

Do not circulate or cite without the author's permission

**Reflecting the Crowd: How the Presumed Influence of Online Opinion Leaders Addresses
Coordination Failures in Autocracies**

Chi Shun (Gary) Fong

PhD, Department of Political Science, Pennsylvania State University

Abstract

While social media aids mobilization by spreading grievances and coordinating dissent, authoritarian governments censor such information and suppress public signals of collective willingness to protest. Citizens' tendency to conceal true preferences to avoid punishment compounds this coordination problem. I argue that online opinion leaders — such as influencers, independent journalists, experts, and activists — mitigate this collective action problem by signaling their online influence and, in turn, the potential for mass persuasion when they endorse the protest. Perceived as knowledgeable and socially influential, these actors raise citizens' expectations that others will participate, thereby increasing individual protest intent. I test this argument using a pre-registered survey experiment that exposes Chinese citizens ($N = 2,461$) to messages about a hypothetical environmental protest and manipulates two influence signals: influence metrics and sender identity. Influence metrics, such as follower count and number of likes and shares received, signal message reach, while sender identity signals expertise for persuasion. Results show that influence signals increase protest intent by raising expectations of others' participation, thus helping explain why autocratic governments increasingly co-opt or repress online opinion leaders.

Keywords: Influence of Presumed Influence, Social Media, Online Opinion Leaders, Collective Action Problems, Authoritarianism

Introduction¹²³

The development of information and communication technologies (ICT), particularly social media, facilitates protest mobilization by enabling faster, less restricted flows of protest information through interpersonal networks (Breuer et al., 2015; Jost et al., 2018; Little, 2016). On the one hand, social media provide an alternative source of information that exposes government wrongdoing and social protests, making citizens doubt the government's legitimacy and more sympathetic to activism. During the Tunisian Revolution in 2010, for instance, activists used Twitter to bypass government censorship, arousing public grievances over the death of a street vendor extorted by police (Lim, 2013). Social media also lowers the cost of coordination by facilitating the horizontal exchange of protest logistics, such as information on protest event details, live updates on protest size, and reports on police presence (Jost et al., 2018; Juris, 2005). Importantly, this information flows not only to those most inclined to join the protest, such as activists, but also to domestic and diasporic citizens who are more distant from the protest community (Steinert-Threlkeld, 2017). Even in authoritarian states with strong capacity for digital repression, like Russia and China, social media reduces the coordination problem by connecting like-minded dissidents without formal organization (Enikolopov et al., 2020) and helps spread isolated protests (Qin et al., 2024).

However, few studies explain how information flows on social media overcome the problem of preference falsification in authoritarian settings, where repression drives citizens to conceal anti-regime attitudes and protest intentions (Kuran, 1991). To decide whether to join costly collective action, citizens rely on information about others' anti-regime attitudes and protest intentions (de Mesquita, 2010; Lohmann, 1994) — what many call second-order information (Chwe, 2013; Edmond, 2013). If citizens think protests are small and unlikely to succeed, the potential for punishment creates a strong selective disincentive to join (Olson, 1971). While dissatisfied citizens may find other like-minded dissenters on social media (Breuer et al., 2015), they are just as likely to find such information environments flooded with pro-regime voices, including indoctrinated regime supporters (Han, 2015a), self-censored citizens

¹ This study has passed the ethical reviews of Penn State (STUDY00028928)

² This study is supported by the Penn State's McCourtney Institute for Democracy, Research in Democracy Support Grants

³ This study's hypotheses, design, and analysis plan were preregistered on the Open Science Framework (OSF)

(Yilmaz et al., 2024), and paid government commenters (Han, 2015b). As digital repression rises, it is therefore increasingly important to understand how social media facilitates protests, especially during the early stages of mobilization when protest size has yet to reach the tipping point for mass mobilization.

In this paper, I argue that online opinion leaders — social media users who command outsized audiences and influence (Besalú et al., 2021; Huffaker, 2010), such as journalists, activists, experts, and influencers — can address this collective action problem. Because citizens perceive these figures as socially influential—and thus capable of reaching and persuading large numbers of people—their endorsement of protest serves as a credible signal of broad participation. New information about broad support for protest, in turn, raises citizens' expectations that others will join and, thus, increases their willingness to do so. This argument builds on “influence of presumed influence” theories in communication studies (Gunther et al., 2006; Gunther & Storey, 2003), which posit that mass media shape personal opinion directly, via persuasion and, indirectly, by demonstrating the presumed influence of information on others. Similarly, scholars in other social science fields theorize that second-order perceptions of others' beliefs shape first-order beliefs, with implications for propaganda (Huang & Cruz, 2022), ideological polarization (Ahler, 2014), foreign policy support (Lo et al., 2022; Todorov & Mandisodza, 2004; Wei et al., 2017), public opinion in autocracies (Buckley et al., 2024), and climate policy attitudes (Mildenberger & Tingley, 2019). My argument echoes these studies in that people first perceive others' beliefs and related behaviors from a piece of information and then react to that perception. In other words, the information indirectly influences the individual through changing their perception of others' beliefs and behaviors.

This study focuses on how citizens' decisions regarding costly protest participation in authoritarian regimes shift in response to information that shapes not only first-order preferences but also second-order beliefs about others' participation. Coordinating behavior requires potential participants to possess “common knowledge” about other people's participation (Chwe, 2013), and the dearth of this second-order information often prevents protest mobilization in authoritarian settings. Endorsement from online opinion leaders (OOLs) serves as a focal signal that helps coordinate these higher-order beliefs by making latent support more visible and

seemingly widespread. When citizens believe these leaders' messages reach and persuade a substantial number of citizens through observable information about these online leaders, the mass public is more likely to believe that others are likely to protest. This belief, in turn, alters citizens' own intention to protest.

In this regard, endorsement of OOLs increases citizens' willingness to protest by reducing uncertainty about others' collective action thresholds and by mitigating fears that their own participation in a small protest will result in repression. These expectations are inherently recursive: individuals form beliefs about others' willingness to act, which in turn depend on beliefs about others' expectations, thereby generating higher-order beliefs, i.e., expectations about others' expectations (Chwe, 2013). In authoritarian contexts where information opacity and strategic silence typically inhibit the formation of the shared expectations necessary for collective mobilization, this mechanism may encourage citizens to protest even when the public does not overwhelmingly support the protest at an early stage but perceives it as such, or as imminent. If, as Kuran's work points out, self-censorship breeds a common belief that everyone supports the regime, then the converse may be true as well: when OOLs help establish a common belief that there is substantial support for protest, this belief can induce mass mobilization.

To test my theory, I conducted a survey experiment with 2,461 Chinese citizens, presenting them with a message endorsing a hypothetical protest event criticizing a controversial waste incineration project on the Chinese social media platform Weibo. All respondents received the same information about the project's environmental hazard and local citizens' intention to protest, and I randomly assigned them to one of five profiles of online leaders: a control group with an ordinary social media user profile and four treatment groups with an influential, platform-verified user profile. These influential users – or BigV, indicated by the big tick on their profile pictures – are generally perceived as OOLs by Chinese netizens (Gallagher & Miller, 2021). I manipulate two dimensions of influence signals: (1) **influence metrics**, which provide information on message reach; and (2) **profile identity**, which enhances the perception that the information is broadly persuasive. **Influence metrics**, such as follower count and the number of shares and likes received (low vs. high), indicate the message's potential reach to other citizens; and **profile identity** (lifestyle influencers vs. investigative journalists) signals

domain expertise on environmental issues and thus the message's perceived persuasiveness to other citizens.

The results show that a protest-endorsing message from an OOL increases respondents' protest intent as well as expectations of others' protest intent. While the treatment effect on others' protest intent is the strongest when the message is posted by an investigative journalist with high influence metrics, the effect of **profile identity** and **Influence metrics** alone is not statistically significant. This suggests the effect in the experiment is mainly driven by the broad perception of the messenger as influential, and the additional effects of these two dimensions of influence alone, at least in this case, are minimal if not present together. Mediation analysis further reveals that the treatment effects on respondents' protest intent are mostly indirectly captured by respondents' increased expectations of others' protest intent. Overall, these results support my argument that, beyond direct persuasion, OOLs could drive citizens to participate in protest by signaling their mass influence.

The findings of this paper broaden the existing discussion on how social media facilitate collective actions. Beyond facilitating the spread of information that shapes pro-protest attitudes among citizens and reduces the cost of tactical coordination, social media also provide cues for potential participants to evaluate the extent to which other citizens will participate, which is a critical determinant of their participation. While such cues are more openly available in democracies, whether online or offline, they are scant in authoritarian settings as citizens falsify their private preferences in public (Kuran, 1991). This paper demonstrates how OOLs could transform diffuse, often hidden discontent into credible expectations of collective action under this repressive context. As such, this paper advances a general theory of digital dissent under authoritarianism by unpacking the belief mechanisms linking first-order persuasion to second-order expectations in ways that existing studies seldom theorize (Breuer et al., 2015; Little, 2016). The findings also help to explain why authoritarian regimes increasingly coopt and repress OOLs (Chen & Gao, 2023; Y. Lu, 2026; Woolley, 2022), as these regimes shift their informational strategies from monitoring and moderating online content (King et al., 2013, 2017) to targeting influential OOLs (Gallagher & Miller, 2021; Yan & Li, 2024).

Obstacles to (Online) Protest Mobilization in Authoritarian Regimes

Initiating protests in authoritarian contexts is difficult, especially when the target of protest is the central political authority.

First, governments invest in indoctrination through institutions such as education (Cantoni et al., 2017), the entertainment industry (Mattingly & Yao, 2022; Xia, 2022), and the media (Edney, 2012; Rozenas & Stukal, 2019; Stockmann & Gallagher, 2011), making these information environments less likely to spread dissent or provide coordinating information for potential mobilization. Second, authoritarian governments repress independent civil society groups and control media access, leaving little room for political organization (Toepler et al., 2020) that provides resources for mobilization and protest coordination (McCarthy & Zald, 1977). Third, and most importantly, authoritarian regimes make protest costly by punishing citizens (Davenport, 2007; Svolik, 2012). On the one hand, repression directly disincentivizes citizens from protesting by both increasing the costs of participation (Earl, 2003; Hassan et al., 2022) and by fostering fear (Pearlman, 2013; Young, 2019). More perniciously, repression indirectly deters protest by inducing preference falsification (Kuran, 1991). Citizens living under authoritarianism often misrepresent their true political preference in public to protect themselves, generating an informational environment in which public support for the regime appears far stronger than it is. This distortion keeps protest thresholds high, as citizens are discouraged from engaging in risky actions that are unlikely to succeed without sufficient public support (de Mesquita, 2010; Lohmann, 1994).

The rise of the internet and social media addressed some of these problems. First, social media undermines authoritarian regimes' information control, exposing citizens to dissenting information (Breuer et al., 2015; Little, 2016). With the increasing penetration of the internet and the digitalization of our everyday lives, citizens, whether living in autocracies or not, now access a wide range of information sources beyond print and broadcast media, including user-curated content. Given the scale of online information, authoritarian regimes, even those with strong digital capacity, cannot censor all negative news about the regime, as it is not cost-effective (Edmond, 2013). Indeed, authoritarian governments may permit the free flow of information about the performance of local state agents (Lorentzen, 2014). For instance, the Chinese

government's censorship efforts focus on information with collective action potential (King et al., 2013) and on opinion leaders (Gallagher & Miller, 2021), rather than on information critical of the government. These efforts facilitate the consensus mobilization process (van Stekelenburg & Klandermans, 2013) –the deliberate attempt by movement organizers to align citizens to share a common understanding of social issues with them and to sympathize with the cause of future-related protests.

Second, social media supports the action mobilization process that turns potential sympathizers into participants by facilitating protest coordination. Social media enables like-minded citizens to organize by drawing on expertise from diverse fields and tapping into a larger social network at lower cost, reducing the need for formal organization (Bennett et al., 2014; Juris, 2005). During the 2019 Hong Kong protests, social media platforms such as Telegram drew in ordinary citizens, offering resources and expertise for mobilization, publicity, and on-site tactical coordination (Fu, 2023; Urman et al., 2021). Longitudinal studies in Russia similarly suggest that social media reduces the cost of protest coordination, especially in larger cities where coordination is harder to achieve (Enikolopov et al., 2020).

However, we still do not understand how social media facilitates protest in an autocratic context rife with preference falsification: How, for example, does information from social media transform sympathizers into active participants during the early periods of mobilization before protest behavior is widely observed in public? While social media can help citizens observe others' protest behavior, this information only diffuses once protests reach a visible scale. During the 2011 Tunisian revolution, for example, 73.8% of survey respondents learned online that others had signed up for local demonstrations (Breuer et al., 2015). Such signals are, however, weak in early-stage protests before widespread support for the protest becomes publicly visible on social media. As Little's (2016) game-theoretic model shows, social media can also demobilize when pro-regime voices dominate— whether driven by self-censorship (Yilmaz et al., 2024), indoctrination (Han, 2015a), or regime-led information campaigns (Han, 2015b; King et al., 2013, 2017). Cross-national evidence reflects this: the link between internet use in autocracies and protest occurrence has reversed (Kuznetsova & Tolbert, 2024) as authoritarian regimes improved their digital capacity (Kendall-Taylor et al., 2020). Weidmann & Rød (2019)

similarly find that internet coverage suppresses protest onset in autocracies while increasing the likelihood that protests continue once started—consistent with Little's (2016) model.

Online Opinion Leader and their Presumed Influence on Protest Mobilization

I argue that OOLs' endorsement of protest increases individuals' protest intent by raising viewers' expectations of others' willingness to protest, even when such information is lacking in authoritarian settings due to preference falsification or censorship.

While the internet and social media are originally thought to permit free and equal access to the public sphere (Albrecht, 2006; Diamond, 2010), recent studies find that online opinion is strongly mediated by particular groups of people, or OOLs. While some of them are traditional opinion leaders, notably professionals such as journalists or politicians, who possess information from privileged sources or their own expertise (Alexandre et al., 2022; Besalú et al., 2021; Dubois & Gaffney, 2014), the lower barriers to information production and dissemination also allow a new form of opinion leadership to emerge from ordinary citizens. These new opinion leaders — often called influencers — typically lack formal credentials or institutional positions, yet gain online influence and credibility through several reinforcing mechanisms: building virtual social relationships with followers (Cheng et al., 2024; Rothut, 2025); adopting attractive and engaging linguistic styles (Harff & Schmuck, 2023; Huffaker, 2010; Schmuck et al., 2022); and accumulating visible engagement metrics such as likes, shares, mentions, and subscriptions, along with official verification badges (Treem et al., 2020). These signals are further amplified by platform algorithms that prioritize and reward high-engagement content.

Many OOLs participate in online political discussions in democracies and autocracies (Goodwin et al., 2023; Riedl et al., 2023; Woolley, 2022) with implications for protest participation. Most studies focus on the role of opinion leaders in consensus mobilization that shapes citizens' understanding of social issues and problems. In democracies, where citizens face information overload, opinion leaders increase attention to political news by lending authenticity and trustworthiness to curated content (Harff & Schmuck, 2023) and by making it more comprehensible (Dekoninck & Schmuck, 2024), thus shaping political attitudes and long-term political participation. In authoritarian regimes, influencers circumvent state control to expose

injustice and counter official narratives (N. Liu & Liu, 2023; Nip & Fu, 2016), thereby undermining government legitimacy. Others examine the role of OOLs in action mobilization — specifically, how they recruit protest participants. During the 2011 Egyptian revolution, activists used marketing-style strategies, posting in first-person, colloquial tones, while avoiding explicit ideological messaging and distancing themselves from formal political organizations (Poell et al., 2016). In Russia, feminist activists cultivated a positive and personalized online image to attract participants to low-risk flash-mob protests (Zhaivoronok, 2026). Further research demonstrates that opinion leaders mobilize action by arousing emotions, such as enthusiasm, anger, and solidarity, that lower risk perception and increase optimism about protest outcomes (Pearlman, 2013). These dynamics can be seen in Occupy Wall Street (Jost et al., 2018), the 2011 Egyptian revolution (Poell et al., 2016), the Spanish Indignados (Gerbaudo, 2016), and the 2019 Hong Kong Protests (Zhu et al., 2022)

Drawing on these studies, I argue that OOLs increase the persuasiveness of their messages by using more fashionable, simplified, and emotional language and bolstering their credibility. Credibility is particularly important when the message challenges authority in nondemocratic settings, where citizens often have high institutional trust. The qualities of OOLs — such as being trustworthy online friends, platform-verified content providers with a high follower count, experts, or people who hold institutional positions — signal credibility. Thus, exposure to messages from OOLs increases citizens' protest intent more than exposure to messages from ordinary online users.

H1 (General persuasion of OOL): Exposure to messages from online opinion leaders increases citizens' protest intent more than exposure to messages from ordinary online users.

Nonetheless, I argue that OOLs mobilize citizens not only through direct persuasion but also by changing their expectations of others' willingness to protest. Protest is a collective endeavor (Granovetter, 1978; Olson, 1971), and thus, citizens' willingness to participate depends not only on their own attitudes but also on their beliefs about others' willingness to join (Lohmann, 1994). However, such signals are difficult to observe in authoritarian contexts due to preference falsification (Kuran, 1991), especially in the early phase of action mobilization when

publicly visible dissent is scarce. OOLs provide such signals by showing support for protest actions, even when they are not the organizers. Their visible endorsement—through posts, shares, or symbolic gestures—serves as a heuristic cue that protest support is broader than what directly observable offline dissent would suggest. While opinion leaders have long played this signaling role, social media amplifies their impact by rapidly scaling the visibility, reach, and perceived consensus of their endorsements, thereby lowering informational barriers that previously constrained collective action.

When potential dissenters believe that OOLs reach and persuade a large online audience, this presumed influence increases individual intent to protest. These beliefs about presumed influence stem from qualities of OOLs, such as linguistic style (Harff & Schmuck, 2023; Huffaker, 2010; Schmuck et al., 2022), trust in virtual relationships (Cheng et al., 2024; Rothut, 2025), and perceived domain expertise (Besalú et al., 2021; Maddox, 2023). Further, this perception is reinforced by visible, objective platform metrics — such as likes, shares, and views (Treem et al., 2020), which provide a more concrete and thus stronger presumed exposure to the message than that in traditional mass media (Cho et al., 2021). As a result, even without direct evidence of mass participation in the early stages of mobilization, an online opinion leader's perceived influence creates the impression of growing collective momentum. In this way, OOLs function as early protest mobilizers, whose endorsements signal the fact that other citizens are likely to protest.

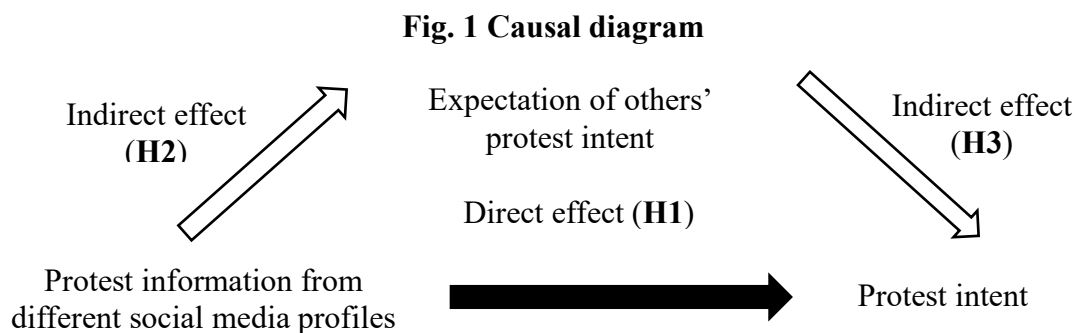
Following this logic, I hypothesize that exposure to messages from OOLs increases citizens' expectations of others' protest intent more than exposure to messages from ordinary online users. Further, stronger influence signals, as proxied by information indicating a larger social media following, should increase protest intent more than weaker influence signals. This expectation, in turn, positively mediates the effect of exposure to OOL messages on protest intent.

H2 (Presumed influence of OOL):

a) Exposure to messages from online opinion leaders increases citizens' expectations of others' protest intent more than exposure to messages from ordinary online users.

b) This presumed influence effect strengthens with stronger influence signals of perceived persuasiveness and reach.

H3 (Influence of Presumed Influence of OOL): The expectation of others' protest intent positively mediates the effect of exposure to messages from online opinion leaders on citizens' protest intent.



This argument draws on the theory of the “influence of presumed influence” in communication studies (Gunther et al., 2006; Gunther & Storey, 2003). Extending persuasive press inference (Gunther, 1998)—the idea that people infer public opinion from mass media—this theory argues that media effects on attitudes and behavior depend on how much individuals believe those messages influence others. Related research examines how first-order beliefs are shaped by second-order beliefs about others’ views. For example, studies show that perceived legitimacy and foreign hostility boost foreign policy support (Lo et al., 2022; Todorov & Mandisodza, 2004; Wei et al., 2017), and that perceived regime popularity shapes regime support (Buckley et al., 2024). The present study advances this research by applying this logic to cooperative yet costly behavior—namely, protest participation. In doing so, it echoes recent work on other cooperative scenarios, such as the perceived mass effect of propaganda on protest intent (Huang & Cruz, 2022) and the perceived foreign public support on climate issues that shape individuals' climate policy preferences (Mildenberger & Tingley, 2019).

Method

To test my argument, I conducted a pre-registered survey experiment in May 2026 with 2,461 Chinese citizens aged 18 or older who passed a basic attention check. Respondents were recruited via a Chinese online survey platform and stratified by province, education level, age, and income, characteristics that largely resembled those of the Chinese online population⁴. The experiment is a 2×2 factorial plus control design to test how two influence signals of protest-sympathetic OOLs — perceived persuasiveness and reach— affect respondents’ own protest intent and their perceived peers’ protest intent. I manipulate reach by altering the influence metrics (low vs. high) and domain expertise by changing the sender identity, either a lifestyle influencer (low expertise) or an investigative journalist (high expertise) – a highly respected occupation in China for exposing government wrongdoings (Svensson, 2017). All respondents first read an identical social media criticism of a waste incineration project in a hypothetical Chinese city and are then randomly assigned a profile, either that of an ordinary user in the control group or that of an online opinion leader with one of the four combinations of the influence signals.

Case Selection

The study uses protests against a waste incineration project in a hypothetical Chinese city. First, this case is ecologically valid. Environmental protests in China over the past few decades, many of which were large-scale, could not have occurred without the internet and social media (Brunner, 2019). While ordinary internet users were critical of government authorities during the protests, OOLs exposed environmental pollution and other related misgovernance (Feng et al., 2020; Wu & Martus, 2021; X. Zhang, 2021; Y. Zhang, 2020); and, on some occasions, participated in organizing protests, including those against incineration facilities (J. Lu & Steinhardt, 2022; Wong, 2016). Second, the Chinese government has tolerated environmental protests, permitting large-scale mobilization of environmental grievances even amid increased repression under Xi Jinping’s rule (Ong & Luo, 2024). Indeed, researchers have documented over 30 anti-waste incineration protests with more than 1,000 participants in the past two decades (Liu & Fong, forthcoming). Third, environmental issues have become increasingly politicized and a source of political legitimacy in China, as the Chinese Communist Party (CCP) has adopted the rhetoric of ecological civilization to promote environmental protection and

⁴ See Appendix A for the balance of demographic data

secure global legitimacy (Arantes, 2024; Lian & Li, 2024). While these protests initially targeted local incineration projects, this political opportunity may enable participants to scale the issue into broader challenges to state policy and regime legitimacy, as occurred in some cases during the 2010s when NGO's, experts, and middle-class participants became (Lu & Steinhardt, 2022; Wong, 2016).

Survey setup

In the survey, respondents are first asked to answer a series of questions about social media usage, political attitudes, and attitudes toward environmental issues. Then, they are randomly assigned to one of the five groups – comprising four treatment groups and one control group – with equal probability. The treatments are posts from Weibo – the Chinese microblogging service with the profile of a user. All respondents initially receive the same post stating that a scientist from the reputable Chinese Academy of Sciences has raised concerns about the environmental risks posed by a newly planned waste-incineration plant in a hypothetical Chinese city. Concerned residents subsequently began organizing a “collective walk,” a euphemism for protest commonly used in China to reduce political sensitivity and to avoid censorship. Importantly, the dissenting activity does not constitute a direct, explicit call for protest, as such appeals are rare in politically sensitive contexts (Yang, 2014). The full translated message is as follows:

An expert from the Chinese Academy of Sciences points out that the approval process for Tongyuan City's waste incinerator involved suspected black-box operations and failed to meet environmental standards. Local residents worry about health impacts and plan a small-scale "group walk" near the city government to protect their rights, demanding the release of environmental assessment data and a public hearing.

Next, respondents received a randomly assigned profile claiming to be the sender of the message. In the control group, respondents received a profile featuring an ordinary user. In the four treatment groups, respondents received a profile of a BigV, an influential user recognized by Weibo, with a tick on their profile picture. Thus, the treatment profiles all signal online influence for Chinese Internet users (Gallagher & Miller, 2021). The four treatment profiles vary along two dimensions, each with two levels: (1) **influence metrics**—high versus low numbers of followers,

reposts, likes, and comments—which signal the message’s reach; and (2) **profile identity**—presented through the profile name and personal biography as either a lifestyle influencer or an investigative journalist—which signals the influencer’s credibility and persuasiveness by conveying domain expertise. Investigative journalists are widely respected in China for exposing social problems and are therefore likely to be perceived by respondents as more professional and trustworthy on social issues, and thus more persuasive in their messaging than lifestyle influencers (Svensson, 2017). I also searched for profiles of real investigative journalists on Weibo to set a reasonable level of influence metrics. Table 1 below provides the details of the treatment vignettes⁵

After exposure to the treatment, respondents answered outcome questions measuring both their own willingness to join collective action against the project (if they were city residents) and their expectations about how likely the average city resident would be to participate. Both questions were answered on a six-point Likert scale. I followed the question wording regarding others’ protest intent from Huang and Cruz (2022), who studied how the presumed effect of propaganda on others influenced Chinese citizens’ willingness to protest. The question does not explicitly ask respondents the extent to which a substantive number of citizens or the exact proportion of others would hold a particular attitude or engage in certain behaviors, as in other studies of presumed influence (Gunther et al., 2006; Gunther & Storey, 2003). Unlike these studies, which refer to real-life events and ask respondents to evaluate their peers, this paper and Huang & Cruz (2022) study a hypothetical situation that makes it hard for respondents to make such an assessment in an out-of-context setting. Thus, asking how respondents think an average person would respond allows them to make more realistic judgments. This question also aligns with the theory because beliefs about how the average person would behave are closely tied to perceptions of how widespread that behavior is. These expectations are inherently recursive: individuals form beliefs about others’ willingness to act based on what they think others expect from one another. In other words, these are higher-order beliefs (Chwe, 2013). The question measuring expectations of others’ participation is asked after the question about respondents’ own protest intentions to avoid priming a potentially spurious mediation relationship (Chaudoin

⁵ See Appendix F for the social media profiles.

et al., 2021). Finally, respondents completed a series of demographic questions covering age, gender, province of residence, household registration type, education, and income.

Table 1 Treatment Vignette – Social Media Profiles			
Experimental Group (N)	Account Type	Influence Metric (follower count)	Biography
1 (495)	Ordinary user	Negligible	Street Musings – A miserable Worker
2 (479)	Lifestyle Influencer	Low (50,000)	Brother Yu’s Life Talk - Lifestyle Blogger, Daily Shares + Social Topics Opinion
3 (503)		High (180,000)	
4 (482)	Investigative Journalist	Low (50,000)	Investigative Journalist-Fangjianguo - Investigative Journalist, Focus on Environment and Public Issues, Support Social Oversight
5 (502)		High (180,000)	

Identification

The primary estimand is the average treatment effect on the treated, where the treatment involves exposure to information about protest mobilization (**H1**). For the primary outcomes, the regression model is

$$Y_i = \alpha + \beta_j Profile_j + \varphi X_i + \varepsilon$$

where (*i*) indexes respondents, (*Y_i*) is the primary outcome measuring respondents’ willingness to participate in a protest, (*Profile_j*) is a factor indicating respondent (*i*)’s treatment assignment to one of the five profiles (*j*), (*β_j*) represents the effect of profile (*j*), and (*X_i*) is a vector of covariates, including *social media use, attitudes on socio and environmental issues, and demographic characteristics*.⁶ Most of these covariates are binarized for more straightforward interpretation and presentation of heterogeneous effects. To estimate presumed influence (**H2**), I re-estimate the model using respondents’ expectations of others’ protest participation as the outcome variable in place of intent to protest. Finally, to assess mediation effects (**H3**), I use the framework developed by Imai et al. (2011) to decompose the treatment effect into the average causal mediation effect (indirect effect) operating through expectations of others’ protest

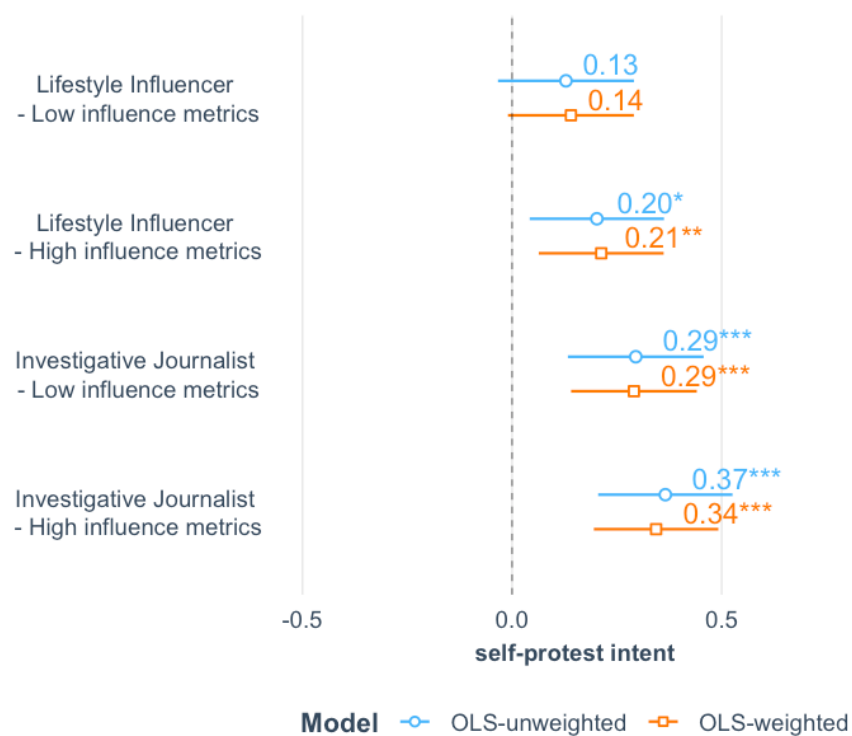
⁶ See Appendix B for the correlation of main covariates, and Appendix E for the survey questions.

participation and the average direct effect on respondents' own protest intentions, estimated via simulation.

Effects of OOLs' Message on Self-Protest Intent and Peer Expectation

Fig. 2 shows the ATEs on the respondent's own willingness to participate in the collective activities against the incineration project in the unweighted and covariate-adjusted OLS models. The regression coefficients are provided on each line, along with the 95% confidence intervals bounded by the two ends. The results match the theoretical expectation. Across both models and various treatment groups, messages from OOLs have positive and statistically significant effects, except for the lifestyle influencer with low influence metrics, supporting **H1**.

Fig. 2 ATEs on self-protest intent



Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

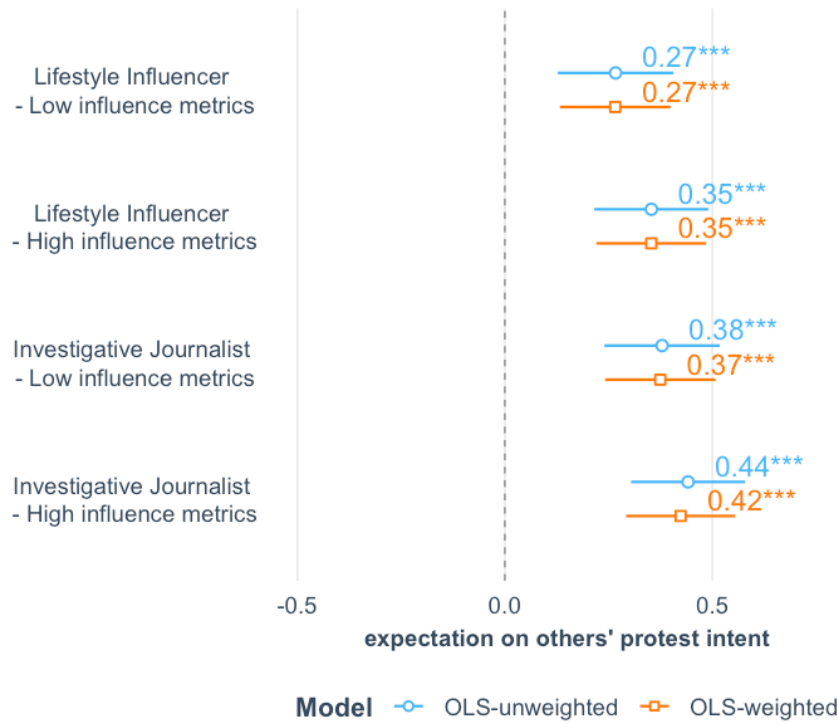
More importantly, the two dimensions of influence proxy work as expected: an investigative journalist could more effectively move respondents to protest than a lifestyle influencer, and the same applies to an OOL with higher influencer metrics than one with lower

influencer metrics. Among the experimental group, the combination of an investigative journalist with high influence metrics achieves the maximum direct impact on self-intent at 0.37 (unweighted) and 0.34 (weighted), with high statistical significance ($p < 0.001$). In this regard, high visibility combined with verified professional integrity, which is associated with stronger perceived influence, is the most effective combination for action mobilization, suggesting that the results are likely driven by respondents' expectations of others' protest intent. These effect sizes, though not particularly large, are substantial in a highly repressive authoritarian regime like China, especially when the protest is hypothetical and unrealized. Consider that the mean value of self-protest intent in the control group is 3.78; the treatment effect is sufficient to move the respondent to answer 4, or "somewhat likely" to join.

Fig. 3 provides clues to probe whether differences in direct treatment effects among treatment groups are driven by respondents' expectations about other citizens' willingness to participate in collective activities opposing the incineration project. Again, the results closely match the theoretical expectation. Across both models and all treatment groups, the message from OOLs has positive and highly statistically significant effects on respondents' expectations of others' protest intent, supporting **H2a**. **Notably**, the treatments shift expectations about others' protest intent (Fig. 3) more strongly than they shift self-protest intent (Fig. 2), consistently across all treatment groups and models. For instance, the high-follower journalist changes peer expectations by 0.42, but self-intent by 0.33 in the covariate-adjusted model. This supports my theory by showing OOLs' calls for action move respondents' assessment of others' intentions more than their assessment of their own protest intent. Additional analysis of heterogeneous treatment effects in Appendix D further supports this argument in that treatment effects on expectations of others' protest intent are consistently higher across all treatment groups among those who perceive OOLs as having a stronger influence on others⁷. However, the heterogeneous effect became saturated when the OOL is an investigative journalist with high influence metrics, perhaps because their perception of being influential is so well-established and independent of personal judgment.

⁷ See Appendix D2 for the analysis on heterogeneous effects of respondents who perceive OOLs with different levels of peer influence.

Fig. 3 ATEs on expectation on others' protest intent⁸



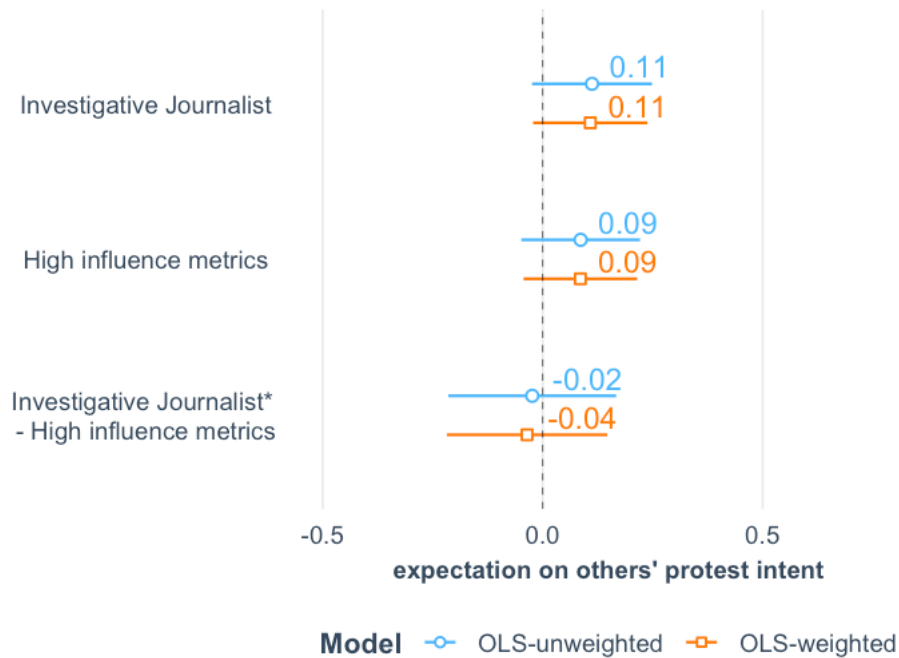
Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Fig. 3 also shows that the dimensions of influence worked as proposed – respondents in treatment groups receiving messages from an investigative journalist or from an influencer with high influence metrics consistently reported higher expectations about others' protest intent than those in treatment groups receiving messages from a lifestyle influencer or from an influencer with low influence metrics. The ATE for the more professional investigative journalist with higher influence metrics is 0.44 (unweighted) and 0.42 (weighted), compared with 0.27 (both unweighted and weighted) for the lifestyle influencer with low influence metrics. This provides partial evidence supporting H2b, indicating that higher visibility combined with verified professional integrity is the most effective combination for shaping second-order beliefs. Fig. 4 further examines whether the differences among the treatment groups are significant by reconfiguring the model in a factorial manner with the control group removed from the sample. However, the results do not detect a statistically significant effect of a single factor, suggesting that treatment effects are mainly driven by a broad impression of the sender's status as

⁸ See Appendix C for the tabulated results.

influential. Profile identity and influence metrics, instead, provide minimal additional updating to peer expectations individually and show a substantive effect only when present together —i.e., citizens require both credibility and social proof simultaneously to update second-order expectations. Overall, these results suggest that the treatment effectively shifts respondents' calculations of the collective environment, laying the empirical foundation for the mediation mechanism.

Fig. 4 ATEs on expectation on others' protest intent – factorial design⁹



Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

OOLs' Perceived Influence Positively Mediates Self-Protest Intent

To causally examine whether the ATEs on self-protest intent are brought about via changes in expectations of others' protest intent, I adopt the causal mediation model in Imai et al. (2010). Causal mediation analysis first models the mediator as a function of the treatment, then models the outcome as a function of the treatment and the previously modeled mediator, and finally inputs these results into a nonparametric bootstrap algorithm for simulation. As such, the results of the two models separate the treatment effects into two components: the Average Causal

⁹ See Appendix C for the tabulated results.

Mediation Effect (ACME) through the mediator (expectation of others' protest intent) and the Average Direct Effect (ADE), which operates independently of the mediator. Table 2 shows the results of the mediation analysis with the standard errors of the effects in brackets.

**Table 2: Mediation analysis: OOLs' message effect on self-protest intent
with the expectation of others' protest intent as the mediator**

Treatment Profile (vs. Control)	Lifestyle Influencer - Low Influence metrics	Lifestyle Influencer – High Influence metrics	Investigative Journalist - Low Influence metrics	Investigative Journalist – High Influence metrics
ACME	0.180*** (0.049)	0.239*** (0.046)	0.254*** (0.049)	0.287*** (0.049)
ADE	-0.040 (0.061)	-0.026 (0.062)	0.037 (0.061)	0.057 (0.059)
Total Effect	0.141· (0.077)	0.213** (0.078)	0.291*** (0.078)	0.344*** (0.075)

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

As shown in Table 2, all treatments have a positive statistically significant ACME that, for each treatment group, is substantively much larger than the ADE. These results suggest that the treatment effects on self-protest intent are mediated by expectations of others' protest intent, thereby supporting **H3**. Moreover, all the ADEs are close to zero and statistically insignificant, suggesting that the total effect is almost entirely mediated by expectations of others' protest behavior. This evidence indicates that digital elites do not mobilize individuals by directly shifting their hearts and minds. Instead, online influencers mobilize by altering citizens' strategic assessment of the social landscape—providing the structural reassurance of safety in numbers and signaling that the threshold for low-risk, coordinated group behavior has been met, even at the earlier stage of mobilization. The robustness of the mediation analysis against omitted confounders between the treatment and the mediator is proven in Appendix D.¹⁰

¹⁰ See Appendix D1 for the results of sensitivity test.

Discussion and Conclusion

This study demonstrates that online opinion leaders (OOLs) serve as critical structural nodes for mobilizing collective action within repressive authoritarian environments. By introducing a pre-registered survey experiment within the context of anti-waste incineration grievances in China, this paper maps the psychological and strategic pathways through which citizens navigate the risks of public dissent. The empirical evidence confirms that exposure to protest-sympathetic messages from OOLs increases individuals' self-protest intent. More importantly, mediation analysis suggests their message works mainly by changing respondents' expectations about others' protest intent, rather than through direct persuasion.

These findings advance our understanding of how digital elites counteract the structural barriers that authoritarian regimes have traditionally engineered to suppress mass mobilization. In repressive settings, collective action operates as a strategic game of coordination (Granovetter, 1978; Olson, 1971): an individual's utility from participating increases with the expected participation of the crowd (Chwe, 2013; Edmond, 2013), which in turn lowers individual risk and dilutes state retaliation. In these restrictive environments, authoritarian regimes maintain a near-monopoly over information (Hassan et al., 2022), even in the digital era (Kendall-Taylor et al., 2020). This control fosters widespread preference falsification (Kuran, 1991), leaving citizens isolated and with few cues to determine whether their anti-regime grievances are private frustrations or widely shared public sentiments (Lohmann, 1994). Yet OOLs may be able to break through this informational monopoly by using the internet—even if only briefly before censorship intervenes—to publicly signal support or sympathy for collective action, sometimes through coded language such as a “group walk.” As such, potential dissenters observe the OOL's potential digital reach and persuasiveness, inferring that large numbers of their peers are simultaneously digesting the same mobilizing signal and likely to be moved, and subsequently project that the wider public is ready to move.

In this regard, OOLs served as the vanguard of collective action (De Mesquita, 2010), even without being the organizers. Such endorsements nonetheless entail risks, shaped by the perceived threat by the regime, the realized outcomes of collective action, and the prevailing political opportunity structure. Given the current political situation in China, it is unlikely that

OOLs would openly endorse collective action targeting the central government. Meanwhile, scholars also observe increasing repression (Gallagher & Miller, 2021; Yan & Li, 2024) and co-optation of digital elites (Chen & Gao, 2023; Y. Lu, 2026). Since Xi Jinping came to power in 2012, many public intellectuals have been purged and are in exile. BigVs are now also required to publicly disclose their identities once they reach a certain follower threshold, prompting many of them to limit their reach to stay below this threshold. Celebrities also face social pressure to forward content from government social media accounts (Chen & Gao, 2023). OOLs in China are increasingly targeted because of their capacity to unleash information cascades that might not only undermine CCP legitimacy but also encourage offline collective action. This study clarifies this logic by demonstrating that OOLs do not directly persuade citizens to protest. Instead, they shape potential participants' beliefs about the likelihood of mass mobilization and, in doing so, help explain why the CCP shifts its informational strategies from monitoring and moderating online content (King et al., 2013, 2017) to targeting influential OOLs (Gallagher & Miller, 2021; Yan & Li, 2024).

While this study provides clear evidence that OOLs influence social mobilization by altering beliefs about mass participation rather than by persuasion, several boundary conditions warrant further investigation. First, the analysis focuses on an environmental grievance—a policy domain that the CCP has partially tolerated and co-opted under the state rhetoric of “ecological civilization” (Arantes, 2024; Lian & Li, 2024). Future research should assess whether this mediation mechanism holds when OOLs endorse more sensitive, explicitly anti-regime issues, where the baseline risk of state repression is higher. Second, the survey design relies on a hypothetical protest scenario. Although this approach mitigates political sensitivity and maintains a degree of ecological validity within an online setting, it leaves unresolved how shifts in beliefs translate into actual participation when real-time censorship and physical repression disrupt the information environment.

Reference

- Ahler, D. J. (2014). Self-Fulfilling Misperceptions of Public Polarization. *The Journal of Politics*, 76(3), 607–620. <https://doi.org/10.1017/S0022381614000085>
- Albrecht, S. (2006). Whose voice is heard in online deliberation?: A study of participation and representation in political debates on the internet. *Information, Communication & Society*, 9(1), 62–82. <https://doi.org/10.1080/13691180500519548>
- Alexandre, I., Jai-sung Yoo, J., & Murthy, D. (2022). Make Tweets Great Again: Who Are Opinion Leaders, and What Did They Tweet About Donald Trump? *Social Science Computer Review*, 40(6), 1456–1477. <https://doi.org/10.1177/08944393211008859>
- Arantes, V. (2024). Towards a Green Nationalism with Chinese Characteristics? *Journal of Contemporary China*, 33(145), 65–82. <https://doi.org/10.1080/10670564.2023.2196514>
- Bennett, W. L., Segerberg, A., & Walker, S. (2014). Organization in the crowd: Peer production in large-scale networked protests. *Information, Communication & Society*, 17(2), 232–260. <https://doi.org/10.1080/1369118X.2013.870379>
- Besalú, R., Pont-Sorribes, C., & Martí, A. (2021). Perceived Credibility of Tweets by Opinion Leaders During the COVID-19 Pandemic in Spain. *International Journal of Communication*, 15(0), Article 0.
- Breuer, A., Landman, T., & Farquhar, D. (2015). Social media and protest mobilization: Evidence from the Tunisian revolution. *Democratization*, 22(4), 764–792. <https://doi.org/10.1080/13510347.2014.885505>
- Brunner, E. (2019). *Environmental Activism, Social Media, and Protest in China*. Bloomsbury Publishing.
- Buckley, N., Marquardt, K. L., Reuter, O. J., & Tertychnaya, K. (2024). Endogenous Popularity: How Perceptions of Support Affect the Popularity of Authoritarian Regimes. *American Political Science Review*, 118(2), 1046–1052. <https://doi.org/10.1017/S0003055423000618>
- Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2017). Curriculum and Ideology. *Journal of Political Economy*, 125(2), 338–392. <https://doi.org/10.1086/690951>
- Chaudoin, S., Gaines, B. J., & Livny, A. (2021). Survey Design, Order Effects, and Causal Mediation Analysis. *The Journal of Politics*, 83(4), 1851–1856. <https://doi.org/10.1086/715166>
- Chen, D., & Gao, G. (2023). Chinese Celebrities' Political Signalling on Sina Weibo. *The China Quarterly*, 254, 466–483. <https://doi.org/10.1017/S0305741022001734>
- Cheng, Z., Chen, J., Peng, R. X., & Shoenberger, H. (2024). Social media influencers talk about politics: Investigating the role of source factors and PSR in Gen-Z followers' perceived information quality, receptivity and sharing intention. *Journal of Information Technology & Politics*, 21(2), 117–131. <https://doi.org/10.1080/19331681.2023.2173700>
- Cho, H., Shen, L., & Peng, L. (2021). Examining and extending the influence of presumed influence hypothesis in social media. *Media Psychology*, 24(3), 413–435. <https://doi.org/10.1080/15213269.2020.1729812>
- Chwe, M. S.-Y. (2013). *Rational ritual: Culture, coordination, and common knowledge* (Reissue with a new afterword by the author). Princeton University Press.
- Davenport, C. (2007). State Repression and Political Order. *Annual Review of Political Science*, 10(1), 1–23. <https://doi.org/10.1146/annurev.polisci.10.101405.143216>
- De Mesquita, E. B. (2010). Regime Change and Revolutionary Entrepreneurs. *The American Political Science Review*, 104(3), 446–466.

- Dekoninck, H., & Schmuck, D. (2024). The “greenfluence”: Following environmental influencers, parasocial relationships, and youth’s participation behavior. *New Media & Society*, 26(11), 6615–6635. <https://doi.org/10.1177/14614448231156131>
- Diamond, L. (2010). Liberation Technology. *Journal of Democracy*, 21(3), 69–83.
- Dubois, E., & Gaffney, D. (2014). The Multiple Facets of Influence: Identifying Political Influentials and Opinion Leaders on Twitter. *American Behavioral Scientist*, 58(10), 1260–1277. <https://doi.org/10.1177/0002764214527088>
- Earl, J. (2003). Tanks, Tear Gas, and Taxes: Toward a Theory of Movement Repression. *Sociological Theory*, 21(1), 44–68.
- Edmond, C. (2013). Information Manipulation, Coordination, and Regime Change*. *The Review of Economic Studies*, 80(4), 1422–1458. <https://doi.org/10.1093/restud/rdt020>
- Edney, K. (2012). Soft Power and the Chinese Propaganda System. *Journal of Contemporary China*, 21(78), 899–914. <https://doi.org/10.1080/10670564.2012.701031>
- Enikolopov, R., Makarin, A., & Petrova, M. (2020). Social Media and Protest Participation: Evidence From Russia. *Econometrica*, 88(4), 1479–1514. <https://doi.org/10.3982/ECTA14281>
- Feng, J., Yang, S., & Hong, J. (2020). Social Media and Online Opinion Leaders in China. In J. Hong (Ed.), *China in the era of social media: An unprecedented force for an unprecedented social change* (pp. 149–170). Lexington Books. <https://doi.org/10.5040/9781666987270>
- Fu, K. (2023). Digital mobilization via attention building: The logic of cross-boundary actions in the 2019 Hong Kong social movement. *The Information Society*, 39(3), 158–170. <https://doi.org/10.1080/01972243.2023.2185717>
- Gallagher, M., & Miller, B. (2021). Who Not What: The Logic of China’s Information Control Strategy. *The China Quarterly*, 248(1), 1011–1036. <https://doi.org/10.1017/S0305741021000345>
- Gerbaudo, P. (2016). Constructing Public Space| Rousing the Facebook Crowd: Digital Enthusiasm and Emotional Contagion in the 2011 Protests in Egypt and Spain. *International Journal of Communication*, 10(0), Article 0.
- Goodwin, A., Joseff, K., Riedl, M. J., Lukito, J., & Woolley, S. (2023). Political Relational Influencers: The Mobilization of Social Media Influencers in the Political Arena. *International Journal of Communication*, 17, Article 0.
- Granovetter, M. (1978). Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6), 1420–1443. <https://doi.org/10.1086/226707>
- Gunther, A. C. (1998). The Persuasive Press Inference: Effects of Mass Media on Perceived Public Opinion. *Communication Research*, 25(5), 486–504. <https://doi.org/10.1177/009365098025005002>
- Gunther, A. C., Bolt, D., Borzekowski, D. L. G., Liebhart, J. L., & Dillard, J. P. (2006). Presumed Influence on Peer Norms: How Mass Media Indirectly Affect Adolescent Smoking. *Journal of Communication*, 56(1), 52–68. <https://doi.org/10.1111/j.1460-2466.2006.00002.x>
- Gunther, A. C., & Storey, J. D. (2003). The Influence of Presumed Influence. *Journal of Communication*, 53(2), 199–215. <https://doi.org/10.1111/j.1460-2466.2003.tb02586.x>
- Han, R. (2015a). Defending the Authoritarian Regime Online: China’s “Voluntary Fifty-cent Army.” *The China Quarterly*, 224, 1006–1025. <https://doi.org/10.1017/S0305741015001216>

- Han, R. (2015b). Manufacturing Consent in Cyberspace: China's "Fifty-Cent Army." *Journal of Current Chinese Affairs*, 44(2), 105–134. <https://doi.org/10.1177/186810261504400205>
- Harff, D., & Schmuck, D. (2023). Influencers as Empowering Agents? Following Political Influencers, Internal Political Efficacy and Participation among Youth. *Political Communication*, 40(2), 147–172. <https://doi.org/10.1080/10584609.2023.2166631>
- Hassan, M., Mattingly, D., & Nugent, E. R. (2022). Political Control. *Annual Review of Political Science*, 25(1), 155–174. <https://doi.org/10.1146/annurev-polisci-051120-013321>
- Huang, H., & Cruz, N. (2022). Propaganda, Presumed Influence, and Collective Protest. *Political Behavior*, 44(4), 1789–1812. <https://doi.org/10.1007/s11109-021-09683-0>
- Huffaker, D. (2010). Dimensions of Leadership and Social Influence in Online Communities. *Human Communication Research*, 36(4), 593–617. <https://doi.org/10.1111/j.1468-2958.2010.01390.x>
- Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4), 309–334. <https://doi.org/10.1037/a0020761>
- Jost, J. T., Barberá, P., Bonneau, R., Langer, M., Metzger, M., Nagler, J., Sterling, J., & Tucker, J. A. (2018). How Social Media Facilitates Political Protest: Information, Motivation, and Social Networks. *Political Psychology*, 39(S1), 85–118. <https://doi.org/10.1111/pops.12478>
- Juris, J. S. (2005). The New Digital Media and Activist Networking within Anti-Corporate Globalization Movements. *The ANNALS of the American Academy of Political and Social Science*, 597(1), 189–208. <https://doi.org/10.1177/0002716204270338>
- Kendall-Taylor, A., Frantz, E., & Wright, J. (2020). The Digital Dictators. *Foreign Affairs*, 99(2), 103–115. JSTOR.
- 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. <https://doi.org/10.1017/S0003055413000014>
- 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. <https://doi.org/10.1017/S0003055417000144>
- Kuran, T. (1991). Now Out of Never: The Element of Surprise in the East European Revolution of 1989. *World Politics*, 44(1), 7–48. <https://doi.org/10.2307/2010422>
- Kuznetsova, D., & Tolbert, C. (2024). Modelling temporal dynamics: Does internet use fuel anti-government protests? *Democratization*, 31(2), 389–410. <https://doi.org/10.1080/13510347.2023.2268019>
- Lian, C., & Li, J. (2024). Legitimacy-seeking: China's statements and actions on combating climate change. *Third World Quarterly*, 45(1), 171–188. <https://doi.org/10.1080/01436597.2023.2216135>
- Lim, M. (2013). Framing Bouazizi: 'White lies', hybrid network, and collective/connective action in the 2010–11 Tunisian uprising. *Journalism*, 14(7), 921–941. <https://doi.org/10.1177/1464884913478359>
- Little, A. T. (2016). Communication Technology and Protest. *The Journal of Politics*, 78(1), 152–166. <https://doi.org/10.1086/683187>
- Liu, J., & Fong, C. S. (forthcoming). Discursive channelling: Comparing Chinese newspapers' responses to anti-waste incineration protests under Hu Jintao and Xi Jinping. *Chinese Journal of Communication*.

- Liu, N., & Liu, J. (2023). Leading with hearts and minds: Emotion contagion in China's online activism. *Social Movement Studies*, 22(5–6), 728–750.
<https://doi.org/10.1080/14742837.2021.2011716>
- Lo, V., Kuo, L., Wei, R., & Li, Z. (2022). Anger Yes, Boycott No: Third-Person Effects and the China–U.S. Trade War. *International Journal of Communication*, 16, 22–22.
- Lohmann, S. (1994). The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91. *World Politics*, 47(1), 42–101.
<https://doi.org/10.2307/2950679>
- Lorentzen, P. (2014). China's Strategic Censorship. *American Journal of Political Science*, 58(2), 402–414. <https://doi.org/10.1111/ajps.12065>
- Lu, J., & Steinhardt, H. C. (2022). Alliance Building among Environmental Nongovernmental Organizations in China: The Emergence and Evolution of the Zero Waste Alliance. *Modern China*, 48(1), 105–133. <https://doi.org/10.1177/0097700420956250>
- Lu, Y. (2026). Performative Propaganda Engagement: How Celebrity Fans Engage with State Propaganda on Weibo. *Political Communication*, 43(1), 99–127.
<https://doi.org/10.1080/10584609.2025.2584999>
- Maddox, J. (2023). Micro-celebrities of information: Mapping calibrated expertise and knowledge influencers among social media veterinarians. *Information, Communication & Society*, 26(14), 2726–2752. <https://doi.org/10.1080/1369118X.2022.2109980>
- Mattingly, D. C., & Yao, E. (2022). How Soft Propaganda Persuades. *Comparative Political Studies*, 55(9), 1569–1594. <https://doi.org/10.1177/00104140211047403>
- McCarthy, J. D., & Zald, M. N. (1977). Resource Mobilization and Social Movements: A Partial Theory. *American Journal of Sociology*, 82(6), 1212–1241.
<https://doi.org/10.1086/226464>
- Mildenberger, M., & Tingley, D. (2019). Beliefs about Climate Beliefs: The Importance of Second-Order Opinions for Climate Politics. *British Journal of Political Science*, 49(4), 1279–1307. <https://doi.org/10.1017/S0007123417000321>
- Nip, J. Y. M., & Fu, K. (2016). Challenging Official Propaganda? Public Opinion Leaders on Sina Weibo. *The China Quarterly*, 225, 122–144.
<https://doi.org/10.1017/S0305741015001654>
- Olson, M. (1971). *The logic of collective action: Public goods and the theory of groups* (21. printing). Harvard Univ. Press.
- Ong, L. H., & Luo, K. W. (2024). Stability Maintenance, Preventive Repression, and Contentious Politics in China under Xi Jinping's Rule. *Issues & Studies*, 60(04), 2450015. <https://doi.org/10.1142/S1013251124500152>
- Pearlman, W. (2013). Emotions and the Microfoundations of the Arab Uprisings. *Perspectives on Politics*, 11(2), 387–409. <https://doi.org/10.1017/S1537592713001072>
- Poell, T., Abdulla, R., Rieder, B., Woltering, R., & Zack, L. (2016). Protest leadership in the age of social media. *Information, Communication & Society*, 19(7), 994–1014.
<https://doi.org/10.1080/1369118X.2015.1088049>
- Qin, B., Strömberg, D., & Wu, Y. (2024). Social Media and Collective Action in China. *Econometrica*, 92(6), 1993–2026. <https://doi.org/10.3982/ECTA20146>
- Riedl, M. J., Lukito, J., & Woolley, S. C. (2023). Political Influencers on Social Media: An Introduction. *Social Media + Society*, 9(2), 20563051231177938.
<https://doi.org/10.1177/20563051231177938>

- Rothut, S. (2025). Promoting Politics: Political Social Media Influencers, Their Online Engagement, and Implications for Democracy. *American Behavioral Scientist*, 00027642251344206. <https://doi.org/10.1177/00027642251344206>
- 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. <https://doi.org/10.1086/703208>
- Schmuck, D., Hirsch, M., Stevic, A., & Matthes, J. (2022). Politics – Simply Explained? How Influencers Affect Youth's Perceived Simplification of Politics, Political Cynicism, and Political Interest. *The International Journal of Press/Politics*, 27(3), 738–762. <https://doi.org/10.1177/19401612221088987>
- Steinert-Threlkeld, Z. C. (2017). Spontaneous Collective Action: Peripheral Mobilization During the Arab Spring. *American Political Science Review*, 111(2), 379–403. <https://doi.org/10.1017/S0003055416000769>
- Stockmann, D., & Gallagher, M. E. (2011). Remote Control: How the Media Sustain Authoritarian Rule in China. *Comparative Political Studies*, 44(4), 436–467. <https://doi.org/10.1177/0010414010394773>
- Svensson, M. (2017). The rise and fall of investigative journalism in China: Digital opportunities and political challenges. *Media, Culture & Society*, 39(3), 440–445. <https://doi.org/10.1177/0163443717690820>
- Svolik, M. W. (2012). *The politics of authoritarian rule*. Cambridge University Press.
- Todorov, A., & Mandisodza, A. N. (2004). Public Opinion on Foreign Policy: The Multilateral Public that Perceives Itself as Unilateral. *Public Opinion Quarterly*, 68(3), 323–348. <https://doi.org/10.1093/poq/nfh036>
- Toepler, S., Zimmer, A., Fröhlich, C., & Obuch, K. (2020). The Changing Space for NGOs: Civil Society in Authoritarian and Hybrid Regimes. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 31(4), 649–662. <https://doi.org/10.1007/s11266-020-00240-7>
- Treem, J. W., Leonardi, P. M., & van den Hooff, B. (2020). Computer-Mediated Communication in the Age of Communication Visibility. *Journal of Computer-Mediated Communication*, 25(1), 44–59. <https://doi.org/10.1093/jcmc/zmz024>
- Urman, A., Ho, J. C., & Katz, S. (2021). Analyzing protest mobilization on Telegram: The case of 2019 Anti-Extradition Bill movement in Hong Kong. *PLOS ONE*, 16(10), e0256675. <https://doi.org/10.1371/journal.pone.0256675>
- van Stekelenburg, J., & Klandermans, B. (2013). The social psychology of protest. *Current Sociology*, 61(5–6), 886–905. <https://doi.org/10.1177/0011392113479314>
- Wei, R., Lo, V., & Golan, G. (2017). Examining the Relationship Between Presumed Influence of U.S. News About China and the Support for the Chinese Government's Global Public Relations Campaigns. *International Journal of Communication*, 11, 18–18.
- Weidmann, N. B., & Rød, E. G. (2019). *The Internet and political protest in autocracies*. Oxford University Press.
- Wong, N. W. M. (2016). Advocacy Coalitions and Policy Change in China: A Case Study of Anti-incinerator Protest in Guangzhou. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 27(5), 2037–2054. <https://doi.org/10.1007/s11266-015-9603-1>
- Woolley, S. C. (2022). Digital Propaganda: The Power of Influencers. *Journal of Democracy*, 33(3), 115–129. <https://doi.org/10.1353/jod.2022.0027>

- Wu, F., & Martus, E. (2021). Contested environmentalism: The politics of waste in China and Russia. *Environmental Politics*, 30(4), 493–512. <https://doi.org/10.1080/09644016.2020.1816367>
- Xia, S. (2022). Amusing Ourselves to Loyalty? Entertainment, Propaganda, and Regime Resilience in China. *Political Research Quarterly*, 75(4), 1096–1112. <https://doi.org/10.1177/10659129211049389>
- Yan, X., & Li, L. (2024). Censoring the Intellectual Public Space in China: What Topics Are Not Allowed and Who Gets Blacklisted? *Perspectives on Politics*, 22(3), 753–770. <https://doi.org/10.1017/S1537592723002815>
- Yang, G. (2014). Internet Activism & the Party-State in China. *Daedalus*, 143(2), 110–123.
- Yilmaz, I., Akbarzadeh, S., Abbasov, N., & Bashirov, G. (2024). The Double-Edged Sword: Political Engagement on Social Media and Its Impact on Democracy Support in Authoritarian Regimes. *Political Research Quarterly*, 10659129241305035. <https://doi.org/10.1177/10659129241305035>
- Young, L. E. (2019). The Psychology of State Repression: Fear and Dissent Decisions in Zimbabwe. *American Political Science Review*, 113(1), 140–155. <https://doi.org/10.1017/S000305541800076X>
- Zhaivoronok, D. (2026). The Birth of Collective Action Out of Influencer Culture? *Communist and Post-Communist Studies*, 59(1), 124–140. <https://doi.org/10.1525/cpcs.2025.2685898>
- Zhang, X. (2021). Anti-Incineration Mobilization on WeChat: Evidence from 12 WeChat Subscription Accounts. *Environmental Communication*, 15(8), 1061–1076. <https://doi.org/10.1080/17524032.2021.1941176>
- Zhang, Y. (2020). BROKER AND BUFFER: WHY ENVIRONMENTAL ORGANIZATIONS PARTICIPATE IN POPULAR PROTESTS IN CHINA*. *Mobilization: An International Quarterly*, 25(1), 115–132. <https://doi.org/10.17813/1086-671X-25-1-115>
- Zhu, Y., Cheng, E. W., Shen, F., & Walker, R. M. (2022). An Eye for an Eye? An Integrated Model of Attitude Change Toward Protest Violence. *Political Communication*, 39(4), 539–563. <https://doi.org/10.1080/10584609.2022.2053915>

Appendix A: Balance of Demographic Data

Table A1 Balance of demographics data by experimental groups ¹¹						
Variable	Group 1	Group 2	Group 3	Group 4	Group 5	p-value (ANOVA)
Male	0.558	0.555	0.567	0.533	0.544	0.805
Age ≤ 25	0.204	0.228	0.207	0.207	0.183	0.473
Age 26-35	0.463	0.399	0.449	0.423	0.444	0.243
Age 36-45	0.190	0.213	0.207	0.214	0.227	0.581
Age 46-55	0.083	0.100	0.074	0.100	0.094	0.817
Age > 56	0.061	0.061	0.064	0.056	0.052	0.931
Monthly Income > 10,000 RMB	0.240	0.248	0.235	0.245	0.233	0.993
College	0.503	0.528	0.531	0.523	0.558	0.502
Urban Household Registration	0.820	0.810	0.779	0.788	0.821	0.448
Coastal	0.507	0.503	0.511	0.517	0.492	0.973
Inland	0.349	0.357	0.340	0.351	0.380	0.840
Border	0.143	0.140	0.149	0.133	0.127	0.872

Table A1 shows that all demographic variables are balanced between experimental groups; ANOVA results suggest that the differences among groups are not statistically significant.

¹¹ **Coastal:** Beijing, Tianjin, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong; **Inland:** Hebei, Shanxi, Anhui, Jiangxi, Henan, Hubei, Hunan, Hainan, Chongqing, Sichua; **Border:** Inner Mongolia, Liaoning, Jilin, Heilongjiang, Guangxi, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang.

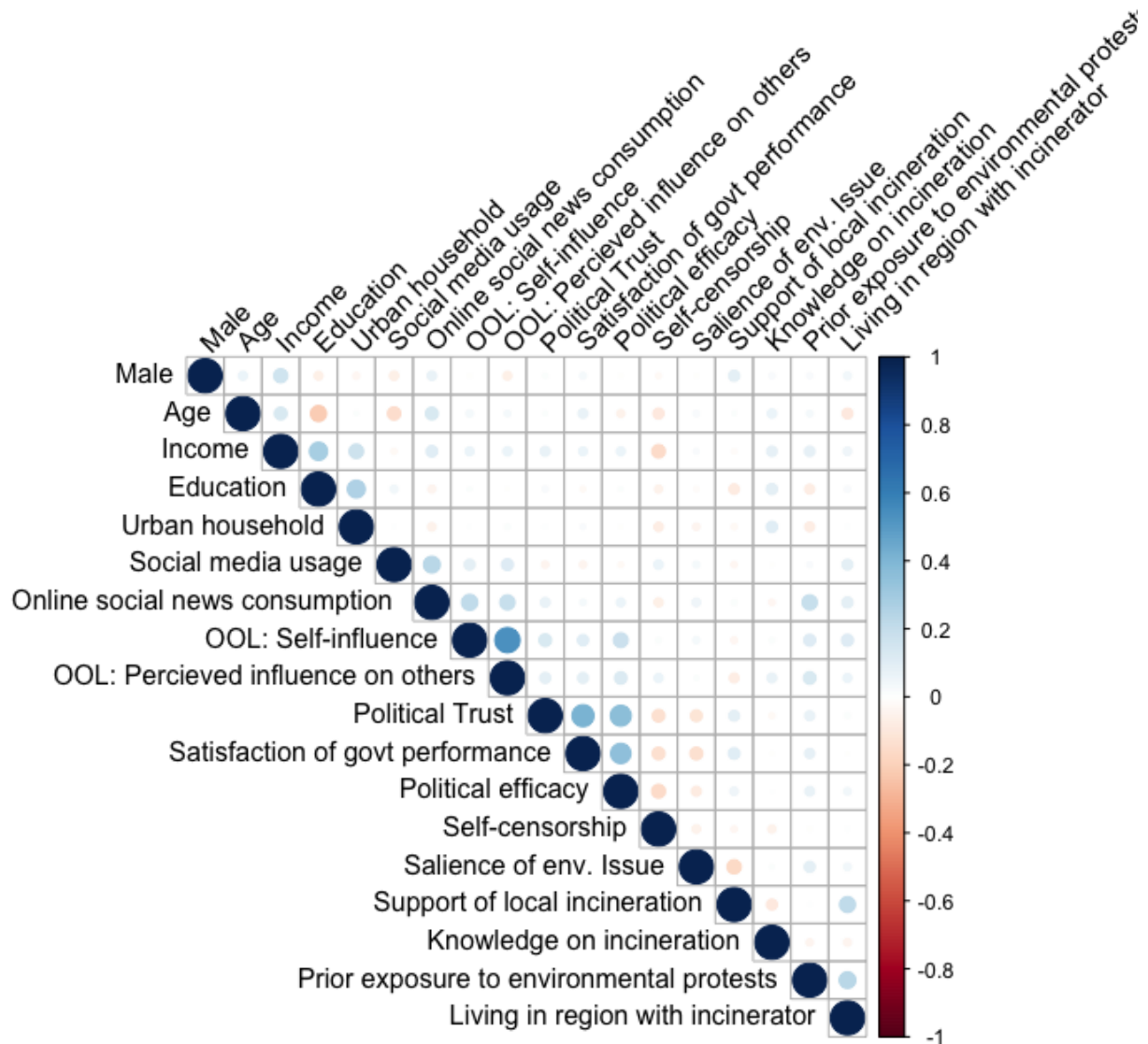
Table A2 Survey Samples vs Chinese Online Population		
Variable	Survey sample	Chinese online population aged above 18 (2024) ¹²
Male	0.551	0.511
Age ≤ 25	0.206	0.124
Age 26-35	0.436	0.231
Age 36-45	0.210	0.228
Age 46-55	0.090	0.210
Age >56	0.059	0.207
Monthly Income > 10,000 RMB	0.240	0.072
College	0.529	0.148
Urban Household Registration	0.804	0.723
Coastal	0.506	0.421
Inland	0.356	0.413
Border	0.139	0.166

While the survey respondents are significantly younger, wealthier, more highly educated, and more concentrated in coastal provinces than the actual online population, these baseline differences do not compromise the empirical validity of our findings. First, this demographic profile aligns closely with the core research focus on the more active and digitally literate users. Second, the study maintains strict internal validity and structural consistency, as there was still sufficient variation within the data to separate the effects of different variables.

¹² The Chinese online population data is sourced from the China Internet Network Information Center (CNNIC) 53rd Statistical Report (published in 2024) and the National Bureau of Statistics (NBS) 7th National Population Census (published in 2021). The report's standard 10-year age brackets were split uniformly based on census weights to match the survey sample cohort. The regional population was computed as the product of the census population and the internet penetration rate reported by CNNIC.

Appendix B: Correlation of covariates

Fig. A1 Correlation of covariates



The bivariate correlation analysis indicates that the dataset is free of critical multicollinearity. Key patterns align perfectly with theoretical expectations: income, education, and urban household registration are linked; self-influence and perceived influence of online opinion are also strongly and positively correlated; and institutional evaluations, i.e., political trust, satisfaction with government, and political efficacy, show similar patterns.

Appendix C: Regression Results

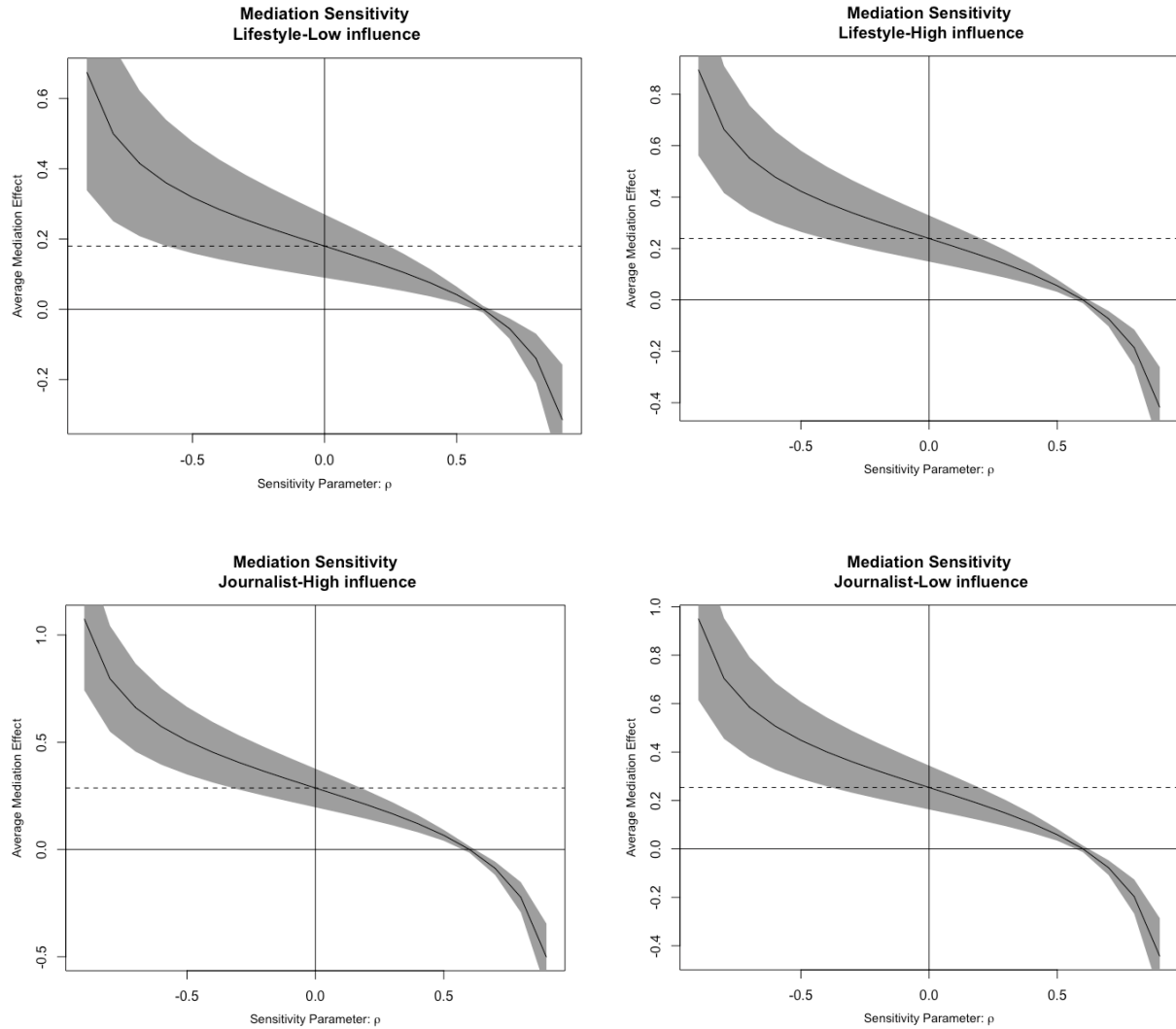
Table A3 Regression results of Fig. 2, Fig. 3, and Fig. 4						
Outcome	Self-protest Intent		Expectation on others'			
Lifestyle Influencer (Low Influence)	0.128 (0.083)	0.140+ (0.077)	0.267*** (0.071)	0.266*** (0.068)		
Lifestyle Influencer (High Influence)	0.202* (0.082)	0.213** (0.076)	0.353*** (0.070)	0.353*** (0.067)		
Journalist (Low Influence)	0.295*** (0.083)	0.291*** (0.077)	0.379*** (0.071)	0.375*** (0.068)		
Journalist (High Influence)	0.366*** (0.082)	0.344*** (0.076)	0.442*** (0.070)	0.424*** (0.067)		
Journalist					0.112 (0.070)	0.108 (0.066)
High Influence					0.0864 (0.069)	0.086 (0.066)
Journalist* High influence					-0.0237 (0.097)	-0.035 (0.093)
Covariate-adjusted	No	Yes	No	Yes	No	Yes
N	2,461	2,461	2,461	2,461	1,966	1,966

Notes: Standard errors in parentheses. + < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

Appendix D: Additional data analysis

D1 Sensitivity Test of Mediation Analysis

Fig. A2 Sensitivity Test of Mediation by treatment groups

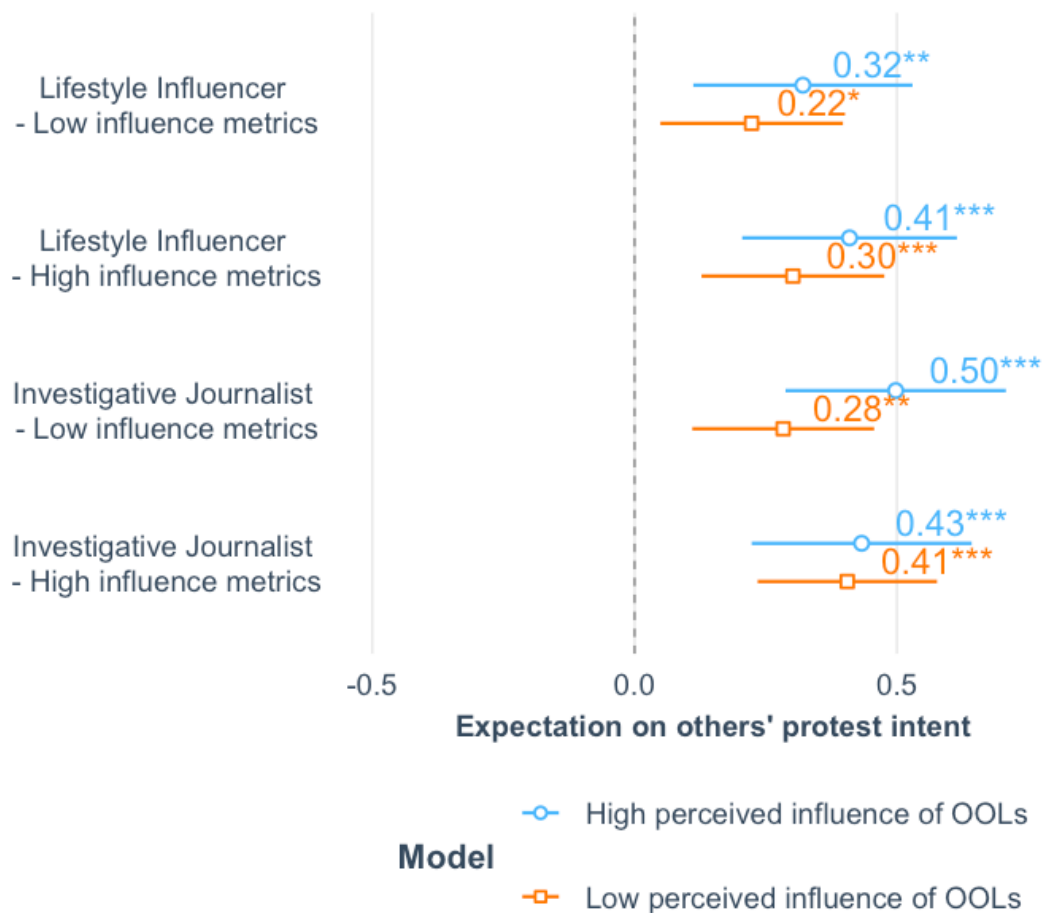


Causal mediation analysis relies on the sequential ignorability assumption, which requires both treatment assignment and the mediator (given the treatment) to be independent of the potential outcomes. In this study, I measured the outcome prior to the mediator, which introduces question-order priming. I employ the sensitivity analysis framework established by Imai et al. (2010) to test the robustness against this risk. In this framework, the parameter ρ quantifies the correlation between the error terms of the mediator and outcome regression equations due to unobserved pre-treatment confounding. Across all four treatment conditions, the

models demonstrate substantial resilience, requiring a uniform threshold of $\rho = 0.60$ to drive the Average Causal Mediation Effect (ACME) to zero. A sensitivity parameter of this magnitude indicates that unobserved confounding would need to induce an exceptionally strong residual correlation between the two equations to attenuate the effect—an unlikely scenario given our comprehensive control for foundational political and psychological baselines (including political trust, government performance, political efficacy, self-censorship, and baseline project support).

D2 Heterogeneous Effects

Fig. A3 Average Treatment Effects on Expectation on others' protest intent by perceived influence of online opinion leaders (OOLs)



Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix E Codebook

Part I: Pre-treatment questions on socio-political attitude and behaviors

1. Social Media Usage

On a typical day when you use social media (like Weibo, Xiaohongshu, Douyin), about how much total time do you spend on it?

- 0 = Low (less than 2 hours or 2 – 3 hours)
- 1 = High (3 – 4 hours or more than 4 hours)

2. Online consumption of social news

On a typical day when you use social media (like Weibo, Xiaohongshu, Douyin), how often do you follow political, economic, or social news (excluding entertainment news)?

- 0 = Low (less than half an hour or half an hour to one hour)
- 1 = High (one to one and a half hours or more than one and a half hours)

3. Views and behaviors related to online opinion leaders

Based on your usage habits of social media (such as Xiaohongshu, Weibo, and Douyin), to what extent do you agree with the following statements? – strongly agree (4); somewhat agree (3); somewhat disagree (2); strongly disagree (1)

3a) Self-influence (Items a & b)

a) *I always follow or share content from influencers*

b) *The opinion of some influencers affects my views on some social issues*

- 0 = Low (Sum score of items ≤ 6)
- 1 = High (Sum score of items > 6)

3b) Perceived influence on others (Items c & d)

c) *I think influencers have a huge influence on public opinion*

d) *My family members and friends always share or discuss content from influencers with me*

- 0 = Low (Sum score of items ≤ 6)
- 1 = High (Sum score of items > 6)

4. Political Trust

Do you think the government gives full consideration to citizens' needs?

- 0 = Low (strongly disagree or somewhat disagree)
- 1 = High (strongly agree or somewhat agree)

5. Satisfaction with government performance

Overall, are you satisfied with the local living conditions (e.g., infrastructure, public services, environmental governance, living conditions, economy and people's livelihood) over the past 5 years?

- 0 = Low (very dissatisfied or somewhat dissatisfied)
- 1 = High (very satisfied or somewhat satisfied)

6. Political Efficacy

In general, do you think ordinary citizens' participation can influence government decisions?

- 0 = Low (not at all or not very much)
- 1 = High (very much or somewhat)

7. Self-Censorship

When discussing public and social affairs, to what extent do you agree with the following statement? – (4) strongly agree; (3) somewhat agree; (3) somewhat disagree; (2) strongly disagree

a) *I worry that expressing different opinions will cause trouble*

b) *I avoid certain topics in public*

c) *I worry that others will judge me because of what I say*

d) *I choose not to express certain opinions online out of fear of the consequences*

- 0 = Low (Sum score of items ≤ 12)
- 1 = High (Sum score of items > 12)

8. Salience of environmental issues

a) How serious do you think the environmental pollution problem is in our country? - very serious (4); somewhat serious (3); not very serious (2); not serious at all (1)

b) Do you agree or disagree with the following statement? "When promoting economic development, bearing a certain level of environmental cost is unavoidable." - strongly agree (1); somewhat agree (2); somewhat disagree (3); strongly disagree (4)

- 0 = Low (Combined sum score ≤ 4)
- 1 = High (Combined sum score > 4)

9. Support for local waste incineration projects

Do you support or oppose the government building a new waste incineration plant near where you live?

- 0 = Low (strongly oppose or somewhat oppose)
- 1 = High (strongly support or somewhat support)

10. Knowledge on Waste Incineration

What is the most concerning emission in the waste incineration process?

- 0 = Incorrect (Carbon Dioxide, Nitrogen or Water Vapor)
- 1 = correct (Dioxin)

11. Attention Check

What is your biggest hobby? We are not concerned with your specific hobby; we simply want to ensure that you have read the questionnaire carefully. Therefore, please choose "Reading the instructions" as your answer.

- 0 = Failed (Exercising; watching movies; gardening; listening to music)
- 1 = Passed (Reading the instructions)

Part II Outcome questions

12. Outcome variable: Self-offline protest intent

Assuming you live in Tongyuan City, how likely are you to participate in the collective activities of Tongyuan citizens opposing waste incineration? (6) extremely likely; (5) very likely; (4) somewhat likely; (3) somewhat unlikely; (2) very unlikely; (1) extremely unlikely

13. Outcome variable: Others' protest intent

Assuming this Weibo post circulates online, how likely do you think an average Tongyuan citizens would participate in collective activities against waste incineration? (6) extremely likely; (5) very likely; (4) somewhat likely; (3) somewhat unlikely; (2) very unlikely; (1) extremely unlikely

Part II Demographics questions

14. Gender

What is your gender

- 0 = Female
- 1 = Male

15. Age

In which year were you born? Please enter a four-digit number. For example, if you were born in 1999, please enter 1999”.

- 1 = 25 years old or under
- 2 = 26 – 35 years old
- 3 = 36 – 45 years old
- 4 = 46 – 55 years old
- 5 = 56 years old or older

16. Education

Which of the following best describes your highest level of education?

- 0 = No College (elementary school or below, junior high school, high school or equivalent, or vocational college or equivalent)
- 1 = College (bachelor's degree or graduate degree)

17. Residence

In which province or municipality do you reside?

- **Coastal** = Beijing, Tianjin, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong
- **Inland** = Hebei, Shanxi, Anhui, Jiangxi, Henan, Hubei, Hunan, Hainan, Chongqing, Sichuan
- **Border** Inner Mongolia, Liaoning, Jilin, Heilongjiang, Guangxi, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang

18. Household registration

Do you currently primarily live in (towns, townships, and villages are considered rural; municipalities, prefecture-level cities, county-level cities, and county seats are considered urban)?

- Rural
- Urban

19. Income

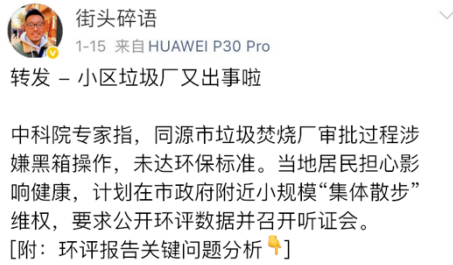
Which of the following best describes your household's monthly income?

- 0 = Low income (Less than 3k RMB, 3k – 6k RMB, or 6k – 10k RMB)
- 1 = High income (10k – 15k RMB, 1.5k – 20k RMB, or 20k RMB or above)

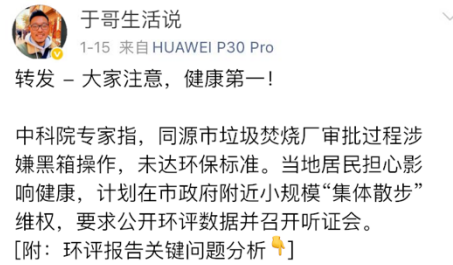
Appendix F Treatment Vignettes

Fig. A4 Treatment Vignettes

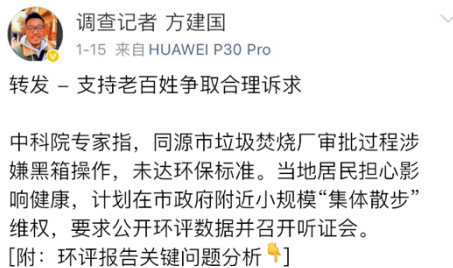
a) Treatment message: control



b) Treatment message: Lifestyle influencer



c) Treatment message: Investigative journalist



d) Treatment profile: control



e) Treatment profile: Lifestyle influencer with low influence metric



f) Treatment profile: Lifestyle influencer with high influence metric



g) Treatment profile: Investigative journalist with low influence metric



h) Treatment profile: Investigative journalist with high influence metric

