satisfaction are measured on a scale of four from 1 “Very dissatisfied” to 4 “Very satisfied.” For description of other variables, see Online Appendix 2. Provincial fixed-effects and cutoff points are not reported but are available upon request. Standard errors clustered at the individual level are in parentheses.

**p < 0.01, *p < 0.05
The results suggest that those who initially chose to view pro-regime comments would not expose themselves repeatedly to pro-regime comments. In other words, we find no evidence that Chinese netizens’ comment-viewing habits are bifurcated, such that pro-regime (critical) netizens would consume exclusively pro-regime (critical) comments. Rather, as our respondents would self-select to expose themselves to critical netizens’ comments over time, despite the fact that their initial choice was pro-regime comments, the results support the hypothesis that Chinese netizens regard online comments as a source of alternative information (Hypothesis 2).

To visually depict respondents’ comment choice variation through the exposure to five pieces of policy news, we apply the multi-state survival model to generate a diagram. The model not only allows us to take respondents’ initial comment choice into account but also considers respondents’ comment choice at each time point and how it may affect their later choice. This model is based on the Markov assumption that the occurrence of event at time $t = 2$ depends only on the previous event at time $t = 1$ (Crowther & Lambert, 2017). The model is often used to examine the change of patients’ conditions under different medical states, as patients’ health condition can switch from the starting state (i.e., post-surgery state) to a transition state (i.e., relapse state) or even directly to the absorbing state (i.e., death) (Crowther & Lambert, 2017). Using the three-state survival model proposed by Crowther and Lambert (2017), we examine respondents’ comment choice switching tendency from recommended comments to netizens’ comments (see Fig. 5). We consider the choice of skip comments as middle state because we are only interested in the transition from recommended comments to netizens’ comments.

We show the predicted transition probabilities in Fig. 6. The white area in each figure represents the probabilities of respondents choosing recommended comments over time, whereas the gray area refers to the probabilities of them choosing netizens’ comments over time. From all four diagrams, we see that immediately after the exposure to the first policy, the white area (probabilities of respondents staying in recommended comments) shrinks precipitously, while the gray area (probabilities of respondents staying in netizens’ comments) increases significantly. This implies that for respondents who chose to view recommended comments at the beginning, the probability of them continuing to choose recommended comments declines over time, while the probability of them changing their comment exposure to netizens’ comments increases over time.

The figures help visualize the importance of Hypothesis 2 that individuals who exposed themselves to pro-regime comments are less likely to continue consuming them but more likely to switch their comment exposure to critical comments. To
conclude, the results reveal that individuals likely consume alternative information that deviates from official discourses, thereby calling into question the effectiveness of soft propaganda orchestrated by the authorities (Dubois & Blank, 2018).

Approval of Policies

To examine whether online comments succeed in shaping people’s policy support, we need to compare our respondents with individuals who have no comment exposure. Recall that Study 2 consists of 652 respondents who were allowed to select comments to view. We included additional 601 respondents who were shown the same policy news but were not given a chance to view any comment. The latter respondents serve as the control group for comparison, while the former respondents the treatment group. By comparing the policy support of these two groups of

---

8. We base our analysis of this section on Study 2 because the survey design of Study 2 is robust to question-order and option-order effects while allowing us to control for the potential confounding effects of political attitudes.

9. We assigned respondents to these groups at random. As such, we can reduce the influences of unobserved factors. For instance, critical netizens may self-select to read more critical comments, while regime-supporting netizens may be more likely to expose themselves to non-critical comments. The random assignment helps reduce such selection effects.
Table 3 Profiles of respondents in the control group and the treatment group

|                        | Control group | Treatment group | Standardized mean difference |
|------------------------|---------------|-----------------|-----------------------------|
|                        | N  | Mean   | Std     | Min | Max | N  | Mean   | Std     | Min | Max    |               |
| Male                   | 601 | 0.474  | 0.498   | 0   | 1   | 652 | 0.440  | 0.497   | 0   | 1      | 0.068          |
| Age                    | 601 | 37.370 | 11.157  | 18  | 80  | 652 | 36.900 | 10.660  | 18  | 78     | 0.043          |
| CCP member             | 601 | 0.226  | 0.419   | 0   | 1   | 652 | 0.253  | 0.435   | 0   | 1      | −0.063         |
| Education level        | 601 | 5.488  | 1.0458  | 1   | 8   | 652 | 5.544  | 1.019   | 1   | 8      | −0.055         |
| Income level           | 601 | 3.411  | 1.717   | 1   | 9   | 652 | 3.460  | 1.784   | 1   | 9      | −0.028         |
| Political interest     | 601 | 4.130  | 0.806   | 1   | 5   | 652 | 4.133  | 0.795   | 1   | 5      | −0.005         |
| Frequency of social media news consumption | 601 | 3.562  | 0.645   | 1   | 4   | 652 | 3.538  | 0.683   | 1   | 4      | 0.036          |
| Frequency of social media use | 601 | 4.483  | 0.851   | 1   | 5   | 652 | 4.469  | 0.913   | 1   | 5      | 0.015          |
| Most recent social media use | 601 | 6.659  | 1.485   | 1   | 8   | 652 | 6.584  | 1.632   | 1   | 8      | 0.048          |
| Frequency of social media discussion | 601 | 2.923  | 0.825   | 1   | 4   | 652 | 2.873  | 0.854   | 1   | 4      | 0.060          |
| Internal political efficacy | 601 | 2.416  | 0.757   | 1   | 4   | 652 | 2.437  | 0.768   | 1   | 4      | −0.028         |
| External political efficacy | 601 | 2.649  | 0.815   | 1   | 4   | 652 | 2.643  | 0.846   | 1   | 4      | 0.008          |
| Resistance against the West | 601 | 3.251  | 0.722   | 1   | 4   | 652 | 3.222  | 0.708   | 1   | 4      | 0.040          |
| Collectivism           | 601 | 3.300  | 0.712   | 1   | 4   | 652 | 3.247  | 0.700   | 1   | 4      | 0.074          |
| National pride         | 601 | 3.785  | 0.502   | 1   | 4   | 652 | 3.781  | 0.489   | 1   | 4      | 0.009          |
| Information transparency | 601 | 2.980  | 0.787   | 1   | 4   | 652 | 3.014  | 0.775   | 1   | 4      | −0.043         |
| Life satisfaction      | 601 | 3.073  | 0.652   | 1   | 4   | 652 | 3.021  | 0.707   | 1   | 4      | 0.076          |
| Society satisfaction   | 601 | 3.063  | 0.658   | 1   | 4   | 652 | 3.064  | 0.660   | 1   | 4      | −0.002         |

Note: For variables descriptions, please see Online Appendix 2.
respondents \((N = 652 + 601 = 1253)\), we will be able to detect the potential effects of comment exposure on policy approval.

Table 3 shows the demographic characteristics of the respondents in the treatment group and the control group. Both groups have a similar number of respondents. Male respondents are slightly underrepresented in both groups, while members of the Chinese Communist Party are overrepresented in the sample. Despite these observational differences, the standardized mean differences of all variables are less than 0.1, suggesting that the control and treatment groups are more or less balanced across the covariates.

Since the dependent variable is a binary variable (agree/disagree with the policy), we apply both the random-effects logistic and the logistic regression model\(^{10}\) to analyze the effects of different comment exposures on respondents’ approval of policies between the treatment group and the control group. The random-effects logistic regression specification is as follows:

\[
\Pr(Z_{it}) = \text{logit}^{-1}(\sigma_i + \rho D_i + \mu N_i + \delta (D_i \times N_i) + P_{it}^\prime \rho + W_{it} \phi + \lambda_{it})
\]  

(2)

where \(Z_{it}\) denotes individual \(i\)’s approval for a policy shown at time \(t\),\(^{11}\) \(\rho\) is the coefficient on the treatment status \(D_i\), which takes a value of 1 if respondent \(i\) is in the treatment group and 0 otherwise, \(\mu\) is the coefficient on policy number \(N_i\), \(\delta\) is the coefficient on the interaction term between treatment status \(D_i\) and policy number \(N_i\), \(\rho\) is a vector of coefficients on policy type dummies \(P_{it}\),\(^{12}\) \(\phi\) is a vector of coefficients on individual-level control variables \(W_{it}\), \(\sigma_i\) is individual \(i\)’s intercept, and \(\lambda_{it}\) is the error term clustered at the individual level which is assumed to be independent and identically distributed.

Table 4 shows the regression results. The coefficient on the treatment status \(D_i\) is significantly negative, which implies that compared with the control group, people in the treatment group who have choices of comment exposure show significantly lower levels of approval of the policies. This finding implies that with the influence of online opinions, netizens are more likely to question and even disapprove of social policies implemented by the state. The coefficients on the policy number variable \(N_i\) and the interaction term of policy number \(N_i\) and treatment status \(D_i\) are insignificant, indicating that individuals’ policy approvals in both groups are not affected by the display order of policies.

Political attitude variables also seem to play some roles in determining individuals’ policy approval. The coefficients on variables *Internal political efficacy*, *Collectivism*, *National pride*, and *Society satisfaction* are significantly positive. The results show that people who think politics is hard to understand, who consider collective

---

\(^{10}\) Individual fixed-effects cannot be added to the logistic regression model because of perfect multicollinearity between individual fixed-effects and individuals’ treatment status.

\(^{11}\) \(Z_{it}\) is a binary variable that takes a value of “1” if a respondent agrees with the policy and “0” otherwise.

\(^{12}\) The policy type dummies \(P_{it}\) are intended to control for idiosyncratic effects of particular policy news to which respondents are exposed.
| Dep. var. policy approval | Random-effects logistic | Logistic |
|--------------------------|-------------------------|----------|
|                          | (1)                     | (2)      |
| $D_i$                    | $-0.438^{*}$            | $-0.417^{**}$ |
|                          | (0.177)                 | (0.156)  |
| $N_t$                    | $-0.050$                | $-0.052$  |
|                          | (0.034)                 | (0.030)  |
| $D_i \times N_t$         | 0.051                   | 0.051    |
|                          | (0.042)                 | (0.037)  |
| Policy 2                 | 1.551^{**}              | 1.452^{**} |
|                          | (0.311)                 | (0.290)  |
| Policy 3                 | $-0.718^{**}$           | $-0.497^{**}$ |
|                          | (0.212)                 | (0.187)  |
| Policy 4                 | $-2.378^{**}$           | $-1.962^{**}$ |
|                          | (0.186)                 | (0.159)  |
| Policy 5                 | $-0.631^{**}$           | $-0.480^{**}$ |
|                          | (0.184)                 | (0.161)  |
| Policy 6                 | $-0.462^{*}$            | $-0.335^{*}$ |
|                          | (0.188)                 | (0.165)  |
| Policy 7                 | 0.742^{**}              | 0.734^{**} |
|                          | (0.201)                 | (0.178)  |
| Policy 8                 | $-0.094$                | $-0.013$  |
|                          | (0.191)                 | (0.169)  |
| Male                     | 0.168                   | 0.164^{*} |
|                          | (0.094)                 | (0.082)  |
| Age                      | 0.011^{*}               | 0.010^{*} |
|                          | (0.005)                 | (0.004)  |
| CCP member               | 0.058                   | 0.055    |
|                          | (0.112)                 | (0.098)  |
| Education level          | $-0.082$                | $-0.072$  |
|                          | (0.054)                 | (0.047)  |
| Income level             | $-0.012$                | $-0.013$  |
|                          | (0.032)                 | (0.029)  |
| Political interest       | 0.073                   | 0.061    |
|                          | (0.077)                 | (0.068)  |
| Frequency of social media news consumption | $-0.042$ | $-0.027$ |
|                          | (0.099)                 | (0.085)  |
| Frequency of social media use | 0.078 | 0.080 |
|                          | (0.092)                 | (0.079)  |
| Most recent social media use | $-0.046$ | $-0.045$ |
|                          | (0.053)                 | (0.046)  |
| Frequency of social media discussion | 0.177^{*} | 0.143^{*} |
|                          | (0.071)                 | (0.063)  |
Table 4 (continued)

| Dep. var. policy approval | Random-effects logistic | Logistic |
|---------------------------|-------------------------|----------|
|                           | (1)                     | (2)      |
| Internal political efficacy | 0.182**                 | 0.161**  |
|                           | (0.064)                 | (0.057)  |
| External political efficacy | 0.010                   | -0.002   |
|                           | (0.057)                 | (0.050)  |
| Resistance against the West | 0.043                   | 0.039    |
|                           | (0.071)                 | (0.063)  |
| Collectivism               | 0.198**                 | 0.189**  |
|                           | (0.073)                 | (0.063)  |
| National pride             | 0.293*                  | 0.261**  |
|                           | (0.116)                 | (0.099)  |
| Information transparency   | 0.074                   | 0.069    |
|                           | (0.059)                 | (0.053)  |
| Life satisfaction          | -0.006                  | -0.001   |
|                           | (0.090)                 | (0.080)  |
| Society satisfaction       | 0.374**                 | 0.325**  |
|                           | (0.096)                 | (0.085)  |
| Number of Obs             | 7512                    | 7512     |
| Log pseudolikelihood      | -3220.336               | -3307.814|
| Wald Chi²                 | 883.570                 |          |
| Pseudo R²                 |                         | 0.154    |

Notes: Dependent variable is respondents’ policy approval at time \( t (t = 1, 2, \ldots, 6) \) that takes a value of “1” if they agree with the policy and “0” otherwise. Independent variables include the treatment status \( D_i \) that takes a value of “1” if respondents are in the treatment group and “0” otherwise; policy number \( N_t (N_t = 1, 2, \ldots, 6) \), and the interaction term of \( D_i \) and \( N_t \). Policy 1 is the baseline category. Provincial fixed-effects and cutoff points are not reported but are available upon request. Standard errors clustered at the individual level are in parentheses.

**p < 0.01, *p < 0.05

Random-effects logistic regression margins

Logistic regression margins

Fig. 7 Marginal effects of policy approval over time between the treatment group and the control group
interests are prior to individual interests, who have greater pride in China, and who are more satisfied with the current society are more likely to support the policies.

Our variable of interest affects policy approval via the standalone term $D_i$ and the interaction term $D_i \times N_t$. It is instructive to visualize the combined effects of these two terms. Figure 7 shows the marginal effects of the treatment on policy approval. As shown in the figure, individuals in the control group tend to disagree with the policies over time while respondents in the treatment group—namely, those who have a choice to view comments—consistently show markedly lower approval of the policies, although the significant difference between the treatment and control groups gradually disappears as $N_t$ increases.

To conclude, once individuals are provided with a chance to view different discourses, they not only tend to expose themselves to critical information but also are influenced by comments and become more critical of the underlying policies. The result is consistent with Hypothesis 3.

**Conclusion**

Some scholars contend that leaders in authoritarian regimes in the twenty-first century no longer rule by fear. Rather, they rule by information manipulation (Chen & Xu, 2017; Guriev & Treisman, 2019; Rozenas & Stukal, 2019). In particular, autocrats use social media to run “soft propaganda” to change citizens’ political preferences (Gunitsky, 2015). One way to achieve this is to intervene in online discussion by producing and amplifying pro-regime comments. Those messages are believed to shape public opinions in favor of the regime, as they help strengthen individuals’ trust in the authorities, thereby promoting regime legitimacy.

While the regime has a strong incentive to spread pro-regime online comments, netizens’ views on such comments remain less understood. Some scholars argue that information manipulation in authoritarian regimes does not always work. At times it may even backfire, because people in those regimes may resist the information controls (Pan & Siegel, 2020; Roberts, 2020). This article provides one of the first systematic investigations into social media comment manipulation in China and aims to contribute to the literature of resilience to information manipulation in authoritarian regimes.

Our findings suggest that the comment section on social media platforms performs an important information role; it helps social media users gain unorthodox perspectives on social events. In other words, it provides Chinese netizens with a source of alternative information under a highly regulated media environment. This illustrates that even a tiny space for critical voices may undermine the public’s policy support. Indeed, manipulating social media comments is a kind of soft propaganda that is more subtle and less direct compared with censorship that aims to withhold information from the public entirely. Nonetheless, our findings indicate that people can be resilient to this subtle manipulation as well. People are not passive recipients of state propaganda (Rosenfeld, 2018). Rather, they have the ability to distinguish different kinds of information and even have an appetite for information that deviates from the official discourse. Extant works on authoritarian media...
controls focus on resilience against censorship (Chen & Yang, 2019; Hobbs & Roberts, 2018; Wong & Liang, 2021). Our findings suggest that resilience against soft propaganda is also worth investigating.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11109-021-09744-4.

Acknowledgements The authors would like to thank Juan Chen, Martin Dimitrov, Christian Göbel, Robert Harmel, Hiroki Takeuchi, Hans Tung, Elsie Yan, Yao-Yuan Yeh, Yu Zeng, and Vivian Zhan for their excellent comments on the previous drafts of this article. All errors are our own.

References

Arceneaux, K., Johnson, M., & Murphy, C. (2012). Polarized political communication, oppositional media hostility, and selective exposure. The Journal of Politics, 74(1), 174–186.

Bremmer, I. (2010). Democracy in cyberspace: What information technology can and cannot do. Foreign Affairs, 89(6), 86–92.

Chase, M. S., & Mulvenon, J. C. (2002). You’ve got dissent! Chinese dissident use of the internet and Beijing’s counter-strategies. Rand Corporation.

Chen, D. (2017). Local distrust and regime support: Sources and effects of political trust in China. Political Research Quarterly, 70(2), 314–326.

Chen, J., & Xu, Y. (2017). Information manipulation and reform in authoritarian regimes. Political Science Research and Methods, 5(1), 163–178.

Chen, J., Zhong, Y., & Hillard, J. W. (1997). The level and sources of popular support for China’s current political regime. Communist and Post-Communist Studies, 30(1), 45–64.

Chen, Y., & Yang, D. Y. (2019). The impact of media censorship: 1984 or brave new world? American Economic Review, 109(6), 2294–2332.

Clifford, S., Jewell, R. M., & Waggoner, P. D. (2015). Are samples drawn from Mechanical Turk valid for research on political ideology? Research & Politics, 2(4), 1–9.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. Journal of Communication, 64(2), 317–332.

Creemers, R. (2017). Cyber China: Upgrading propaganda, public opinion work and social management for the twenty-first century. Journal of Contemporary China, 26(103), 85–100.

Crowther, M. J., & Lambert, P. C. (2017). Parametric multistate survival models: Flexible modelling allowing transition-specific distributions with application to estimating clinically useful measures of effect differences. Statistics in Medicine, 36(29), 4719–4742.

Deibert, R., & Rohozinski, R. (2010). Liberation vs. control: The future of cyberspace. Journal of Democracy, 21(4), 43–57.

Diamond, L. (2010). Liberation technology. Journal of Democracy, 21(3), 69–83.

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. Information, Communication & Society, 21(5), 729–745.

Egorov, G., Guriev, S., & Sonin, K. (2009). Why resource-poor dictators allow freer media: A theory and evidence from panel data. American Political Science Review, 103(4), 645–668.

Fang, K., & Repnikova, M. (2018). Demystifying “Little Pink”: The creation and evolution of a gendered label for nationalistic activists in China. New Media & Society, 20(6), 2162–2185.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. Journal of Consumer Research, 20(2), 303–315.

Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among internet news users. Journal of Computer-Mediated Communication, 14(2), 265–285.

Geddes, B., & Zaller, J. (1989). Sources of popular support for authoritarian regimes. American Journal of Political Science, 33(2), 319–347.
Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, 26(3), 213–224.

Greitens, S. C. (2013). Authoritarianism Online: What can we learn from internet data in democracies? *PS Political Science & Politics*, 46(2), 262–270.

Guriev, S., & Treisman, D. (2019). Informational autocrats. *Journal of Economic Perspectives*, 33(4), 100–127.

Guintsky, S. (2015). Corrupting the cyber-commons: Social media as a tool of autocratic stability. *Perspectives on Politics*, 13(1), 42–54.

Han, R. (2015a). Defending the authoritarian regime online: China’s “voluntary fifty-cent army.” *The China Quarterly*, 224, 1006–1025.

Han, R. (2015b). Manufacturing consent in cyberspace: China’s “fifty-cent army.” *Journal of Current Chinese Affairs*, 44(2), 105–134.

Hobbs, W. R., & Roberts, M. E. (2018). How sudden censorship can increase access to information. *American Political Science Review*, 112(3), 621–636.

Huang, H. (2015). Propaganda as signaling. *Comparative Politics*, 47(4), 419–444.

Huang, H. (2018). The pathology of hard propaganda. *The Journal of Politics*, 80(3), 1034–1038.

Hung, C. F. (2010). China’s propaganda in the information age: Internet commentators and the Weng’an incident. *Issues & Studies*, 46(4), 149–180.

Hyun, K. D., & Kim, J. (2015). The role of new media in sustaining the status quo: Online political expression, nationalism, and system support in China. *Information, Communication & Society*, 18(7), 766–781.

Hyun, K. D., Kim, J., & Sun, S. (2014). News use, nationalism, and internet use motivations as predictors of anti-Japanese political actions in China. *Asian Journal of Communication*, 24(6), 589–604.

Jiang, Y. (2012). Cyber-nationalism in China: Challenging Western media portrayals of internet censorship in China. University of Adelaide Press.

Kennedy, J. J. (2009). Maintaining popular support for the Chinese Communist Party: The influence of education and the state-controlled media. *Political Studies*, 57(3), 517–536.

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, 107(2), 326–343.

King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *American Political Science Review*, 111(3), 484–501.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47(4), 2025–2047.

Li, L. (2004). Political trust in rural China. *Modern China*, 30(2), 228–258.

Li, L. (2016). Reassessing trust in the central government: Evidence from five national surveys. *The China Quarterly*, 225, 100–121.

Li, X., Shi, W., & Zhu, B. (2017). The face of internet recruitment: Evaluating the labor markets of online crowdsourcing platforms in China. *Research & Politics*, 5(1), 1–8.

Ludwig, J. Z. (2020). Chernobyl, coronavirus, and the art of the authoritarian cover-up: Who is to blame, and what is to be done? *SAGE International Australia*, 1(1), 1–7.

Lü, X. (2014). Social policy and regime legitimacy: The effects of education reform in China. *American Political Science Review*, 108(2), 423–437.

Lynch, M. (2011). After Egypt: The limits and promise of online challenges to the authoritarian Arab state. *Perspectives on Politics*, 9(2), 301–310.

Miller, B. (2016). *Automated detection of Chinese government astroturfers using network and social metadata*. Available at SSRN 2738325.

Nip, J. Y., & Fu, K. W. (2016). Challenging official propaganda? Public opinion leaders on Sina Weibo. *The China Quarterly*, 225, 122–144.

Pan, J., & Siegel, A. A. (2020). How Saudi crackdowns fail to silence online dissent. *American Political Science Review*, 114(1), 109–125.

Pan, J., & Xu, Y. (2018). China’s ideological spectrum. *The Journal of Politics*, 80(1), 254–273.

Qin, B., Strömberg, D., & Wu, Y. (2017). Why does China allow freer social media? Protests versus surveillance and propaganda. *Journal of Economic Perspectives*, 31(1), 117–140.

Qiu, J. L. (2006). The changing web of Chinese nationalism. *Global Media and Communication*, 2(1), 125–128.
Repnikova, M., & Fang, K. (2018). Authoritarian participatory persuasion 2.0: Netizens as thought work collaborators in China. *Journal of Contemporary China, 27*(113), 763–779.

Roberts, M. E. (2020). Resilience to online censorship. *Annual Review of Political Science, 23*(1), 401–421.

Rosenfeld, B. (2018). The popularity costs of economic crisis under electoral authoritarianism: Evidence from Russia. *American Journal of Political Science, 62*(2), 382–397.

Rozenas, A., & Stukal, D. (2019). How autocrats manipulate economic news: Evidence from Russia’s state-controlled television. *The Journal of Politics, 81*(3), 982–996.

Shi, T. (2001). Cultural values and political trust: A comparison of the People’s Republic of China and Taiwan. *Comparative Politics, 33*(4), 401–419.

Tang, M., & Huhe, N. (2014). Alternative framing: The effect of the internet on political support in authoritarian China. *International Political Science Review, 35*(5), 559–576.

Teixeira, L. (2020, April 17). Thais show how to beat China’s online army: A pop culture battle turned into a nationalist frenzy and a propaganda problem. *Foreign Policy*. Retrieved from https://foreignpolicy.com/2020/04/17/nnevvy-bright-firewall-thailand-china-online-army/

Tong, Y., & Lei, S. (2013). War of position and microblogging in China. *Journal of Contemporary China, 22*(80), 292–311.

Vuori, J. A., & Paltemaa, L. (2015). The lexicon of fear: Chinese internet control practice in Sina Weibo microblog censorship. *Surveillance & Society, 13*(3/4), 400–421.

Wang, Z. (2005). Before the emergence of critical citizens: Economic development and political trust in China. *International Review of Sociology, 15*(1), 155–171.

Wong, S. H. W., & Liang, J. (2021). Dubious until officially censored: Effects of online censorship exposure on viewers’ attitudes in authoritarian regimes. *Journal of Information Technology & Politics*. https://doi.org/10.1080/19331681.2021.1879343

Wu, A. X. (2013). Ideological polarization over a China-as-superpower mindset: An exploratory charting of belief systems among Chinese internet users, 2008–2011. *International Journal of Communication, 8*, 2243–2272.

Yang, G. (2014). The return of ideology and the future of Chinese internet policy. *Critical Studies in Media Communication, 31*(2), 109–113.

Yang, P., & Tang, L. (2018). “Positive energy”: Hegemonic intervention and online media discourse in China’s Xi Jinping era. *China: An International Journal, 16*(1), 1–22.

Yang, Q., & Tang, W. (2010). Exploring the sources of institutional trust in China: Culture, mobilization, or performance? *Asian Politics & Policy, 2*(3), 415–436.

Zhang, C. (2020). Covid-19 in China: From ‘Chernobyl moment’ to impetus for nationalism. *Made in China Journal*, 5(2), 162–165.

Zhao, S. (1998). A state-led nationalism: The patriotic education campaign in post-Tiananmen China. *Communist and Post-Communist Studies, 31*(3), 287–302.

Zhao, S. (2016). The ideological campaign in Xi’s China: Rebuilding regime legitimacy. *Asian Survey, 56*(6), 1168–1193.

Zhou, Y. (2005). Informed nationalism: Military websites in Chinese cyberspace. *Journal of Contemporary China, 14*(44), 543–562.

Zhu, J., Lu, J., & Shi, T. (2013). When grapevine news meets mass media: Different information sources and popular perceptions of government corruption in mainland China. *Comparative Political Studies, 46*(8), 920–946.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.