The Social Costs of Digital vs. In-person Surveillance

Xu Xu*

Abstract

The world is witnessing an explosion of digital surveillance in recent years. This paper examines the social costs of digital surveillance versus in-person surveillance in dictatorships. I argue that both types of surveillance deter political participation because citizens fear targeted repression. However, digital surveillance does not entail human-agent intrusion into private lives and therefore is less likely to undermine interpersonal trust and regime legitimacy than in-person surveillance. I manipulate information about surveillance operations in an in-the-field survey experiment on college students in two regions of China and find supportive evidence. I further establish the external validity of the experimental findings on digital surveillance using a nationally representative survey and a natural experiment caused by the 2015 Tianjin explosion. Overall, the findings that digital surveillance deters political participation without some of the costly byproducts of in-person surveillance have board implications for surveillance in dictatorships and even in democracies.

*Postdoctoral Research Fellow, Department of Political Science and Department of Communication, Stanford University. xuxu@stanford.edu, https://xu-xu.net. For helpful comments, I would like to thank Xun Cao, Volha Charnysh, Martin Dimitrov, Matt Golder, Sona Golder, Sheena Greitens, Luke Keele, Genia Kostka, Jennifer Pan, Molly Roberts, Suzanne Scoggins, Joseph Wright, Yiqing Xu, Boliang Zhu; participants at the MPSA conference, the annual meeting of the Society for Political Methodology, and department seminars at Penn State.

"WIR SIND ÜBERALL (WE ARE EVERYWHERE)"

— Stasi, German Democratic Republic

1 Introduction

Rarely in history could any autocratic regimes surveil citizens at a scale as large as those achieved by today's digital surveillance states. For instance, China has used cutting-edge surveillance technology to monitor hundreds of millions of netizens. It has also built the world's largest video surveillance networks, some are AI-powered, to monitor citizens' everyday activities (Liu and Wang 2017). The Iranian and Syrian governments have developed sophisticated digital surveillance systems to identify and track opposition members (Gunitsky 2015). By 2015, at least 30 countries have employed digital surveillance to spy on citizens, and more than half of them are autocracies (Valentino-DeVries, Vo and Yadron 2015). This trend is rapidly evolving through the export of surveillance technology from China, Israel, the US, and the UK to less developed countries such as Uganda, Zimbabwe, Angola, Bahrain, Kazakhstan, Mozambique, Nicaragua, and Saudi Arabia, among others (Feldstein 2019). Further, the COVID-19 pandemic is ushering in a new era of digital surveillance as numerous countries are imposing surveillance tools to track individuals. Despite the omnipresence of digital surveillance in today's dictatorships and even in democracies, it remains unclear how digital surveillance shapes interpersonal trust, political participation, and regime legitimacy, as well as how it differs from in-person surveillance in terms of these social consequences.

Traditionally, autocratic regimes rely on secret police and informers to spy on citizens. East Germany's Stasi (short for *Staatssicherheitsdienst*, or State Security Service) at its peak employed over 90 thousand employees and nearly 170 thousand informers who collected vast amounts of information used to intimidate the citizens and eliminate regime opposition. Counting part-time informers, the Stasi had 1 collaborator per 80 to 160 inhabitants depending on the region (Gieseke 2014). Unlike traditional surveillance, digital surveillance does not rely on human agents to collect information directly from citizens. Instead, citizens

transmit information in an electronic form on the internet and social media that can be accessed and analyzed by automatic algorithms and machines with minimal human assistance. For example, government agencies across China have invested heavily in software to track and analyze online activities (Qin, Strömberg and Wu 2017). The difference in information collection techniques suggests that digital surveillance may have different social consequences compared with in-person surveillance.

In this paper, I argue that both types of surveillance operations deter free speech and political participation. Surveillance helps dictators selectively repress citizens who are most likely to pursue anti-regime activities. When faced with the prospect of repression, citizens under surveillance have an incentive to refrain from expression (Kuran 1991). Surveillance also deters protest coordination because, under repression threat, individuals expect a higher cost of political participation – for themselves and for others – and therefore anticipate fewer people to participate. A key feature of the Stasi's surveillance strategy was "silent" methods of repression rather than overt persecution by the police (Knabe 1999).

Despite these similarities, digital surveillance and in-person surveillance have different implications for interpersonal trust and regime legitimacy. For traditional, in-person surveillance to work, surveillance agents must penetrate citizens' close social networks to collect information and identify dissidents. In societies penetrated by human agents who gather information directly from observing and interacting with their targets, citizens are atomized from one another. They hide their true anti-regime sentiments and exhibit low levels of interpersonal trust (Blaydes 2018). In-person surveillance also undermines regime legitimacy. Legitimacy is the belief on the part of citizens that the dictates of the state are right and proper (Hechter 2009). Citizens may consider a regime less legitimate when in-person surveillance conducted by the regime fosters betrayal, sabotage, and unethical exchanges of information for personal gains. On the contrary, digital surveillance does not entail human-agent intrusion into citizens' private lives. Although some of its data layers still rely on human inputs, digital surveillance largely avoids the intervention of subjective,

self-interested human informers. Thus, digital surveillance is *less* likely to decrease trust and regime legitimacy than in-person surveillance.

I use an in-the-field survey experiment with a sample of over 500 students in two universities in North and West China to examine the social costs of digital surveillance and in-person surveillance. As students are the most active social group concerning protests and political movements, finding a deterrence effect of surveillance on a student sample implies an even stronger effect on the general population. In the experiment, respondents are randomly assigned to read information about a digital surveillance scenario, an in-person surveillance scenario, and a control scenario without surveillance.

I find that both types of surveillance negatively influence political expression and intent to protest. In addition, both types of surveillance reduce respondents' beliefs about how many other individuals would participate in "anti-regime" collective action. Moreover, causal mediation analysis indicates that both types of surveillance reduce individuals' protest intention through influencing their beliefs about others' participation in protest rather than through influencing interpersonal trust. These findings suggest that both in-person and digital surveillance aid authoritarian survival by deterring anti-regime coordination.

Moreover, I find that in-person surveillance significantly reduces interpersonal trust and regime legitimacy whereas digital surveillance has a small, negative effect on trust and a negative but statistically insignificant effect on legitimacy. These differences suggest that digital surveillance may not yield the same potential costs to the regime – in the form of decreased interpersonal trust and regime legitimacy – that in-person surveillance entails. Thus, from the perspective of regime control, digital surveillance yields many of the same benefits as in-person surveillance (decreasing citizen coordination) without some of the costly byproducts of in-person surveillance, namely decreased interpersonal trust and regime legitimacy.

To establish the external validity of the experimental findings on digital surveillance, I use the 2015 Chinese General Social Survey with a nationally representative sample of nearly 11 thousand respondents and an interrupted time series design that exploits an exogenous shock

to digital surveillance caused by the 2015 Tianjin Explosions in China. At the midnight of August 12, 2015, a series of massive explosions in the port of Tianjin killed 173 people and injured nearly a thousand (Merchant 2017). Online surveillance and censorship increased tenfold in China immediately afterward (Dou 2015). By comparing respondents surveyed just before the explosions with those surveyed just afterward, I show that the intensified digital surveillance after the explosions decreases citizens' confidence in free speech and petitioning the government but has no effect on interpersonal trust. There is also a decrease in regime legitimacy, but this is likely due to the accident itself rather than the aftermath surveillance. To further confirm that digital surveillance is the mechanism underpinning these results, I explore provincial-level differences in surveillance capacity, using the number of pilot counties for China's digital surveillance system (the Golden Shield Project, developed by the Ministry of Public Security). I show that citizens feel less secure about expressing free speech and petitioning the government in provinces with greater surveillance capacity.

This paper is the first to compare the social costs of digital surveillance with those of in-person surveillance. It contributes to a growing body of literature on information technology and authoritarian survival. Many early proponents of Internet development believed that information technologies would spread freedom and spur democratization. Yet, two decades after the advent of the digital era, we have not observed widespread authoritarian collapse. Existing studies explore how authoritarian governments use the Internet and ICT to censor and repress online expressions (King, Pan and Roberts 2013), collect information about citizen preferences (Gunitsky 2015), monitor local politicians (Qin, Strömberg and Wu 2017), distract or guide public opinion (King, Pan and Roberts 2017; Roberts 2018), and identify political opponents for targeted repression (Xu 2020). My paper contributes to the literature by emphasizing the role of digital surveillance in deterring mass mobilization. Indeed, truthful communication and beliefs in others' participation may be necessary for citizens to coordinate a successful protest (Chwe 2013; Edmond 2013). For this reason, this paper explores how digital surveillance prevents mobilization in authoritarian countries by

discouraging expression, reducing willingness to protest, and decreasing beliefs about others' willingness to participate in anti-regime collective action.

This paper also examines how surveillance influences interpersonal trust and political participation – two cornerstones of free societies and the building blocks of open access markets. Interpersonal trust is essential for facilitating trade and well-functioning market economies (Greif 1989; Algan and Cahuc 2010). Political participation, particularly as the basis for coordinating citizens' anti-government behavior, is a central part of toppling dictatorships peacefully and establishing free and democratic societies (Kendall-Taylor and Frantz 2014). Though studies of both topics are numerous (e.g., Almond and Verba 1963; Putnam, Leonardi and Nanetti 1994), few have considered surveillance, especially digital surveillance, as a determinant of trust and political participation. In addition, recent studies find that autocratic rule decreases interpersonal trust and civic engagement generations after individuals migrate to democratic countries (Xu and Jin 2018), but the exact sources of social distrust and isolation in autocracies are not well understood. This paper contributes to this literature by identifying how autocratic rule damages the fabric of society: through government in-person and digital surveillance.

Finally, this paper complements historical research on traditional, informer-based surveillance, especially the Stasi era of mass surveillance in East Germany (e.g., Bruce 2010; Gieseke 2014; Blaydes 2018). While the pernicious social consequences of in-person surveillance are well-documented, quantitative research on state surveillance is rare. This paper adds to the thin literature on the causal analysis of the consequences of in-person surveillance (e.g., Lichter, Loeffler and Siegloch 2019). More importantly, it compares the social costs of digital surveillance to those of in-person surveillance. My findings, therefore, have important implications for understanding authoritarian control in the digital age. In particular, digital surveillance, like in-person surveillance, deters political expression and protest participation, but this digital repression, unlike human informers, does so without undermining interpersonal trust and regime legitimacy.

This paper compares digital surveillance with in-person surveillance in terms of their social consequences, but the findings do not suggest that digital surveillance is a dominant strategy. Evidently, authoritarian governments can use both tools to control society. For example, on China's social media, there is a growing trend of reporting and attacking each other on the grounds that their views are anti-Party or anti-China (Yan 2019). Such in-person surveillance is encouraged as many Chinese websites and social media platforms provide shortcuts to facilitate user reporting. Such evidence implies that dictators may intentionally allow citizens to spy on each other to "divide and rule" or to substitute for costly digital surveillance in poor regions. Moreover, these two types of surveillance can serve different purposes. Digital surveillance is quicker and more comprehensive than human-agent based surveillance. It is useful for screening out suspicious opponents from a very large population. Once the targets are identified, dictators may rely on human agents and informers to collect more detailed information. Nevertheless, a thorough cost-benefit analysis of both tools is beyond the scope of this paper.

2 Surveillance and Its Social Impact

Surveillance is a common tool for information collection and political control in dictatorships (Greitens 2016). Unlike democratic leaders, dictators are inherently uninformed because citizens in dictatorships often hide their true anti-regime sentiments when faced with the prospect of state repression (Kuran 1991). To gauge public opinion for policymaking and to contain threats before they spread, dictators historically rely on human security agents and/or informers to collect information from citizens. In the digital age, information technology expands dictators' information-collection toolkit. As citizens move to online media for socializing, networking, communicating, shopping, and expressing opinions, computers and algorithms replace human agents as tools for governments to collect information from citizens. This shift from human to digital surveillance has important implications for interpersonal trust and political participation in modern dictatorships.

2.1 In-person Surveillance

Human societies have a long tradition of in-person surveillance. In BC 839, King Li of Zhou Empire in China asked his wizards to spy on the people and kill those who criticized his tyranny (Zuo 1998). The infamous Jinyiwei (Embroidered Uniform Guard) was founded in the 1360s by the Hongwu Emperor of the Ming Dynasty and served as the dynasty's secret police until the collapse of Ming in 1644. In Europe, secret police organizations emerged after the French Revolution in the 18th-century. Hitler's regime in Germany (1933–1945) utilized the Gestapo to eliminate opponents. East Germany (1945–1990) created the Stasi with unparalleled social penetration. Other dictatorships such as Iraq under Saddam Hussein, Chile under Pinochet, Peru under Fujimori, Philippine under Marcos, and North Korea under the Kims also used large secret police organizations to control society (Greitens 2016; McMillan and Zoido 2004).

In-person surveillance in dictatorships reduces interpersonal trust and regime legitimacy because it encourages betrayal, sabotage, and unethical exchanges of information for personal gains.

Societal Penetration and Betrayal. To obtain precise information about opposition groups, a traditional surveillance apparatus needs a large body of security agents and informers to penetrate citizens' social networks and private lives. For example, the Stasi, at its peak, employed nearly 90 thousand employees and over 170 thousand informers, comprising roughly 1.6 percent of the population in East Germany (Gieseke 2014). Other regimes, such as the Kim regime in North Korea and the Republic of China (under Chiang Kai-shek and Chiang Ching-kuo), surveilled their citizens with massive secret police organizations (Greitens 2016). Informers are ordinary citizens but secretly gather information within their professional and social networks by betraying the trust of friends, neighbors, colleagues, relatives, and even family members (Ash 1998). Societal penetration and betrayal thus generate widespread suspicion and a deep sense of mistrust within society (Blaydes 2018). Relating surveillance operations to the government, citizens may blame the regime for their suffering

under surveillance, thereby lowering their approval of the regime.

Subjective and Self-interested Human Agents. In-person surveillance relies on self-interested human agents and informers who use subjective assessment of other citizens' loyalty to the regime. To provide information, these self-interested informers may demand benefits from the regime, such as government jobs, opportunities to travel abroad, or monetary compensation. They may also maliciously target "innocent" people to resolve personal disputes (Kalyvas 2006). Furthermore, potential informers may be "tricked" by the government to provide others' information to clear up their own "blemishes" (Ash 1998). To gain rewards or prove their innocence, self-interested agents may misreport or sabotage their fellow citizens. The potential power abuse by agents and informers associated with in-person surveillance further foments distrust and anti-regime sentiments in society.

Unethical Exchanges. In-person surveillance often involves unethical exchanges as informers trade others' secrets to the regime in exchange for material or non-material gains. These unethical exchanges could further upset people, thereby reducing trust and regime legitimacy.

Based on the above discussion, I derive the following testable implications.

Trust hypothesis (human): In-person surveillance reduces interpersonal trust.

Legitimacy hypothesis (human): In-person surveillance reduces regime legitimacy.

2.2 Digital Surveillance

Since the advent of the information era, dictatorships increasingly adopt digital surveillance for social control. By 2015, more than fifteen authoritarian countries were conducting domestic surveillance (Valentino-DeVries, Vo and Yadron 2015). Relatively cheap spyware available on international markets has made digital surveillance tools accessible to the poorest dictatorships. Authoritarian governments use malware to spy on opposition leaders and journalists (Deibert 2017), collect metadata from social media to keep tabs on political opponents (Qin, Strömberg and Wu 2017), and employ high-resolution digital cameras and

facial recognition technologies to identify dissidents (Liu and Wang 2017). Recent advances in artificial intelligence detect suspicious movements in crowds, identify thousands of people at once, and recognize citizens who attempt to conceal their identities by wearing hats, sunglasses, or scarves to cover their faces (Intel 2017; Singh et al. 2017).

Unlike in-person surveillance, digital surveillance does not entail human intrusion into citizens' private lives and thus may be less recognizable to citizens and less likely to cause a backlash. In the digital age, citizens communicate and spend a substantial amount of time online, leaving personal digital information for governments to access and analyze. Digital surveillance relies on Internet infrastructures such as computers, software, algorithms, cameras, cables, routers, servers, and data storage centers, rather than human informers. Further, digital surveillance allows governments to monitor a large population and reach the most private part of people's lives with minimal human assistance.[1] Computers and algorithms also yield more accurate and objective data than human agents who often misreport or intentionally sabotage "loyal" citizens for personal gains. Thus, the absence of human betrayal and sabotage in the operation of digital surveillance suggests that it is less likely to reduce interpersonal trust than in-person surveillance.

Moreover, although digital surveillance can deter political participation and therefore reduce individuals' regime approval, human betrayal and sabotage associated with in-person surveillance are more direct sources of regime disapproval.[2] Digital surveillance involves less

[1]For example, digital cameras allow governments to analyze population who do not use the Internet; Data from search engines reveal people's private preferences such as pornography that they would not even tell their most close friends.

[2]Note that both digital and in-person surveillance may lead to targeted repression that reduces regime legitimacy. However, as long as individuals can hide from surveillance to avoid repression, they would not clash with the government. In addition, since digital surveillance can be disguised as public security tools, individuals might not blame the government for watching them. Indeed, a recent study finds that individuals are willing to sacrifice privacy

or little human interactions, and therefore should not reduce regime legitimacy as much as in-person surveillance does.

To sum up, we expect the following testable implications.

Trust hypothesis (digital): Digital surveillance is less likely to reduce interpersonal trust than in-person surveillance.

Legitimacy hypothesis (digital): Digital surveillance is less likely to reduce regime legitimacy than in-person surveillance.

2.3 Surveillance and Political Participation

Surveillance, no matter digital or in-person, discourages political participation because it entails preemptive, targeted repression against regime opponents (Dimitrov and Sassoon 2014; Xu 2020). In dictatorships where meaningful elections and other representative channels of political expression are often unavailable, political participation takes the forms of petitions, protests, and even violent revolts. These actions disrupt social order and may threaten autocratic survival. Dictators thus use repression as a strategy to ensure political stability and avoid revolution (Wintrobe 2000). Surveillance enables dictators to find dissidents for targeted repression, thereby discouraging citizens' anti-regime political expression and protest participation.This suggests that both types of surveillance discourage political expression and protest participation.

Expression hypothesis: Both in-person and digital surveillance deter political expression.

Protest hypothesis: Both in-person and digital surveillance deter protest participation.

To mount a protest, participants must coordinate their actions (Chwe 2013). Surveillance may deter protest coordination via two channels. First, interpersonal trust may help induce protest participation by increasing individuals' belief that protest participation will be *safe* and *worthwhile* (Benson and Rochon 2004). Surveillance lowers interpersonal trust, thereby reducing individuals' willingness to participate. Second, strategic considerations are

for security even in democracies (Dietrich and Crabtree 2019).

another important determinant of protest participation: an individual's behavior is shaped by beliefs about the participation of others (Edmond 2013). Thus, surveillance may also deter participation by influencing individuals' beliefs about how many others will participate. The following hypotheses examines these two potential channels through which surveillance deters protest participation.

Trust-protest hypothesis: Surveillance deters protest participation by lowering inter-personal trust among citizens.

Coordination-protest hypothesis: Surveillance deters protest participation by influencing individuals' belief about others' participation.

3 Experimental Design

To test the hypotheses, I pursue two strategies. First, I use an in-the-field survey experiment to compare the social consequences of in-person surveillance with those of digital surveillance. Second, I address the external validity of the experimental findings on digital surveillance by analyzing a nationally representative survey in an interrupted time-series (ITS) setting. This section explains the details of the experimental design. The ITS design comprises Section 5.

3.1 Design, Randomization, and Implementation

I conducted an in-the-field survey experiment with a sample of 539 Chinese university students in March 2019. Surveying respondents on a potentially sensitive topic in the field circumvents state censorship that may be present in China-based online survey platforms. It also helps create trust and cooperation between enumerators and respondents. More importantly, the *digital-surveillance* treatment in this study may induce respondents' self-censorship in *online* surveys. An in-the-field survey experiment avoids this problem because respondents answer questions on paper questionnaires. Students were recruited in dining halls and on main roads in universities. Online Appendix A1 discusses survey implementation and ethics in more detail.

I recruited students in two universities in North and West China to broaden sample representativeness. As Figure 1 shows, the home provinces of the student sample cover most regions in China. I choose university students because they are the most active social group in political participation. For example, in the 1960s, students initiated the anti-war movement and actively participated in the Vietnam War protests in the U.S. (Moore 1999). In 1989, hundreds of thousands of student protesters occupied the Tiananmen square to demand democracy in China, sparking large-scale student protests throughout the country (Zhao 2004). During the 2014 Hong Kong protests, students were also at the heart of the "Umbrella Revolution" protests (Cantoni et al. 2019). If we can find a deterrence effect of surveillance on political participation in the demographic group most likely to protest, we are likely to observe a stronger effect in less active demographics. Thus, examining a student sample can shed light on a much larger population.

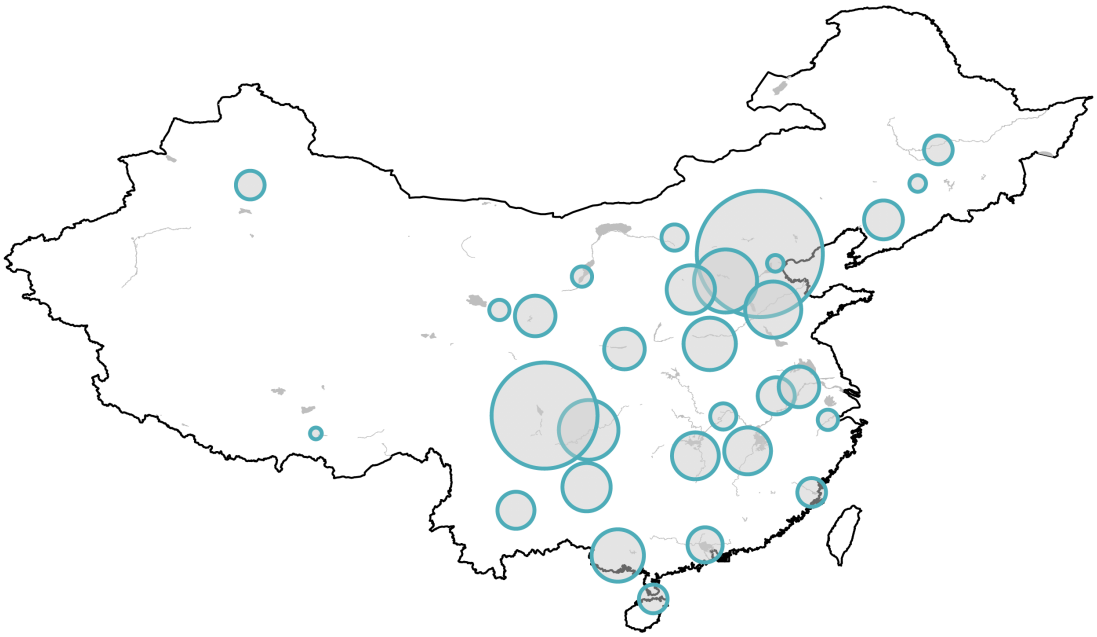


Figure 1: Distribution of Respondents by Home Province

I use a within-subject design to compare interpersonal trust, political participation, and regime support among three groups: a treated group with in-person surveillance, a treated

group with digital surveillance, and a control group without surveillance.[3] To assess possible experimenter demand effects of the within-subject design, I also embed a between-subject design for comparison. This two-by-three design yields six experimental groups. Each student has a 2/3 chance to be assigned to the within-subject design group and a 1/3 chance to be assigned to the between-subject design group.[4] Within each of the two design groups, students have an equal chance to be assigned to the control condition or one of the two treatment conditions. In total, 353 students were randomly assigned to the three experimental groups in the within-subject design, and 186 students were randomly assigned to the three experimental groups in the between-subject design.

During the survey, all respondents first answer background questions about age, gender, income, party affiliation, social distrust, general civil participation, and media usage. They then read a descriptive vignette about an issue concerning their campus life. Respondents in the within-subject design group first answer questions about willingness to make public political expressions and protest, trust toward fellow students, as well as approval of the university authority over this issue. After receiving information about the school authority's in-person surveillance, digital surveillance, or no surveillance operation, the respondents answer the same set of questions for the second time. This within-subject design allows me to differentiate out respondents' intrinsic attitudes that are difficult to manipulate in a short experiment. Alternatively, respondents in the between-subject design group directly receive information about surveillance without the pretest questions. Table 1 shows the structure

[3]A pilot study on a sample of 214 college students revealed large standard deviations in trust and participation variables. As trust and political participation are intrinsic values, it would require a very large sample to observe significant effects of manipulations from a between-subject design. On a relatively small sample, I thus implement a within-subject design that differentiates respondents' intrinsic values, following Wiswall and Zafar (2014).

[4]To balance students' gender, I use a block random assignment procedure whereby complete random assignment occurs within the male block and the female block.

of the design. Table A.3 in the Online Appendix shows that covariates are balanced across control and treatment groups.

Table 1: Experimental Design for In-person vs. Digital Surveillance

	Within-Subjects Design			Between-Subjects Design		
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Groups:	Control	Treatment 1	Treatment 2	Control	Treatment 1	Treatment 2
Background Qs:	Yes	Yes	Yes	Yes	Yes	Yes
Scenario:	Yes	Yes	Yes	Yes	Yes	Yes
Pre-test:	Yes	Yes	Yes	No	No	No
Treatment:	No Surveil.	In-person	Digital	No Surveil.	In-person	Digital
Post-test:	Yes	Yes	Yes	Yes	Yes	Yes
N of Obs.:	124	103	126	61	64	61

3.2 Treatments and Measures

Protest Scenario and Surveillance Treatments

To reduce sensitivity and protect respondents, I design a hypothetical scenario under which students confront the university authority for an unfair housing policy. In this scenario, students were forced to move from a new dormitory to an old one with bad living conditions, and the university refused to refund the price differentials. Students were discussing the means to fight for their rights, including filing a petition to the Ministry of Education. This scenario corresponds to a typical real-world confrontation between citizens and the government but with less political risk. Students' appeals to the Ministry of Education also mimic citizens' petitions to upper-level administrations for justice – a common phenomenon in authoritarian countries (Lorentzen 2013). To simulate real-world protest coordination, I remind the respondents that more participants lead to a higher chance of petition success. In addition, I made punishment upfront by reminding the respondents of the costly consequences of protest participation. This punishment reminder also helps respondents relate the surveillance scenarios to repression.

One may argue that protesting against a university authority is different from confronting a political authority. But the fact is that most universities in China are state entities, and

they often directly engage in political repression. For example, during the 2018 Jasic labor rights conflict in Shenzhen, several universities including Peking University and Renmin University were involved in cracking down on student activists who supported the movement (Yang 2019). In addition, the hypothetical scenario mentions interrogation and recording of a demerit as punishment for protest participation, which are very serious to college students. This level of threat to a student is very close to the level of a typical repression threat to a citizen. More importantly, even if protesting a university authority is less risky than protesting a political authority, the survey results concerning campus protests have general implications for political protests. This is because if surveillance can deter less risky protests, it will have even stronger deterrence effects on riskier political participation. In short, this experimental scenario is very relevant to campus life to elicit students' truthful responses, while representing a common situation of contentious politics in authoritarian countries.

In the experiment, I carefully differentiate the scenarios for in-person surveillance, digital surveillance, and no surveillance. The in-person surveillance scenario mimics the traditional, Stasi-style surveillance that relies on human informers to spy on fellow citizens. It reminds correspondents of the key features of in-person surveillance: societal penetration, betrayal, and the unethical exchanges of information for personal gains. The digital surveillance scenario is similar to real-world online surveillance conducted by authoritarian governments with no human informers involved. Because respondents' prior experience of surveillance may influence their responses even if they do not receive any new information about surveillance, I specify in the control condition that "the university does not know who participate in the protest" to "reset" respondents' prior beliefs about surveillance operations.

Measurement

In the survey, I ask respondents' willingness to express their discontent both in front of their fellow students and online. For protest participation, I elicit respondents' willingness to file an online complaint to the Ministry of Education and their beliefs regarding other students' participation (percentage points). The responses for willingness to express and

protest are recorded on a scale from 1 to 4, with 4 indicating the most affirmative answer. Interpersonal trust is measured by respondents' trust toward other students in the same residential hall on a scale from 0 to 10, with 10 indicating the highest level of trust.[5] Regime legitimacy is citizens' belief about the right and acceptance of an authority. One basis for the belief is government performance (Levi, Sacks and Tyler 2009), which is particularly important in China since satisfying people's needs for a decent livelihood has roots in Chinese traditional political culture (Perry 2008). Thus, I measure legitimacy by the extent to which respondents *generally* approve what the university authority does concerning student affairs, also on a scale from 0 to 10. See Table A.3 in Online Appendix for detailed questions concerning these outcome variables.

4 Experimental Findings

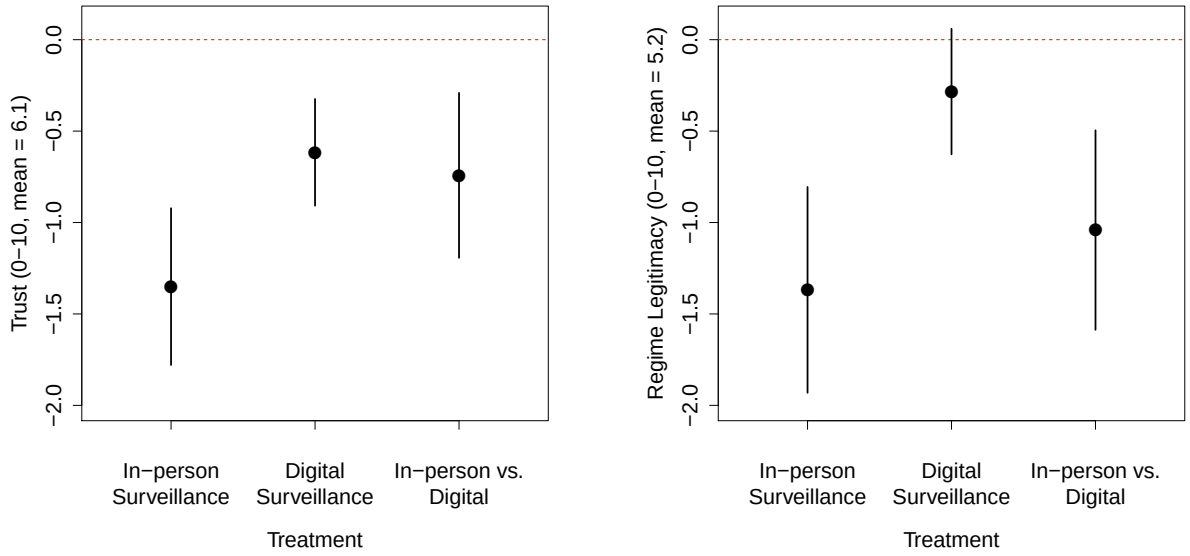
This section presents the results of the survey experiment based on the within-subject design. I take the differences between post-test answers and pre-test answers to generate the outcome variables of trust, regime legitimacy, expression, and protest participation. For each outcome, I first examine the effects of in-person surveillance and digital surveillance using no-surveillance as the comparison group. I then compare in-person surveillance with digital surveillance using the latter as the comparison group. The statistical inferences are based on standard comparisons of means using OLS estimation.[6]

[5]Scholars find that the 11-point scale consistently outperforms the dichotomous counterpart for measuring trust in surveys (Lundmark, Gilljam and Dahlberg 2015).

[6]As Table A.4 in Online Appendix shows slight differences (though statistically insignificant) in social distrust between three groups, I control for this variable in all specifications. The results are robust with the randomization inference approach. In addition, I fit Ordered Probit models for expression and protest participation. The effects are similar and statistically more significant (Table A.9 in Online Appendix). Nevertheless, I use the more conservative results from OLS models.

4.1 Trust and Regime Legitimacy

I begin by presenting the effects of in-person and digital surveillance on interpersonal trust and regime legitimacy. My theory suggests that in-person surveillance reduces trust and legitimacy, while digital surveillance is less likely to have impacts. In addition, in-person surveillance is more likely to reduce trust and legitimacy than digital surveillance. The evidence from the survey experiment is consistent with these predictions.



(a) Interpersonal Trust (b) Regime Legitimacy

(Notes: OLS estimates with 95% Confidence Intervals. See Online Appendix B.1 for the regression results underlying these figures.)

Figure 2: Trust and Legitimacy

As Figure 2a shows, in-person surveillance largely reduces interpersonal trust. Recall that interpersonal trust is measured on a scale of 0 to 10, with 10 indicating the highest level of trust. Given that the sample mean is 6.1 and the standard deviation is 2.1, a 1.4 decrease in trust scale is quite substantial (23 percent of the mean). Digital surveillance only slightly reduces trust (0.6 or 10 percent of the mean). Both effects are statistically significant at the 0.001 level. More importantly, compared with digital surveillance, in-person surveillance further reduces trust by 13 percent, and the effect is statistically significant. Figure 2b examines regime legitimacy (on a scale of 0 to 10) and shows that in-person surveillance

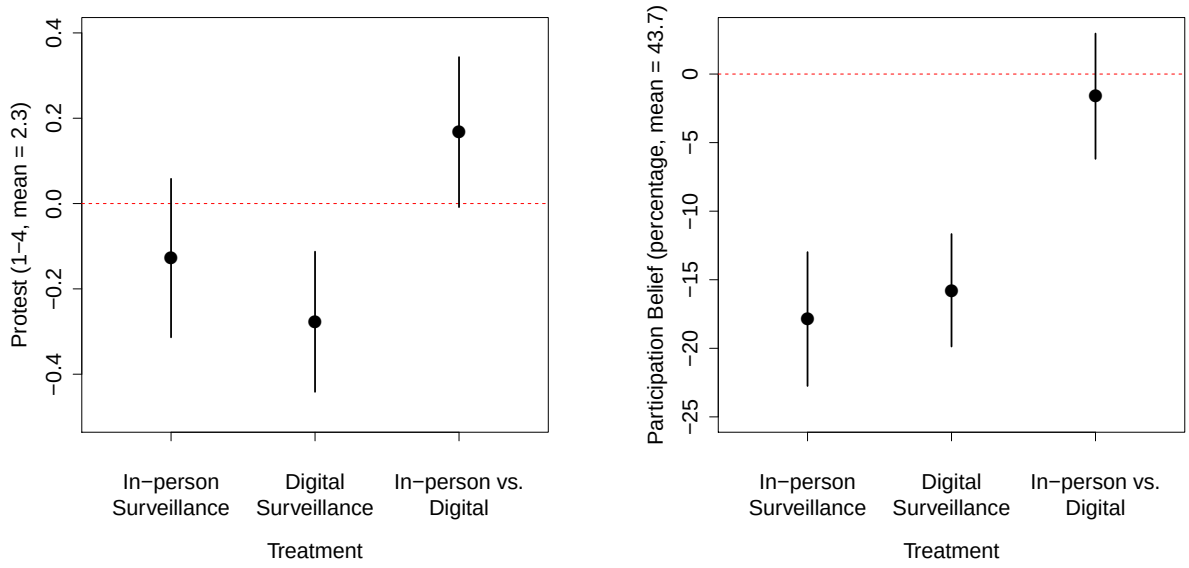
reduces legitimacy by a large margin (1.4 or 26 percent, significant at the 0.001 level) whereas digital surveillance has a negative but statistically insignificant effect. The negative effect of in-person surveillance is 19 percent larger than that of digital surveillance and is statistically significant at the 0.001 level.

4.2 Expression and Political Participation

Next, I examine how in-person and digital surveillance influence various aspects of political participation. As discussed in the theoretical part, I expect both types of surveillance to discourage expression and protest participation. Figure 3a shows that both digital surveillance and in-person surveillance deters protest participation and the effect of digital surveillance is stronger. In particular, digital surveillance decreases respondents' willingness to participate by 0.28. Given that the sample mean is 2.3 and the standard deviation is 1, the negative effect is substantial. In-person surveillance also negatively affects protest participation, but the effect is statistically insignificant in OLS models.[7] This is likely due to my framing of the protest as an online protest instead of a street protest. In theory, digital surveillance deters online protest whereas in-person surveillance may not because citizens can hide their online activities from their friends, colleagues, or family members. Nevertheless, the negative effects of both types of surveillance are consistent with the theoretical predictions.

Figure 3b presents the effects of surveillance on individuals' beliefs about others' protest participation. On average, respondents believe that 43.7 percent of students in the residence hall will participate in the protest. As we can see, in-person surveillance and digital surveillance reduce the belief by 17.9 and 15.8 percentage points respectively, which are about 41 percent and 36 percent decreases. These large, negative effects are highly significant, and there is no statistically significant difference between two types of surveillance in affecting

[7]Note that the effect is statistically significant in Ordered Probit models, see Table A.9 in Online Appendix.



(a) Protest Participation (b) Blief in Others' Participation

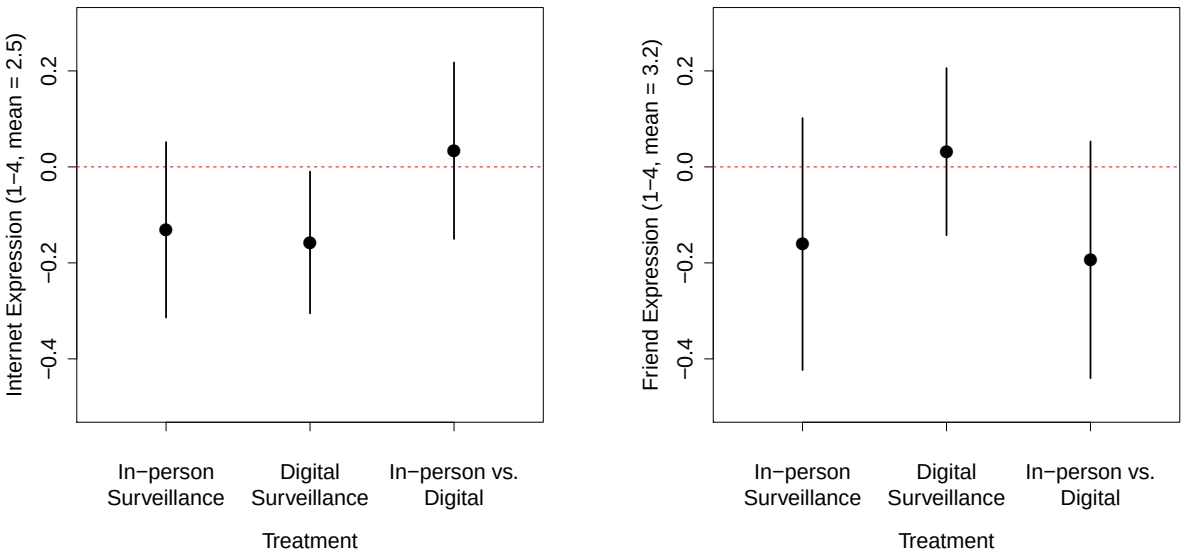
(Notes: OLS estimates with 95% Confidence Intervals. See Online Appendix B.1 for the regression tables underlying these figures.)

Figure 3: Protest Participation and Beliefs

respondents' beliefs.

Figure 4 plots the effects of digital and in-person surveillance on the online and offline expression of discontent. Both types of surveillance reduce respondents' willingness to express their discontent online and the effect of digital surveillance is statistically significant. The difference between the two types of surveillance is statistically insignificant. With regard to expression in front of fellow students, in-person surveillance has a negative but insignificant effect whereas the effect of digital surveillance is close to zero. The findings that digital surveillance discourages online expression but not offline expression meet common expectations.[8]

[8]Note that the results from Ordered Probit models (Table A.9 in Online Appendix) suggest the effects of both types of surveillance on online expression are statistically significant, and the effect of human surveillance on offline expression is also statistically significant.



(a) Online Expression (b) Expression to Fellow Students

(Notes: OLS estimates with 95% Confidence Intervals. See Online Appendix B.1 for the regression tables underlying these figures.)

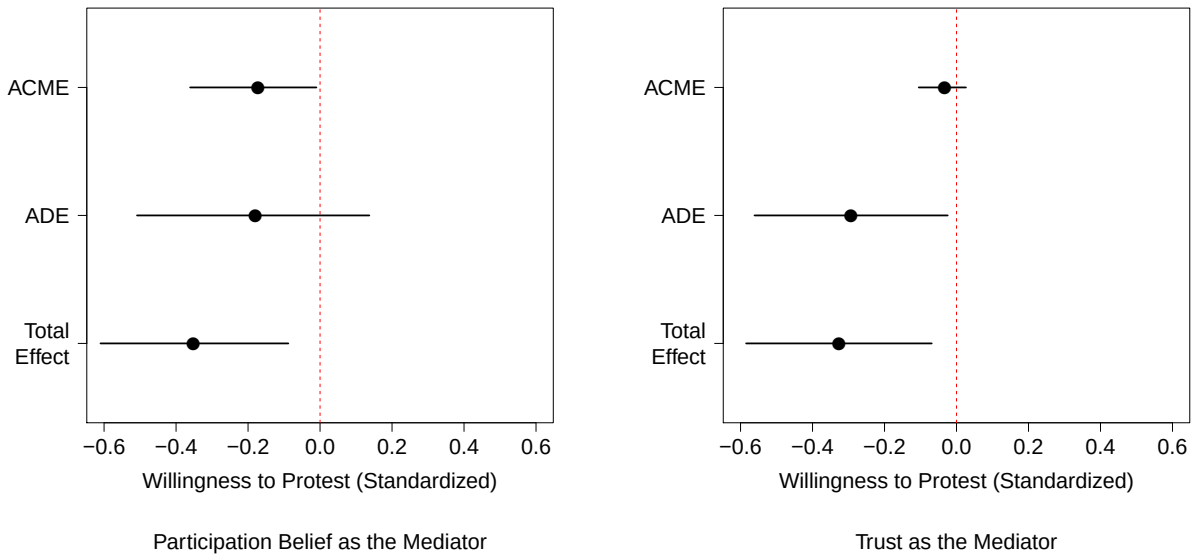
Figure 4: Online and Offline Expression

4.3 Causal Mediation Tests

I further examine the mechanisms through which surveillance deters protest participation using causal mediation analysis developed by Imai, Keele and Tingley (2010). As I mentioned in the theory section, both interpersonal trust and beliefs about others' participation could reduce individuals' willingness to protest. Thus, I use interpersonal trust and respondents' beliefs about others' turnout as mediators.

There are two major assumptions underlying the causal mediation test in identifying the mediation effect. First, there should be no unmeasured confounders between surveillance and the willingness to protest. Second, there should be no unmeasured confounders between the mediator and willingness to protest. The first assumption holds since the surveillance treatments are randomized. To address the second assumption, I control for as many covariates as possible, including age, gender, family income, income satisfaction, party affiliation, membership in official school organization, membership in student organizations, interest in discussing politics, media usage, and social distrust.

Figure 5 plots the results from the causal mediation analysis concerning digital surveillance. Individuals' belief about others' participation has a positive and statistically significant average causal mediation effect (ACME) on their own willingness to participate in the protest. In contrast, the ACME of trust is close to 0 and statistically insignificant.

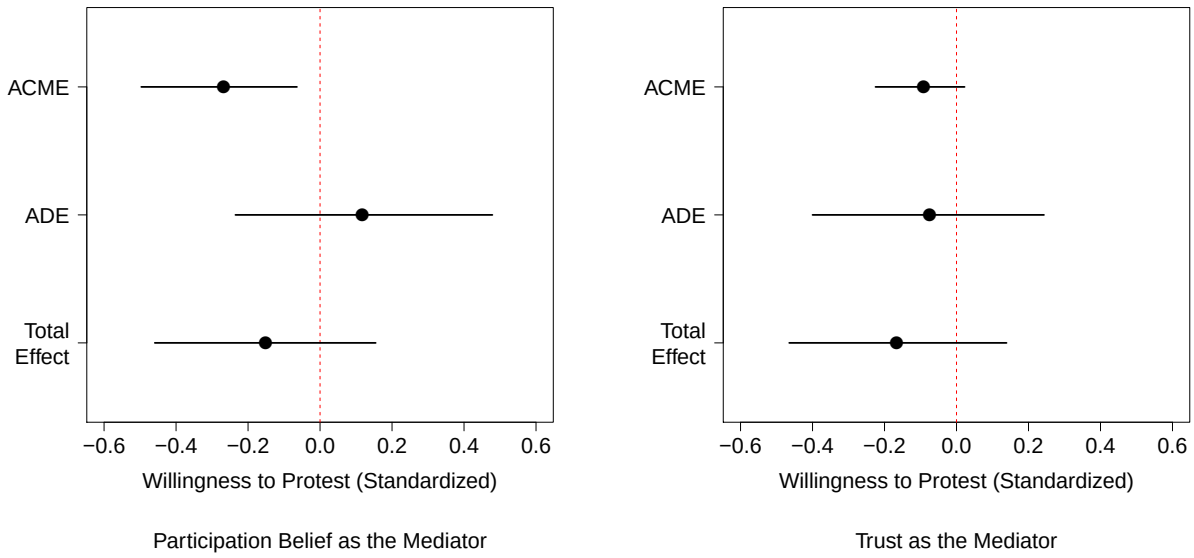


(Notes: 95% Confidence Intervals obtained via bootstrapping with 1000 resamples.)

Figure 5: Mediators between Digital Surveillance and Protest Participation

Figure 6 presents the causal mediation analysis concerning in-person surveillance. Although the total effect is not statistically significant, the average causal mediation effect of coordination beliefs on protest participation is still negative and statistically significant. The results from Figure 5 and Figure 6 suggest that surveillance discourages protest participation mainly through influencing individuals' strategic considerations about others' participation rather than their trust levels.[9]

[9]Note that belief about others' turnout is different from generalized trust. An individual may trust others but still think others would not participate under repression threat. The fact that participation belief serves as a mediator rather than trust suggests that the relationship between participation beliefs and trust is weak.



(Notes: 95% Confidence Intervals obtained via bootstrapping with 1000 resamples.)

Figure 6: Mediators between In-person Surveillance and Protest Participation

4.4 Discussion

Experimenter demand effects may bias estimates from a within-subject design because experimental subjects may tailor their responses to conform to their perceptions of the researcher's hypothesis due to answering the same question twice (before and after the treatment). However, recent literature (e.g., Mummolo and Peterson 2019) demonstrates that demand effects are typically modest and usually do not alter the treatment effects in survey and field experiments since research participants often exhibit a limited ability to infer researchers' expectations. In addition, experimenter demand effects caused by the within-subject design would likely produce similar outcomes for both the digital and the in-person surveillance treatment groups since respondents in both groups answer the repeated questions in the same way. The different findings between these two groups suggest the demand effects from the within-subject design are not very likely to be a concern. Moreover, for demand effects, if existed, to bias the results, respondents would have to be able to infer the researcher's intentions. But most of the respondents spent only about 5 minutes to complete

the survey. It would be difficult for them to infer the researcher's intentions in such a short period, especially given that respondents in one group do not know the treatment condition of the other group.

I formally assess potential demand effects by comparing post-treatment responses between the within-subjects design group and the between-subjects design group. Because demand effects are less likely to present in the between-subjects design that has no repeated questions, if there are any demand effects, we would observe systematic differences in post-treatment responses between these two design groups. Figure A.1 in Online Appendix presents the details of the comparison, which shows that the differences in regime legitimacy, online expression, and protest participation between the two design groups are statistically *insignificant*. The difference in beliefs about others' participation between the two design groups is statistically insignificant for the in-person surveillance treatment group but significant for the digital surveillance treatment group, which are inconsistent. Only does trust show systematic difference between the two design groups. But this is also likely due to the fact that asking the trust question again makes respondents think about others' trustworthiness more carefully.

Information spillovers of the treatments are also unlikely to bias the survey results. First, the survey was not conducted in classrooms and dormitories where respondents would more likely be classmates or roommates. The survey enumerators randomly approached individual students on campus roads or in dining halls. Occasionally, respondents came in groups, but the survey enumerators made sure that the respondents did not communicate during the survey. This strategy reduces the likelihood of spillovers. Practically, if there were information spilling over from the treatment groups to the control group, such spillovers would bias estimates towards 0. Moreover, information spillovers between treatment groups would lead to similarities in the effects of in-person and digital surveillance. However, we see significant differences between the two treated groups, suggesting that information spillovers are unlikely to be a concern.

Finally, one might be concerned about the statistical inferences based on standard comparisons of means using the OLS estimation. As an alternative, I use the Randomization Inference approach to examine the statistical significance of the results (Gerber and Green 2012). I randomly assign (fictional) treatment status and estimate treatment effects 1,000 times. I then calculate the p-values of the estimated treatment effects from the actual treatment assignment based on the sampling distribution of the fictional treatment assignments. Two-tailed tests find very similar p-values for the treatment effects of surveillance on trust, legitimacy, and protest participation as those from the OLS estimation. See Table A.8 in Online Appendix for details.

5 Interrupted Time Series Design

To establish the external validity of my experimental results concerning digital surveillance, I provide additional evidence using a nationally representative sample of Chinese citizens. To be specific, I use the 2015 Chinese General Social Survey with a sample of 10,968 respondents and an interrupted time series design that exploits an exogenous shock to the Chinese government's digital surveillance operation caused by the Tianjin explosion in 2015. Due to the limitation of the observational data, I am not able to compare digital surveillance with in-person surveillance in a real-world setting. Nevertheless, recent empirical studies find that traditional, Stasi-style surveillance has long-lasting negative effects on interpersonal trust, institutional trust (i.e., regime legitimacy), and election participation (e.g., Lichter, Loeffler and Siegloch 2019), which lend external validity to my experimental findings on in-person surveillance.

5.1 2015 Tianjin Explosions and Government Surveillance

On August 12, 2015, a series of blasts in a Sinochem subsidiary's warehouse in the port of Tianjin killed 173 people and injured nearly a thousand (Merchant 2017). More than 17,000 housing units were damaged by the explosion, and 779 businesses suffered property damages. The two major explosions were caused by combustible fertilizer ammonium nitrate,

detonated by fire and small explosions due to the misuse of firewater sprinklers on some chemicals (Huang and Zhang 2015). According to the earthquake waveform records, the first major explosion occurred at 11:34:06 pm, and the local earthquake magnitude (ML) was about 2.3. The second major explosion occurred 30 seconds later, and the ML was about 2.9. The resulting fireballs reached hundreds of meters high. The second explosion was estimated to be 336 tons TNT equivalent (Huang and Zhang 2015). Days later, local authorities ordered the evacuation of residents within a 3-kilometer (1.9-mile) radius of the blast site, prompted by the threat of "toxic substances", including sodium cyanide (Ryan 2015).

Immediately after the explosions, information on the event, including blast videos, was released over social media platforms like Weibo and WeChat. This accident drew a great deal of attention among Chinese netizens, with the topic racking up more views on Weibo than the country's total population of nearly 1.4 billion (Dou 2015). Figure 7 shows the temporal distribution of the Google search Index from Mainland China using "Tianjin" as the keyword. We can see the search intensity peaked in the two weeks immediately following the accident.[10]

The devastating explosions raised serious questions about corruption, industrial safety, and emergency responses in China (Merchant 2017; Dou 2015). As soon as discussions and rumors went viral on the Internet and social media, the country's Internet surveillance and censorship machines operated at full capacity to control information and silence discussions. Data from the censorship tracker Weiboscope, developed by the Department of Journalism at the University of Hong Kong, shows that surveillance and censorship rates on Weibo were up tenfold after the explosions compared with earlier in the month (Dou 2015). A large number of posts and discussions were rapidly deleted on the Internet and social media.

[10]Although Google is blocked in Mainland China, people can use virtual private networks (VPNs) to circumvent the firewall. Thus, Google search index still reflects the temporal trend of public interests.

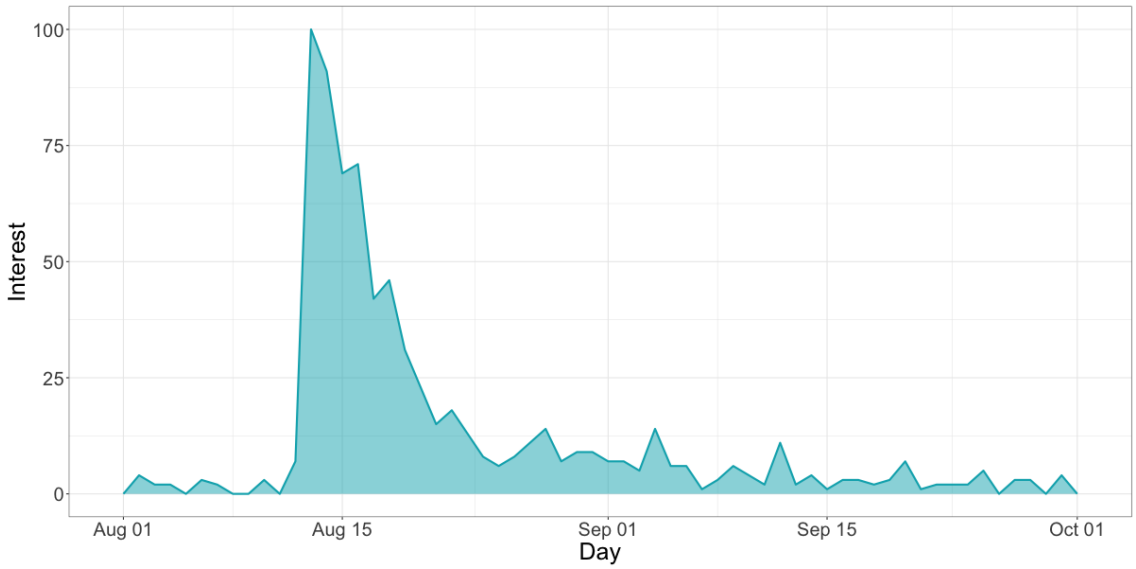


Figure 7: Google Trends on Tianjin Explosions

These include facts about the casualties, pictures, and videos of the explosions and the site afterward, investigative reports, as well as comments that question the political ties of the warehouse owners, criticize the government's responses, and discuss the chemicals inside the warehouse (Dou 2015; Hanrahan 2015).

One may argue that the aftermath of this event is a manifestation of government censorship instead of surveillance. However, conceptually, censorship is the combination of digital surveillance and online repression – censors need to identify the objectionable posts first and then delete them. Netizens who see posts being censored, even if they did not re-post or comment on them, certainly know that the government is monitoring the Internet and repressing online expression. As I discussed in the theory section, surveillance deters political participation because surveillance is associated with targeted repression. This is also the logic behind online censorship. The Tianjin accident creates an exogenous shock to online censorship and hence can be used to identify the causal effects of digital surveillance (and online repression) on individuals' trust and political participation.

5.2 Empirical Strategy and Data

My empirical strategy takes advantage of the coincidence that the 2015 Chinese General Social Survey (CGSS) was being conducted across China around the time of the Tianjin blasts. The CGSS is a nationwide survey ran every other year by the China Survey and Data Center at Renmin University – one of the top research universities in China. It participates in the International Social Survey Program and is the most reliable social survey in China. Figure 8 presents the time, location, and the number of individuals surveyed in the 2015 CGSS, with the dashed line indicating the time of Tianjin blasts.

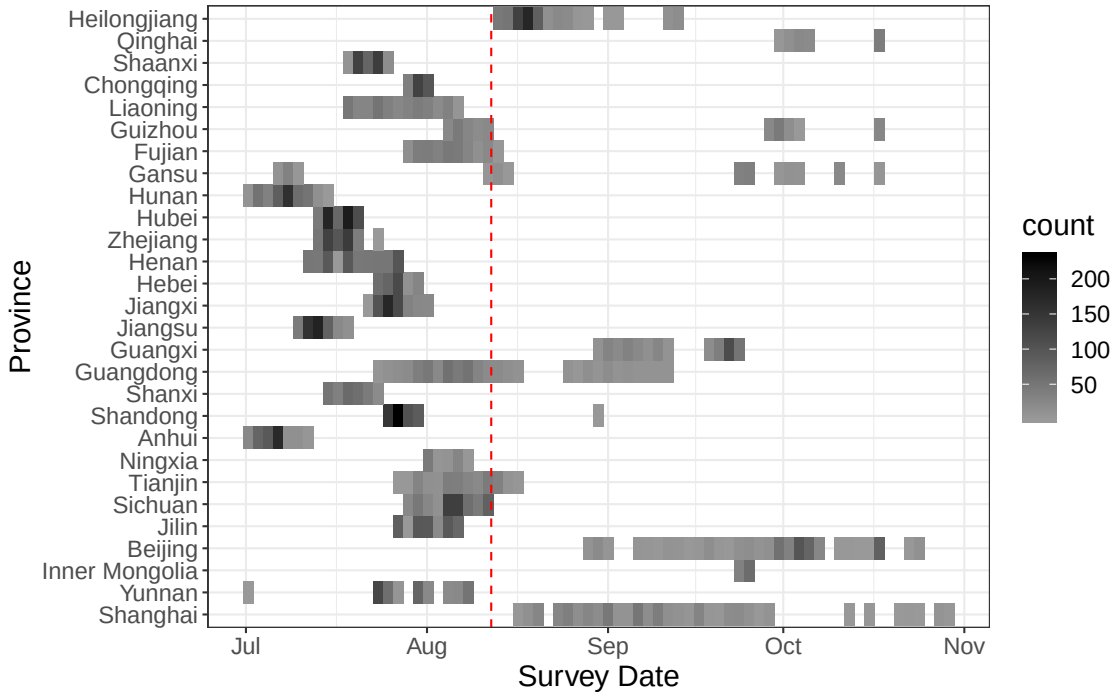


Figure 8: Distribution of Survey Respondents in the 2015 CGSS

Following Mummolo (2018), I use the interrupted time series (ITS) approach to compare individuals surveyed right before the Tianjin accident with individuals surveyed right after it.[11] As shown in Figure 8, some provinces do not have observations either before or after

[11]Scholars also refer this approach to a Regression Discontinuity in Time design (e.g., Hausman and Rapson 2018).

the accident. Because China is a large country where people in different regions may hold systematically different beliefs, the ITS estimate could be biased by regional differences. Thus, I include province fixed effects in the model. Specifically, I estimate the following equation:[12]

$$Y_{ip} = \alpha + \delta\, cutoff_{ip} + \pi\, min_{ip} + \lambda\, cutoff_{ip} \cdot min_{ip} + X'_{ip}\,\Psi + province_p + \epsilon_{ip} \qquad (1)$$

where i indexes the respondent and p the province; $cutoff_{ip}$ is a binary variable that takes the value of 1 if the respondent was interviewed after 11:59 pm on August 12, 2015 and 0 otherwise; min_{ip} is the running variable – the minute when the respondent was interviewed; X'_{ip} is a set of individual controls; $province_p$ is the province fixed effects. For Y_{ip}, I use four questions measuring interpersonal trust, regime legitimacy, views about expression, and views on petitioning (See Table B.1 in Online Appendix for the survey questions).

Non-netizens might not feel the intensified online surveillance after the Tianjin accident. I use individuals' Internet usage (including WAP phone services) to identify and exclude those who do not use the Internet (approximately half of the sample).[13] In addition, the identifying assumption of the ITS approach requires the groups of individuals surveyed before and after the time cutoff to be identical. As other concurrent events might influence respondents surveyed much earlier or much later, I further restrict the samples to one-week, two-week,

[12]I use a parametric linear model as the main specification because this model allows me to add interactions for mechanism tests (see Section 5.3). It also allows selecting appropriate sample windows around the cutoff time to avoid the events that could bias the results. I also use the local polynomial RD models developed by Calonico, Cattaneo and Titiunik (2014) and find similar results (Panal A in Table B.8 in Online Appendix). Panel B and C in Table B.8 also show that the results are largely robust to nonlinear global polynomial specifications.

[13]After sample restriction, it is still much more representative than the student sample in the survey experiment.

and three-week windows (i.e., one week before and one week after, so on and so forth). I use the two-week window sample for the main analysis because people's interests peaked within two weeks after the Tianjin accident (Figure 7). Different time windows also serve as robustness checks for the ITS estimates. Table B.2 in Online Appendix shows the summary statistics of the samples.

The ITS design would be weakened if there were "precise" sorting of the survey respondents around the cutoff time to imbalance the treated and control groups. This is unlikely the case since, first, respondents could not decide when they were interviewed, and, second, survey organizers could not change the predetermined sampling scheme to select a biased sample after the accident. In addition, survey interviewers all over the country had neither incentives nor capabilities to select a systematically biased sample in terms of trust and participation right after the event. I further test the assumption of local randomization by looking at whether baseline covariates are balanced. Table B.3 in Online Appendix shows that a number of covariates have no statistically significant changes around the cutoff time.[14]

5.3 Specification for Mechanism Testing

One concern is whether the effects identified by the ITS approach are due to surveillance caused by the Tianjin Explosions or just due to the accident itself. Other unknown concurrent events could also bias the results. To examine the mechanism of digital surveillance, I construct a measure of provincial-level surveillance intensity using the number of pilot counties for China's Golden Shield Project (GSP)[15] – a domestic digital surveillance and con-

[14]I check for discontinuities in other covariates at the cutoff instead of using the McCrary density test to examine the sorting problem because the density of the running variable (time) is uniform, which renders the test for discontinuities in its conditional density irrelevant (Hausman and Rapson 2018).

[15]I aggregate the measure at the provincial level because the CGSS only provides province identifiers.

tent filtering system that integrates online government databases with an all-encompassing surveillance network developed by the Ministry of Public Security (Walton 2001).

The phase-in GSP was implemented in small scale in some prefectures in 2000s but carried over in large scale in early 2010s, especially the "3111" Initiative that built local networks of digital surveillance tools with integrated street surveillance cameras. Under the "3111" Initiative, a total of 660 pilot counties/districts were selected by provincial governments between 2008 and 2012 to build the surveillance camera systems (Li and Hikvision Digital Technology Co. 2015). The ratio of pilot counties to the total counties in a province reflects the strength of digital surveillance in that province because the operation of surveillance camera systems requires well-developed surveillance infrastructures, integrated surveillance platforms, and sufficient security personnel. It also reflects the effort in surveillance operations made by security agencies in a province. In addition, although digital surveillance and censorship were mainly operated online, local security force is required to enforce punishment such as intimidation, harassment, and detention (Mozur 2019). Thus, despite that the "3111" Initiative aimed to build surveillance camera systems, the ratio of "3111" counties reflects the capacity of digital surveillance in a province.

Following Fitzpatrick (2010), I include an interaction term into Equation (1) to examine how the effects of surveillance and censorship caused by the Tianjin accident vary at different level of surveillance capacity. This method is an extension of Model (1) since it assumes the effects of digital surveillance on civil participation are conditional on surveillance capacity. In particular, I estimate the following specification.

$$Y_{ip} = \alpha + \delta\, cutoff_{ip} + \pi\, min_{ip} + \beta\, intensity_p + \gamma\, cutoff_{ip} \cdot intensity_p + \lambda\, cutoff_{ip} \cdot min_{ip} + X'_{ip}\,\Psi + province_p + \epsilon_{ip} \quad (2)$$

where $intensity_p$ is the ratio of "3111" pilot counties to the total counties in a province and $cutoff_{ip} \cdot intensity_p$ the interaction term. I expect γ to be negative and δ to be negative and statistically significant within certain ranges of $intensity_p$.

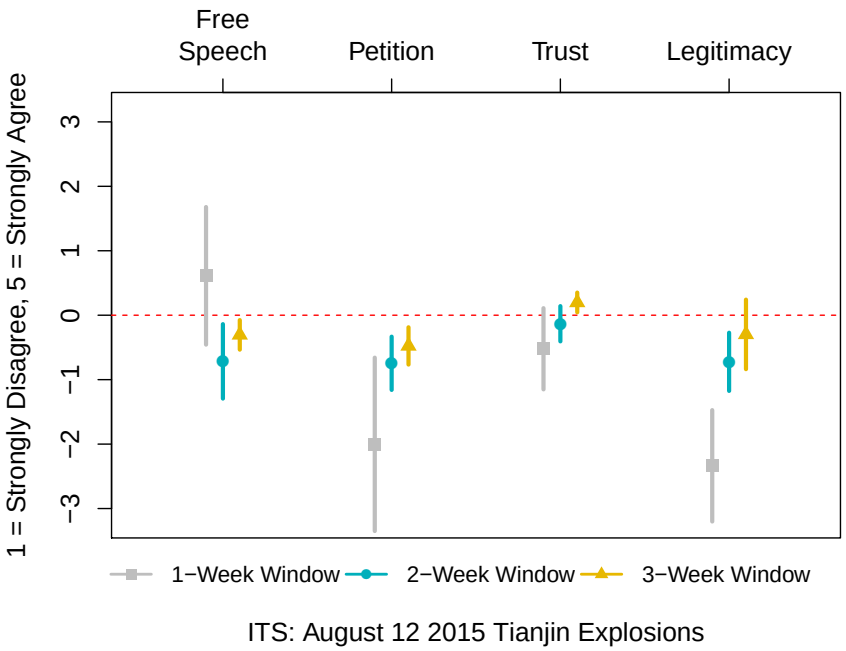
6 Findings from the 2015 CGSS Sample

This section presents the results of the Interrupted Time Series design using the 2015 CGSS data. I first present the main effects of the Tianjin Explosions on trust and political participation. Then, I show how the effects vary with the level of surveillance capacity.

6.1 Main Effects

Figure 9 plots the effects of digital surveillance (caused by the Tianjin Explosions) on individuals' view of free speech, perceived risk of petitioning to the government, interpersonal trust, and regime support. Under intensified digital surveillance, individuals are less likely to think that the right to criticize the government publicly is protected by law. They are more likely to think that petitions would be obstructed by the government. The effect on the view of petitioning is statistically significant on samples within the one-week, two-week, and three-week windows. The effect on the view of free speech is statistically significant on samples within the two-week and three-week windows. On the other hand, the effect of digital surveillance on trust is negative but statistically insignificant. These findings are consistent with the experimental results.

The theory and the experimental findings suggest that digital surveillance should not reduce regime legitimacy much. However, I find a negative effect of the Tianjin accident on regime legitimacy in the one-week and two-week windows. This inconsistency is likely due to the negative effect of the accident itself: after the horrifying explosions, people blamed the government for its failures to contain corruption, ensure industrial safety, and respond to emergencies, which, in turn, lower regime legitimacy.



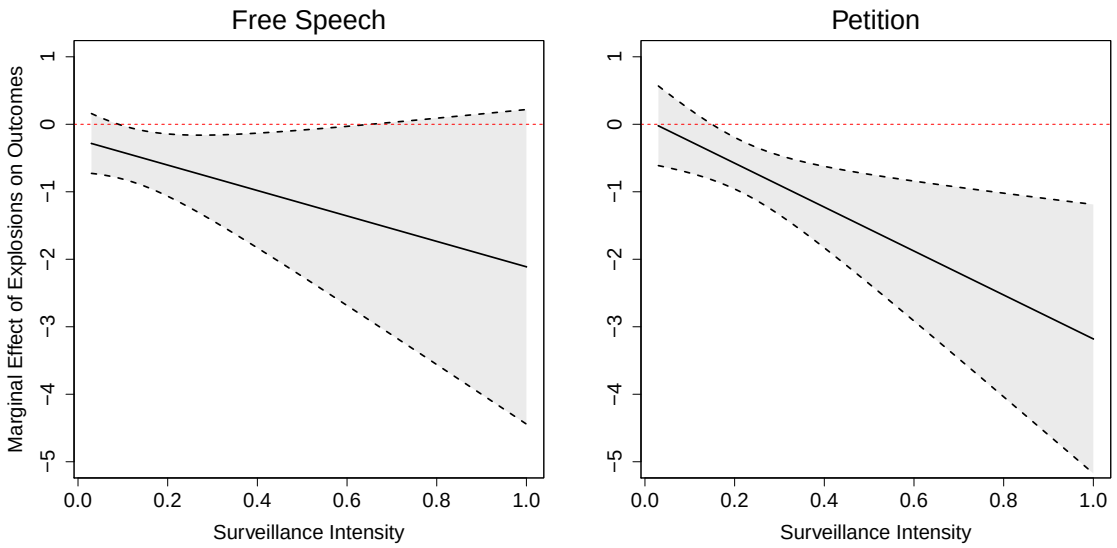
(Notes: This figure shows OLS estimates with 95% Confidence Intervals. Standard errors are clustered on prefectures. See Online Appendix B.3 for the regression tables underlying this figure.)

Figure 9: Tianjin Explosions on Political Participation, Trust, and Legitimacy

6.2 Mechanism Testing

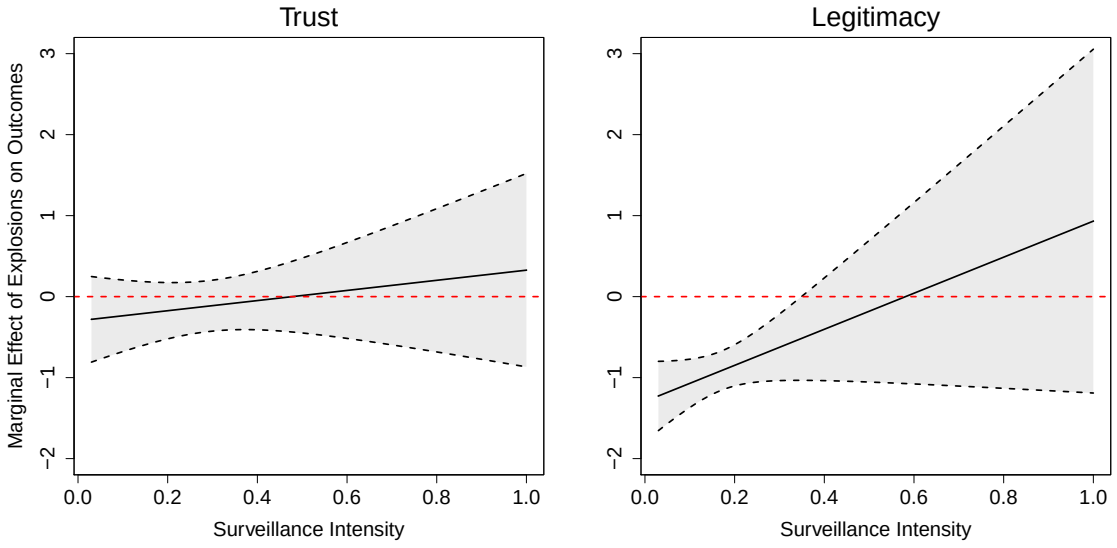
I further examine whether the negative effects of Tianjin Explosions on individuals' beliefs are conditional on the capacity of digital surveillance. In Figure 10, I plot the results of the ITS model with interaction between the time cutoff and surveillance capacity. The left panel shows that the marginal effect of Tianjin Explosions on individuals' views of free speech decreases with higher surveillance capacity and the effect is statistically significant when the capacity is within the 0.1 – 0.7 range. The right panel shows that the marginal effect of explosions on the view of petitioning also decreases with higher surveillance capacity and is statistically significant at most of the surveillance capacity levels. These findings provide strong evidence that digital surveillance is the driving force behind the decreased confidence in political participation after the Tianjin Explosions.

Figure 11 plots the marginal effects of the Tianjin accident on interpersonal trust and regime legitimacy. Consistent with our expectation, the marginal effect on trust is statis-



(Notes: OLS estimates with 95% Confidence Intervals. Standard errors clustered at the prefecture level. See Online Appendix B.3 for the regression results underlying these figures.)

Figure 10: Marginal Effects of Explosions on Political Participation



(Notes: OLS estimates with 95% Confidence Intervals. Standard errors clustered at prefecture level. See Online Appendix B.3 for the regression results underlying these figures.)

Figure 11: Marginal Effects of Explosions on Trust and Legitimacy

tically insignificant. Interestingly, the marginal effect of the accident on regime legitimacy increases with higher surveillance capacity and is statistically significant when this capacity is within the 0 – 0.4 range. This provides further evidence that the reduced legitimacy is

due to the Tianjin accident itself because stronger surveillance and censorship prevent the information about the accident from spreading to decrease citizens' regime approval.

The above findings from the 2015 CGSS sample are largely consistent with the experimental findings, which lend further weight to the main theoretical claims.

6.3 Robustness Tests

Some concurrent events could potentially bias the estimates of the ITS approach. Fortunately, during the four-week period, there were no big events or incidents that could invalidate the surveillance treatment caused by the Tianjin Explosions. On July 26, the death of a 31-year-old mother on an escalator set off a furor online, but this incident did not trigger surveillance and censorship. Even if there were censorship, it would lead to underestimation because censorship on this incident would have occurred before Tianjin Explosions. There were a few small incidents such as a chemical explosion in Shandong (August 31, 5 death) and landslide in Shaanxi (August 12, 26 missings), but none of them caught nationwide attention in China. There are two relatively large events around Tianjin Explosions: a stock market rollercoaster (June 12 – July 10) and the 70th Anniversary Parade of China's Victory over Japan Day (September 3). But both events were not in the four-week treatment window. Even if they had any influence on citizens during the four-week window, the stock market crash would have reduced trust and regime legitimacy before the treatment time, and the anniversary parade would have increased regime legitimacy after the treatment time. Thus, both events would lead to underestimation of the treatment effects. Figure B.1 in Online Appendix plots Google search trends for a number of events; none of them could threat the validity of the ITS design.

Another concern is that the results could be driven by the differences between individuals surveyed before and after the Tianjin incident. Individuals from different regions may hold different opinions to bias the ITS estimates if the pre- and post-treatment samples were drawn from different regions. I control for province fixed effects to reduce this potential bias.

Further, I conduct placebo tests on a series of outcomes that should not be affected by Tianjin Explosions. The logic is that, if it were regional differences that drive the differences in the outcomes of interest, we would observe similar patterns in other attitudes and behavior. However, I did not find significant differences between the pre- and post-treatment groups in terms of randomly picked questions such as attitudes and behavior concerning inequality, gender role, housework, voting, one-child policy, and homosexuality (Table B.7 in Online Appendix). Another concern is that the Tianjin incident could cause some people to become new Internet users afterward, changing the post-treatment sample. Although I cannot rule out this possibility, the fact that the covariates and other outcomes are balanced between the pre- and post-treatment groups suggests that self-selection is not a concern.

Moreover, the mechanism test that the Tianjin Explosions event has stronger effects in areas with higher surveillance capacity provides further evidence for the theoretical argument, which suggests that concurrent events and sample selection are unlikely to invalidate the findings.

7 Conclusion

Dictatorships have a long history of using surveillance to collect information and control society. Traditional, in-person surveillance was the main form until the onset of the information age. While the free flow of digital information has had transformative effects on economic well-being for citizens throughout the world, it is increasingly important to understand how dictatorial governments exploit digital information for surveillance, and ultimately political control. This paper compares the social costs of digital surveillance with those of in-person surveillance. Using an in-the-field survey experiment in China, I show that both types of surveillance deter expression and protest participation, but digital surveillance is less likely to lower interpersonal trust and regime legitimacy. Evidence from a natural experiment using the 2015 Chinese General Social Survey is consistent with the experimental findings on digital surveillance. The theory and results imply that digital surveillance is

en effective tool of authoritarian control: it deters expression and protest participation and, meanwhile, does not yield the same potential costs for dictators that human surveillance entails.

The empirical findings in this paper are based on individuals' short-term attitudes toward state surveillance in China, but the theory and implications are applicable to long-run civic culture in other authoritarian countries. For example, Lichter, Loeffler and Siegloch (2019) find long-run negative impacts of the Stasi's in-person surveillance on trust, civic culture, regime support, and economic outcomes in post-reunification Germany. In Iraq, low levels of social trust and political expression appear in places under heavier state surveillance (Blaydes 2018). My findings on in-person surveillance are largely consistent with the findings from previous work. The findings that digital surveillance deters expression and protest coordination should hold in the long run given its strong deterrence effects even in a short-run survey experiment. Although we have limited evidence on the long-term effects of digital surveillance on trust and regime legitimacy, the experimental findings, particularly the digital-in-person comparisons, suggest that the impact will be weaker than that of in-person surveillance. It would be helpful if future research could explore the long-run impacts of digital surveillance on trust, regime legitimacy, and other social-economic outcomes.

This paper draws on evidence from China, but the theory and findings have board implications for surveillance in other authoritarian countries and even in democracies. The deterrence effect of surveillance on political participation exists in most authoritarian surveillance states since surveillance and targeted repression often coexist in dictatorships (Gunitsky 2015; Sullivan 2016; Xu 2020). This deterrance effect is observed even in advanced democracies despite the institutional safeguard against human rights violations (Sidhu 2007). In addition, the pernicious effects of in-person surveillance on trust and regime legitimacy are due to destructive social relationships and should exist in any regimes that encourage citizens to spy on each other. Future research on the impact of surveillance in democracies is worth exploring.

References

Algan, Yann and Pierre Cahuc. 2010. "Inherited Trust and Growth." *American Economic Review* 100(5):2060–92.

Almond, Gabriel Abraham and Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton university press.

Ash, Timothy Garton. 1998. *The File: a Personal History*. Vintage.

Benson, Michelle and Thomas R Rochon. 2004. "Interpersonal Trust and the Magnitude of Protest: a Micro and Macro Level Approach." *Comparative Political Studies* 37(4):435–457.

Blaydes, Lisa. 2018. *State of Repression: Iraq Under Saddam Hussein*. Princeton University Press.

Bruce, Gary. 2010. *The Firm: the Inside Story of the Stasi*. Oxford University Press.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-discontinuity Designs." *Econometrica* 82(6):2295–2326.

Cantoni, Davide, David Y Yang, Noam Yuchtman and Y Jane Zhang. 2019. "Protests as Strategic Games: Experimental Evidence From Hong Kong's Antiauthoritarian Movement." *The Quarterly Journal of Economics* 134(2):1021–1077.

Chwe, Michael Suk-Young. 2013. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton University Press.

Deibert, Ron. 2017. "Evidence That Ethiopia is Spying on Journalists Shows Commercial Spyware is Out of Control." *Wired* . December 06.

Dietrich, Nick and Charles Crabtree. 2019. "Domestic Demand for Human Rights: Free Speech and the Freedom-Security Trade-Off." *International Studies Quarterly* 63(2):346–353.

Dimitrov, Martin K. and Joseph Sassoon. 2014. "State Security, Information, and Repression: A Comparison of Communist Bulgaria and Ba'thist Iraq." *Journal of Cold War Studies*

16(2):3–31.

Dou, Eva. 2015. "China's Censors Scramble to Contain Online Fallout After Tianjin Blast." *The Wall Street Journal* . August 16.

Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *Review of Economic Studies* 80(4):1422–1458.

Feldstein, Steven. 2019. "The Global Expansion of AI Surveillance." *Carnegie Endowment for International Peace* . September 17, Available at: https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847. Accessed June 15, 2020.

Fitzpatrick, Maria Donovan. 2010. "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten." *Journal of Labor Economics* 28(1):51–85.

Gerber, A.S. and D.P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. W. W. Norton.

Gieseke, Jens. 2014. *The History of the Stasi: East Germany's Secret Police, 1945-1990*. Berghahn Books.

Greif, Avner. 1989. "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders." *The journal of economic history* 49(4):857–882.

Greitens, Sheena Chestnut. 2016. *Dictators and Their Secret Police: Coercive Institutions and State Violence*. Cambridge University Press.

Gunitsky, Seva. 2015. "Corrupting the Cyber-commons: Social Media as a Tool of Autocratic Stability." *Perspectives on Politics* 13(1):42–54.

Hanrahan, Mark. 2015. "Tianjin Explosions: Disasters In China Prompt Wave Of Media Censorship." *International Business Times* . August 13.

Hausman, Catherine and David S Rapson. 2018. "Regression Discontinuity in Time: Considerations for Empirical Applications." *Annual Review of Resource Economics* 10:533–552.

Hechter, Michael. 2009. "Legitimacy in the modern world." *American Behavioral Scientist* 53:279–288.

Huang, Ping and Jingyuan Zhang. 2015. "Facts Related to August 12, 2015 Explosion

Accident in Tianjin, China." *Process Safety Progress* 34(4):313–314.

Imai, Kosuke, Luke Keele and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological methods* 15(4):309.

Intel, Public Relations. 2017. "Intel Movidius Helps Bring Artificial Intelligence to Video Surveillance Cameras." *Intel Newsroom* . April 5. Available at: https://newsroom.intel.com/news/intel-movidius-helps-bring-artificial-intelligence-video-surveillance-cameras/#gs.unl5w3. Accessed August 9, 2019.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge University Press.

Kendall-Taylor, Andrea and Erica Frantz. 2014. "How Autocracies Fall." *The Washington Quarterly* 37(1):35–47.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107(2):326–343.

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

Knabe, Hubertus. 1999. "Die feinen Waffen der Diktatur. Nicht-strafrechtliche Formen politischer Verfolgung in der DDR." *Die DDR–Erinnerungen an einen untergegangenen Staat* 88:191–220.

Kuran, Timur. 1991. "Now Out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44(1):7–48.

Levi, Margaret, Audrey Sacks and Tom Tyler. 2009. "Conceptualizing legitimacy, measuring legitimating beliefs." *American behavioral scientist* 53(3):354–375.

Li, Yanxiang and Ltd. Hikvision Digital Technology Co. 2015. "The History and Prospects of Safe City (in Chinese)." *China Public Security* 09. Available at: http://www.cnki.com.cn/Article/CJFDTOTAL-GGAZ201509013.htm. Accessed July 4, 2019.

Lichter, Andreas, Max Loeffler and Sebastian Siegloch. 2019. "The Economic Costs

of Mass Surveillance: Insights From Stasi Spying in East Germany." Working Paper. Available at: https://www.dropbox.com/s/dm253x8bh8nspgo/Stasi_onlineversion_aug2019.pdf?dl=0. Accessed August 3, 2019.

Liu, Joyce and Xiqing Wang. 2017. "In Your Face: China's all-seeing state." *BBC News* . December 10.

Lorentzen, Peter L. 2013. "Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime." *Quarterly Journal of Political Science* 8(2):127–158.

Lundmark, Sebastian, Mikael Gilljam and Stefan Dahlberg. 2015. "Measuring Generalized Trust: An Examination of Question Wording and the Number of Scale Points." *Public Opinion Quarterly* 80(1):26–43.

McMillan, John and Pablo Zoido. 2004. "How to Subvert Democracy: Montesinos in Peru." *Journal of Economic perspectives* 18(4):69–92.

Merchant, Nomann. 2017. "China Investigates Former Local Party Boss for Bribery." *The Associated Press News* . January 22.

Moore, Kelly. 1999. "Political Protest and Institutional Change: the Anti-vietnam War Movement and American Science." *How social movements matter* 10:97–118.

Mozur, Paul. 2019. "Twitter Users in China Face Detention and Threats in New Beijing Crackdown." *New York Times* . January 10.

Mummolo, Jonathan. 2018. "Modern Police Tactics, Police-citizen Interactions, and the Prospects for Reform." *The Journal of Politics* 80(1):1–15.

Mummolo, Jonathan and Erik Peterson. 2019. "Demand Effects in Survey Experiments: an Empirical Assessment." *American Political Science Review* 113(2):517–529.

Perry, Elizabeth J. 2008. "Chinese conceptions of "rights": From Mencius to Mao—and now." *Perspectives on politics* 6(1):37–50.

Putnam, Robert D, Robert Leonardi and Raffaella Y Nanetti. 1994. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton university press.

Qin, Bei, David Strömberg and Yanhui Wu. 2017. "Why Does China Allow Freer Social

Media? Protests Versus Surveillance and Propaganda." *Journal of Economic Perspectives* 31(1):117–40.

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

Ryan, Fergus. 2015. "China Explosions: Police Order Mass Evacuations Amid Further Blasts." *The Guardian* . August 15.

Sidhu, Dawinder S. 2007. "The chilling effect of government surveillance programs on the use of the internet by Muslim-Americans." *U. Md. LJ Race, Religion, Gender & Class* 7:375.

Singh, Amarjot, Devendra Patil, Meghana Reddy and SN Omkar. 2017. Disguised Face Identification (DFI) With Facial Keypoints Using Spatial Fusion Convolutional Network. In *Proceedings of the IEEE International Conference on Computer Vision*. pp. 1648–1655.

Sullivan, Christopher M. 2016. "Political Repression and the Destruction of Dissident Organizations: Evidence from the Archives of the Guatemalan National Police." *World Politics* 68(4):645–676.

Valentino-DeVries, Jennifer, Lam Thuy Vo and Danny Yadron. 2015. "Cataloging the World's Cyberforces." *The Wall Street Journal* . December 28.

Walton, Greg. 2001. *China's Golden Shield: Corporations and the Development of Surveillance Technology in the People's Republic of China*. Rights & Democracy.

Wintrobe, Ronald. 2000. *The Political Economy of Dictatorship*. Cambridge University Press.

Wiswall, Matthew and Basit Zafar. 2014. "Determinants of College Major Choice: Identification Using an Information Experiment." *The Review of Economic Studies* 82(2):791–824.

Xu, Xu. 2020. "To Repress or to Co-opt? Authoritarian Control in the Age of Digital Surveillance." *American Journal of Political Science* .

Xu, Xu and Xin Jin. 2018. "The Autocratic Roots of Social Distrust." *Journal of Comparative Economics* 46(1):362–380.

Yan, Alice. 2019. "Chinese Teacher Suspended for 'Belittling' Great Inventions in Online Chat." *South China Morning Post* . https://www.scmp.com/news/china/society/article/3024372/chinese-teacher-suspended-belittling-great-inventions-online. August 26. Date accessed: September 3, 2019.

Yang, Yuan. 2019. "Inside China's crackdown on young Marxists." *Financial Times* . https://www.ft.com/content/fd087484-2f23-11e9-8744-e7016697f225 February 13. Date accessed: May 20, 2020.

Zhao, Dingxin. 2004. *The Power of Tiananmen: State-society Relations and the 1989 Beijing Student Movement*. University of Chicago Press.

Zuo, Qiuming. 1998. *Discourses of the State (BC 500)*. Shanghai Ancient Books Publishing House.