To Repress or To Co-opt? Authoritarian Control in the Age of Digital Surveillance*

Xu Xu†

(Forthcoming at the *American Journal of Political Science*)

Abstract

This paper studies the consequences of digital surveillance in dictatorships. I first develop an informational theory of repression and co-optation. I argue that digital surveillance resolves dictators' information problem of not knowing individual citizens' true anti-regime sentiments. By identifying radical opponents, digital surveillance enables dictators to substitute targeted repression for non-exclusive co-optation to forestall coordinated uprisings. My theory implies that as digital surveillance technologies advance, we should observe a rise in targeted repression and a decline in universal redistribution. Using a difference-in-differences design that exploits temporal variation in digital surveillance systems among Chinese counties, I find that surveillance increases local governments' public security expenditure and arrests of political activists but decreases public goods provision. My theory and evidence suggest that improvements in governments' information make citizens worse off in dictatorships.

Keywords:: digital surveillance; dictatorship; China; information; repression; co-optation

*For helpful comments, I would like to thank Xun Cao, Martin Dimitrov, Christopher Fariss, Matt Golder, Sheena Greitens, Erin Hartman, Ji Yeon Hong, John Huber, Luke Keele, Hanzhang Liu, Zhaotian Luo, Xiao Ma, Melanie Manion, Daniel Mattingly, Ricardo Piqué, Adam Przeworski, Suzanne Scoggins, David Siegel, Wenfang Tang, Rory Truex, Joseph Wright, Tianyang Xi, Yiqing Xu, Yao-Yuan Yeh, Changdong Zhang, Boliang Zhu; participants at the APSA annual meeting, the MPSA conference, the annual meeting of the Society for Political Methodology, the Quantitative China Studies Seminar at New York University, the New Faces in China Studies conference at Duke University, and department seminars at Penn State. Three anonymous referees and Editors Sarah Brooks, Kathleen Dolan, Jennifer L. Lawless, Jan Leighley, and Graeme Robertson provided constructive criticisms that greatly improved the paper. Finally, I would like to thank Xin Jin for her helpful comments and unwavering support. All remaining errors are my own. Online Appendix and replication materials are available at https://xu-xu.net/research.html

†Ph.D. Candidate, Department of Political Science, Penn State University. xux112@psu.edu, https://xu-xu.net. 210 Pond Lab, University Park, PA 16802

1 Introduction

In early June 2013, former NSA contractor Edward Snowden revealed top-secret documents concerning the U.S. government's surveillance operations. In Western democracies, the Snowden leaks fueled intense controversy among the public and caused a significant retreat in government surveillance.[1] However, in dictatorships where the governments pay less heed to public opinion, courts, and civil rights, digital surveillance has steadily increased in the past decade. Many authoritarian countries have purchased, developed, or enhanced surveillance tools in recent years (Valentino-DeVries, Vo and Yadron 2015). More importantly, unlike U.S. surveillance that mainly focuses on international communications (Gellman and Poitras 2013), government surveillance in authoritarian regimes primarily targets domestic activity. Why are authoritarian governments so enthusiastic about domestic surveillance? How does digital surveillance influence government behavior and citizen well-being? The rapid rise in digital surveillance, as I show, has implications for our understanding of how dictatorships use repression and co-optation to address the risk of popular uprisings and to enhance regime stability.

Dictators control their populace through repression and co-optation, whereby citizens benefit from privileges doled by the regime but also are subject to government suppression and even violence (Svolik 2012). Dictators' strategies to mitigate the risk of popular revolt and maintain regime stability must weigh the costs and benefits of both tools. Though a large body of literature emphasizes dictators' repression-cooptation trade-off (e.g., Wintrobe 2000; Frantz and Kendall-Taylor 2014), few studies have addressed how changes in the informational environment shapes this tradeoff.[2] The literature on information in authoritarian regimes largely focus on the *horizontal information problem* faced by dissenting citizens who do not know each other's anti-regime sentiment, which prevents them from coordinating uprisings successfully (e.g., Kuran 1991; Edmond 2013). In such settings, communication technologies facilitate horizontal information exchanges,

[1] For example, the U.S. Congress passed the USA Freedom Act on June 2, 2015, which significantly curtailed the government's sweeping surveillance of phone calls and data gathering; on May 25, 2018, the European Union passed the strictest privacy law in EU history – the General Data Protect Regulation.

[2] See Blaydes (2018, Chapter 2) for a recent exception.

enabling better coordination among citizens (e.g., Edmond 2013). A few recent studies argue that dictators allow *vertical information flows* through media or multiparty elections to address social grievances and to monitor local officials (Huang, Boranbay-Akan and Huang 2019; Lorentzen 2014; Miller 2015). However, an important question in vertical information exchange remains unexplored: dictators do not have refined information on each citizen's anti-regime sentiment and thus cannot optimally allocate resources between repression and co-optation to deter collective action. This article argues that increasing digital surveillance shapes the repression-cooptation trade-off in substantively important ways by mitigating this *vertical information problem*, thereby allowing authoritarian governments to substitute targeted, preventive repression for more costly universal co-optation.

All dictators face threats from citizens excluded from the power, especially organized opposition groups (Svolik 2012). Even if disenfranchised citizens cannot threaten regime survival, dictators still have an incentive to prevent everyday protests and maintain social order for economic growth so that they can extract more rents from the economy (Pan 2015). Radical opponents – citizens with stronger anti-regime sentiments – pose greater threats to social order than moderate ones because they have a stronger tendency to protest (Lust-Okar 2005) and they are crucial to forming a critical mass necessary for anti-regime mobilization (Oliver and Marwell 1988). In *stable dictatorships*, large-scale, indiscriminate repression is rare and usually not a preferable option to dissuade mass mobilization in the first place. Ideally, dictators could selectively repress radicals to prevent mass mobilization by addressing threats when the opposition is still organizing collective action (Ritter and Conrad 2016). However, dictators are inherently uninformed. In contrast to democratic leaders, who obtain information about citizens from free speech, critical media, and multiparty elections, dictators face the difficulty of obtaining accurate information about citizens even if they hold elections because citizens have an incentive to misrepresent their anti-regime sentiments when faced with the prospect of repression (Kuran 1991).[3] This authoritarian informa-

[3]Miller (2015), among others, argue that multi-party elections in dictatorships allow citizens to signal dissatisfaction with the regime. But this type of information is much blunter than information collected in a society with political freedom. Further, many authoritarian regimes, such as China, Saudi Arabia, and North Korea, do not have multi-party elections.

tion problem undermines dictators' ability to identify radicals for targeted repression. As a result, dictators resort to expensive co-optation policies such as *universal* welfare provision to prevent moderates from coordinating with radicals to form an organized opposition.

The development of information and communications technology (ICT), especially surveillance technology, alleviates the vertical information problem in dictatorships. ICT facilitates citizens' digital communication, leaving substantial information in digital formats for governments to access and analyze. For example, the Iranian and Syrian regimes employ an array of digital surveillance tools to spy on citizens, especially those they deem threatening to regime survival (Gunitsky 2015; Gohdes 2014). Chinese government agencies also invest in digital tools to track and analyze online activities for containing threats before they spread (Qin, Strömberg and Wu 2017). Moreover, recent technology advancements in high-resolution cameras, facial recognition, and big data processing empower governments to monitor citizens and identify dissidents in a timely manner (Liu and Wang 2017). When digital surveillance enables dictators to identify radical opponents, they substitute less expensive targeted repression for non-exclusive co-optation to lower regime survival cost. Thus, as digital surveillance technologies advance, we should observe increased selective repression but less non-exclusive co-optation.

This paper focuses on surveillance technology rather than ICT penetration,[4] using China as a testing ground. In 1998, the Ministry of Public Security in China proposed the Golden Shield Project – a digital surveillance system (Walton 2001) – to improve the efficiency and effectiveness of the police. In 2001, the central government approved and started to fund the project, which was then implemented through several phases. The first phase of the project mainly built population databases, ID tracking systems, and Internet surveillance tools. The second phase of the project – the 3111 Initiative – focused on street surveillance camera systems. Recent advancements on this project started to integrate facial recognition, big data, and artificial intelligence technologies into the system. Due to the central ministry's phase-in strategy in experimenting this project, the

[4]This paper considers increases in digital surveillance for a given level of Internet penetration. See Shapiro and Siegel (2015) for an example of how changes in ICT access alter insurgent mobilization and government surveillance together.

timing of implementation varied across counties. This variation in the timing allows comparison of counties with and without completed Golden Shield systems, using difference-in-differences (DiD) designs.

Using unique archival data on Chinese county government expenditures and county-level political prisoners, I find that the implementation of the Golden Shield project increases local governments' targeted repression, as measured by local public security expenditure and the number of political prisoners. I also find that the Golden Shield project has an negative impact on government welfare provision, agriculture investments, and provision of other public goods. These findings are consistent with the theoretical predictions that governments substitute preventive, targeted repression for non-exclusive co-optation. The results are robust to lagged dependent variable models that address endogeneity problem between the local implementation of the Golden Shield project and a need for repression, as well as matching and trajectory balancing models that address time-variant omitted variable biases.

This paper directly contributes to a large body of literature on repression and co-optation in dictatorships (e.g., Gandhi and Przeworski 2006; Davenport 2007), especially work that focuses on the repression-cooptation trade-off. For example, Wintrobe (2000) presents a general theory of dictatorships that hold power through a combination of repression and loyalty, a concept similar to co-optation. Gershenson and Grossman (2001) analyze the changing mix of repression and co-optation that Soviet elites used to retain power in response to exogenous changes in domestic and international threats. Frantz and Kendall-Taylor (2014) develop a framework to show how co-optation institutions influence different types of repressive strategies in dictatorships. My work differs in two ways. First, I argue that though both repression and co-optation are important tools of authoritarian control, the information requirements to employ common forms of repression and co-optation differ – (targeted) repression requires more refined information. Second, I derive informational conditions under which dictators substitute repression for co-optation.

The theory and evidence in this paper address a common phenomenon regarding information and repression in dictatorships. Sullivan (2016) examines police repression in Guatemala to show

that when the government accumulates sufficient information about organizers to repress mobilization repeatedly, it can destroy challenger organizations by targeting the clandestine activities; but when repression is directed at ongoing overt challenges, it motivates backlash that escalates dissent. Dimitrov and Sassoon (2014) compare the operation of state security organs in Bulgaria and Iraq and find that, in Bulgaria, massive repression was replaced by surveillance and targeted repression because State Security was able to collect more fine-grained information on citizens. Truex (2019) also shows that the Chinese government strategically anticipates events that create focal points for protest coordination and then uses preventive repression to reduce dissent beforehand. My paper contributes to this literature by highlighting the role of digital technologies in selective repression and examining the *trade-off* between targeted repression and non-exclusive co-optation when governments' information improves.

This paper also contributes to a growing body of literature on ICT and authoritarian survival. Early studies posit that information technologies increase political freedom since they spread democratic values and empower ordinary citizens to mobilize (e.g., Diamond 2010). However, the Internet and ICT also provide authoritarian governments with new opportunities for political control (e.g., Lynch 2011). 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), and identify demonstrators and political opponents (Gunitsky 2015; Lynch 2011). Recent scholarship also finds that many regimes, including Russia, Turkey, Egypt, Bahrain, and Syria, employ domestic digital surveillance to monitor citizens (e.g., Gohdes 2014, p.91). This paper contributes to this literature by emphasizing the information advantages of digital surveillance for dictators to better target repression and decrease costly co-optation. As important, the theory and findings highlight that improvements in government information likely induce more repression instead of more policy concessions, as suggested in prior literature (e.g., Gunitsky 2015; Miller 2015).

2 An Informational Theory of Repression and Co-optation in Stable Dictatorships

Dictators rely on repression and co-optation to control citizens excluded from power (Svolik 2012). Repression subdues opposition to the ruling elite by force or violence (e.g., Davenport 2007). It is the defining characteristic of all states and is particularly salient in dictatorships. Co-optation, on the other hand, provides benefits to a group of citizens in exchange for political support (e.g., Gandhi and Przeworski 2006). While co-optation often involves using formal institutions such as parties and legislatures to cement enduring commitments to provide benefits to co-opted groups (Gandhi and Przeworski 2006), even dictatorships without strong formal institutions appease popular discontent with policy concessions or material handouts (Kitschelt and Wilkinson 2007). To forestall dissent before coordinated protest manifests, dictators choose the optimal levels of repression and co-option by weighing the costs and benefits of both.

2.1 The tradeoff between targeted repression and non-exclusive co-optation

Repression and co-optation can be indiscriminate or targeted, but most of today's *stable, non-rentier dictatorships* seldom use indiscriminate repression or targeted co-optation as the main tools for everyday social control.

Large-scale, indiscriminate repression can incur international sanctions (Way and Levitsky 2006), lowers economic productivity (Nafziger and Auvinen 2002), causes mass backlash and mobilization (Sullivan 2016), and potentially leads to elite splits (O'Donnell, Schmitter and Whitehead 1986). For a regime on the brink of collapse, large-scale, indiscriminate repression may be a drastic remedy but it is not a long-term solution for sustaining a stable dictatorship by dissuading mass mobilization in the first place. Further, indiscriminate repression is less effective than targeted repression in deterring participation,[5] although it requires less information on citizens types

[5] Siegel (2011*b*) demonstrates that targeted repression is more likely to deter mobilization than random repression. Although social networks and the distribution of preferences influence the relative advantages of targeted repression, it is still better at preventing mobilization than random repression for all levels of repression intensity, and especially at low-intensity levels.

(Dimitrov and Sassoon 2014). Scholars find that targeted repression against individual opponents and weakening challenger organizations is far more frequently employed as an everyday tool of repression in dictatorships than indiscriminate repression (Sullivan 2016; Truex 2019; Ritter and Conrad 2016).

On the other hand, targeted co-option is expensive because its cost increases with the number of individuals co-opted, whereas non-exclusive co-optation such as public goods provision is far more efficient for buying mass support (see De Mesquita et al. [2003]'s selectorate theory building on this intuition). In most of today's stable dictatorships, citizens' anti-regime mobilization is low and regime stability high because the regimes provide the two most basic public goods all citizens value – social stability and broad-based long-term economic growth, which increase citizen satisfaction with the regime (Magaloni 2006). Therefore, dictators seldom use exclusive co-optation to buy mass support, especially in non-rentier states where governments face resource constraints.

Moreover, targeted co-optation of radicals is less efficient in deterring mobilization than targeted repression because the cost to buy support from radicals can be significantly higher than that to imprison them. Further evidence suggests that selective co-optation is ineffective in preventing protests because the unequal distribution of benefits among citizens intensifies contention and catalyzes collective action (Pan 2015).[6] More importantly, selective co-optation incurs a commitment problem that undermines the effectiveness of this strategy: dictators can renege on the promised rewards once the mobilization threat is gone, and co-opted opponents may also renege by mobilizing later, ignoring the reward they received (Stokes 2005). This commitment problem intensifies if radicals have strong anti-regime sentiment and thus are likely to renege on their promises.[7] Thus, to deal with radicals, dictators prefer employing targeted repression to targeted co-optation.

This discussion underscores an important scope condition for my theory: recent stable, non-rentier dictatorships where the governments attempt to deter mass mobilization. Given that most dictatorships, most of the time, are relatively stable,[8] my theory has general implications for mod-

[6]Targeted co-optation can be as or more effective than repression when it has a positive spillover effect on insurgents' social network (Siegel 2011*a*), but this evidence suggests the spillover effect could be negative.

[7]Non-exclusive co-optation is less prone to the commitment problems since it can still buy off honest moderates.

[8]Military juntas are less stable but they are increasingly rare in the past thirty years (Geddes, Wright and Frantz

ern dictatorships.

As *non-exclusive co-optation* and *targeted repression* become common practices in recent stable dictatorships, dictators face a tradeoff between these two strategies for everyday social control. Ideally, dictators prefer targeted repression to forestall citizen mobilization by *identifying* radical opponents and then *stopping* them through harassment, intimidation, or detention (Truex 2019). However, dictators face a vertical information problem that citizens tend to hide their true anti-regime preferences when they feel threatened by the prospect of state repression (Kuran 1991). Due to this information problem, dictators cannot accurately target radicals. Non-exclusive co-optation is most useful when the regime has sufficient geographic, occupational, or racial/ethnic information about a large group of potential opponents to offer public or club goods but does not have accurate individual-level information to distinguish radical from moderate opponents. Because buying support from a moderate is cheaper than buying support from a radical, and public goods are non-exclusive regardless of the number of recipients, the regime gains support from a large number of moderates. Although this strategy does not win over radicals, whose price of support is higher than that for moderates, if protests require coordination between radicals and moderates, non-exclusive co-optation deters moderates from participating, thereby reducing the likelihood of protest mobilization.

Although moderates typically outnumber radicals, the latter remains the primary threat to social order because radicals are more likely to organize a protest, and a protest is more likely to succeed when radicals participate (Oliver and Marwell 1988). Because successful anti-regime protests require coordination between moderates and (particularly) radicals, dictators selectively repress radicals to prevent coordinated protests rather than providing universal welfare to all opponents when the regime has accurate individual-level information about citizens' types. Thus, when the probability of finding radicals is low, dictators prefer co-optation to repression. In contrast, when the likelihood of finding radicals is sufficiently high, dictators opt for selective repression over non-exclusive co-optation.

2018).

2.2 Digital surveillance, co-optation, and repression

In the past two decades, the development of the Internet and the spread of social media have altered the way of communication, which provides authoritarian governments with new opportunities for surveillance. Citizens transmit information in electronic forms through an Internet infrastructure that can be controlled or interrupted by the government. Digital surveillance technologies, such as spying malware and automated mass-detection systems, are widely used in authoritarian countries (Gohdes 2014). In addition, recent advances in automated text analysis, machine learning techniques, and high-powered computing have reduced the costs of identifying critical users and censoring messages (Edmond 2013). Dictators use these tools to track and analyze online activities, to gauge public opinion, and to contain threats before they spread (Qin, Strömberg and Wu 2017). Moreover, street surveillance cameras, combined with facial recognition and big data technologies, help dictators identify individual radical opponents and monitor early protests (Liu and Wang 2017). Thus, when information through digital surveillance increases dictators' *probability* of detecting radical opponents, dictators will substitute targeted repression for co-optation to pre-empt anti-regime mobilization. This logic informs the following expectations:

Repression hypothesis: Government digital surveillance increases targeted, preventive repression in authoritarian regimes.

Co-optation hypothesis: Government digital surveillance decreases non-exclusive co-optation in authoritarian regimes.[9]

2.3 Discussion

My theory focuses on the role of information in shaping authoritarian repression and co-optation. It applies to a majority of today's stable dictatorships. Even in societies where regime ideology or cultural legacies motivate high-intensity indiscriminate repression (e.g., genocide) against certain religious or ethnic groups, states still rely on surveillance and targeted repression to achieve

[9]I examine repression and co-optation separately because they are different outcomes of surveillance. Yet, since repression and co-optation are correlated due to the tradeoff discussed above, I also use a Seemingly Unrelated Regression approach to fit models of repression and co-optation together to allow for correlated errors.

social control (Blaydes 2018; Dimitrov and Sassoon 2014; Gohdes 2014).

Preventive repression as a result of increased government surveillance might further deter citizens from revealing true preferences, exacerbating the authoritarian information problem in the long run. However, if governments employ less visible, low-intensity repression, citizens will be less likely to hide their regime preferences. Thus, authoritarian governments have an incentive to make repression less visible to prevent backlash and to create uncertainty about their repression targets so as not to exacerbate the information problems in the future.[10]

Although my theory does not endogenize the adoption of digital surveillance, it is helpful to think about why digital surveillance arises in the first place. First, dictators may develop digital surveillance to improve repression efficiency as long as its benefits outweigh costs. For example, states with high technology capacity (i.e., lower surveillance costs) may be more likely to engage in digital surveillance. Second, dictators may switch to targeted repression because they cannot afford non-exclusive co-optation or they need to direct the resources towards more productive investments. An economic downturn, an interstate war, a growing threat of social uprising, or emerging sectionalism may force dictators to develop digital surveillance for targeted repression. Further, dictators may adopt digital surveillance to control hard-to-penetrate populations such as religious or ethnic groups. Nevertheless, even if digital surveillance is endogenous to repression, the prediction that digital surveillance increases targeted repression and decreases non-exclusive co-optation still holds. My empirical strategy addresses endogeneity concerns to identify the causal effect of digital surveillance on repression and co-optation.

3 Data and Empirical Strategies

I use a full sample of about three thousand Chinese counties (and districts) to examine how digital surveillance influences government repression and co-optation. These counties, as analogs to states in the international system, along with Chinese government's phase-in digital surveillance

[10] Although dictators may sometimes make preemptive repression visible to deter potential protesters, there is no need to publicize all cases since acknowledging that many people are protesting against the regime signals regime vulnerability (Gueorguiev 2017). Instead, dictators can make a few representative cases of repression visible for deterrence purpose while hiding their surveillance practice from the public.

programs allow me to adopt a difference-in-differences (DiD) design to identify the effects of digital surveillance on repression and co-optation. Moreover, the large number of counties allows matching and trajectory balancing methods to handle potential time-varying confounders that could weaken the DiD estimation.

3.1 Fiscal Arrangements in Chinese Counties

China is an authoritarian country with a multilevel government incorporating about three thousand counties/districts (for simplicity, hereafter, I use county to indicate both) nested in prefectures/cities (about 335) and provinces/municipalities (31) under the central government. On average, a county has about 500,000 residents.

I focus on county governments because they are directly responsible for social welfare provision and repression at the local level. Similar to leaders in authoritarian countries, county leaders in China worry about popular uprisings because their career advancements are tied to local stability (Edin 2003). County leaders, therefore, employ co-optation and repression (e.g., providing social welfare and arresting dissidents) to prevent protests. Due to fiscal decentralization, county governments keep a share of the tax income and allocate their budgets without direct central control (Jia, Guo and Zhang 2014). Scholars have long attributed China's economic growth to local governments' fiscal autonomy — see, e.g., Montinola, Qian and Weingast (1995)'s "federalism, China style" and Oi (1995)'s "entrepreneurial local government". Further, China's 1994 tax reform – the famous "Separating Tax System" — further expanded county governments' expenditure obligations, including support for agricultural production, financing compulsory education, provision of public health services, and management of social security programs (Jia, Guo and Zhang 2014). In other words, each county has a largely constrained budget (from tax and revenue) to allocate across different categories. If local officials save money from some redistribution obligations, they can spend more on other activities that benefit their political careers such as large-scale development projects that signal economic and political achievements (Guo 2009). Thus, Chinese county leaders' career incentives, fiscal responsibilities, and budget constraints are similar to those of state

leaders in the international system.

3.2 Measures and Data

I use county-level public security expenditure and the number of political prisoners as two complementary measures for targeted repression. For investment in non-exclusive co-optation, I use local welfare and public goods expenditure as well as local public goods provision. The data are collected from multiple archives kept in Beijing, Hong Kong, and other cities in China.

Public Security Expenditure Measure

Local public security expenditure is the total annual spending of local police forces, procuratorates, and courts. This expenditure, therefore, includes spending on the local government organizations that implement targeted repression: detentions, monitoring, harassments, beatings, and torture of local activists and dissidents. It does not include the spending of high-intensity repressive bodies such as the Armed Police Force and the People's Liberation Army, which the central government and provincial governments fund. Moreover, foreign observers have long used local security spending as an indicator for domestic repression in China (e.g., Martina 2014); and archival evidence suggests that the regime invests in local security spending to maintain everyday social stability.

Local public security budgets exclude expenses of the Golden Shield Project, which was mostly funded by the central government. While local security expenditure may also be used to fight crime, scholars find that Chinese security agents are better trained and equipped to control protests than to combat crimes (Scoggins 2018). Thus, the local public security expenditure is a reasonable proxy for targeted repression. See Online Appendix A.1 for details about this measure.

Political Prisoners Measure

While security spending measures investment in repressive capacity, which closely hews to the theoretical expectations, it does not measure actual repression. Therefore, I use the number of political prisoners as a complement to the spending measure.

The political prisoner data draws on information contained in the Congressional-Executive

Commission on China's Political Prisoner Database (CECC-PPD). The CECC lists individuals who have been detained or imprisoned in China for non-criminal reasons. Human rights organizations, such as the Human Rights in China, the Network of Chinese Human Rights Defenders, Human Rights Watch, and Dui Hua Foundation compile lists of the prisoners' names and dates of incarceration. Several recent studies employ this political prisoner data to measure preventive repression (e.g., Truex 2019; Gueorguiev 2017). Data on detentions are available starting in 1981 but poor data quality during the early years, when entries were recorded retrospectively, makes them unsuitable for analysis. In addition, I use the second-phase Golden Shield Project to examine this prisoner data. Considering that the first-phase Golden Shield project (completed in 2005) could contaminate the effect of the second-phase project, I then exclude all entries prior to 2006. After excluding entries with missing information, 4,531 (out of 5,007) detentions remain between 2006 and 2017. Among them, 82 percent of arrests were related to political crimes such as association, speech, and the spread of information – precisely the types of actions that facilitate protest coordination. Unlike combating crime that is loosely organized, controlling social unrest is well-disciplined in China with specific government protocols (Scoggins 2018), which makes prisoner data a comparable measure of targeted repression across different localities.

I generate a measure of local repression by aggregating every prisoner's time and location of detention (complemented by location of residence) in each county-year unit. This yields a 12-year panel for about 2,860 county-level units between 2006 and 2017. On average, each county arrested about 1.52 activists during the time period (the maximum is 256 and the minimum is 0).[11] See Online Appendix A.2 for details about this measure.

Discussion of Repression Measures

Both measures of targeted repression have their drawbacks: security spending includes all expenses of local security bodies; the prisoner data has missingness. Nevertheless, using both measures together mitigates the drawbacks of each. Moreover, while the prisoner data may have reporting biases because some regions are more heavily repressed traditionally, the difference-in-

[11] I address the concern of under-reporting in Online Appendix B.6.

differences design exploits *temporal* variation within counties rather than cross-sectional variation, mitigating bias from spatial missingness.

Co-optation Measures

I use local welfare expenditure – the total expense of subsistence allowances, pension, endowment, medical and unemployment insurance policies – as a measure of investment in co-optation. This measure is a reasonable proxy for co-optation because scholars find that welfare policies in authoritarian countries effectively co-opt groups of people who pose a high risk to regime stability (Knutsen and Rasmussen 2018). I also use the expenditures and quantities of local public goods, such as the numbers of welfare centers and beds and the numbers of students in elementary and middle schools, as measures of co-optation. Note that I include agriculture production as a proxy for government spending on agriculture because this spending is one of local government obligations. These measures of local public goods therefore capture the non-exclusive nature of co-optation.

Data Sources

I construct two datasets for these measures. First, I compile an original panel of county-level annual fiscal data using multiple sources. The 1994 - 2007 county-level public security, welfare, education, and agriculture expenditures come from the Fiscal Statistical Yearbook of Chinese Cities and Counties from 1995 to 2008.[12] Data on other covariates such as population and GDP comes partly from the China County Socio-Economic Statistical Yearbook. The final dataset contains about 2,860 counties between 1994 and 2007, roughly 36,180 observations. Second, I matched the 2006-2017 county-level political prisoner data with the social economic and public goods data from the China County Socio-Economic Statistical Yearbooks.[13] Summary statistics of these two datasets are reported in Online Appendix A.3.

[12]The national fiscal yearbooks are only available by 2008 with data by 2007.

[13]The socio-economic yearbooks contain no data for city districts – same administrative-level units as counties. I address this missing data problem in Online Appendix B.2.

3.3 The Golden Shield Project

My main empirical strategies rely on difference-in-differences comparisons that exploit local *temporal* variation in the implementations of a surveillance system — the Golden Shield Project (GSP). In 1998, the Ministry of Public Security initiated the GSP, a domestic surveillance and filtering system that integrates on online government databases with an all-encompassing surveillance network (Walton 2001). This project was constructed through several phases.

The first phase of the project, completed in 2005, mainly built population databases, ID tracking systems, and Internet surveillance tools. A local GSP platform includes a local population database that records residents' photo, ID number, address, household composition, and other basic information; further, it connects to a network of population databases nationwide. This local-level database also communicates with a database for high-priority monitoring and control（重点管控人口）that records residents who potentially threaten social and regime stability. Based on these two population databases, the platform further integrates several surveillance programs or tools. In addition, the system has ID scanning and tracking terminals installed in hotels, Internet cafes, train stations, bus terminals, airports, and train and air ticket offices. These tracking systems allow local police to track local and migrant populations, especially the population for high-priority monitoring and control. For example, once a blacklisted petitioner books a bus, train, or flight ticket to the province capital or Beijing, the system automatically notifies the local police bureau, and then the bureau will send officers to stop the petitioner.[14]

Moreover, the local GSP platform incorporates Internet surveillance modules that monitor important websites, online forums, and social media. Based on keywords searching software, these modules automatically discover and record public sentiments, and then report signs of social instability to local police (Du 2013). There are also Internet control modules that can conduct remote attacks (e.g., denial-of-service attacks) to disrupt websites, fabricate posts in threads, and guide online discussions (Du 2013). Combining IP address tracking with population databases enables the police to identify individuals who post sensitive information and track their location.

[14] An interview with a sub-district office director in Sichuan Province in June, 2015.

In the late 2000s, the Ministry of Public Security started the second phase of the GSP by integrating street surveillance cameras into the system. The ministry initiated a series of projects, including the Plan "3111" initiative.[15] With recently developed artificial intelligence and facial recognition technologies, the surveillance camera system identifies individuals in real time (Liu and Wang 2017).

Abundant anecdotal evidence suggests that the GSP facilitates preventive repression. For example, two famous lawsuits filed by Chinese political prisoners against Cisco reveals that the GSP, equipped with Cisco networking hardware, can detect, identify and track political dissidents, who were later imprisoned by police (Moses 2011). There are also many individuals monitored and detained for their online political writings.[16]

3.4 Empirical Specifications

Based on the local implementation of the GSP, I employ two main estimation strategies to analyze the effect of digital surveillance on local repression: a difference-in-differences (DiD) approach that compares differential changes in repression before and after the completion of this project, and a lagged dependent variable estimation that addresses endogeneity concerns that arise if past repression influences the timing of implementation of local GSPs in different cross-section units. Because different datasets cover different time ranges, I use the implementations of the first-phase GSP (2005) as an identification strategy for the expenditure dataset and the second phase ("3111" initiative, three waves) for the prisoner dataset.

I also use the same DiD specifications to estimate the effect of digital surveillance on government co-optation. Because measures of co-optation and repression are different budget expenditures, which are correlated with each other given local government budget constraints, I also use Seemingly Unrelated Regression (SUR) models to fit specifications with different budget expenditures together that allows errors to be correlated across equations.

Difference-in-Differences Design

[15] Other projects include the "Safe Cities" project, the "Skynet" project, and the Rural "Sharp Eyes" project.

[16] See China's Political Prisoner Database by the *Congressional-Executive Commission on China* at https://www.cecc.gov/resources/political-prisoner-database. Accessed November 5, 2017.

I first use a difference-in-differences approach that compares changes in outcome Y_{it} in places with and without completed Golden Shield systems:

$$Y_{it} = \alpha_1 + \rho_1 D_{it} + \tau_1 time_t + \gamma_i + \lambda_t + X'_{it}\Psi_1 + \epsilon_{it}; \tag{1}$$

where Y_{it} can be county i's repression or co-optation at year t; γ_i and λ_t are county and year fixed-effect dummies; X'_{it} is a set of time-varying county-level controls, and ϵ_{it} the error term. $time_t$ is a dummy variable indicating the period that the local GSP system has been completed. D_{it} is an indicator variable that is one if county i has a completed system at year t and zero otherwise. I am interested in estimating ρ_1, the DiD estimate.[17]

This DiD model estimates the over-time changes in "control" and "treatment" groups and then takes the difference of the two over-time differences. With exogenous controls X_{it}, the key identification assumption is $E(\epsilon_{it}|\gamma_i, \lambda_t, D_{it}) = 0$. That is, a county with a completed GSP system should have to maintain the same difference to a county without a completed system had the former not completed the system. In other words, there are underline parallel trends on the outcomes between the two groups.

Note that I define treatment and time differently for the two datasets because they span different phases of the Golden Shield project as the following:

Golden Shield First Phase: The expenditure dataset covers a period between 1994 and 2007. I use the *first phase* of the GSP as an identification strategy. The GSP was initialized in 1998 and began operation in 2003. The first phase of this project, including the population databases and most of the ID tracking systems, was completed in 2005.[18] Some counties were even equipped with an Internet surveillance system by 2005. In 2006, the central Ministry of Public Security awarded about 40 prefectural-level Bureaus of Public Security for their excellent work in completing the first phase of this project, which suggests that these 40 prefectures have finished

[17] I use cluster-bootstrap variance matrix developed by Bertrand, Duflo and Mullainathan (2004) to adjust for serially correlated standard errors.

[18] China.com, 2003, The Golden Shield Project, http://www.china.com.cn/chinese/zhuanti/283732.htm. Accessed July 4, 2019.

the 1st-phase surveillance systems by 2005 while other 280 prefectures did so only after 2005. The DiD estimator for this dataset exploits this temporal variation: $time_t$ is a dummy variable that is equal to one after 2005, and D_{it} is an indicator variable that is equal to one if county i is located in a prefecture that has completed GS after 2005 and zero otherwise.

Golden Shield Second Phase – the "3111" Initiative: The political prisoner dataset covers a period between 2006 and 2017. I use the second phase of the GSP — the City Alarm and Surveillance Camera Pilot Project ("3111" Initiative) — as an identification strategy. Since 2006, the Ministry of Public Security and provincial governments jointly conducted three waves of the "3111" Initiative (Ding 2010). The first wave included four cities (consisting of about 50 county-level administrations), where the systems took effect in 2008. The second wave included 22 cities (consisting of about 130 counties/districts), with the systems completed in 2010. During the last wave, another 470 counties/districts piloted "3111" projects, with completion in 2012 (Li and Hikvision Digital Technology Co. 2015). A total of 660 pilot counties/districts were selected with a relatively even distribution within 31 provinces to *demonstrate* the surveillance camera systems to all other counties in China. Thus, local security concerns were not a major reason for selecting these counties, which mitigates the endogeneity problem. I also find that past political arrests are not correlated with pilot county selection. The phase-in nature of the "3111" Initiative permits a staggered DiD design with multiple time periods and multiple groups — a stronger identification strategy than a single treatment period DiD. For the political prisoner dataset, D_{it} is an indicator variable that is equal to one if county i is a "3111" pilot county after the time of completion and zero otherwise.[19]

Lagged Outcome Variable Design

The causal identification of the DiD method assumes *time-invariant confounders* that can be captured by the fixed effects. However, it is possible that repression and surveillance are endogenous. For example, local pressure for repression may directly affect the implementations of local GS systems or the self-selections of the "3111" pilot counties. It is also possible that counties

[19]See Online Appendix A.4 for details about the "3111" Initiative as well as how those pilot counties were selected.

spend more money on repression tend to spend more on surveillance. To address this endogeneity problem, I use a lagged outcome model to account for pressure for repression or other omitted variables that lead to repression (Angrist and Pischke 2008):

$$Y_{it} = \alpha_2 + \rho_2 D_{it} + \pi_2 Y_{it-1} + \lambda_t + X'_{it}\Psi_2 + \epsilon_{it}, \quad t = 1, ...T; \tag{2}$$

where Y_{it-1} refers to one-year lagged outcome variable (i.e., repression). For different datasets, the variable D_{it} indicates different counties in different time periods as mentioned above.

DiD, Lags, and Trajectory Balancing

While both DiD and Lagged Outcome models are solutions of causal inference, each requires relatively strong assumptions for identification that are often hard to justify. Fortunately, Angrist and Pischke (2008) point out an interesting bounding property of these two approaches. They prove that if the DiD model is correct but the analyst uses a Lagged Outcome model, this will generate a correlation between the treatment and the lagged outcome which will bias the treatment effect downward. Conversely, if the Lagged Outcome model generates the data whereas the analyst estimates a DiD model, the estimated treatment effect will be too large since the unestimated lag parameter will be additive with the treatment effect through the error term. Therefore, we can view the estimates from a DiD model and a Lagged Outcome model as the upper bound and the lower bound for the causal effect of interest.

Another way to address the endogeneity problem is a recently developed trajectory balancing approach, which uses kernel balancing to weight control observations to match the treated observations according to the pretreatment trends (i.e., high-order "trajectory") of outcome and covariates (Hazlett and Xu 2018). This approach handles time-vary confounders by assuming some time-fixed linear combination of the time-varying confounders. Because pretreatment outcomes contain information on those confounders, balancing control observations with the treated on pretreatment outcomes differences out the time-varying confounders.

4 Results

In this section, I first present the results of surveillance on repression using those two repression measures. I then show that repression substitutes for co-optation after surveillance implementation.

4.1 Golden Shield and Local Public Security Expenditure

My argument posits that digital surveillance increases preventive repression. Figure 1 provides initial evidence from security spending data. The plot in the left panel shows that since 2006, the year that the GSP's first phase was completed, counties in prefectures with completed surveillance systems increase local public security expenditure more so than counties without completed surveillance systems, especially in 2007. The trends between the two groups are relatively parallel. To mitigate time-varying omitted variable biases, I further use the trajectory balancing approach to weight the control group. As shown in the right panel, the weighted control group – counties without completed surveillance systems – have almost the same pretreatment trend as that of the treated group, and the treated group has a larger increase in security expenditure since 2006.

Unweighted Trends

Trends with TjBal Weighted Control Group

Figure 1: Trends in Security Expenditure, 1st-Phase Golden Shield Project

Table 1 presents the estimated results using the unweighted data. Column (1) reports the DiD estimate (*GS Completed County*), controlling for local total fiscal expenditure, population, and

urbanization. Because the first-phase GS systems are completed at the prefecture level, I cluster standard errors by prefectures. The estimated effect is positive, large, and statistically significant, indicating that county governments in prefectures with completed GS systems spent 11.6 percent more on public security than counties in prefectures with uncompleted GS systems after 2005.

Table 1: 1st-Phase Golden Shield Project and Public Security Expenditure

	Fixed Eff. (1)	Lagged DV (2)	Fixed Eff. (3)	Lagged DV (4)	Fixed Eff. (5)	Lagged DV (6)
VARIABLES	Security (log)	Security (log)	Security (log)	Security (log)	Security (log)	Security (log)
Golden Shield×Time	0.116***	0.0363**	0.0995***	0.0297**	0.128***	0.0338**
	(0.0287)	(0.0141)	(0.0303)	(0.0122)	(0.0369)	(0.0139)
Time: Post-2005	0.784***		0.0387***		0.552***	
	(0.0833)		(0.0128)		(0.102)	
Lagged DV		0.754***		0.773***		0.755***
		(0.0118)		(0.0126)		(0.0138)
Expenditure(log)	0.638***	0.263***	0.570***	0.237***	0.531***	0.223***
	(0.0360)	(0.0132)	(0.0492)	(0.0146)	(0.0315)	(0.0145)
Population(Log)	0.0431	-0.00982*	0.0445	-0.0153***	0.0450	-0.00173
	(0.0290)	(0.00536)	(0.0401)	(0.00550)	(0.0298)	(0.00587)
Urb. Ratio	-0.00628	0.0460***	-0.00247	0.0411***	-0.00100	0.0619***
	(0.0130)	(0.00996)	(0.0163)	(0.00974)	(0.0132)	(0.0118)
GDP (log)			0.00205	0.0112***		
			(0.00588)	(0.00373)		
Admin. Exp. (log)					0.149***	0.0465**
					(0.0531)	(0.0215)
County FEs	Yes	No	Yes	No	Yes	No
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.121	-0.590***	1.363***	-0.741***	0.000178	-0.771***
	(0.310)	(0.0751)	(0.488)	(0.0785)	(0.407)	(0.103)
Observations	34,757	31,632	25,340	24,827	27,329	24,489
R-squared	0.889	0.953	0.826	0.948	0.858	0.948
N of County	2,745		2,731		2,742	

Robust standard errors clustered by prefectures. The number of observations are different between fixed effects models and lagged DV models because the lags automatically drop observations in the first year. Table B.4 shows that the results remain similar and robust when matching the sample of the fixed effects models with that of the lagged DV models.
*** p<0.01, ** p<0.05, * p<0.1

Columns (2) reports the results of the lagged DV model using the unweighted data; the estimate suggests the lower bound for the causal effect. The coefficient for *GS Completed County* remains positive and statistically significant: county governments in surveillance prefectures spent roughly

3.6 percent more on public security since 2005. Columns (3) - (6) show that the results are stable when including additional controls such as GDP and administrative expenditure.[20] In addition, the result from weighted data using Trajectory Balancing approach indicates a 6.9 percent increase in security expenditure after surveillance.[21] Together, these results suggest that the GSP has a large, positive effect on local repression. Given that China's internal security spending was larger than the army's budget in 2011, an 4 - 12 percent increase in security spending is substantial. I also conduct a DiD falsification test based on the weighted data and find no pre-trend difference (Online Appendix B.5).

4.2 "3111" Initiative and Political Prisoners

Figure 2 plots the number of political prisoners between the last wave of the "3111" counties and the never-treated counties before and after 2012. Pilot counties with completed surveillance camera systems have a sharper increase in the number of prisoners than non-pilot counties. Figure 2 also shows that the pre-surveillance trends between these two groups of counties are almost identical, which gives us more confidence to interpret DiD results as causal because it suggests common underlying trends between these two groups in the absence of the "3111" Initiative.

Column (1) in Table 2 reports the DiD estimates from Equation (1), with standard errors clustered on counties because the "3111" Initiative use counties as pilot project units.[22] The result indicates that pilot counties with advanced surveillance camera systems arrest 0.1 more political activists than non-pilot counties. Given that the average number of prisoners per county-year is 0.13, political prisoners almost double in surveillance pilot counties relative to non-pilot counties. Columns (2) reports the lagged DV model estimate from Equation (2). The result is also substantial and statistically significant. A DiD falsification test finds no pre-trend difference between these two groups (Online Appendix B.5).[23]

[20]The main models (Column [1] and [2]) do not include these two controls because the GDP variable misses all counties/cities in 1994, 1995, 1996, and 2007, and the administrative expenditure variable misses all counties/cities in 2004 and 2005.

[21]Online Appendix B.4 presents detailed results from the Trajectory Balancing approach.

[22]The results are robust to clustering on prefectures.

[23]The models do not include social-economic controls because city districts'(more than one-third of county-level units) controls are missing. Nevertheless, the results are robust with controls based on a subsample excluding city dis-

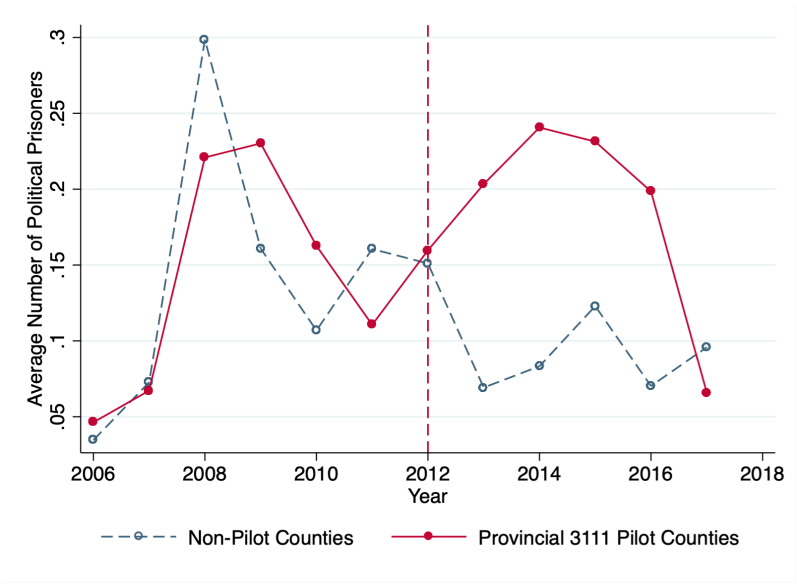

Figure 2: Trends in Political Prisoners, 3rd-Wave "3111" Initiative

Table 2: "3111" Initiative (2nd-Phase GSP) and Political Prisoners

	All Provinces		Tibet, Xinjiang & Beijing Excluded	
	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV
	(1)	(2)	(3)	(4)
VARIABLES	Prisoners	Prisoners	Prisoners	Prisoners
"3111" County	0.0949***	0.0844***	0.109***	0.0422***
	(0.0283)	(0.0215)	(0.0247)	(0.0143)
Lagged Prisoners		0.157***		0.142***
		(0.0114)		(0.0104)
County FEs	Yes	No	Yes	No
Year FEs	Yes	Yes	Yes	Yes
Constant	0.0381***	0.0647***	0.0347***	0.0648***
	(0.00715)	(0.0100)	(0.00694)	(0.0105)
Observations	34,328	31,460	32,033	29,359
R-squared	0.002	0.025	0.002	0.021
Number of County	2,868		2,674	

Robust standard errors clustered by counties. The results remain similar and robust when matching the sample of fixed effects models with that of lagged DV models.
*** p<0.01, ** p<0.05, * p<0.1

tricts (Online Appendix B.2). In addition, I use Fixed-Effects Negative Binomial models, Negative Binomial models, and Zero-inflated Negative Binomial models to address the discrete and non-negative values of the prisoner variable. The results remain robust to all model specifications. See Online Appendix B.2 for details.

The Tibet and Xinjiang autonomous regions rely more heavily on repression than other provinces due to ethnic tensions and secession. In addition, petitioners and protesters are more likely to gather in Beijing, China's political center, and therefore preventive repression is more intensive there. It is possible that the positive results are driven by outlier counties in Beijing, Tibet, and Xinjiang. The prisoner data also shows that governments in these three regions arrest disproportionally more activists than governments in other areas. To address this concern, I replicate the analyses using a subsample that excludes counties from those three provinces. As shown in Column (3) and (4), the results remain large and statistically significant.

Note that the average numbers of political prisoners in "3111" pilot and non-pilot counties converge in 2017, possibly suggesting that local governments may have arrested most dissidents so that extensive imprisonment is no longer necessary. It is also possible that citizens have learned to hide from government surveillance. Unfortunately, the data does not allow me to test these possibilities. Nevertheless, the evidence suggests that digital surveillance is effective for at least four years.

4.3 Surveillance and the Repression-cooptation trade-off

Scholars argue that Internet and social media surveillance enables authoritarian governments to collect accurate information on mass opinion and to tailor policies accordingly (Gunitsky 2015). While the Chinese government may use the information to adjust policy, this practice does not preclude the government from employing surveillance to identify political opponents. If, however, surveillance is primarily a tool to inform policy, we could expect surveillance to increase welfare spending in areas of potential unrest. In contrast, my theory suggests that surveillance for identifying individual opponents should result in more repression and less redistribution because the former is a more cost-efficient way to maintain social stability.

To test this trade-off, I examine how surveillance influences redistribution and security expenditures together using Seemingly Unrelated Regression models with a DiD setup. I use levels of expenditure rather than log scales to compare spending in different categories.[24] Table 3 shows

[24]See Table B.5 in Online Appendix B.1 for estimation in log scales.

large, negative, and statistically significant effects of surveillance on welfare and agriculture spending. Further, I examine how the "3111" Initiative influences local public goods provision. Table 4 shows that most of the effects are negative and the effects on the number of beds in hospitals and agriculture products are large, negative, and statistically significant (2.4 - 12 percentage decreases). Overall, these results suggest that surveillance decreases non-exclusive co-optation. Note that the negative effects do not mean an absolute reduction in public goods and welfare provision. Instead, these DiD estimates indicate that welfare spending and public goods have smaller increases in surveillance counties than in non-surveillance counties.

Table 3: 1st-Phase GSP, Public Security, and Redistribution Expenditures

VARIABLES	(1) Security (10,000 CNY)	(2) Welfare (10,000 CNY)	(3) Education (10,000 CNY)	(4) Agriculture (10,000 CNY)
Golden Shield×Time	2,880***	-3,075***	-432.5	-993.2**
	(679.5)	(1,171)	(1,191)	(430.7)
Time: Post-2005	-1,195***	1,266**	1,816***	1,356***
	(301.9)	(520.1)	(528.9)	(191.3)
Total Expenditure (10,000 CNY)	0.0731***	0.0379***	0.131***	0.0292***
	(0.000229)	(0.000394)	(0.000401)	(0.000145)
Population (10,000 Person)	-1.377***	3.608***	32.77***	0.817***
	(0.223)	(0.383)	(0.390)	(0.141)
Urbanization Ratio (0-1)	345.8	992.5**	-3,131***	-1,089***
	(279.1)	(480.9)	(489.1)	(176.9)
County FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Constant	-379.2*	115.8	3,416***	1,345***
	(214.2)	(369.1)	(375.3)	(135.8)
Observations	19,755	19,755	19,755	19,755
R-squared	0.923	0.567	0.950	0.839

Seemingly unrelated regressions with two-way fixed effects. Robust standard errors clustered by prefectures.
*** p<0.01, ** p<0.05, * p<0.1

4.4 Alternative Mechanisms

One concern is that digital surveillance lowers the thresholds of what dictators perceive to be radical opponents to include individuals' browsing preferences. This means increased repression might result from "redefining" radicals instead of detecting radicals. However, information on

Table 4: "3111" Initiative (2nd-Phase GSP) and Public Goods Provision

	Welfare			Education		Agriculture		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Welfare Centers (log)	Welfare Beds (log)	Hospital Beds (log)	Primary Stdts (log)	Middle Stdts (log)	Grain Prodct (log)	Cotton Prodct (log)	Oil Crop Prodct (log)
"3111" County	0.00614	-0.00352	-0.0243**	0.00643	0.00143	-0.047***	-0.121*	-0.108***
	(0.0241)	(0.0326)	(0.0121)	(0.00935)	(0.0124)	(0.0111)	(0.0716)	(0.0263)
Land Area(Log)	-0.0625	0.00403	0.0244	0.0589**	0.0317	0.0745	0.276	0.0632
	(0.0806)	(0.0641)	(0.0225)	(0.0266)	(0.0307)	(0.0514)	(0.183)	(0.0910)
Population(Log)	0.140*	0.312**	0.309***	0.447***	0.503***	0.124**	0.349*	0.599***
	(0.0760)	(0.125)	(0.0718)	(0.0813)	(0.0892)	(0.0526)	(0.211)	(0.150)
1st Industry(Log)	0.0405	0.117**	-0.0129	-0.09***	-0.0130	0.319***	0.266**	0.253***
	(0.0420)	(0.0541)	(0.0195)	(0.0210)	(0.0220)	(0.0516)	(0.113)	(0.0685)
2nd Industry(Log)	0.0652**	0.0722**	0.036***	-0.041***	0.0137	0.0152	0.226***	0.128***
	(0.0265)	(0.0322)	(0.0102)	(0.00878)	(0.0150)	(0.0103)	(0.0687)	(0.0290)
Fiscal Bgt.(Log)	0.0355	0.114***	0.055***	0.026***	0.089***	-0.00566	0.0586	0.134***
	(0.0273)	(0.0383)	(0.0143)	(0.0101)	(0.0153)	(0.0123)	(0.0906)	(0.0328)
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.649	1.129	4.284***	9.376***	6.923***	7.113***	-3.407*	-0.0467
	(0.912)	(1.064)	(0.378)	(0.400)	(0.426)	(0.632)	(1.937)	(0.998)
Observations	19,474	19,447	20,065	20,107	20,099	18,361	8,542	19,275
R-squared	0.018	0.351	0.613	0.234	0.258	0.118	0.104	0.045
N of County	2,081	2,081	2,086	2,086	2,086	2,066	1,093	2,041

Two-way fixed effects models. Robust standard errors clustered by counties.
*** p<0.01, ** p<0.05, * p<0.1

the prisoners arrested for their online activities in the CECC-PPD database suggests that most of them used the Internet as a platform for anti-regime mobilization, indicating they are likely radical regime opponents. I further exclude all internet-related arrests (about 10.6%) from the sample. The results remain large and statistically significant at 0.01 level (See Online Appendix B.6 for details).

If surveillance allows the government to more efficiently govern by improving redistribution accuracy and monitoring local officials, then increased surveillance should mean a fixed amount of government spending on local public goods will yield more gains in citizen welfare and more gains in social stability. Thus, while surveillance may allow the government to substitute repression for (costly) local public goods provision, a decline in spending on public goods provision could simply

be the result of more efficient governance brought about by increased surveillance. However, we can rule out this mechanism for several reasons. First, my empirical strategy is based on the Golden Shield project that is designed for police departments to monitor citizens. It is unlikely that this surveillance system improves government efficiency in delivering public goods. Second, the 3111 Initiative (the 2nd phase of the GSP) is designed to install street camera systems in cities. It is unlikely that city camera systems improve redistribution accuracy in rural areas, and thus influence agricultural production (Table 4, Column [6], [7], and [8]).

A reduction in welfare spending after surveillance might also result from local government budget constraints: they spend more money on repression with less left for redistribution. However, if this were the case, we would observe an accounting balance between increases in repression spending and decreases in co-optation. As shown in Table 3, the reduced co-optation spending is about 60 percent larger than the increased security spending, suggesting that the tradeoff between repression and co-optation is unlikely to stem from local governments' accounting balance. A final concern is that increased surveillance might better inform a government strategy of individual-level targeted benefits, which might influence welfare expenditure. Though I cannot rule out this possibility, the increased political arrests with surveillance suggest that governments use more targeted repression.

5 Conclusion

Today, roughly half of the world's population lives under some form of non-democratic government. Although social scientists, policymakers, and human rights advocates celebrated the dawn of the Internet Era in the hopes that better communication technology would become a powerful tool to ensure and encourage freedom and democracy (Diamond 2010), we have not observed widespread authoritarian collapse in the two decades since the advent of this era. In this paper, I argue that digital surveillance empowers dictators by resolving their information problem in identifying radical opponents and enabling them to substitute preventive repression for co-optation to prevent social unrest. The evidence from analyzing the effects of the Golden Shield Project on

Chinese subnational governments' repression and co-optation is consistent with this theory.

The findings of this paper are based on the practices of surveillance and repression conducted by the police in a non-electoral dictatorship, but the theory developed here is generalizable to other types of surveillance and repression as well as to electoral autocracies. For example, in 2004, the Hugo Chávez regime in Venezuela was able to identify several million voters who had attempted to remove him from office. The regime then used a user-friendly software program known as Maisanta to distribute the list throughout the government bureaucracy for punishment. Scholars find that voters who were identified as Chávez opponents experienced a 5 percent drop in earnings and a 1.3 percentage point drop in employment rates after the voter list was released (Hsieh et al. 2011). This case suggests that dictators have many ways to surveil and repress opponents, and surveillance can be effective when it is used to collect individual-level information on citizens' political preferences during elections.

The theory and findings in this paper have several implications for explaining authoritarian survival in the digital age. First, digital surveillance helps dictators repress radical opponents to forestall large-scale social protests that might threaten their survival. Second, contemporary dictatorships seldom employ large-scale in-person surveillance (e.g., Stasi in East Germany) to control society due to its detrimental effects on social trust and economic growth (Blaydes 2018) as well as the costs and risks of creating and managing a large bureaucracy comprised of human agents who may threaten the dictator's survival. Digital surveillance relies less on human agents and can more easily be disguised. The reduced social and political costs facilitate using surveillance to control society in dictatorships.[25] Moreover, surveillance enables dictators to use preventive repression to address threats before overt mass protests mobilize, which reduces the necessity to engage in reactive, often more violent, state-led repression. As a consequence, surveillance likely reduces dictators' reliance on military or paramilitary forces for retaining power and therefore lowers the risk of military takeover. Future research in these directions is worth exploration.

[25]Digital surveillance is quicker and more comprehensive than human-agent based surveillance, and thus may be more accurate if human agents are subjective or even malicious against "innocent" people.

References

Angrist, Joshua D. and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-differences Estimates?" *The Quarterly Journal of Economics* 119(1):249–275.

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

Davenport, Christian. 2007. "State Repression and Political Order." *Annual Review of Political Science* 10:1–23.

De Mesquita, Bruce Bueno, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. MIT Press.

Diamond, Larry JF. 2010. "Liberation Technology." *Journal of Democracy* 21(3):69 – 83.

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.

Ding, Zhaowei. 2010. "The Progress of National-wide "Safe City" Projects in 2010." *China Public Security* 04.

Du, Xinwei. 2013. Design and Implementation of Taiyuan Comprehensive Information Platform for a Bureau of Police (in Chinese). Master's thesis. University of Electronic Science and Technology of China. Available at http://cdmd.cnki.com.cn/Article/CDMD-10614-1013330231.htm. Accessed July 4, 2019.

Edin, Maria. 2003. "State Capacity and Local Agent Control in China: CCP Cadre Management From a Township Perspective." *The China Quarterly* 173:35–52.

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

Frantz, Erica and Andrea Kendall-Taylor. 2014. "A Dictator's Toolkit: Understanding How Cooptation Affects Repression in Autocracies." *Journal of Peace Research* 51(3):332–346.

Gandhi, Jennifer and Adam Przeworski. 2006. "Cooperation, Cooptation, and Rebellion Under Dictatorships." *Economics & Politics* 18(1):1–26.

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

Gellman, Barton and Laura Poitras. 2013. "U.S., British Intelligence Mining Data from Nine U.S. Internet Companies in Broad Secret Program." *The Washington Post* . July 03.

Gershenson, Dmitriy and Herschel I Grossman. 2001. "Cooption and Repression in the Soviet Union." *Economics & Politics* 13(1):31–47.

Gohdes, Anita R. 2014. Repression in the Digital Age: Communication Technology and the Politics of State Violence. Phd dissertation. Universität Mannheim. Available at: https://ub-madoc.bib.uni-mannheim.de/37902/. Accessed July 4, 2019.

Gueorguiev, Dimitar D. 2017. "In Public or in Private: Repressing Dissidents in China." Working Paper. Available at SSRN: https://ssrn.com/abstract=3038812. Accessed July 4, 2019.

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

Guo, Gang. 2009. "China's Local Political Budget Cycles." *American Journal of Political Science* 53(3):621–632.

Hazlett, Chad and Yiqing Xu. 2018. "Trajectory Balancing: A General Reweighting Approach to Causal Inference with Time-Series Cross-Sectional Data." Working Paper, UCLA and UCSD. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3214231. Accessed July 4, 2019.

Hsieh, Chang-Tai, Edward Miguel, Daniel Ortega and Francisco Rodriguez. 2011. "The Price of Political Opposition: Evidence from Venezuela's Maisanta." *American Economic Journal: Applied Economics* 3(2):196–214.

Huang, Haifeng, Serra Boranbay-Akan and Ling Huang. 2019. "Media, Protest Diffusion, and Authoritarian Resilience." *Political Science Research and Methods* 7(1):23–42.

Jia, Junxue, Qingwang Guo and Jing Zhang. 2014. "Fiscal Decentralization and Local Expenditure Policy in China." *China Economic Review* 28:107–122.

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.

Kitschelt, Herbert and Steven I Wilkinson. 2007. *Patrons, Clients and Policies: Patterns of Democratic Accountability and Political Competition*. Cambridge University Press.

Knutsen, Carl Henrik and Magnusn Rasmussen. 2018. "The Autocratic Welfare State: Old-Age Pensions, Credible Commitments, and Regime Survival." *Comparative Political Studies* 51(5):659–695.

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

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.

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

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

Lust-Okar, Ellen. 2005. *Structuring Conflict in the Arab World: Incumbents, Opponents, and Institutions*. Cambridge University Press.

Lynch, Marc. 2011. "After Egypt: The Limits and Promise of Online Challenges to the Authoritarian Arab State." *Perspectives on Politics* 9(2):301–310.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and Its Demise in Mexico*. Cambridge University Press.

Martina, Michael. 2014. "China Withholds Full Domestic-security Spending Figure." *Reuters* .

March 04.

Miller, Michael K. 2015. "Elections, Information, and Policy Responsiveness in Autocratic Regimes." *Comparative Political Studies* 48(6):691–727.

Montinola, Gabriella, Yingyi Qian and Barry R Weingast. 1995. "Federalism, Chinese Style: The Political Basis for Economic Success in China." *World Politics* 48(1):50–81.

Moses, Asher. 2011. "Fighting China's Golden Shield: Cisco Sued Over Jailing and Torture of Dissidents." *The Sydney Morning Herald* . August 16.

Nafziger, E Wayne and Juha Auvinen. 2002. "Economic Development, Inequality, War, and State Violence." *World Development* 30(2):153–163.

Oi, Jean C. 1995. "The Role of the Local State in China's Transitional Economy." *The China Quarterly* 144:1132–1149.

Oliver, Pamela E and Gerald Marwell. 1988. "The Paradox of Group Size in Collective Action: A Theory of the Critical Mass. II." *American Sociological Review* 53(1):1–8.

O'Donnell, Guillermo, Philippe C Schmitter and Laurence Whitehead. 1986. *Transitions from Authoritarian Rule: Comparative Perspectives*. Vol. 3 JHU Press.

Pan, Jennifer. 2015. Buying Inertia: Preempting Social Disorder with Selective Welfare Provision in Urban China. Phd dissertation. Harvard University. Available at: https://dash.harvard.edu/handle/1/17467174. Accessed July 4, 2019.

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.

Ritter, Emily Hencken and Courtenay R. Conrad. 2016. "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression." *American Political Science Review* 110(1):85–99.

Scoggins, Suzanne E. 2018. "Policing Modern China." *China Law and Society Review* 3(2):79–117.

Shapiro, Jacob N and David A Siegel. 2015. "Coordination and Security: How Mobile Communications Affect Insurgency." *Journal of Peace Research* 52(3):312–322.

Siegel, David A. 2011*a*. "Non-disruptive Tactics of Suppression Are Superior in Countering Terrorism, Insurgency, and Financial Panics." *PloS One* 6(4):e18545.

Siegel, David A. 2011*b*. "When Does Repression Work? Collective Action in Social Networks." *The Journal of Politics* 73(4):993–1010.

Stokes, Susan C. 2005. "Perverse Accountability: A Formal Model of Machine Politics With Evidence from Argentina." *American Political Science Review* 99(3):315–325.

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.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule*. Cambridge University Press.

Truex, Rory. 2019. "Focal Points, Dissident Calendars, and Preemptive Repression." *Journal of Conflict Resolution* 63(4):1032–1052.

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.

Way, Lucan A. and Steven Levitsky. 2006. "The Dynamics of Autocratic Coercion After the Cold War." *Communist and Post-Communist Studies* 39(3):387–410.

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

Online Appendix

To Repress or To Co-opt? Authoritarian Control in the Age of Digital Surveillance

Xu Xu

xux112@psu.edu
https://xu-xu.net

Contents

A Data and Summary Statistics **1**
A.1 Public Security Expenditure Measure 1
A.2 Political Prisoner Measure 2
A.3 Summary Statistics 3
A.4 Golden Shield Award Prefectures and "3111" Initiative Pilot Counties 4

B Additional Empirical Analyses **5**
B.1 Public Security Expenditure Data 5
B.2 Political Prisoner Data 6
B.3 Currently Detained Prisoners with Arrest Information Searched Online 9
B.4 Trajectory Balancing and Matching for Security Spending Data 9
B.5 DiD Falsification Tests 10
B.6 Alternative Mechanisms 11

C A Formal Model **15**
C.1 Payoff Structure 15
C.2 Timing of the Game 18
C.3 Analysis 18
C.4 Discussion 19

A Data and Summary Statistics

A.1 Public Security Expenditure Measure

I use local public security expenditure as a measure for preventive repression because Chinese local governments intensively use local police force, procuratorates, and courts for repression. In China, Career advancement is one of the most important concerns for local officials (e.g., Edin 2003). Among all criteria for cadre evaluation, local instability, such as large-scale protests and riots, serves as "one-strike" veto to local officials' promotion (Edin 2003; Li 2014). Thus, local officials have a strong incentive to repress dissidents, petitioners, and protesters to prevent mass demonstrations (Li 2014). During my fieldwork in China in 2015, a local Street Office Director stated that their office often helps local police prevent petitioners and protesters from congregating on the street or traveling to the province or Beijing to appeal to upper-level governments. Further, local security officials boost preventive coercion during sensitive times, such as the Tiananmen Square Memorial Day (Truex 2019).

Archival documents show that police spending usually constitutes more than 60 percent of total public security expenditure, far greater than other repressive bodies' spending combined (e.g., Yin and Shandong Provincial Department of Finance 2011). Importantly, local public security expenditure does not include the spending on the Armed Police, allowing the following tests to isolate investment in preemptive repression conducted by local police rather than mass repression perpetrated by the armed police at the behest of provincial governments or the central government.

Abundant archival evidence suggests that local public security expenditure is used for maintaining social stability (Zhou and Anhui Provincial Department of Finance 2006). For example, the Anqin City Bureau of Finance mentions in their 2005 annual report that "... we further increased *the 'stability maintenance funding'* by providing an annual public security expenditure of 17,050,000 Yuan (about 2.12 million USD in 2005), which is an increase of 17.6 percent from 2004. This funding effectively helped maintain social stability and build a harmonious society..." (Zhou and Anhui Provincial Department of Finance 2006).

Further, local public security expenditure does not include spending on the GS Project, which is mainly funded by the central government through the National Planning Commission (currently, the National Development and Reform Commission).[1] In addition, evidence from provincial fiscal yearbooks shows that the expense of Golden Shield constitutes less than 0.5 percent of provincial public security expenditure (e.g., 0.4 percent in Gansu Province and 0.3 percent in Shandong Province in 2010). And data from prefecture fiscal yearbooks further shows that prefecture public security expenditure does not include the expense of the local GS system. See, Gansu Provincial Department of Finance (2011), Yin and Shandong Provincial Department of Finance (2011), and Bai and Heichi City Bureau of Finance (2014).

One concern is that cross-sectional measurement errors may occur because local governments' capacities to fund their coercive agents vary across cities and regions (Greitens 2017). However, most of scholars agree that the Chinese Communist Party has indeed attempted to strengthen its coercive capacity in recent decades, as reflected by an increase in security expenditure (Greitens

[1] China.com (2003), The Golden Shield Project, http://www.china.com.cn/chinese/zhuanti/283732.htm. Accessed July 4, 2019.

2017), the empowering of the public security chiefs (Wang 2014), and a strong correlation between these two developments (Wang and Minzner 2015). Thus, temporal changes in security funding over time within counties can reflect changes in local repression. Since the DiD approach in my paper exploits temporal variation rather than cross-sectional variation, it is less prone to measurement errors caused by local variation in funding capacity.

A.2 Political Prisoner Measure

The date of detention started in 1981 but the data quality appears worse for early years because the Commission started to record Chinese prisoners in 2004 and data entries in earlier years were retrospective. To get a longer pretreatment period and avoid using early-year entries, my empirical strategy for this prisoner data is based on the second phase of the Golden Shield project finished in 2012, whereas the first phase of the project was finished in 2005. Considering the possibility that the effect of the first-phase Golden Shield project could contaminate the effect of the second-phase project, I exclude all entries prior to 2006.

Between 2006 and 2017, there were 5,007 entries in the CECC-PPD. 52 percent of all arrests are related to association, 62 percent of them are related to speech, and 11 percent are related to spread of information (categories are not mutually exclusive). Excluded double counted entries, 82 percent of all arrests are related to association, speech, and spread of information that often facilitates protests. The rest 18 percent of arrests are related to purely religious and ethnic issues, which often trigger protests or even violent actions. Thus, the political prisoner measure well captures local preventive repression in China.

One may be concerned that reporting of prisoners can be affected by investments in public security or in the Golden Shield Project since political imprisonment is a sensitive topic where media report is suppressed in China (but many NGOs and human right groups still make great efforts to collect and report information on political imprisonment) (Gueorguiev 2017). However, if the reporting of imprisonment is affected by investments in local public security, counties with more security spending will have stronger media repression and hence less reporting on political arrests, which will bias the DiD estimate downward. In other words, the positive, significant effect of surveillance on political prisoners will be underestimated. Thus, removing reporting bias can make the results even stronger. I further address the concern of under-reporting in Section B.6.

In the main text, I code the number of county-level political prisoners using the locations of detention complimented by their counties of residence. Though most of the dissidents are detained in prisons or detention centers close to the places where they were arrested and live, it is possible that the places of arrest are different from the places of detention. To address this concern, I focus on a subsample of 1,414 dissidents currently in prisons and search every prisoner's arrest location online because the information of current prisoners are more available online than ex-prisoners. The current prisoner data is obtained from the Political Prisoner Database collected by the Congressional-Executive Commission on China.[2] Among 1,193 prisoners whose locations are identifiable in both the main sample and the subsample, 759 of them (63%) have same locations of arrests and detentions, and 944 of them (80%) are located in same prefectures. I aggregate the number of current prisoners in each county-year unit based on the searched locations of arrests. Appendix B.3 reports the results obtained from using this current prisoner measure. The estimates

[2]Data available at: https://www.cecc.gov/resources/political-prisoner-database.

remain statistically significant at 0.01 level.

A.3 Summary Statistics

Table A.1: Summary Statistics, Public Security Dataset

Variable	Unit	Obs	Mean	Std. Dev.	Min	Max
Expenditure - Public Security	10,000 Yuan	36,185	1,641.9	3,453.7	0	139,932
Expenditure - Social Welfare	10,000 Yuan	25,911	1,530.9	4,405.9	0	265,206
Expenditure - Education	10,000 Yuan	36,247	6,161.4	8,297.1	0	331,798
Expenditure - Agriculture	10,000 Yuan	27,426	797.9	1,511.0	0	44,578
Expenditure - Administration	10,000 Yuan	28,376	2,598.0	3,220.1	17	163,536
Expenditure - Total	10,000 Yuan	36,282	26,259.9	47,398.2	162	2,225,040
GDP	10,000 Yuan	26,929	364,974.0	652,030.2	0	29,200,000
Population	10,000 Persons	37,018	45.1	48.1	0	5,022
Urbanization Ratio	0-1	35,722	0.25	0.24	0	1
Golden Shield	(County-year)	37,349	0.10	0.30	0	1
Year	(Year)	37,349	2001	4.0	1994	2007

Table A.2: Summary Statistics, Political Prisoners Dataset

Variable	Unit	Obs	Mean	Std. Dev.	Min	Max
Arrests of Political Activists	Person	34,325	0.132	1.630	0	165
Surveillance Camera System	(County-year)	35,275	0.139	0.346	0	1
Year	(Year)	35,275	2011	3.4	2006	2017
Primary Industry Value Added	10,000 Yuan	20,130	177,816.0	160,460.0	84	1,399,516
Secondary industry Value Added	10,000 Yuan	20,132	624,798.1	1,064,072.0	13	17,000,000
Population	10,000 Persons	20,154	70.0	1,456.1	1	115,453
Land Area	Square km	20,150	4,283.3	9,970.2	56	202,298
Expenditure - Total	10,000 Yuan	20,143	168,598.2	168,685.4	214	3,889,833
Welfare Center	Number	19,510	16.1	24.0	0	1,474
Welfare Bed	Bed	19,483	1,164.3	1,461.1	0	20,790
Hospital Bed	Bed	20,112	1,269.6	1,183.0	10	37,846
Elementary School Students	Person	20,153	35,632.6	33,324.9	482	1,600,272
Middle School Students	Person	20,140	25,888.0	22,265.2	106	224,076
Grain	Ton	18,392	256,490.7	287,939.5	7	3,349,885
Cotton	Ton	8,552	6,948.3	21,132.6	1	390,177
Oil Crop	Ton	19,312	14,653.3	26,095.5	1	381,336

A.4 Golden Shield Award Prefectures and "3111" Initiative Pilot Counties

Awarded Prefectures for the 1st phase of the GS Project

In 2006, the Ministry of Public Security awarded about 40 prefectural-level Bureaus for their excellent work in completing the first phase of this project. As shown in Figure A.1, the distribution of these prefecture is even in China.

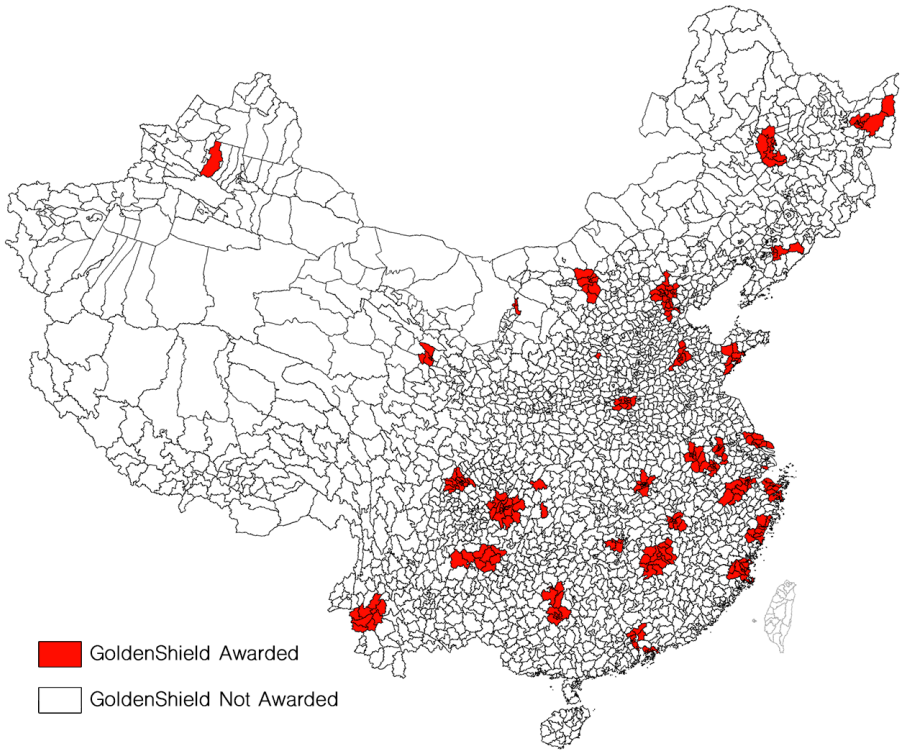

Figure A.1: The Distribution of Golden Shield Award Prefectures
Source: Ministry of Public Security Official Website

Criteria for Selecting "3111" pilot counties/cities – 2nd Phase of the GS Project

As discussed in the main text, about 660 pilot cities/counties all over China were selected in three waves to install and operate the "3111" street surveillance camera and alarm systems. The materials collected during my fieldwork also provide some information regarding how those "3111" pilot counties were selected. In particular, in the interviews with the managers of several security product companies conducted by the China Public Security journal, one manager mentioned that "many '3111' projects were constructed based on the current telecommunication networks provided by major Telecom operators. Thus, it is very important for us to cooperate with those operators"(Ding 2007). Because the *Notice on Construction of City Security Alarm System (2005)* states that "Pilot counties start first ... provinces should promote security alarm system in pilot areas with *suitable conditions* and use those areas as examples for other cities/counties to replicate the security alarm system", we can infer that areas with *suitable conditions* are counties with better telecommunication infrastructures, better economic conditions, and larger populations, etc. We can also infer that the selected pilot counties are evenly distributed within China's 31 provinces to serve as examples for other counties to replicate their success, which means the endogeneity between repression and surveillance should not be a major concern. Figure A.2 shows the distribution of the pilot counties/cities.

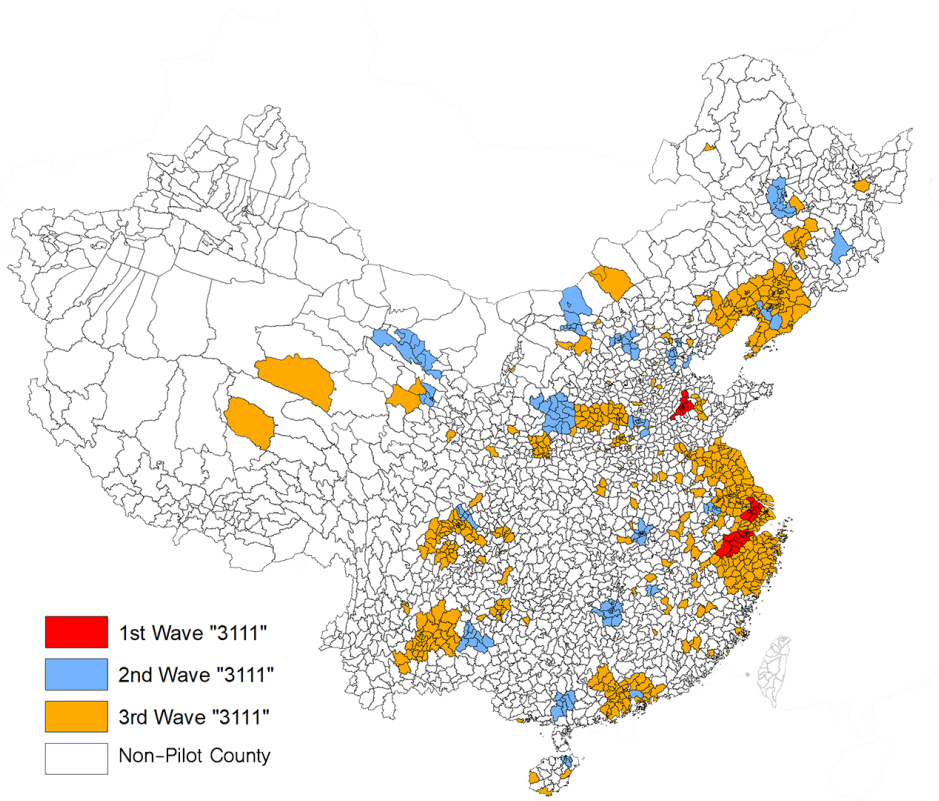

Figure A.2: The Distribution of "3111" Pilot Counties/Cities
(Xinjiang, Tibet, and Beijing Excluded)
Source: Ministry of Public Security No.1 Research Institute

Correlation Matrix for Past Repression and Surveillance Counties Selection

Table A.3 shows that past public security expenditure and past political prisoners are not correlated with the selection of Golden Shield prefectures and the 3111 counties.

Table A.3: Correlations Between Past Repression and Selections of Surveillance Counties

	Public Security Expenditure		Political Arrests	
	1-year Lag	2-year Lag	1-year Lag	2-year Lag
Golden Shield Counties	0.0479	0.0471	-	-
3111 Initiative Counties	-	-	0.0165	0.0133

B Additional Empirical Analyses

B.1 Public Security Expenditure Data

I include control variables such as local GDP and local governments' other administrative expenditure. The former control for local economic development level, although this is addressed by controlling for total government expenditure and urbanization ratio in the main models; the latter addresses measurement problem since non-repressive operating expenses of local public security bureaus are highly correlated with other administrative expenditure of local governments. In Table 1 of the main text, the number of observations are different between Fixed Effect models and

Lagged DV models because the lags automatically drop observations in the first year. Table B.4 shows that the results remain similar and robust when matching the sample of the Fixed Effect model with the Lagged DV model. Note that I do not restrict all the model specifications on the same sample because the GDP variable is missing for all counties/cities in 1994, 1995, 1996, and 2007, and the administrative expenditure variable is missing for all counties/cities in 2004 and 2005.

Table B.4: 1st-Phase GSP and Public Security Expenditure, Sample Size Matched

	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Security (log)	Security (log)	Security (log)	Security (log)	Security (log)	Security (log)
Golden Shield×Time	0.109***	0.0363**	0.0941***	0.0297**	0.121***	0.0338**
	(0.0283)	(0.0141)	(0.0298)	(0.0122)	(0.0362)	(0.0139)
Time: Post-2005	0.228***		0.0366***		0.170***	
	(0.0235)		(0.0129)		(0.0363)	
Lagged DV		0.754***		0.773***		0.755***
		(0.0118)		(0.0126)		(0.0138)
Exp. (log)	0.608***	0.263***	0.567***	0.237***	0.501***	0.223***
	(0.0393)	(0.0132)	(0.0502)	(0.0146)	(0.0339)	(0.0145)
Pop. (Log)	0.0431	-0.00982*	0.0556	-0.0153***	0.0449	-0.00173
	(0.0302)	(0.00536)	(0.0403)	(0.00550)	(0.0324)	(0.00587)
Urb. Ratio	-0.00644	0.0460***	-0.00703	0.0411***	-0.00121	0.0619***
	(0.0129)	(0.00996)	(0.0161)	(0.00974)	(0.0136)	(0.0118)
GDP (log)			0.00162	0.0112***		
			(0.00590)	(0.00373)		
Admin. Exp. (log)					0.135**	0.0465**
					(0.0590)	(0.0215)
County FEs	Yes	No	Yes	No	Yes	No
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Constant	1.004**	-0.590***	1.365***	-0.741***	0.832	-0.771***
	(0.410)	(0.0751)	(0.500)	(0.0785)	(0.532)	(0.103)
Observations	31,632	31,632	24,827	24,827	24,489	24,489
R-squared	0.877	0.953	0.828	0.948	0.843	0.948
N of County	2,728		2,717		2,728	

Model (1), (3), and (5) use two-way fixed effects models with cluster-bootstrap variance matrix developed by Bertrand, Duflo and Mullainathan (2004). Model (2), (4), and (6) use lagged dependent variable models. Robust standard errors clustered at the prefecture level.
*** p<0.01, ** p<0.05, * p<0.1

Table B.5 reports the results on repression-cooptation comparison using log-transformed variables. The results are similar to those of Table 3 in the main text.

B.2 Political Prisoner Data

Additional Controls (Sample with Missing Data)

The models in Table 2 in the main text do not include social-economic control variables because

Table B.5: 1st-Phase GSP, Public Security, and Redistribution Expenditures (Log Transformation)

	Fixed Eff.	Fixed Eff.	Fixed Eff.	Fixed Eff.
	(1)	(2)	(3)	(4)
VARIABLES	Security (log)	Welfare (log)	Education (log)	Agriculture (log)
Golden Shield×Time	0.181***	-0.131**	0.0194	-0.162***
	(0.0226)	(0.0649)	(0.0166)	(0.0496)
Time: Post-2005	-0.0501***	0.107***	0.0413***	-0.0142
	(0.0103)	(0.0297)	(0.00759)	(0.0227)
Total Expenditure (log)	0.971***	0.910***	0.704***	1.102***
	(0.00454)	(0.0130)	(0.00333)	(0.00994)
Population (Log)	0.0599***	0.316***	0.288***	-0.0599***
	(0.00456)	(0.0131)	(0.00335)	(0.00999)
Urbanization Ratio	0.200***	0.171***	-0.228***	-0.811***
	(0.00941)	(0.0270)	(0.00690)	(0.0206)
County Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
Constant	-2.882***	-4.151***	0.543***	-3.806***
	(0.0333)	(0.0956)	(0.0245)	(0.0730)
Observations	19,714	19,714	19,714	19,714
R-squared	0.922	0.705	0.952	0.754

SUR Models. Robust standard errors clustered at the prefecture level.
*** p<0.01, ** p<0.05, * p<0.1

the social-economic data from the County Statistical Yearbooks misses all the county-level districts in cities (about two-fifths of the total county/district-level units, Figure B.3).

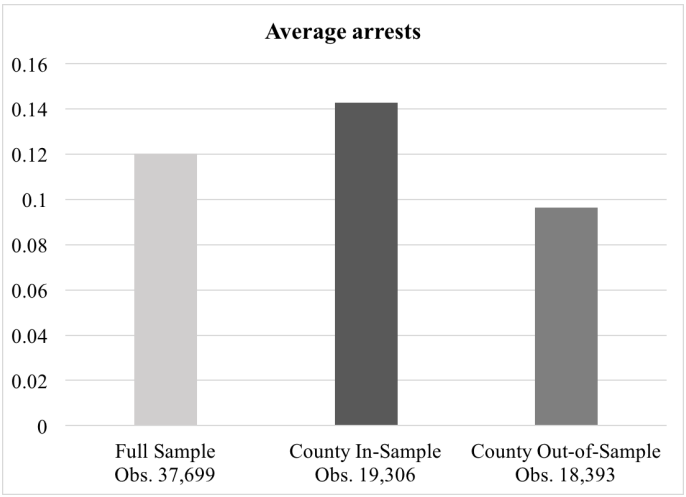

Figure B.3: Average Arrests in Samples With and Without Controls

I further address potential selection biases by controlling for the number of telephone users (a proxy for telcommunication infrastructures), land area, population, primary industry value increase, secondary industry value increase, and local fiscal expenditure. These controls along with the lagged outcome variable model can address potential endogeneity between surveillance and

repression. As Table B.6 shows, the effect of "3111" Initiative on political prisoners remains large and statistically significant, and including the telecommunication variable does not change the magnitude of the estimate much.

Table B.6: "3111" Initiative (2nd-Phase GSP) and Political Prisoners, with Controls

	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV
	(1)	(2)	(3)	(4)
VARIABLES	Prisoners	Prisoners	Prisoners	Prisoners
3111 Pilot Counties	0.0687**	0.0555**	0.0703**	0.0572**
	(0.0268)	(0.0247)	(0.0278)	(0.0245)
Lagged Prisoners		0.131***		0.131***
		(0.0107)		(0.0107)
Tel User (Log)			0.0198	-0.00518
			(0.0376)	(0.0224)
Land Area (Log)	0.0612	0.0571***	0.0693	0.0587***
	(0.0877)	(0.0182)	(0.0937)	(0.0172)
Population (Log)	0.0987	-0.160***	0.0786	-0.153***
	(0.344)	(0.0517)	(0.362)	(0.0540)
1st Industry (Log)	0.0256	0.0308	0.00195	0.0306
	(0.0656)	(0.0247)	(0.0719)	(0.0244)
2nd Industry (Log)	-0.114*	-0.163**	-0.106*	-0.160**
	(0.0671)	(0.0659)	(0.0638)	(0.0664)
Fiscal Income (Log)	0.0415	0.282***	0.0279	0.277**
	(0.0781)	(0.108)	(0.0799)	(0.109)
County Fixed Effect	Yes	No	Yes	No
Year Fixed Effect	Yes	Yes	Yes	Yes
Constant	-0.195	-1.285***	0.720	-1.233***
	(1.265)	(0.469)	(1.566)	(0.465)
Observations	17,918	16,172	17,761	16,023
R-squared	0.002	0.031	0.002	0.031
Number of County	1,847		1,844	

Robust standard errors clustered at the county level. Samples exclude counties in Beijing, Tibet, and Xinjiang.
*** p<0.01, ** p<0.05, * p<0.1

Count Data Models

Due to the panel structure of the prisoner data, I fit linear fixed effects models as main analyses. Since the outcome variable is a count of political arrests with discrete and non-negative values, I also fit the data using Fixed-Effects Negative Binomial models for difference-in-differences estimations and Negative Binomial (NB) models for lagged DV estimations. The effects remain substantial and statistically significant. Moreover, I use Zero-inflated Negative Binomial (ZINB) models for the lagged DV estimation to account for the excessive zeros in the political arrests variable. The results remain large and statistically significant. Surveillance increases political arrests by 1.5 to 1.7 more persons (Table B.7). Given that the average number of prisoners in counties with non-zero arrests is 3.1, 1.5 - 1.7 increase is substantial. Akaike's information criterion (AIC) and Bayesian information criterion (BIC) tests suggest that ZINB models have slightly smaller AICs

and BICs than NB models. Thus, excessive zeros in the dependent variable is not of big concern.

Table B.7: "3111" Initiative (2nd-Phase GSP) and Political Prisoners, Count Data Models

	Fixed Eff. Neg. Bino.	Neg. Bino. Lag-DV	ZINB Lag-DV
	(1)	(2)	(3)
VARIABLES	Prisoners	Prisoners	Prisoners
3111 Pilot Counties	0.415***	0.531***	0.377***
	(0.111)	(0.118)	(0.146)
Lagged Prisoners		0.652***	0.119***
		(0.107)	(0.0261)
County Fixed Effect	Yes	No	No
Year Fixed Effect	Yes	Yes	Yes
Constant	-3.116***	-2.744***	-0.717***
	(0.151)	(0.145)	(0.191)
Observations	9,744	31,460	31,460
Number of County	812		

Robust standard errors are clustered at the county level.
*** p<0.01, ** p<0.05, * p<0.1

"3111" Initiative and Co-optation, Original Units

Table B.8 reports the effect of the "3111" Initiative on public goods provision using the original units of those goods. Note that log transformation in Table 4 in the main text is preferable to using the original unit due to data skewness and large outliers.

B.3 Currently Detained Prisoners with Arrest Information Searched Online

As mentioned in Section A.2, I focus on a subsample of 1,414 dissidents currently under detention and search every prisoner's arrest location online to address the concern that the places of arrest might be different from the places of detention. The results from this new data are very close to the main results. Table B.9 Column (1) reports the difference-in-differences estimates. The result indicates that counties with advanced surveillance camera system currently detain 0.04 more activists than counties without surveillance. The average number of current prisoners per county is about 0.038 persons, so the number of arrests is more than doubled. Column (2) reports the result of the lagged DV model, which is also large and statistically significant. Excluding counties from Beijing, Tibet, and Xinjiang does not change the statistically significance.

B.4 Trajectory Balancing and Matching for Security Spending Data

I use recently developed trajectory balancing and matching approaches to address potential time-varying omitted biases. The trajectory balancing approach (Hazlett and Xu 2018) uses kernel balancing to weight control observations to match the treated observations in terms of the pre-treatment trends (i.e., high-order "trajectory") of outcome and covariates. This approach handles time-vary confounders by assuming some time-fixed linear combination of those time-varying confounders. Because pretreatment outcomes contain information of those confounders, balancing

Table B.8: "3111" Initiative (2nd-Phase GSP) and Public Goods Provision, Original Units

	Welfare			Education		Agriculture		
	Fixed E.	Fixed E.	Fixed E.	Fixed E.	Fixed E.	Fixed E.	Fixed E.	Fixed E.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Welfare Centers	Welfare Beds	Hospital Beds	Primary Stdts	Middle Stdts	Grain Prodct	Cotton Prodct	Oil Crop Prodct
3111 Counties	-0.103	61.95	-14.74	1,276***	323.4	-5,033*	-1,189**	-1,445***
	(0.537)	(47.16)	(25.54)	(459.2)	(365.1)	(2,885)	(522.9)	(336.3)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Eff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Eff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	13.13	357.7	53.97	38,808	29,024	201,328	-28,977	9,215
	(2.536)	(124.3)	(226.4)	(2,272)	(2,034)	(5,932)	(22,386)	(1,395)
Observations	19,477	19,450	20,068	20,109	20,101	18,362	8,544	19,277
R-squared	0.020	0.273	0.572	0.047	0.314	0.153	0.093	0.047
N of admcode	2,081	2,081	2,086	2,086	2,086	2,066	1,093	2,041

Model (1) - (8) use two-way fixed effects models with robust standard errors clustered at the county level.
*** p<0.01, ** p<0.05, * p<0.1

control observations with the treated on pretreatment outcomes helps difference out those time-varying confounders. Table B.10 shows the results from the Trajectory Balancing approach for the security expenditure data. In total, surveillance leads to an 6.6 percent increase in public security spending and the effect is statistically significant at 0.01 level. The yearly average effects show a 5 percent increase in 2006 and an 8.3 increase in 2007. The pretreatment effects are not statistically significant whereas the post treatment effects are significant at 0.01 level.

I further employ a matching method for panel data developed by Imai, Kim and Wang (2018). I select 5 control observations for each treatment observation among observations with identical treatment history for one year ahead, based on the outcome variable, public security expenditure, and counties' population (i.e., controlling for city size). Table B.11 checks the balance between the treated group and the control group in the matched sample. Balance is assessed by taking the difference between the values of the outcome variable in the treated unit and the weighted average of that across all the control units in each matched set, divided by one standard deviation of the values of the outcome variable across all treated units of all matched sets. A smaller value indicates better balance. The results suggests that the treated group and the control group are rather close.

Table B.12 presents the matching results. Comparing with counties without completed GS system, counties with completed GS systems increase public security expenditure by 4.2 percent in 2006 and 6.4 percent in 2007. The effect in 2006 is statistically significant at 0.05 level.

B.5 DiD Falsification Tests

To examine whether other possible (but unidentified) differences might drive the repression level higher in counties with a completed surveillance system, I examine if there exist some pre-trend differences. This exercise provides a general falsification test to check whether the differential changes in repression link to surveillance projects or other confounders. Figure B.4 plots

Table B.9: "3111" Initiative (2nd-Phase GSP) and Political Prisoners, In-prison Sample

	All Provinces		Tibet, Xinjiang & Beijing Excluded	
	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV
	(1)	(2)	(3)	(4)
VARIABLES	Prisoners	Prisoners	Prisoners	Prisoners
3111 Pilot Counties	0.0398***	0.0412***	0.0330***	0.0198**
	(0.0136)	(0.0115)	(0.00914)	(0.00881)
Lagged Prisoners		0.151***		0.0868
		(0.0474)		(0.0534)
County Fixed Effect	Yes	No	Yes	No
Year Fixed Effect	Yes	Yes	Yes	Yes
Constant	0.00165*	0.000837	0.000836	0.000684
	(0.000870)	(0.000622)	(0.000550)	(0.000533)
Observations	34,326	31,460	32,031	29,359
R-squared	0.007	0.028	0.007	0.013
Number of County	2,866		2,672	

Robust standard errors clustered at the county level. The lagged DV model has slightly larger coefficient than the fixed-effect model (Column [1]). This is likely due to the fact that currently detained prisoners only constitute a small subsample of total prisoners, so that outlier repressive counties may have a greater influence on the coefficient. Column (3) and (4) show that the fixed effect model has a larger coefficient than the lagged DV model when excluding repressive counties.
*** p<0.01, ** p<0.05, * p<0.1

the effects of pre- and post-GS project on public security expenditure based on the Trajectory Balancing approach, running from two years ahead and eleven years behind. The estimates show no discernible differences in pretrends.

Figure B.5 plots the effects of pre- and post-3111 Initiative treatments on the number of political prisoners. The estimates show no effects in the six years before the completion of the "3111" Initiative, with sharply increasing effects on arrests in the years following the completion. The common pretreatment trends in both figures provide evidence consistent with the parallel-trends assumption necessary for a causal interpretation of the surveillance estimate.

B.6 Alternative Mechanisms

"Redefining" radicals vs. "identifying" radicals?

It is possible that digital surveillance could lower dictators' repression thresholds to include citizens' browsing preferences. Thus, the positive effect of surveillance on repression could be due to an increase in "defining" radicals instead of improved information for identifying radicals. To address this concern, I examine the 5007 prisoners to identify and exclude all Internet and social media related arrests (hereafter, ISM-related). The CECC-PPD prisoner data includes information on the reasons for imprisonment. I identify ISM-related cases by the following keywords: "internet", "wechat", "qq", "online", "weibo", "msn", "tencent", "skype", "google", "twitter", "facebook", "whatsapp", "messenger", "website", "webpage", "web", "bbs", and "forum", etc. Among the 5007 cases, 10.6% are ISM-related. I then examine the effect of surveillance on political pris-

Table B.10: Trajectory Balancing Approach

Total Average Treatment Effect on the Treated:

ATT	S.E.	N.Treated	N.Control
0.06624***	0.01983	328	2734

Yearly Average Treatment Effect on the Treated:

Year	ATT	S.E.	N.Treated	N.Control
2007	0.083***	0.024	164	1367
2006	0.049***	0.020	164	1367
2005	0.008	0.012	164	1367
2004	0.003	0.012	164	1367
2003	0.018	0.013	164	1367
2002	-0.007	0.013	164	1367
2001	-0.018	0.013	164	1367
2000	0.002	0.013	164	1367
1999	0.032	0.013	164	1367
1998	0.015	0.012	164	1367
1997	0.005	0.011	164	1367
1996	-0.005	0.012	164	1367
1995	-0.001	0.013	164	1367
1994	0.010	0.011	164	1367

Table B.11: Balance Table by Panel Matching Approach

Mean Balance of Security Expenditure

Time to Treatment	-1 year	0 year
Log(Lagged Security Spending) (SD)	0.132	0.183
Log(Population) (SD)	-0.048	0.0452
Log(Budget Expenditure) (SD)	0.103	0.115
Log(Urbanization Ratio) (SD)	0.006	0.014

Table B.12: Estimate of Average Treatment Effect on the Treated (ATT) by Period

	t+0	t+1
Point Estimate(s)	0.042	0.064
Standard Error(s)	0.016	0.039
Lower Limit of 95 % Regular Confidence Interval	0.010	-0.018
Upper Limit of 95 % Regular Confidence Interval	0.071	0.133

Weighted Difference-in-Differences with Mahalanobis Distance. Matches created with 1 lags. Standard errors computed with 500 Weighted bootstrap samples

oners based on a subsample that excludes all ISM-related arrests. As Table B.13 shows, the results based on this subsample remain large, positive, and statistically significant. The significant effect suggests that even if dictators might lower their repression thresholds under digital surveillance, better information for identifying radicals still leads to an increase in repression in dictatorships.

It should be noted that arresting online activists does not necessarily means the government has

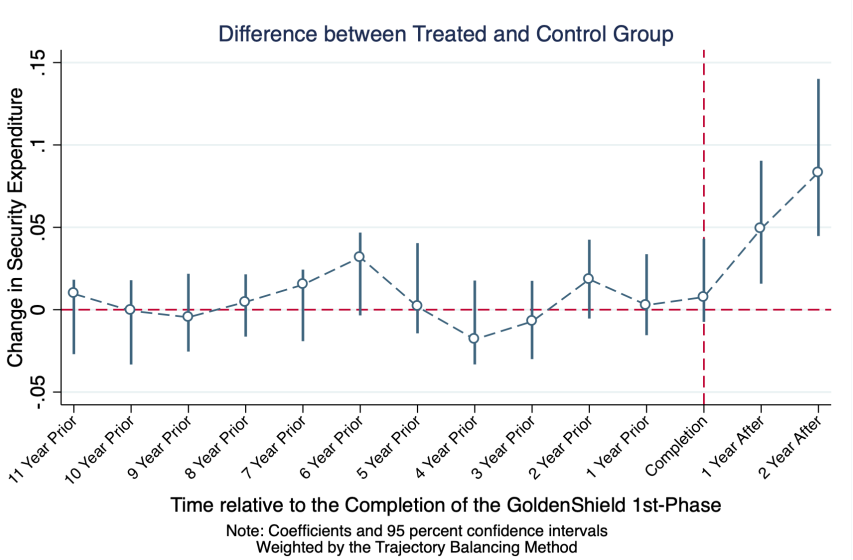

Figure B.4: Pre- and Post-GS Estimates, Public Security Expenditure, DiD, 1994-2007

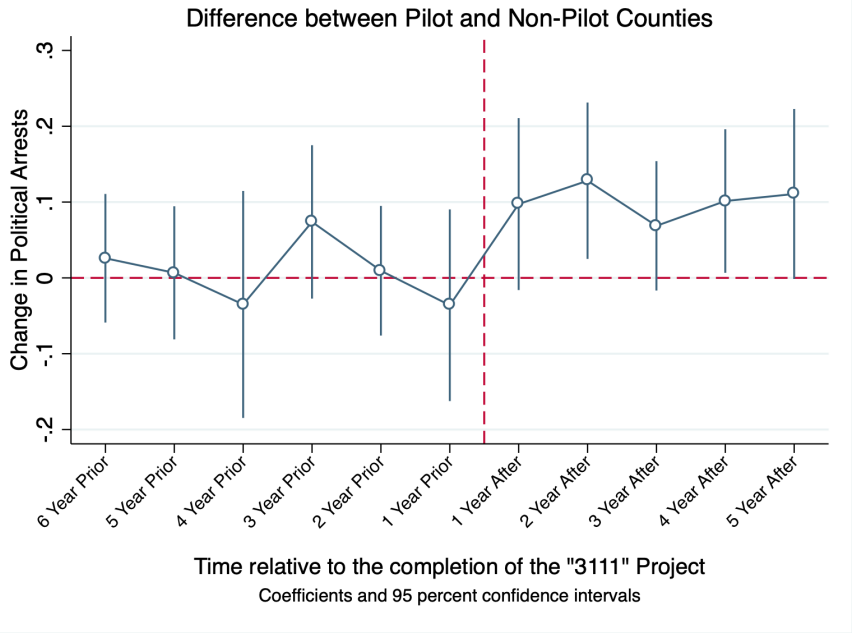

Figure B.5: Pre- and Post-3111 Estimates, Political Prisoners, DiD specification, 2006-2017

lower repression thresholds. Careful examination of the ISM-related cases suggests that most of the online activists used the Internet as a platform for anti-regime mobilization. For example, Ou Quanjiang (Case 2016-00330) along with two others were arrested for "using a WeChat group to organize protests in Ningxiang" during a large June 27 protest in front of Ningxiang government offices against the planned construction of a waste incineration plant.[3] Had the Internet or social media been unavailable, those radicals would use other means to mobilize or just hide. Although

[3]See China's Political Prisoner Database, *Congressional-Executive Commission*. Nov. 5, 2017.

Table B.13: "3111" Initiative (2nd-Phase GSP) and Political Prisoners, No Internet-Related Arrests

	All Provinces		Tibet, Xinjiang & Beijing Excluded	
	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV
	(1)	(2)	(3)	(4)
VARIABLES	Prisoners	Prisoners	Prisoners	Prisoners
3111 Pilot Counties	0.0833***	0.0627***	0.101***	0.0305**
	(0.0273)	(0.0185)	(0.0238)	(0.0127)
Lagged Prisoners		0.150***		0.140***
		(0.0107)		(0.0117)
County Fixed Effect	Yes	No	Yes	No
Year Fixed Effect	Yes	Yes	Yes	Yes
Constant	0.0354***	0.0629***	0.0319***	0.0634***
	(0.00682)	(0.00992)	(0.00635)	(0.0104)
Observations	34,324	31,460	32,030	29,359
R-squared	0.002	0.023	0.002	0.020
Number of County	2,864		2,671	

Robust standard errors clustered at the county level.
*** p<0.01, ** p<0.05, * p<0.1

it is difficult to draw a clear boundary between detecting radicals and "redefining" radicals under digital surveillance, the above evidence suggests that the former mechanism is clearly in play.

More arrests because of increased security expenditure?

One may be concerned that the increase in political prisoners is related to increased security expenditures in the 1st Phase of Golden Shield Project. This is not likely the case since, as discussed in Section A.2, I exclude all prisoners data prior to 2006 when the 1st-phase Golden Shield project had completed. I further address this concern by separating the 2006-2017 sample (with political prisoner measure) by 1st-phase Golden Shield *treated* and *non-treated* counties to examine the effect of the "3111" Initiative on political prisoners in each sample. The logic of this test is that, if the increases in political prisoners were mainly driven by the increase in security expenditure in the 1st phase of the Golden Shield project, we would expect the "3111" Initiative to have little effect in a sample of counties without expenditure increase or in a sample of counties that had all experienced expenditure increase.

Table B.14 shows that a very small proportion of counties had experienced an increase in security expenditures in the 1st phase (315 versus 2,553). It further shows that the effect of the "3111" Initiative on the number of political prisoners is statistically significant across different models and different samples (Column [1] - [4]), which suggests that the increase in political prisoners is unlikely to be driven by the increase in security expenditures during the 1st phase Golden Shield Project.

Under-reporting on political prisoners?

One concern is that the political prisoner data might be under-reported. However, the staggered DiD design based on three waves of "3111" pilot counties are less sensitive to reporting bias: it is unlikely that the systematic bias in reporting, if there were any, would overlap exactly with the multiple time periods and multiple groups of counties to bias the results. I further address

Table B.14: "3111" Initiative (2nd-Phase GSP) and Political Prisoners
(By Treated, Non-treated Counties in the 1st Phase, and GuangDong Province)

	1st Phase GS Counties Excluded		1st Phase GS Counties Only		Guangdong Province	
	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV	Fixed Eff.	Lagged DV
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Prisoners	Prisoners	Prisoners	Prisoners	Prisoners	Prisoners
3111 Pilot Counties	0.0735**	0.0385*	0.106**	0.206***	0.267**	0.292***
	(0.0323)	(0.0209)	(0.0485)	(0.0465)	(0.107)	(0.103)
Lagged Prisoners		0.153***		0.252***		0.132**
		(0.0109)		(0.0622)		(0.0547)
County Fixed	Yes	No	Yes	No	Yes	No
Year Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.0364***	0.0600***	0.0519***	0.0992***	0.131**	0.0564*
	(0.00759)	(0.0104)	(0.0194)	(0.0345)	(0.0652)	(0.0305)
Observations	30,581	28,028	3,747	3,432	1,464	1,342
R-squared	0.002	0.024	0.020	0.087	0.024	0.046
N of County	2,553		315		122	

Robust standard errors clustered at the county level.
*** p<0.01, ** p<0.05, * p<0.1

this under-reporting concern by analyzing the effect of surveillance on political prisoners using a sample of counties in Guangdong Province only. Adjacent to Hong Kong and Macau, Guangdong Province is known for its economic prosperity and its openness to international influences. Due to the impact of NGOs and activists from nearby Hong Kong, voluntary organizations in Guangdong have developed earlier and more rapidly than in any other areas of China. More importantly, many Hong Kong-based organizations closely observe human rights violations in Guangdong Province. Thus, the under-reporting problem is likely less severe in Guangdong than in the entire country. Column (5) and (6) in Table B.14 show that the effect of surveillance on political prisoners is about 2 times larger than that from the full sample. Although this larger effect could be due to heavier repression in Guangdong, the statistical significance suggests that the effect is robust when there are less reporting biases.

C A Formal Model

I use a dynamic game with incomplete information to explain the repression-cooptation tradeoff under information asymmetries. Although simple in structure, it confirms that dictators will choose repression over co-optation when their information on citizen types improves.

C.1 Payoff Structure

Citizen Payoffs without Dictator's actions

The game has four players. There are three citizens, two moderates and one radical, who can challenge a dictator by mounting a protest. This setup captures the fact that moderates often largely outnumber radicals in a society. Citizens differ in their anti-regime sentiments: a radical

revolutionary – an individual with the stronger anti-regime sentiment – obtains greater value from challenging the dictator than a moderate, even if she challenges alone. That is $v_r > v_m$ and $p_r v_r > 0$, where p_r denotes the probability of success if the radical challenges alone. If a citizen does not challenge, she receives a payoff l, $l < v_m$. Assume that a successful protest needs the participation of the radical. Thus, if one or two moderates challenge without the radical, moderates' payoff is 0. I assume that the radical always want to challenge to rule out the possibility of multiple equilibria.[4] That is, $p_r v_r > l$. Thus, in the game without the dictator's actions, the only Nash equilibrium is (challenge, challenge, challenge). Figure A.1 shows the payoffs in this game.

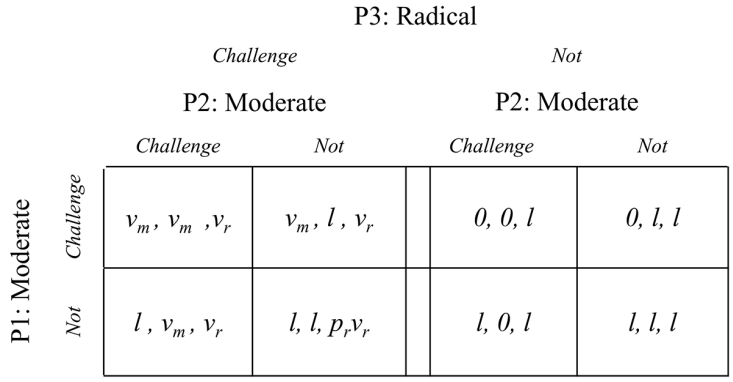

Figure C.1: Citizen Payoffs without Dictator's actions

Citizen Payoffs under Universal Co-optation

I assume that co-optation as public or club goods can distribute benefits w to all three opponents. Following the dictator's co-optation strategy, the moderate has no incentive to challenge the dictator. That is, $l + w > v_m$. I further assume that providing welfare to the radical reduces her sentiment so that she will not challenge alone: $v_r > l + w > p_r v_r$.[5] Thus, the unique Nash equilibrium in this coordination game under universal co-optation is (not challenge, not challenge, not challenge). Figure A.2 shows the payoffs in this game.

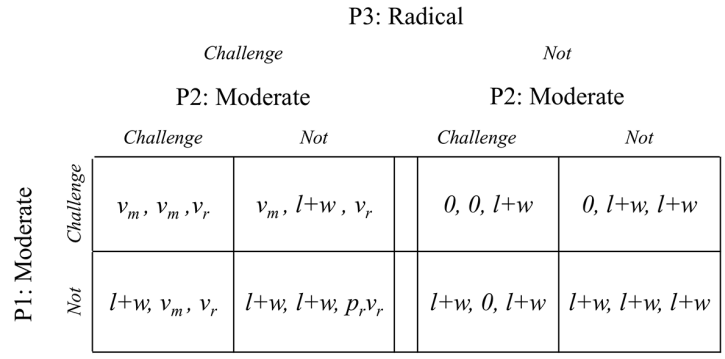

Figure C.2: Citizen Payoffs under Universal Co-optation

Citizen Payoffs when Repressing a Moderate

[4]Otherwise, this game will be similar to the Battle of the Sexes game.

[5]We can relax this assumption without changing the properties of the entire game.

Assume that preventive repression only targets one citizen by imposing cost r to the citizen, and repressing any individual will prevent her from challenging the dictator: $r > v_r$. If the dictator represses a moderate, the other moderate and the radical can still launch a protest to challenge the dictator. The unique Nash equilibrium in this game is (not challenge, challenge, challenge), as shown in Figure A.3.[6]

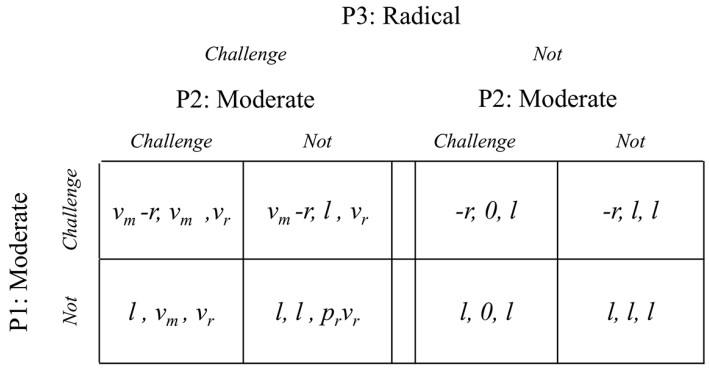

Figure C.3: Citizen Payoffs When Repressing a Moderate

Citizen Payoffs when Repressing a Radical

If the dictator accurately represses a radical ($v_r - r < l$), two moderates cannot launch a protest since as we assumed above, a successful protest requires the participation of the radical. The unique Nash equilibrium in this game is (not challenge, not challenge, not challenge), as shown in Figure A.4.

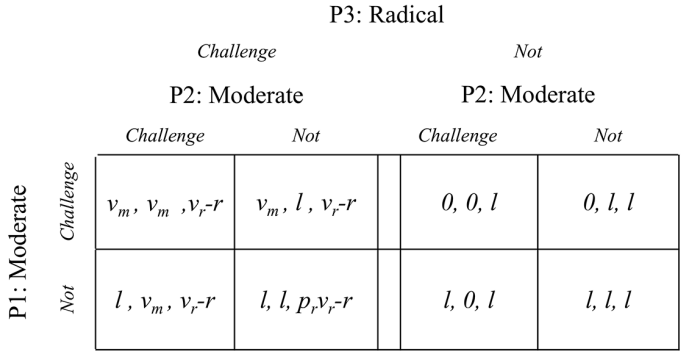

Figure C.4: Citizen Payoffs When Repressing the Radical

Note that in the above games citizens do not receive w when they challenge under co-optation, nor do they bear the cost of r when they do not challenge under repression. These payoff structures make not-challenge a dominant strategy for any repressed or co-opted citizens. This does not mean that the dictator makes co-optation or repression decision after observing citizens' actions. We can consider such payoff changes as dictators' preventive actions: harassing or co-opting a citizen before the stage of protest makes protesting a dominated strategy for this individual so that she

[6]Repressing another moderate is symmetric to this game.

will not participate during the protest.[7]

Dictators' Payoffs

The dictator bears the cost of repression or co-optation and suffers a loss from citizens' challenge. For simplicity, we assume that the loss is equal to citizens' gains from challenging. Assume that the cost of providing universal welfare to all three opponents is greater than the cost of repressing one opponent: $c_w > c_r$. If the dictator uses co-optation, no citizens will challenge him. Thus, the dictator's payoff is $-c_w$. If the dictator represses a moderate, he suffers the cost of repression and a loss from a challenge launched by one moderate and one radical: $-c_r - v_m - v_r$. If the dictator represses the radical, no challenge is possible, so that his payoff is $-c_r$. $c_w < c_r + v_m + v_r$. Thus, for a dictator with perfect information, repressing the radical is the optimal strategy.

C.2 Timing of the Game

The timing of the game is as follows. First, the dictator chooses co-optation or repression before the stage of citizen coordination. If he co-opts citizens, citizens then simultaneously decide to challenge him or not. If he represses one citizen, nature then determines the probability of accurately targeting the radical opponent, denoted by q, $q \in [0, 1]$. This follows by citizens' simultaneous moves on challenging him or not. Figure A.5 shows the sequence of the game. Following node 1, 2, 3, and 4 are the simultaneous games discussed above. Note that even if the random probability of repressing a radical is $1/3$, q can be smaller than $1/3$ give that the radical may hide herself and reduce the chance of being targeted.

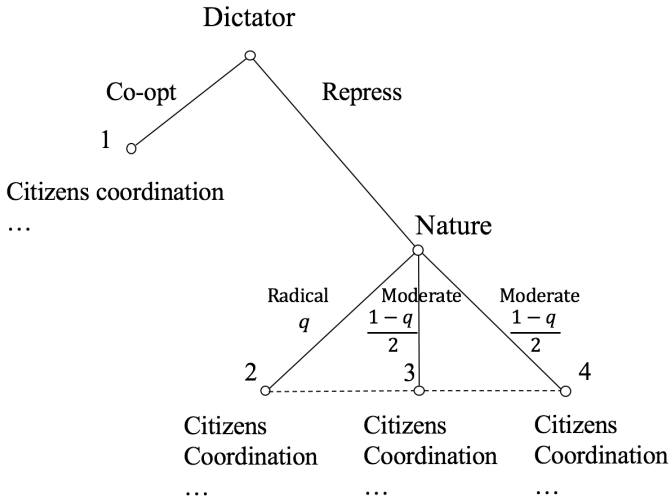

Figure C.5: Sequence of the Repression-Cooptation Game

C.3 Analysis

The Perfect Bayesian Nash Equilibrium will be characterized by backward induction. With repression, the dictator's expected payoff is:

[7] If payoffs for the repressed citizens contain $-r$ for every cell of the coordination game, then this means repression before protest does not affect the repressed citizen's behavior at all. This is usually not the case in the real world: dictators use preventive repression to detain a citizen so that she cannot participate in protests later.

$$- qc_r - (1-q)(c_r + v_m + v_r) \tag{C.1}$$

With co-optation, the dictator's expected payoff is:

$$- c_w \tag{C.2}$$

Comparing the dictator's payoffs from repression and co-optation ends up with comparing two quantities: $(1-q)(v_m+v_r)$ and $c_w - c_r$. If $(1-q)(v_m+v_r) > c_w - c_r$, the dictator will choose co-optation in equilibrium; if $(1-q)(v_m+v_r) < c_w - c_r$, he will choose represion over co-optation. The critical value of q is: $q^* = 1 - \frac{c_w - c_r}{v_m + v_r}$.

This result suggests that if the dictator believes that the loss from citizens' challenge given the probability $(1-q)$ he incorrectly represses a moderate instead of a radical is less than the saved cost from using repression rather than costly universal co-optation, the dictator will choose repression over co-optation; and vice versa. Thus, when the probability (q) of finding the radical is very low, co-optation is better than repression. In contrast, when the probability of targeting the radical is sufficiently high, $q > 1 - \frac{c_w - c_r}{v_m + v_r}$, repression becomes a better option.

C.4 Discussion

This model illustrates the tradeoffs between a dictator's co-optation and repression strategies. It clarifies the conditions under which the dictator prefers co-optation to repression and vice versa. The dictator chooses co-optation when the probability of accurately targeting radicals is low but chooses repression when this probability is sufficiently high. Information about citizens' types thus plays a crucial role in dictators' strategic choice of co-optation and repression.

For simplicity, the model assumes away regime supporters and considers only moderate and radical opponents. Qualitatively, the results will not change if we include regime supporters in the game because their inclusion does not affect the moderates' and radicals' co-optation benefits given the assumption of non-exclusive universal distribution. Including regime supporters only decreases the probability of finding radicals for repression. Thus, the dictator still prefers co-optation to repression when the probability of accurately targeting radicals is sufficiently low. The result regarding co-optation may change if including regime supporters renders the cost of co-optation unaffordable to the dictator. Then, surveillance will lead to no changes in co-optation but only help the dictator identify radicals for repression. Nevertheless, repression still increases after surveillance.

To illustrate the tradeoff between co-optation and repression under information asymmetry, the model treats co-optation and repression as a binary choice for the dictator. The purpose of using this binary choice is to show that if vertical information asymmetry makes repression less efficient, co-optation will be a dominant strategy; whereas if the dictator has improved information for targeting radicals, repression then becomes a dominant strategy. In addition, we only consider two citizens and assume that the moderates would not challenge by themselves. In reality, the relationship between co-optation and repression is not perfectly substitutable: co-optation is better at deterring mobilization from moderates while repression is better at targeting radicals. If there are many moderates, and they are better off by challenging together without the participation of radicals, the dictator may use both co-optation and repression to prevent moderates and radicals from challenging the regime, as long as the dictator has not spent all their budgets on co-optation

and repression, as (Wintrobe 2000) demonstrates in his model of Tin-pot dictators. Nevertheless, when dictators can identify radicals, the optimal strategy is to repress them.

The model is a static game that assumes citizens' types constant. Scholars argue that repressions and especially violent and highly visible repression increases the likelihood that a Moderate becomes a Radical in future (Chenoweth and Lawrence 2010). We are less concerned with such change here because we focus on low-intensity and less visible forms of repression rather than large-scale violence such as violent rebellions. Previous models of authoritarian information focus on how individuals' preference falsification impedes collective actions (e.g., Kuran 1991; Lohmann 1994), or how governments manage the Internet or political participation to allow upward information flow between citizens and political leaders but prevent horizontal information flow among citizens that could facilitate political challenges (Lorentzen 2013, 2014; Chen and Xu 2017). The authoritarian information problem itself is their main focus. Guriev and Treisman (2015)'s model on censorship, propaganda, co-optation, and repression mainly focuses on government information manipulation in dictatorships. My model focuses on how dictators' information influence the repression-cooptation tradeoff, and it focuses on preventive repression rather than reactive repression emphasized by Guriev and Treisman (2015).

References

Bai, Hua and Hechi City Heichi City Bureau of Finance. 2014. *2011 Hechi Fiscal Statistial Yearbook (in Chinese)*. Guangxi People's Publishing House.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-differences Estimates?" *The Quarterly Journal of Economics* 119(1):249–275.

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

Chenoweth, Erica and Adria Lawrence. 2010. *Rethinking Violence: States and Non-state Actors in Conflict*. MIT press.

Ding, Zhaowei. 2007. "How Do Practitioners in Security Industry View the "3111" Initiative." *China Public Security* 03.

Edin, Maria. 2003. "Remaking the Communist Party-State: The Cadre Responsibility System at the Local Level in China." *China: An International Journal* 01(01):1–15.

Gansu Provincial Department of Finance, Gansu Province. 2011. *Gansu Province Fiscal Yearbook 2011 (in Chinese)*. Gansu Research Institute of Fiscal Science Press.

Greitens, Sheena Chestnut. 2017. "Rethinking China's Coercive Capacity: an Examination of PRC Domestic Security Spending, 1992–2012." *The China Quarterly* 232:1002–1025.

Gueorguiev, Dimitar D. 2017. "In Public or in Private: Repressing Dissidents in China." Working Paper. Available at SSRN: https://ssrn.com/abstract=3038812. Accessed July 4, 2019.

Guriev, Sergei M. and Daniel Treisman. 2015. How Modern Dictators Survive: Cooptation, Censorship, Propaganda, and Repression. SSRN Scholarly Paper ID 2572452 Social Science Research Network Rochester, NY: .

Hazlett, Chad and Yiqing Xu. 2018. "Trajectory Balancing: A General Reweighting Approach to Causal Inference with Time-Series Cross-Sectional Data." Working Paper, UCLA and UCSD. Link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3214231. Accessed July 4, 2019.

Imai, Kosuke, In Song Kim and Eric Wang. 2018. "PanelMatch: Matching Methods for Causal

Inference with Time-Series Cross-Section Data." Working Paper. Link: https://imai.fas.harvard.edu/research/tscs.html. Accessed July 4, 2019.

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

Li, Kejun. 2014. *The Governing Strategies of County Party Secretaries (Xian Wei Shu Ji Men de Zhu Zheng Mou Lue)*. Guangdong People's Press.

Lohmann, Susanne. 1994. "The Dynamics of Informational Cascades: the Monday Demonstrations in Leipzig, East Germany, 1989–91." *World Politics* 47(1):42–101.

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

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

Truex, Rory. 2019. "Focal Points, Dissident Calendars, and Preemptive Repression." *Journal of Conflict Resolution* 63(4):1032–1052.

Wang, Yuhua. 2014. "Empowering the Police: How the Chinese Communist Party Manages Its Coercive Leaders." *The China Quarterly* 219:625–648.

Wang, Yuhua and Carl Minzner. 2015. "The Rise of the Chinese Security State." *The China Quarterly* 222:339–359.

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

Yin, Huimin and Shandong Province Shandong Provincial Department of Finance. 2011. *Shandong Province Fiscal Yearbook 2011 (in Chinese)*. Economic Science Publishing House.

Zhou, Chunyu and Anhui Province Anhui Provincial Department of Finance. 2006. *Anhui Province Fiscal Yearbook 2006 (in Chinese)*. China Financial and Economic Publishing House.