AI-tocracy: the Symbiosis of Autocrats and Innovators

Martin Beraja
Andrew Kao
David Y. Yang
Noam Yuchtman*

August 26, 2021

Abstract

Can frontier innovation be promoted and sustained under autocracy? We argue that a symbiotic relationship between frontier innovation and entrenched autocracy can exist. Symbiosis arises when *(i)* innovative output increases autocrats' probability of maintaining power, and *(ii)* autocrats' spending on innovative output to maintain power generates further innovation. We evaluate these two conditions of symbiosis in China's facial recognition AI sector. We gather comprehensive data on firms and government procurement contracts in this sector, as well as on social unrest across China during the last decade. We show that, first, autocrats benefit from AI: local unrest leads to greater government procurement of facial recognition AI, and increased AI procurement suppresses subsequent unrest. Second, the AI sector benefits from autocrats' suppression of unrest: the contracted AI firms innovate more both for the government and commercial markets. Taken together, these results establish the conditions for sustained AI innovation under the Chinese regime: AI innovation entrenches the regime, and the regime's investment in AI for political control stimulates further frontier innovation.

Keywords: artificial intelligence, autocracy, innovation, data, China, surveillance, political unrest
JEL Classification: O30, P00, E00, L5, L63, O25, O40

*Beraja: MIT and NBER. Email: maberaja@mit.edu. Kao: Harvard University. Email: andrewkao@fas.harvard.edu. Yang: Harvard University, NBER, and CIFAR. Email: davidyang@fas.harvard.edu. Yuchtman: LSE, CEPR, and CESifo. Email: n.yuchtman@lse.ac.uk.
Many appreciated suggestions, critiques and encouragement were provided by Sergei Guriev, Andrei Shleifer, and many seminar and conference participants. Yang acknowledges financial support from the Harvard Data Science Initiative; Yuchtman acknowledges financial support from the British Academy under the Global Professorships program.

1 Introduction

Can frontier innovation be promoted and sustained under autocracy? Political economists have long viewed autocratic political institutions as fundamentally misaligned with technological change: autocrats' political and economic rents are eroded by innovation and economic growth; and incentives to innovate are stifled by threats and acts of expropriation under autocracy.[1]

In this paper, we argue that a symbiotic relationship between autocracy and frontier innovation can exist even in the presence of expropriation risk and other economic distortions. Such a symbiotic relationship entails: *(i)* innovative output increasing the autocrats' probability of maintaining power — for example, by enhancing repressive capacity, military strength, or political legitimacy; and *(ii)* autocrats' spending on innovative output to maintain power generating broader innovation beyond that used merely for political repression — for example, through economies of scale or scope, the production of intangible assets, and externalities. These conditions allow for an equilibrium in which autocracy is entrenched and frontier innovation is sustained. They appear to have been present in prominent episodes of frontier innovation under non-democratic regimes, including the development of aerospace technology in the USSR and chemical engineering innovation in Imperial Germany.

Recent research has suggested that Artificial Intelligence (AI) innovation may possess characteristics that precisely satisfy these two conditions of symbiosis. First, it has been argued that as a technology of prediction (Agrawal et al., 2019), AI may be particularly effective at enhancing autocrats' social and political control (Zuboff, 2019; Tirole, 2020). Second, because government data is an input into developing AI prediction algorithms and can be shared across multiple purposes (Beraja et al., 2021), autocracies' collection and processing of data for purposes of political control may directly stimulate AI innovation for the commercial market, far beyond government applications. More general forms of spillovers may also be at work, as in Moretti et al., 2019. Up to now, these possibilities remain untested empirically. It is unclear to what extent autocrats invest in AI to maintain political control, and whether AI is indeed an effective technology to reduce political unrest. Moreover, data collected and shared with firms primarily out of motives

[1] The effects of economic growth on political institutions have been studied by Lipset (1959), Barro (1996), and Glaeser et al. (2007) (see Treisman, 2020 for a recent review). The effects of political institutions on economic growth and frontier innovation have been studied by, among others, North and Weingast (1989), Acemoglu and Robinson (2006), Aghion et al. (2007), North et al. (2009), and Acemoglu and Robinson (2012). Autocracies may also exhibit reduced innovation due to corruption and the misallocation of talent (Murphy et al., 1989; Shleifer and Vishny, 2002).

for political control may not be conducive to — and may even crowd out — commercial AI innovation.[2]

We test for the conditions of a symbiotic relationship in the context of facial recognition technology in China. This context possesses political and economic characteristics making it particularly suitable for studying frontier innovation under autocracy. First, maintaining political control is a paramount objective of the ruling Chinese Communist Party (see, among others, Shirk, 2007). All citizens, even China's most successful entrepreneurs, are threatened by an unconstrained autocrat's ability to violate their property rights — and at times civil rights.[3] Second, facial recognition is one of the most important fields of AI technology,[4] and China is among the world's leading producers of commercial AI innovation.

To conduct our empirical analyses, we combine data on: *(i)* local public security agencies' procurement and of facial recognition AI (and the deployment of complementary surveillance technology and police) primarily from China's Ministry of Finance; *(ii)* episodes of local political unrest in China from GDELT; and *(iii)* China's facial recognition AI firms' software innovation from China's Ministry of Industry and Information Technology, classified into government or commercial intended uses using machine learning (as in Beraja et al. (2021)). Linking datasets *(i)* and *(ii)* allows us to test whether autocracies procure facial recognition AI for purposes of political control, whether facial recognition AI is effective in suppressing unrest, and whether AI procurement is associated with complementary changes in the technology of political control (such as the procurement of surveillance cameras). Then, linking these two datasets to *(iii)* enables us to test the extent to which *commercial* facial recognition AI innovation (our indicator of frontier innovation in AI beyond political uses) benefits from politically motivated procurement and data provision.

Our empirical analyses begin by examining the first condition of symbiosis: whether AI technology can effectively enhance autocrats' political control. We first test whether autocrats respond to political unrest by procuring facial recognition AI technology. Using a difference-in-differences strategy, we find that indeed they do: locations experiencing episodes of political unrest increase their public security procurement of facial recognition AI. Importantly, these same locations do not increase their procurement of facial recogni-

[2]While Beraja et al., 2021 find that government contracts providing access to government data stimulate commercial innovation, the contracts they study are not issued explicitly for the purpose of political control.

[3]For example, Jack Ma, the founder of Alibaba, was detained for months upon arousing the ire of the Chinese Communist Party. See, for example, from the *Wall Street Journal*, `https://on.wsj.com/3rhtD0l`.

[4]For example, in 2020, computer vision was the second largest field of study in AI by publications on arXiv, accounting for 31.7% of the total publications (Zhang et al., 2021).

tion AI for non-public security purposes, indicating that the occurrence of political unrest neither induces nor coincides with a general adoption of AI technology in the public sector. One might still wonder whether the procurement of public security AI was already on a different trend in locations experiencing political unrest (e.g., because of different rates of economic growth). However, we find no evidence of differential pre-trends. One might also wonder whether time and space varying shocks are correlated with the occurrence of political unrest and with public security AI procurement. To address this concern, we implement an IV strategy exploiting variation in the occurrence of political unrest arising from local weather conditions, and we find qualitatively and quantitatively similar results. In addition to increased procurement of AI technology for public security, we also find that locations experiencing political unrest purchase more high resolution video cameras which provide the crucial data input for facial recognition technology. Moreover, public security agencies that have procured more facial recognition AI technologies not only reduce their subsequent hiring of police staff, but also shift the composition of the police force towards higher skilled desk jobs that complement the deployment of AI technology.

Local governments' purchases of AI technology for public security purposes in response to the occurrence of political unrest reflect their revealed preference and suggest a belief in the effectiveness of such technology in curbing future unrest. We next study whether the increased public security AI procurement does indeed enhance autocrats' political control. Precisely because AI is procured endogenously in locations susceptible to political unrest, rather than examining the relationship between AI procurement and subsequent local protests, we examine how past investment in public security AI mitigates the impact of exogenous shocks that tend to instigate political unrest. Following a Bartik-style empirical strategy, we find that weather conditions conducive to protests have smaller effects on contemporaneous unrest in prefectures that have previously invested in public security AI. Such a relationship is not observed in response to past AI procurement for non-public security purposes, suggesting that our results are driven by the deployment of public security AI *per se*, rather than by differing socioeconomic conditions in politically sensitive contexts. Importantly, our results are not due to the time-varying effects of past protests that are associated with public security AI investment: local experience of past protest is not associated with differential unrest arising from current weather conditions.

Having established that AI *does* strengthen autocrats' political control, we then examine the second condition of symbiosis: whether politically motivated AI procurement stimulates commercial AI innovation. We study the effects of AI procurement contracts issued by local governments that experienced above median levels of political unrest in the preceding quarter. We compare the effects of public security contracts issued in this politically sensitive environment — these contracts are most plausibly politically motivated — with the effects of non-public security contracts issued in the same environment. This allows us to isolate the effects of politically motivated contracts beyond the consequences arising from generic contracts issued in a politically sensitive environment. Using a triple-differences empirical strategy, we find that receipt of a politically motivated public security contract is associated with significantly greater innovation of commercial (as well as government) software, relative to the receipt of a contract with a non-public security arm of the government. We find no evidence of differential pre-contract trends in software innovation, supporting a causal interpretation of our findings. To address the concern that political unrest is more likely to occur in economically dynamic locations where commercial AI innovation is also greater, we instead identify politically sensitive environments and classify politically motivated procurement contracts using *predicted* political unrest based on weather conditions, and the results are qualitatively unchanged. In other words, plausibly exogenous episodes of political unrest promote commercial AI innovation through increased local public security demand for AI.

Finally, we investigate whether autocrats' politically motivated AI demands distort the trajectory of innovation. We find that the effects of politically motivated public security contracts on commercial AI innovation are not smaller than other, politically neutral, public security contracts (if anything, we find the effects are larger), suggesting that the political motivation does not diminish the value of a procurement contract for AI firms. Moreover, we find no increase in surveillance oriented commercial software development after the receipt of politically motivated public security contracts. The absence of evidence of significant distortions arising from politically motivated public security AI procurement suggests that the AI-tocracy equilibrium may be able to sustain continuous commercial AI innovation.

Taken together, these results imply that the autocratic political regime and the rapid innovation in China's AI sector are not in conflict, but are symbiotic. We do not interpret our findings as indicating that China's political stability is primarily achieved through AI technology (yet), nor that China's AI innovation is primarily rooted in political repression. Rather, our findings suggest that a component of China's coercive capacity is derived from the application of AI technology, and China's political repression in turn contributes to AI innovation and potentially economic growth.

Beyond the AI innovation in China, our analyses may also provide an illustration of forces that generated frontier innovation in other non-democratic contexts such as the

USSR and Imperial Germany, shedding light on episodes that are difficult to be accounted for by a large political economy literature that highlights forces that tend to limit innovation and growth in autocracies.[5]

Our work extends a recent literature that emphasizes the importance of state capacity for development (e.g., Besley and Persson, 2009) and allows for the possibility of growth under extractive institutions (e.g., Acemoglu and Robinson, 2020, Dell and Olken, 2020). We demonstrate that even frontier innovation can be sustained in autocracy. In fact, our finding that autocracy and frontier innovation are mutually reinforcing implies a different political economy trajectory: the Chinese case suggests a stable equilibrium exhibiting sustained frontier innovation and further entrenched autocracy.[6]

Our results also contribute to a growing literature on the socioeconomic consequences of AI technology. Much of the literature focuses on economic consequences of AI, from its impact on the labor market (Acemoglu and Restrepo, 2018, 2019), to socioeconomic inequality (Korinek and Stiglitz, 2017), to economic growth (Aghion et al., 2017). Some recent research has considered the social consequences of AI, in particular, discrimination arising from the potential biases in its algorithms (Kleinberg et al., 2018; Cowgill and Tucker, 2020). Our paper provides direct evidence on the political consequences of AI technology: it can produce more effective political control, potentially entrenching autocratic government.

In so doing, we also contribute to a large literature on the relationship between technology and political mobilization. Recent papers find that advances in information and communication technologies, and the diffusion of social media, have supported protest movements and populist parties in a broad range of settings (Campante et al., 2018; Enikolopov et al., 2020; Qin et al., 2020; Guriev et al., 2020). We, on the other hand, contribute to a literature that documents how technological change can *repress* political unrest, thus strengthening autocracies and incumbents more generally. This literature describes the evolution of repressive technology: from Autocracy 1.0 — the state as a monopolist of violence using the threat of brute force to produce compliance out of fear (Olson Jr., 1993); to Autocracy 2.0 — the state as manipulator of information using propaganda and censorship to produce compliance out of persuasion (Cantoni et al., 2017; Roberts, 2018; Chen and Yang, 2019; Guriev and Treisman, 2019); and finally, to Autocracy

[5] In addition to works cited above, a large empirical literature identifies negative effects of extractive institutions on long-run development (e.g., Acemoglu et al., 2002, Nunn, 2008, Dell, 2010, Lowes and Montero, 2020). There has been, however, a small strand of the literature that documents the positive economic consequences of colonial investments, particularly in transportation infrastructure and human capital (e.g., Hullery, 2009, Cagé and Rueda, 2016, Donaldson, 2018, Valencia Caicedo, 2019).

[6] It is important to note this political economy equilibrium is not inevitable, because forces for symbiosis may be offset by autocratic distortions (e.g., risks of expropriation).

3.0 — the state (and its AI) as monitor, predictor, and shaper of behaviors to to produce compliance using targeted behavioral incentives (Tirole, 2020).

Finally, we contribute to the large literature on the political economy of growth in China. While much of the literature emphasizes factors that promote China's growth *despite* its autocratic politics (Lau et al., 2000; Brandt and Rawski, 2008; Song et al., 2011), we join an emerging strand of the literature that highlights the autocratic institutional features that facilitate growth (Bai et al., 2020; Beraja et al., 2021). Importantly, we demonstrate that the stimulation of China's facial recognition AI innovation is not due to marginal improvements in institutional features such as protection of property rights and rule of law, nor to the enhancement of infrastructure or state capacity more generally; but rather, AI innovation is spurred directly by the application of political repression itself.

The paper proceeds as follows. In Section 2, we describe prominent episodes of frontier innovation under autocracy that suggest the existence of symbiosis, and we present the case of AI technology that guides the empirical inquiry. We describe the data sources used for the empirical analyses in Section 3. We present the empirical analyses of the effects of AI technology on autocratic political control in Section 4; we then present the analysis of the effects of politically motivated procurement of AI on innovation in Section 5. Finally, we conclude with a discussion of the geopolitical implications of symbiosis between AI innovation and the Chinese regime in Section 6.

2 Symbiosis between frontier innovation and autocracy

2.1 Historical episodes

We first consider the success of scientific innovation in the Soviet Union, which was world leading in areas such as physics, mathematics, and aerospace and nuclear engineering. A striking feature of Soviet politics is the role of scientific advancement in legitimizing the Communist regime.[7] Science served as an effective propaganda tool, both internally and externally, to enhance the prestige and legitimacy of the regime. For example, following the launch of Sputnik (the first satellite), *Pravda* celebrated "how the freed and conscientious labor of the people of the new socialist society makes the most daring dreams of mankind a reality" (Pravda, 1957). Scientific advancement also generated military technology that strengthened the regime against both internal and external threats: from nuclear warheads to Intercontinental Ballistic Missiles to fighter jets. The Soviet state's

[7] The importance of science to Communist ideology is seen in the Soviet government's "official view that science and Soviet socialism are mutually supportive" (Graham, 1989; see also Ings, 2017 and Slezkine, 2017).

financial and institutional support of science produced the world's largest community of scientists and engineers (Graham, 1989).[8] It also produced remarkable technological achievements, most famously in the space program, which launched the first satellite, sent the first human into space, constructed the first space station, and captured the first image of the far side of the moon, among other accomplishments.

A second case of frontier innovation under a non-democratic regime is the Second German Empire, which emerged as a powerhouse of science, industrialization, and innovation in the late 19th century.[9] Scientific and engineering innovation in many sectors were considered critical to ensure that Germany had a leading position among the imperial powers of Europe, not least because such innovation directly strengthened German military and naval capacity. For example, when describing the aim of the soon-to-be-established Imperial Institute of Physics, an imperial official stated that "there can be no doubt that our navy, telegraph system, survey organization, army and even the railways will [...] to a considerable degree be dependent on the results of the research for which this Imperial Institute of Physics is intended." Such imperial research institutes combined the expertise of German scientists with large amounts of state funding, producing not only military technology, but also general (even Nobel Prize-winning) scientific and industrial innovations. The eminent industrialist Von Siemens credited these institutes with Germany's industrial development, writing, "we have only the high quality of scientific education in Germany to thank for the fact that German industry, despite unfavorable circumstances, has somehow managed to retain its prominent position."

Apart from these two prominent episodes, one observes other instances of frontier innovation taking place in non-democratic regimes. In some cases, frontier technology enhances the legitimacy of the state, as in the Soviet example described above. For example, in Socialist Cuba, the remarkable success of the health care sector (e.g., developing vaccines and cancer treatments) served to bolster the regime's claim of political legitimacy (Geloso et al., 2020). In other cases, frontier innovation strengthens the regime through stimulating the economy and developing military technologies, as in the German example described above. Much like Germany, Imperial Japan post-Meiji Restoration heavily invested in frontier innovation in order to industrialize and strengthen its military capacity (Morris-Suzuki 1994). Singapore has since its independence actively supported export-oriented industrial innovation, the success of which fueled its growth miracle and helped entrench its one-party rule (Yue, 2005).

[8]We do not claim that the Soviet's support of science and innovation was without distortion. Ings (2017) describes costly political distortions to science under Stalin.

[9]We rely on Pfetsch (1970) throughout this case study.

The two conditions of symbiosis between frontier innovation and autocracy appear to be shared across these episodes. First, the non-democratic regimes appear to derive political power from frontier innovation. Second, recognizing the political benefits of innovation, the regimes provide financial and institutional support that may be instrumental to technical development.[10]

2.2 AI-tocracy

AI technologies are fundamentally about prediction, as highlighted by Agrawal et al. (2018). Predictions are extraordinarily valuable for an autocrat trying to maintain social and political control. They can serve to enhance monitoring (e.g., using prediction algorithms to identify and track individuals), to project human behaviors (e.g., identifying individuals who are more likely to engage in political unrest), and to shape behaviors (e.g., providing targeted sticks and carrots, as studied by Tirole, 2020 and Zuboff, 2019). These political applications of AI technology to suppress and prevent political instability thus contribute to the first condition of a symbiotic relationship.

At the same time, autocratic governments' procurement of AI technologies for purposes of political control can stimulate AI innovation beyond mere political purposes. This can occur through particular channels related to AI innovation being data-intensive. Firms providing AI services to the state may gain access to valuable government data. To the extent that such government data or algorithms trained with it are shareable within the firm, they can be used to develop AI products for commercial markets Beraja et al. (2021). Moreover, government procurement may increase private data collection, which can then be shared across firms due its non-rivalry (Aghion et al., 2017; Jones and Tonetti, 2018). Moreover, procurement of AI technologies could also stimulate innovation through traditional "crowding-in" channels, including the production of non-tangible assets (e.g., ideas) and technological spillovers across government and commercial applications, both within a firm and between firms.[11] Public procurement also provides resources to firms that may allow them to cover fixed costs of innovation and overcome financial constraints. These economic consequences of government procurement of AI technology (in particular, by an autocrat) thus contribute to the second condition of a symbiotic relationship.

[10]One also observes examples of symbiosis between democratic regimes and innovation. One prominent case is the military innovation developed by DARPA in the US, and its well-known commercial innovation consequences (e.g., the internet). We do not argue that innovation *only* supports autocratic regimes; but rather, that such a regime-enhancing effect of technology may be particularly relevant in non-democracies due to their otherwise unfavorable environment for innovation.

[11]These channels have been shown to be important in the context of space exploration (Alic et al., 1992; Azoulay et al., 2018), the internet (Greenstein, 2015), and military technology (Moretti et al., 2019; Gross and Sampat, 2020).

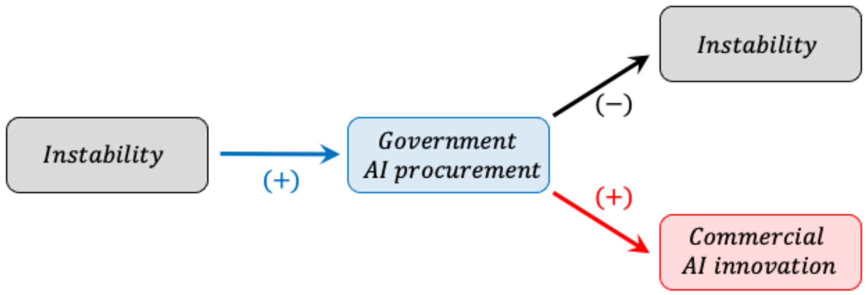


Figure 1: AI-tocracy

When these two conditions are sufficiently strong to overcome distortions in autocracies that discourage innovation (e.g., risk of expropriation), the symbiotic relationship could support an equilibrium — "AI-tocracy" — where an autocratic regime is entrenched, and frontier AI innovation is sustained. They do so by generating a self-reinforcing cycle in which autocrats are strengthened by AI innovation, and their procurement of this innovation stimulates further innovation, which in turn further strengthens the autocrats.

Empirical context: facial recognition AI in China We test for the two conditions of symbiosis in the context of facial recognition AI technology in China. The potential symbiotic relationship between frontier facial recognition AI innovation and the Chinese regime is illustrated in Figure 1; the key causal links that we empirically test for are indicated in the solid arrows in the figure.

To begin, we test for AI procurement that is motivated by the regime's desire for political control. The Chinese regime is particularly concerned with protests and unrest (Shirk, 2007; King et al., 2013). It thus may procure facial recognition AI technology in response to unrest (blue arrow in the figure), which could allow the Chinese government to identify, crack down on, and deter the participants to the unrest.

We next test whether this politically motivated procurement of AI technology satisfies the two conditions of symbiosis: first, whether procurement of AI technology enhances the regime's political control (black arrow in the figure); and second, whether politically motivated procurement of AI technology stimulates further frontier AI innovation (red arrow).

3 Data

To conduct our empirical analyses, we combine data on: *(i)* episodes of local political unrest in China; *(ii)* local governments' procurement of facial recognition AI technology; and *(iii)* facial recognition AI firms' software innovation.[12] We describe, in addition, auxiliary data sources used for various empirical exercises in Appendix A.

3.1 Political unrest

We collect data on political unrest from the Global Database of Events, Language, and Tone (GDELT) project. The GDELT project records instances of real-world events based on articles from a global, comprehensive set of news feeds.[13] We restrict our analysis to events taking place in China between 2014 to 2020.[14] In sum, we find 9,267 events indicating political unrest, corresponding to three broad categories: protests, demands, and threats. Figure 2, Panel A, presents the spatial distribution of the political unrest occurred during the period of 2014 to 2020; and Table 1, Panel A, presents basic summary statistics of these political unrest events.[15]

Local weather conditions used to construct instruments for political unrest We use historical weather data originally collected by the World Meteorological Organization (WMO) and hosted by the National Oceanic and Atmospheric Administration (NOAA). Data is reported at the weather station-day level. These weather stations provide a wide variety of data at the daily level, including mean temperature, amount of precipitation, presence of fog or hail or thunder, maximum windspeed recorded, and visibility.[16] We

[12]Several of these datasets are used and also described in detail in our prior work Beraja et al. (2021).

[13]Text analysis and machine learning methods are applied to the contents of these articles to identify salient characteristics, such as event location (which we geocode at the prefecture level), date of the event, and the nature of these events. See `https://www.gdeltproject.org` for a detailed description of the GDELT project and its methodology.

[14]The GDELT project greatly expanded their scope of sources and text analysis capabilities in 2014, making coverage before 2014 less complete and reliable. From 2014 to 2020, there are over one hundred news sources that provide coverage on China. When multiple news sources cover the same event, GDELT records only one event.

[15]Each event is classified under the Conflict and Mediation Events Observations (CAMEO) event and actor codebook, in which protests (e.g., demonstrations, hunger strikes for leadership change), demands (e.g. demands for material aid, leadership change, or policy change), and threats (e.g., threats to boycott, political dissent) are three of twenty top-level "verbs" that an event can be classified under. We exclude a small amount of events that occur at a national or international level.

[16]This weather data ranges from 2012 to 2020. There are a small number of observations for which weather data is missing (less than 1% of the total). For these, we impute data from the geographically nearest weather station, or in the one instance when all stations are missing data on a given day, we take data from the following day and the same station instead.

assign data to prefectures using the closest weather station to the given prefecture. For the 344 prefectures in our dataset, this results in 260 unique weather stations whose data we use.

3.2 Procurement of AI and the technology of political control

In order to observe the Chinese government's demand for AI technology, we extract information on 2,997,105 procurement contracts issued by all levels of the Chinese government between 2013 and 2019 from the Chinese Government Procurement Database, maintained by China's Ministry of Finance.[17] The contract database contains information on the good or service procured, the date of the contract, the monetary size of the contract, the winning bid, as well as, for a subset of the contracts, information on bids that did not win the contract.

To narrow our focus on the subset of contracts that procure AI technology such as data processing services or platform solutions, we match the contracts with a list of facial recognition AI firms. We identify (close to) all active firms based in China producing facial recognition AI using information from *Tianyancha*, a comprehensive database on Chinese firms licensed by China's central bank.[18] We extract firms that are categorized as facial recognition AI producers by the database, and we validate the categorization by manually coding firms based on their descriptions and product lists. We collect an array of firm level characteristics such as founding year, capitalization, major external financing sources, as well as subsidiary and mother firm information. Overall, we identify 7,837 Chinese facial recognition AI firms.[19]

Our empirical exercises in particular concern the AI procurement contracts awarded by public security agencies of the Chinese government. As an example from our dataset, consider a contract signed between an AI firm and a municipal police department in Heilongjiang Province to "increase the capacity of its identity information collection system" on August 29th, 2018. The contract specifies that the AI firm shall provide a facial recognition system that should cover at least *30 million* individuals, suggesting the large scale of data collection and processing that are required. In total, we identify 28,023 public se-

[17] A primary source of firms' information compiled by Tianyancha is the National Enterprise Credit Information Publicity System, maintained by China's State Administration for Industry and Commerce. See Appendix Figure A.1 for an example contract.

[18] See Appendix Figure A.2 for an example entry. We complement the *Tianyancha* database with information from *Pitchbook*, a database owned by Morningstar on firms and private capital markets around the world. See Appendix Figure A.3 for an example entry.

[19] These firms fall into 3 categories: *(i)* firms specialized in facial recognition AI (e.g., Yitu); *(ii)* hardware firms that devote substantial resources to develop AI software (e.g., Hik-Vision); and *(iii)* a small number of distinct AI units within large tech conglomerates (e.g., Baidu AI).

curity related procurement contracts on AI technology.[20] They include the following four types of public security contracts from the Chinese Government Procurement Database: *(i)* all contracts for China's flagship surveillance/monitoring projects — *Skynet Project*, *Peaceful City Project*, and *Bright Transparency Project*; *(ii)* all contracts with local police departments; *(iii)* all contracts with the border control and national security units; and, *(iv)* all contracts with the administrative units for domestic security and stability maintenance, the government's political and legal affairs commission, and various "smart city" and digital urban management units of the government. Importantly, each of these contracts is linked to a specific prefectural government buyer, and for the baseline analysis, we exclude those signed with the central or provincial government. Many firms receive multiple public security contracts; overall, 1,095 facial recognition AI firms in our dataset receive at least one contract. Figure 2, Panel B, presents the spatial distribution of the facial recognition AI contracts issued by public security units of the prefectural government; and Table 1, Panel B, presents basic summary statistics of the facial recognition AI procurement contracts.[21]

Parts of our empirical strategy to pin down the role of access to government data requires us to compare public security procurement contracts of AI to those award by non-public security units in the public sector, such as (public) banks, hospitals, and schools. There are a total of 6,557 non-public security related procurement contracts of AI technology.[22]

3.3 Innovation of AI firms

We collect all software registration records for our facial recognition AI firms from China's Ministry of Industry and Information Technology, with which Chinese firms are required to register new software releases and major upgrades. We are able to validate our measure of software releases (using a single large firm), by cross-checking our data against the IPO Prospectus of MegVii, the world's first facial recognition AI company to file for an

[20]We present the cumulative number of AI procurement contracts in Appendix Figure A.4 (top panel), as well as the flow of new contracts signed in each month (bottom panel). Both public security and non-public security AI contracts have steadily increased since 2013.

[21]Some public security AI contracts are issued at the provincial level: for example, almost 40% of the public security AI contracts in Xinjiang are issued by the provincial government. Appendix Figure A.5 plot the spatial distribution of public security AI contracts issued by either provincial or prefectural governments.

[22]In the following empirical analysis, we define politically motivated contracts as contracts issued in times of above-median unrest. Public security contracts may or may not be politically motivated contracts — the classification of a public security contract is only dependent on the agency issuing the contract. We exploit both types of variation in Section 5.

IPO.[23] We find that our records' coverage is comprehensive (at least in the case of MegVii): MegVii's IPO Prospectus contains 103 software releases, all of which are included in our dataset.

The count of new software releases (and major upgrades) represents *product innovation*.[24] Reflecting the economic value of such innovation, we observe that facial recognition AI firms that develop more software have significantly and substantially higher market capitalization (see Appendix Figure A.6).

We use a Recurrent Neural Network (RNN) model with tensorflow — a frontier method for analyzing text using machine learning — to categorize software products according to their intended customers and (independently) by their function. Our categorization by customer distinguishes between software products developed for the government (e.g., "smart city — real time monitoring system on main traffic routes") and software products developed for commercial applications (e.g., "visual recognition system for smart retail"). We allow for a residual category of general application software whose description does not clearly specify the intended user (e.g., "a synchronization method for multi-view cameras based on FPGA chips"). By coding as "commercial" only those products that are specifically linked to commercial applications, and excluding products with ambiguous use, we aim to be conservative in our measure of commercial software products.

Our categorization by function first identifies software products that are directly related to AI (e.g., "a method for pedestrian counting at crossroads based on multi-view cameras system in complicated situations"). Within the category of AI software, we also separately identify a subcategory of software that involve components related to surveillance (e.g., "tool that allows parents to locate and track lost children"). Moreover, we identify a separate category of non-AI software products that are data-complementary, involving data storage, data transmission, or data management (e.g., "a computer cluster for webcam monitoring data storage").

To implement the two dimensions of categorization using the RNN model, we manually label 13,000 software products to produce a training corpus. We then use word-embedding to convert sentences in the software descriptions into vectors based on word frequencies, where we use words from the full dataset as the dictionary. We use a Long Short-Term Memory (LSTM) algorithm, configured with 2 layers of 32 nodes. We use 90% of the data for algorithm training, while 10% is retained for validation. We run

[23]Source: Hong Kong Stock Exchange, `https://go.aws/37GbAZG`.

[24]The National Science Foundation defines product innovation as "the market introduction of a new or significantly improved good or service with respect to its capabilities, user-friendliness, components, or subsystems" in its Business Enterprise Research and Development Survey (see `https://www.nsf.gov/statistics/srvyberd`). See also Bloom et al. (2020).

10,000 training cycles for gradient descent on the accuracy loss function. The categorizations perform well in general: we are able to achieve 72% median accuracy in categorizing software customer and 98% median accuracy in categorizing software function as AI or data-complementary in the validation data. Appendix Figure A.7 shows the summary statistics of the categorization output by customers and by function; and, Appendix Figure A.8 presents the confusion matrix (Type-I and Type-II errors) of the predictions relative to categorization done by humans.[25] Table 1, Panel C, presents basic summary statistics of the software innovation of the AI firms.

4 The role of AI in autocrats' political control

4.1 The effect of political unrest on AI procurement and the technology of political control

Our empirical analyses begin by examining whether AI technology can effectively entrench autocrats. We first test whether local police forces respond to episodes of local political unrest by procuring more facial recognition AI. We estimate the following baseline model:

$$AI_{i,t+1} = \beta Unrest_{it} + \alpha_t + \gamma_i + \epsilon_{it},$$

where the explanatory variable of interest is $Unrest_{i,t}$, the local political unrest in prefecture i in quarter t, and $AI_{i,t+1}$ is the public security facial recognition AI procurement per capita of prefecture i in the subsequent quarter (the lag reflects the time needed to issue a contract in response to an event). The baseline model controls for time period and prefecture fixed effects, and in robustness specifications we include time-varying controls.

We present the results in Table 2, Panel A, column 1. One can see that across specifications, political unrest in a prefecture in one quarter is followed by a significantly greater amount of AI procurement in the following quarter. The results remain qualitatively and quantitatively very similar if we control for the prefecture GDP interacted with a full set of (quarterly) time fixed effects (column 2), the prefecture's population interacted with a full set of time fixed effects (column 3), the prefectural government's tax revenue in-

[25] Appendix Table A.1 presents the top words (in terms of frequency) used for the categorization. Appendix Figure A.9 presents the density plots of the algorithm's category predictions. The algorithm is very accurate in categorizing software for government purposes. The algorithm is relatively conservative in categorizing software products for commercial customers, and relatively aggressive in categorizing them as general purpose. In setting our categorization threshold for commercial software we again aim to be conservative in our measure of commercial software products.

teracted with a full set of time fixed effects (column 4), or all sets of controls (column 5). These results suggest that the patterns that we observe are unlikely to be driven by changing local economies. Appendix Table A.2 shows results on political unrest in the separate subcategories of protests, public demands, and threats, with results remaining qualitatively the same.

We next consider the possibility that AI procurement may already have been increasing in locations with political unrest *prior* to the unrest itself. We thus estimate a modified version of the baseline model, but now examining the relationship between unrest in period t and AI procurement in periods $t-1$ and $t-2$. Figure 3 plots the estimated coefficient on unrest, with a stacked regression containing each lead and lagged period. As one can see, political unrest is not associated with preceding levels of AI procurement, indicating that AI procurement did not anticipate but rather respond to the unrest, consistent with a causal effect of unrest on subsequent AI procurement. In Figure 3, we further plot the effect of unrest on the AI procurement in the same period t, as well as in future periods ($t+1$, $t+2$, and $t+3$). The series of coefficients follows a sensible pattern: unrest has small same-quarter effects, with a much larger effect in the following quarter, and fading effects thereafter.

As an alternative empirical strategy, we implement an IV specification that exploits variation in political unrest arising from daily local weather variations (similar in spirit to Madestam et al. (2013) and Larreboure and Gonzalez (2021)). Implementing a weather-based IV strategy in our setting requires overcoming three challenges. The first challenge is high-dimensionality: in a country as vast as China, one must consider a wide range of potentially relevant and interacting weather conditions. To address this, we implement a LASSO regression to select predictors of unrest events among 30 weather variables and their interactions (e.g., temperature, precipitation, and windspeed). The second challenge is the need to consider both the extensive and intensive margins of political unrest. Over a relatively long period of time, there are many days on which no unrest takes place (presumably because of the absence of mobilized political demands on those days), implying no elasticity between weather conditions and unrest occurrence. On certain days, unrest occurs across multiple prefectures, and local weather conditions plausibly would influence the likelihood of unrest occurrence in a specific location. To address this challenge, we allow for the LASSO-selected weather predictors to affect the probability of unrest occurrence heterogeneously depending on whether unrest occurs in at least one prefecture on a given day. A final challenge is the need to aggregate unrest occurrence to match the time frame over which AI procurement decisions are made (several months, which we operationalize as quarterly observations). To resolve this challenge, we follow the literature on 2SLS with different aggregation across stages (Angrist and Krueger, 1992; Inoue and Solon, 2010), and aggregate our first stage estimates to the quarterly level, adjusting accordingly for the statistical inference in the aggregated second stage.

In Appendix Table A.3, we present the first stage estimates using the weather-based LASSO instrument to predict unrest occurrence.[26] In Table 2, Panel B, and Appendix Figure A.10, we replicate our previous analyses using the LASSO instrument and find very similar results. Weather-induced variation in political unrest causes an increase in AI procurement in the following quarter, and the results are robust to controlling for time-varying effects of local economic conditions. Consistent with the exogeneity of the instrument, weather-induced unrest is not statistically significantly associated with AI procurement in previous quarters.

To the extent that one may be worried that the increased procurement of AI technology by public security units of the government may reflect a general shift in policies toward AI technology, potentially even triggered by the occurrence of political unrest, we can examine whether political unrest leads to AI technology procurement by non-public security units in the public sector, such as schools, hospitals, and banks. In Table 2, columns 5 to 8, we present the results replicating our previous analyses, but instead examining the effects of political unrest on non-public security AI procurement. We find no evidence that political unrest leads to increased demand for AI technology beyond the public security sector, indicating that the occurrence of political unrest neither induces nor coincides with a general adoption of AI technology in the public sector.[27]

Upgraded technology of political control Our interpretation of AI procurement as a government response to political unrest suggests that firms receiving public security contracts issued following periods of political unrest should produce AI software for the government oriented towards surveillance. Indeed, we find a significant increase in the production of AI software intended for the government with surveillance functions (see Appendix Figure A.12 for details; the full specification is outlined under Section 5).

[26]To provide a transparent depiction of the operation of the LASSO first stage, we present, in Appendix Table A.4, the weights assigned by LASSO to each of the selected weather predictors. In Appendix Table A.5, we present similar (but less precise) first stage results using an indicator for fine temperature (below 97 and above 0 degrees Fahrenheit), which significantly, positively predicts unrest on days that at least one episode of unrest takes place in China.

[27]As a final exercise, we consider an alternative proxy for the underlying (perceived) risks of political unrest — variation in the minority Uyghur population share, who have been the focus of CCP's repeated expressions of concerns due to separatist and occasionally violent political actions. We present, in Appendix Figure A.11, a binned scatter plot, showing the cross-sectional relationship between the total public security AI procurement by prefecture and the share of the local population that belongs to the Uyghur ethnicity. One observes a strongly positive, significant relationship, consistent with the interpretation that public security AI procurement is motivated by desire to maintain political control.

Moreover, one would also expect that the local government should invest in complementary technology such as high resolution surveillance cameras. In Appendix Table A.6 and Appendix Figure A.13, we replicate the exercises in Table 2 and Figure 3, but instead examine the local public security procurement of surveillance cameras. We find that following the occurrence of political unrest, the local public security units also increase their procurement of high resolution surveillance cameras, which complement the increased deployment of AI technology by increasing the government's ability to collect greater amount of data. Consistent with a causal interpretation, we do not observe increased procurement of surveillance cameras leading up to the occurrence of political unrest.

A final question is whether the increase in AI procurement is associated with changes in other elements of the political control apparatus — in particular, the labor component. It has been argued by Acemoglu and Restrepo (2019) and Agrawal et al. (2019) that AI technology is one that is often labor-saving and likely to be skill-biased. Consistent with this literature, we find that local police hiring is significantly lower one year after the corresponding police department procures AI technology, and the share of desk (as opposed to street) police significantly increases among the new hires (see Appendix Table A.7 for details). This suggests the facial recognition AI deployed in public security replaces labor, in particular the low skilled type.

Taken together, these results suggest that the autocrat views AI technology as potentially useful and actively purses AI as an advanced method for political control. Moreover, as we demonstrate, the increased procurement of AI represents a component of a coherent technological bundle along with high resolution surveillance cameras and skilled labor in the police force which could complement AI and help the autocrat to maintain political control in the face of unrest.

4.2 The effect of AI procurement on suppressing unrest

We next examine whether greater AI procurement by the local government's public security agencies effectively suppresses political unrest. Anecdotally, local governments appear to deploy facial recognition AI to reduce unrest through means such as identifying new faces in a protest, tracking suspicious persons in their daily life, or through simple deterrence.[28]

Importantly, having just demonstrated that AI procurement is endogenous to political unrest, we *cannot* direct estimate the impact of such endogenous AI procurement on subsequent political unrest. Estimating such a relationship is further challenged by the

[28] For example, see "the Panopticon is Already Bere" from the *Atlantic*, source: `https://bit.ly/3aWC1gB`.

strong autocorrelation in local political unrest.

To evaluate the impact of public security AI procurement on autocrats' political control, we thus examine how past public security AI procurement shapes the effects of external shocks on local political unrest. We estimate a Bartik-style model in which exogenous time-varying shocks — specifically, the good weather shocks identified as drivers of local protests in the first stage of our LASSO specification — may have heterogeneous effects depending on the *ex ante* local stock of AI capacity. We expect that good weather will increase the likelihood of local political unrest, but past AI procurement in a prefecture may temper this relationship.

To determine whether past public security AI procurement affects the relationship between local weather conditions and local political unrest, we estimate:

$$Unrest_{it} = \beta_1 AI_{i,t-1} + \beta_2 FineWeather_{it} + \beta_3 FineWeather_{it} \times AI_{i,t-1} + \alpha_t + \gamma_i + \epsilon_{it}. \quad (1)$$

We estimate the effects of contemporaneous weather shocks (as captured by the LASSO first stage described above) in prefecture i at time t on local political unrest, allowing this effect to vary depending on the stock of local public security procurement of AI up to period $t-1$, controlling for prefecture and time period (quarter) fixed effects. Table 3, columns 1-5, present the results, first without additional controls, and then gradually adding time-varying controls. As can be seen, the estimated effect of fine weather is consistently positive, indicating that good weather is conducive to political unrest (as we have seen in previous section). However, the estimated effect of fine weather interacted with past public security AI procurement is negative: AI procurement significantly weakens the positive relationship between good weather and unrest occurrence, suggesting a role of AI in maintaining political control. A one standard deviation increase in the stock of past public security AI procurement halves the effect of fine weather on local political unrest. Crucially, the effect of past AI procurement only appears for the contracts issued by public security agencies. Local AI procurement by non-public security agencies does not mitigate the effects of fine weather on political unrest, as shown in columns 6-10.

It is also important to note that the cross-prefecture variation in previous AI procurement is not exogenous to sequences of political unrest, in particular to the past unrest occurrence as we demonstrate in the previous section. If the past unrest is associated with heterogeneity in the locality's responses to political shocks, this could confound the interpretation that our estimates in Table 4 capture the effects of public security AI procurement. To assess this possibility, we examine whether exogenous political shocks have heterogeneous effects on unrest occurrence depending on the past unrest in the locality. Specifically, we estimate specifications analogous to those described above, replacing

$AI_{i,t-1}$ with $unrest_{i,t-1}$ or $unrest_{i,t-2}$. As shown in Appendix Table A.8, we do not find a noticeable pattern of heterogeneous effects of fine weather depending on past unrest in the locality. This suggests that the pattern of heterogeneity we observe is likely due to public security AI procurement, rather than other mechanisms arising from past unrest.

5 The role of autocratic political control in AI innovation

We now turn to the question of whether politically motivated procurement of AI stimulates AI innovation. Specifically, we focus on AI procurement contracts issued by public security agencies in prefectures that experienced above median levels of political unrest in the quarter prior to the contracts' issuance. As shown in the previous section, these contracts are plausibly issued for purposes of political control.

We use a triple differences design to identify the effects of accessing government data collected for purposes of political control on facial recognition AI firms' subsequent product development and innovation. The empirical strategy exploits variation across time and across firms in the receipt of a government contract, and across types of government contracts that firms receive.

As in an event study design, we compare firms' outcomes — their software releases — before and after they receive their first government contracts, controlling for firm and time period fixed effects.[29] To distinguish the effects of politically motivated contracts from the effects of generic procurement contracts issued in a politically sensitive environment (defined as municipalities with above median political unrest in the previous quarter), we compare the effects of public security contracts with those of non-public security contracts issued in the same environment.[30]

Specifically, among firms receiving their first government contracts in a prefecture that recently experienced political unrest, we estimate the following specification:

$$y_{it} = \sum_T \beta_{1T} T_{it} + \sum_T \beta_{2T} T_{it} PublicSecurity_i + \alpha_t + \gamma_i + \epsilon_{it}$$

where T_{it} equals 1 if, at time t, T quarters have passed before/since firm i received its first contract; $PublicSecurity_i$ is an indicator that the firm's first government contract is issued by a public security agency; α_t are a full set of quarter fixed effects; and γ_i are a full set of firm fixed effects. The coefficients β_{1T} describe software production of a firm around the

[29] We only examine firms' first contracts because subsequent contracts could be endogenous to firms' performance in the initial contracts.

[30] For example, firms receiving *any* government contract in a context of political sensitivity (i.e., following local unrest) may be specifically selected for their potential post-contract productivity and innovation capacity.

time when it receives its first government contract when this contract is issued by a non-public security agency; the sums of coefficients $\beta_{1T} + \beta_{2T}$ describe software production around the time when the firm receives its first government contract when this contract is issued by a public security agency; and the sequence of coefficients β_{2T} thus captures the differential software production before and after a firm receives a public security contract in a politically sensitive environment.

In Figure 4, we plot the series of β_{2T} coefficients, considering different categories of software output. In Panel A, one can see that firms receiving a public security contract issued following episodes of political unrest develop approximately 1.5 additional government software products over the subsequent 2 years, compared to firms receiving a non-public security contract issued in the same local political environment. We present the full set of event study coefficients in Table 4, column 1, and present coefficients from a weighted event study specification, following Borusyak et al. (2017), in column 2. One naturally wonders whether firms receiving the public security contract were already following a different trend of software production before the receipt of the contracts. However, we do not observe differential pre-contract software production levels or trends among firms that would go on to receive a public security procurement contract.

In Panel B, one observes that firms receiving public security procurement contracts following episodes of political unrest also differentially increase their *commercial* software development, compared to firms receiving non-public security contracts in the same local political environment. Differential increase in commercial software development totals around 2.5 additional software products over the course of 2 years after the contract receipt. We present the full set of event study coefficients, using baseline and the weighted specification, in Table 4, columns 5 and 6, respectively. Our findings indicate the role of (politically motivated) government data collection in stimulating commercial innovation. Again we observe no differential commercial software production level or trend prior to the receipt of the public security contracts, suggesting a causal interpretation.[31]

One concern with this analysis is that our definition of politically motivated contracts relies on the endogenous occurrence of political unrest. Factors that shape political unrest may be associated with production of AI software specifically among firms that select into public security contracts. To address this concern, we alternatively define a politically

[31] One may wonder what are the overall effects of government contracts that underly the differential effects in Figure 4. In Appendix Figure A.14, we plot the coefficient β_{1T}, describing software production around the time when a non-public security contract was received, and the sum of the coefficient, $\beta_{1T} + \beta_{2T}$, describing software production around the time when a public security contract was received a politically sensitive environment. We find that government software and commercial software both significantly increase after receipt of both non-public security and public security contracts, with effects being significantly greater in the latter.

motivated contract as a public security contract issued just after a period with above median *predicted* level of political unrest, based on our weather-based LASSO instruments as described in Section 5. Again we difference out the effects of non-public security contracts in the same political environment. The estimated coefficients from this alternative definition of politically motivated contracts are plotted in darker-shaded dots in Figure 4, and presented in Table 4, columns 3-4 and 7-8. One can see the differential effects of public security contracts in politically sensitive environments on software innovation for both the government and commercial sectors are very similar following episodes of plausibly exogenous political unrest.

As an auxiliary test of the role of access to large quantities of government data collected out of political motivation, we examine whether firms receiving public security contracts in a politically sensitive environment develop data-complementary tools (e.g., software supporting data storage) to manage the large quantities of data that they receive access to. Importantly, these data-complementary software products are distinct from the AI software studied above. Again, we compare the effects of public security contracts issued following political unrest to non-public security contracts issued in the same local political environment. In Appendix Figure A.15, Panel A, we present estimates from the same specification as in Figure 4, but now considering the outcome of data-complementary software products. One can see that data-complementary software production differentially increases after the receipt of a public security contract in a politically sensitive environment, relative to the receipt of a non-public security contract.

Robustness and ruling out alternative hypotheses The results presented thus far do not appear to be the result of differential selection by firms into politically motivated public security procurement contracts. We find no evidence of pre-contract differences in software production levels or trends, which one would expect if firms selected into these contracts as a function of their productivity trends. As an additional check, we flexibly control for the time-varying effects of firms' age and pre-contract software production, in order to address concerns about firms selecting into contracts as a function of their potential production growth (see Appendix Table A.9, Panels A.2 and A.3). Moreover, by flexibly controlling for the time-varying effects of firms' pre-contract capitalization as well as the dollar value of the contracts, we also account for selection into these contracts on firms' potential benefit from the capital that the contracts provide (see Panels A.4 and A.5). The results are qualitatively and quantitatively similar across these alternative specifications.

Given the complex process of constructing our dataset, it is important to note that our findings are robust to varying several salient dimensions of our analysis (see Appendix Table A.9). First, our results are robust to adjusting our classification of public security contracts to exclude any government agencies ambiguously related to public security (e.g., contracts with the government headquarters, and smart city management and administrative bureaux could be meant to provide security services just for the government office building; see Panel B). Second, the results are robust to adjustments of the parameters of the machine learning algorithm used to classify software — timestep, embedding, and nodes of the RNN LSTM model (see Panel C). Third, our results are robust to considering a balanced panel of firms within a narrow window, and to expanding the window of time around the receipt of the first contract that we study (see Panel D).

Our results are also maintained under specifications that help us address a range of alternative hypotheses. One concern is that contracts with the public security agencies within the powerful, high-surveillance local governments of Beijing or Shanghai may offer a range of political and economic opportunities that go beyond access to data. To rule out the possibility that our findings are distorted by contracts with these two local governments, we estimate our baseline specification, but add fixed effects for contracts from Beijing and Shanghai governments interacted with a full set of quarter to/from contract fixed effects (see Panel E.1). Results are also robust to dropping contracts from the potential outlier province of Xinjiang (see Panel E.2). We additionally account for a firm's home-prefecture/province government potentially giving the firm a commercial advantage beyond the effects of data by estimating the baseline model excluding contracts signed between firms and any government in their home prefecture/province (see Panels E.3 and E.4). Moreover, to address a broader set of concerns about time and space varying shocks that may drive firms' commercial activities, we control for prefecture by quarter fixed effects and show that results are qualitatively similar (see Panel F). Finally, to address the concern that the differential increase in commercial software production is due to more precise customer targeting by the firms, we include the un-categorized general AI software products to the commercial software counts, and we find qualitatively similar and quantitatively even larger effects (see Panel G).

Distortions due to politically motivated contracts? To the extent that politically motivated public security contracts may be accompanied by additional, non-commercial demands from the local government, or may be associated with greater specialization, such contracts could differentially crowd out firms' commercial activities relative to the non-politically motivated contracts.[32] As discussed in Beraja et al. (2021), the greater the effects of politically motivated contracts on software production for the more general commercial market, the greater the impact these contracts would have on the trajectory of innovation in the AI sector.

To evaluate whether politically motivated contracts are associated with differential crowding out of commercial innovation, we compare the effects of politically motivated public security contracts to the effects of non-politically motivated public security contracts. This analysis is analogous to the exercise conducted in Figure 4 and Table 4, except for the types of procurement contracts whose effects we compare.

We now limit our analysis only to public security contracts, and compare those granted out of political motivation with those that are politically neutral. We define politically motivated contracts as those issued following a quarter with above median political unrest (as we did above), and politically neutral contracts as those issued following a quarter with below median political unrest. We estimate the following triple differences specification:

$$y_{it} = \sum_T \beta_{1T} T_{it} + \sum_T \beta_{2T} T_{it} PoliticallyMotivated_i + \alpha_t + \gamma_i + \epsilon_{it}$$

where T_{it} equals 1 if, at time t, T quarters have passed before/since firm i received its first public security contract that is politically neutral; $PoliticallyMotivated_i$ is an indicator that the firm's first public security contract is proceeded by above median level of political unrest; α_t are a full set of quarter fixed effects; and γ_i are a full set of firm fixed effects. The coefficients β_{1T} describe commercial software production of a firm around the time when it receives its first public security contract when this contract is proceeded by below median level of political unrest; and the sums of coefficients $\beta_{1T} + \beta_{2T}$ describe commercial software production around the time when the firm receives its first government contract when this contract is proceeded by above median level of political unrest. If there exists differential crowding out due to the public security contracts' underlying political motivation, one would see negative β_{2T} coefficients following contract receipt.

Figure 5, Panel A, presents the coefficient on the differential effect of politically motivated public security contracts (β_{2T}), for the AI firms' software production for commercial purposes. We do not observe noticeable crowd-out of politically motivated contracts. In fact, if anything, one sees that politically motivated contracts tend to induce firms to produce more commercial software especially towards the later periods of the sampling frame.

[32]This could arise from fixed costs associated with developing products specifically for politically sensitive and demanding environments.

Another potential margin of distortion is the function of the commercial software produced, following politically motivated public security contracts. To explore this margin, we examine the production of commercial software products containing surveillance components such as monitoring and tracking (identified from the registered software descriptions).[33] In Figure 5, Panel B, we present estimates from the same specification as in Panel A, but now considering the outcome of commercial software products containing surveillance components. We find no increase in surveillance oriented commercial software development after the receipt of politically motivated public security contracts.

While these tests are not absolutely conclusive, the absence of evidence of significant distortions arising from politically motivated public security AI procurement suggests that the AI-tocracy equilibrium may be able to sustain continuous commercial AI innovation.

6 Concluding thoughts: the global political consequences of AI-tocracy

Our analysis provides an existence result, documenting an AI-tocracy equilibrium that sustains China's rapid, frontier facial recognition AI innovation and its increased autocratic political control. This equilibrium has direct implications both for China's economic and political trajectories. First, China's autocratic politics may not constrain its ability to continue to move out the technological frontier in AI: rather, frontier innovation in AI may be stimulated precisely because of China's autocratic politics. Second, continued frontier innovation and economic development in China may not be associated with more inclusive political institutions: rather, such innovation may further entrench the autocratic regime.

The AI-tocracy equilibrium could also have important implications beyond China. Such an equilibrium could emerge in other contexts: Russia, in particular, has already deployed facial recognition AI for purpose of political control, and (not coincidentally), alongside China, is among the world's leading producers of frontier facial recognition AI technology.[34] Moreover, autocrats well inside the technological frontier may wish to

[33] Commercial applications of surveillance include parental monitoring of children's location and activities.

[34] Regarding frontier facial recognition technology, Appendix Figure A.16 presents the 2018 ranking of the companies across the world who have the top 10 facial recognition algorithms in terms of prediction accuracy, as ranked by the Face Recognition Vendor Test (FRVT), organized by the National Institute of Standards and Technology (NIST, an agency of the US Department of Commerce) and considered as one of the most authoritative AI industry competitions. Chinese firms occupy all of the top 5 positions and 6 out of

import the AI technology China develops for purposes of political control. Indeed, anecdotal evidence suggests that China's autocratic-supporting AI technology has already been exported to other autocracies.[35] One naturally worries that AI-tocracy may beget more autocracies. The implications of China's AI innovation for the global political and economic landscape are worthy of further, rigorous investigation.

---

the top 10 developers in facial recognition, with Yitu, SenseTime, and MegVii, the 3 largest firms in China, ranking 1st, 3rd, and 8th, respectively; 8 out of the top 10 positions are occupied by Chinese and Russian firms. Regarding Russia's use of facial recognition for political control, see, for example, "In Moscow, Big Brother Is Watching and Recognizing Protesters" by Bloomberg, source: `https://bloom.bg/3tmtsSG`.

[35]For example, according to an Atlantic article, "Xi Jinping is using artificial intelligence to enhance his government's totalitarian control — and he's exporting this technology to regimes around the globe [...] China is already developing powerful new surveillance tools, and exporting them to dozens of the world's actual and would-be autocracies." Source: `https://bit.ly/3ujqj7g`.

References

Acemoglu, Daron and James A Robinson, "Economic Backwardness in Political Perspective," *American Political Science Review*, February 2006, *100* (1), 1–17.

_ **and** _ , *Why Nations Fail* The Origins of Power, Prosperity, and Poverty, New York: Crown Business, August 2012.

_ **and** _ , *The Narrow Corridor* States, Societies, and the Fate of Liberty, Penguin Books, September 2020.

_ **and Pascual Restrepo**, "Artificial Intelligence, Automation and Work," *NBER Working Paper*, January 2018, pp. 1–43.

_ **and** _ , "The wrong kind of AI? Artificial intelligence and the future of labour demand," *Cambridge Journal of Regions, Economy and Society*, December 2019, *13* (1), 25–35.

_ **, Simon Johnson, and James A Robinson**, "Reversal of fortune: Geography and institutions in the making of the modern world income distribution," *The Quarterly Journal of Economics*, 2002, *117* (4), 1231–1294.

Aghion, Philippe, Alberto Alesina, and Francesco Trebbi, "Democracy, Technology, and Growth," Working Paper 13180, National Bureau of Economic Research June 2007.

_ **, Benjamin F Jones, and Charles I Jones**, "Artificial Intelligence and Economic Growth," *NBER Working Paper*, October 2017, pp. 1–57.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, *Prediction Machines* The Simple Economics of Artificial Intelligence, Harvard Business Press, April 2018.

_ **,** _ **, and** _ , "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction," *Journal of Economic Perspectives*, April 2019, *33* (2), 1–22.

Alic, John A, Lewis M Branscomb, Harvey Brooks, and Ashton B Carter, *Beyond Spinoff* Military and Commercial Technologies in a Changing World, Harvard Business Press, 1992.

Angrist, Joshua D. and Alan B. Krueger, "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 1992, *87* (418), 328–336.

Azoulay, Pierre, Erica Fuchs, Anna Goldstein, and Michael Kearney, "Funding Breakthrough Research: Promises and Challenges of the "ARPA Model"," *NBER Working Paper*, June 2018, pp. 1–32.

Bai, Chong-En, Chang-Tai Hsieh, and Zheng Song, "Special Deals with Chinese Characteristics," *NBER Macroeconomics Annual*, January 2020, *34*, 341–379.

Barro, Robert J, "Democracy and Growth," *Journal of Economic Growth*, 1996, *1* (1).

Beraja, Martin, David Y. Yang, and Noam Yuchtman, "Data-intensive Innovation and the State: Evidence from AI Firms in China," *NBER Working Paper*, January 2021, pp. 1–84.

Besley, Timothy and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, August 2009, *99* (4), 1218–1244.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?," *American Economic Review*, April 2020, *110* (4), 1104–1144.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event Study Designs: Robust and Efficient Estimation," *Working Paper*, 2017, pp. 1–35.

Brandt, Loren and Thomas G Rawski, *China's Great Economic Transformation*, Cambridge University Press, April 2008.

Cagé, Julia and Valeria Rueda, "The Long-Term Effects of the Printing Press in Sub-Saharan Africa," *American Economic Journal: Applied Economics*, June 2016, *8* (3), 1–32.

Caicedo, Felipe Valencia, "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America," *The Quarterly Journal of Economics*, January 2019, *134* (1), 1–50.

Campante, Filipe R, Ruben Durante, and Francesco Sobbrio, "Politics 2.0: the Multifaceted Effect of Broadband Internet on Political Participation," *Journal of the European Economic Association*, 2018, *16* (4), 1–59.

Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y Jane Zhang, "Curriculum and Ideology," *Journal of Political Economy*, 2017, *125* (2).

Chen, Yuyu and David Y. Yang, "Media Censorship in China: 1984 or Brave New World?," *American Economic Review*, 2019, *109* (6), 1–133.

Cowgill, Bo and Catherine Tucker, "Algorithmic Fairness and Economics," *Working Paper*, September 2020, pp. 1–31.

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78* (6), 1863–1903.

_ **and Benjamin A Olken**, "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java," *The Review of Economic Studies*, 2020, *87*, 1–40.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, March 2018, *108* (4), 1–36.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova, "Social Media and Protest Participation: Evidence From Russia," *Econometrica*, 2020, *88* (4), 1479–1514.

Geloso, Vincent, Gilbert Berdine, and Benjamin Powell, "Making sense of dictatorships and health outcomes," *BMJ Global Health*, May 2020, *5* (5), e002542.

Glaeser, Edward L, Giacomo A M Ponzetto, and Andrei Shleifer, "Why does democracy need education?," *Journal of Economic Growth*, May 2007, *12* (2), 77–99.

Graham, Loren, "Russian & Soviet Science and Technology," *History of Science Society Newsletter*, 1989, *18* (4).

Greenstein, Shane, *How the Internet Became Commercial* Innovation, Privatization, and the Birth of a New Network, Princeton University Press, October 2015.

Gross, Daniel P and Bhaven N Sampat, "Inventing the Endless Frontier: the Effects of the World War Ii Research Effort on Post-War Innovation," *NBER Working Paper*, June 2020, pp. 1–58.

Guriev, Sergei and Daniel Treisman, "Informational Autocrats," *Journal of Economic Perspectives*, November 2019, *33* (4), 100–127.

_, Nikita Melnikov, and Ekaterina Zhuravskaya, "3G Internet and Confidence in Government," *Working Paper*, 09 2020, pp. 1 – 150.

Hullery, Elise, "History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa," *American Economic Journal: Applied Economics*, March 2009, *1* (2), 1–45.

Ings, Simon, *Stalin and the scientists: a history of triumph and tragedy 1905-1953*, paperback edition ed., London: Faber & Faber, 2017.

Inoue, Atsushi and Gary Solon, "Two-Sample Instrumental Variables Estimators," *The Review of Economics and Statistics*, 2010, *92* (3), 557–561.

Jones, Charles I and Christopher Tonetti, "Nonrivalry and the Economics of Data," *Working Paper*, October 2018, pp. 1–43.

Jr., Mancur Olson, "Dictatorship, Democracy, and Development," *American Political Science Review*, 1993, *87* (3).

King, Gary, Jennifer Pan, and Margaret E Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," *American Political Science Review*, May 2013, *107* (02), 326–343.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, "Human Decisions and Machine Predictions," *The Quarterly Journal of Economics*, 2018, *133* (1), 1–57.

Korinek, Anton and Joseph E Stiglitz, "Artificial Intelligence and Its Implications for Income Distribution and Unemployment," *NBER Working Paper*, December 2017, pp. 1–45.

Larreboure, Magdelena and Felipe Gonzalez, "The Impact of the Women's March on the U.S. House Election," *Working Paper*, April 2021, pp. 1–38.

Lau, Lawrence J, Yingyi Qian, and Gerard Roland, "Reform without Losers: An Interpretation of China's Dual-Track Approach to Transition," *Journal of Political Economy*, February 2000, *108* (1), 120–143.

Lipset, Seymour Martin, "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *American Political Science Review*, 1959, *53* (1), 69–105.

Lowes, Sara and Eduardo Montero, "Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State," *Working Paper*, September 2020, pp. 1–150.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott, "Do Political Protests Matter? Evidence from the Tea Party Movement," *The Quarterly Journal of Economics*, October 2013, *128* (4), 1633–1685.

Moretti, Enrico, Claudia Steinwender, and John Van Reenen, "The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers," *NBER Working Paper*, November 2019, pp. 1–76.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny, "Industrialization and the Big Push," *Journal of Political Economy*, October 1989, *97* (5), 1–25.

North, Douglass C and Barry R Weingast, "Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England," *Journal of Economic History*, 1989.

_, John Joseph Wallis, and Barry R Weingast, *Violence and Social Orders* A Conceptual Framework for Interpreting Recorded Human History, Cambridge: Cambridge University Press, February 2009.

Nunn, Nathan, "The Long-Term Effects of Africa's Slave Trades," *The Quarterly Journal of Economics*, January 2008, *123* (1), 1–38.

Pfetsch, Frank, "Scientific Organisation and Science Policy in Imperial Germany, 1871-1914: The Foundation of the Imperial Institute of Physics and Technology," *Minerva*, 1970, *8* (4), 557–580.

Pravda, "Pravda Newspaper Article, 'Announcement of the First Satellite'," *History and Public Policy Program Digital Archive*, October 1957, pp. 311–312.

Qin, Bei, David Strömberg, and Yanhui Wu, "Social Media and Protests in China," *Working Paper*, March 2020, pp. 1–53.

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

Shirk, Susan L, *China: The Fragile Superpower*, Oxford University Press, April 2007.

Shleifer, Andrei and Robert W Vishny, *The Grabbing Hand* Government Pathologies and Their Cures, Harvard University Press, 2002.

Slezkine, Yuri, *The House of Government: a saga of the Russian Revolution*, Princeton ; Oxford: Princeton University Press, 2017.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing Like China," *American Economic Review*, February 2011, *101* (1), 196–233.

Tirole, Jean, "Digital Dystopia," *Working Paper*, February 2020, pp. 1–36.

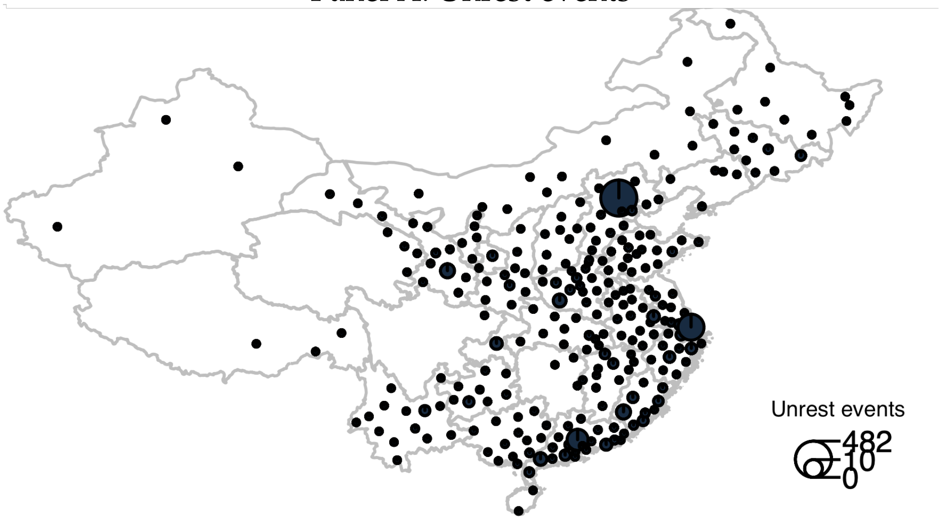
Treisman, Daniel, "Economic Development and Democracy: Predispositions and Triggers," *Annual Review of Political Science*, 2020, *23*, 1–19.

Yue, Chia Siow, "Singapore Model of Industrial Policy-Past and Present," in "Second LAEBA Annual Meeting, Buenos Aires, Argentina" 2005.

Zhang, Daniel, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, James Manyika, Juan Carlos Niebles, Michael Sellitto, Yoav Shoham, Jack Clark, and Raymond Perrault, "The AI Index 2021 Annual Report," Technical Report, AI Index Steering Committee, Human-Centered AI Institute, Stanford University March 2021.

Zuboff, Shoshana, *The Age of Surveillance Capitalism* The Fight for a Human Future at the New Frontier of Power, PublicAffairs, January 2019.

Panel A: Unrest events



Panel B: Public security AI contracts

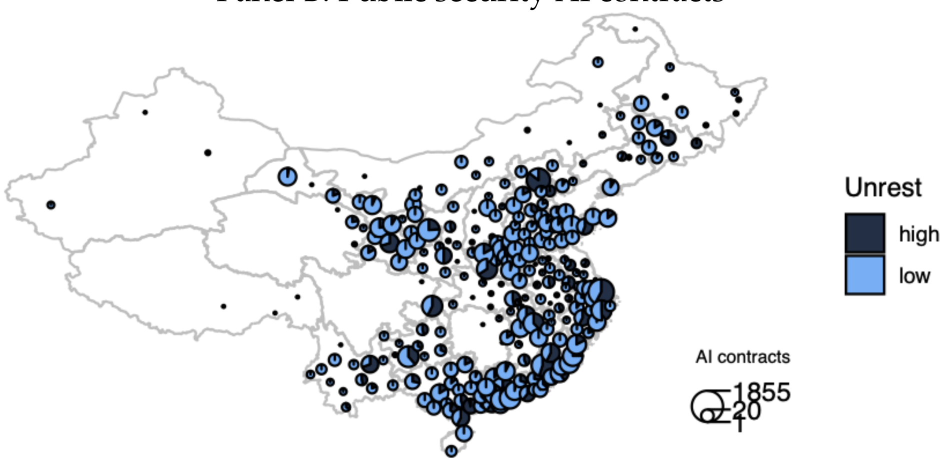


Figure 2: Each circle represents a prefecture in our dataset that has at least one public security AI contract that is an AI firm's first government contract. In Panel A, circle size indicates the number of unrest events in a prefecture, while in Panel B, circle size indicates the number of public security AI contracts awarded in the prefecture (larger circles indicate more, log scale). Circle shading in Panel B indicates the fraction of first AI contracts that were procured during high or low unrest periods, where the within-prefecture variation comes from changes in the number of unrest events in a prefecture over time (a larger fraction of dark shading indicates a larger fraction of prefecture contracts procured during high unrest periods).

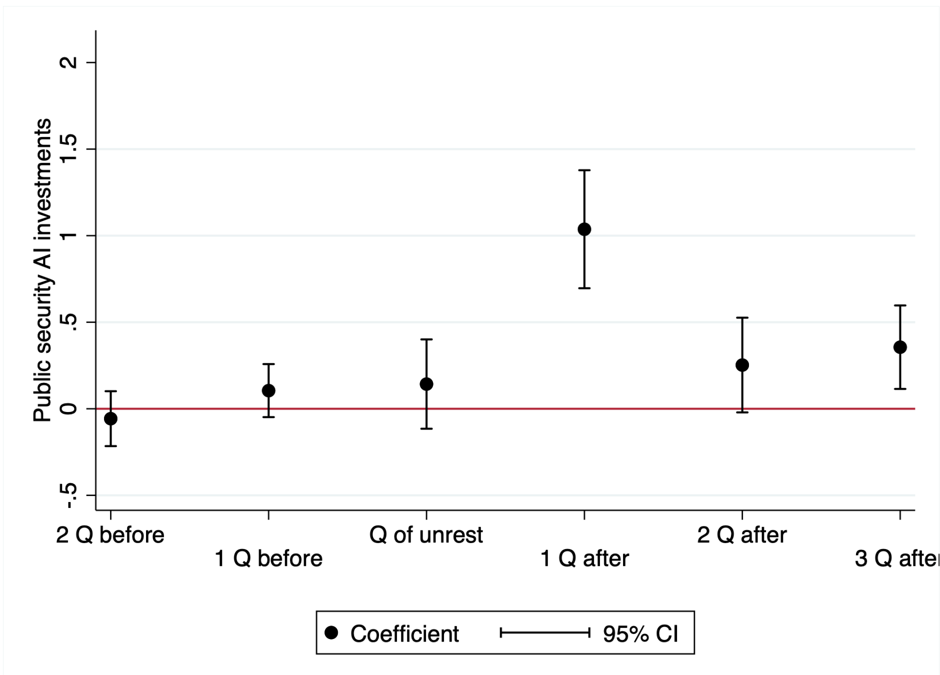


Figure 3: Public security AI investments relative to the quarter of political unrest. Coefficients and confidence intervals displayed are from a regression following the specification in Table 2, Panel A, column 5 (all controls), but additionally stacking observations containing (residualized) unrest from multiple quarters. Public security AI investments are per million residents.

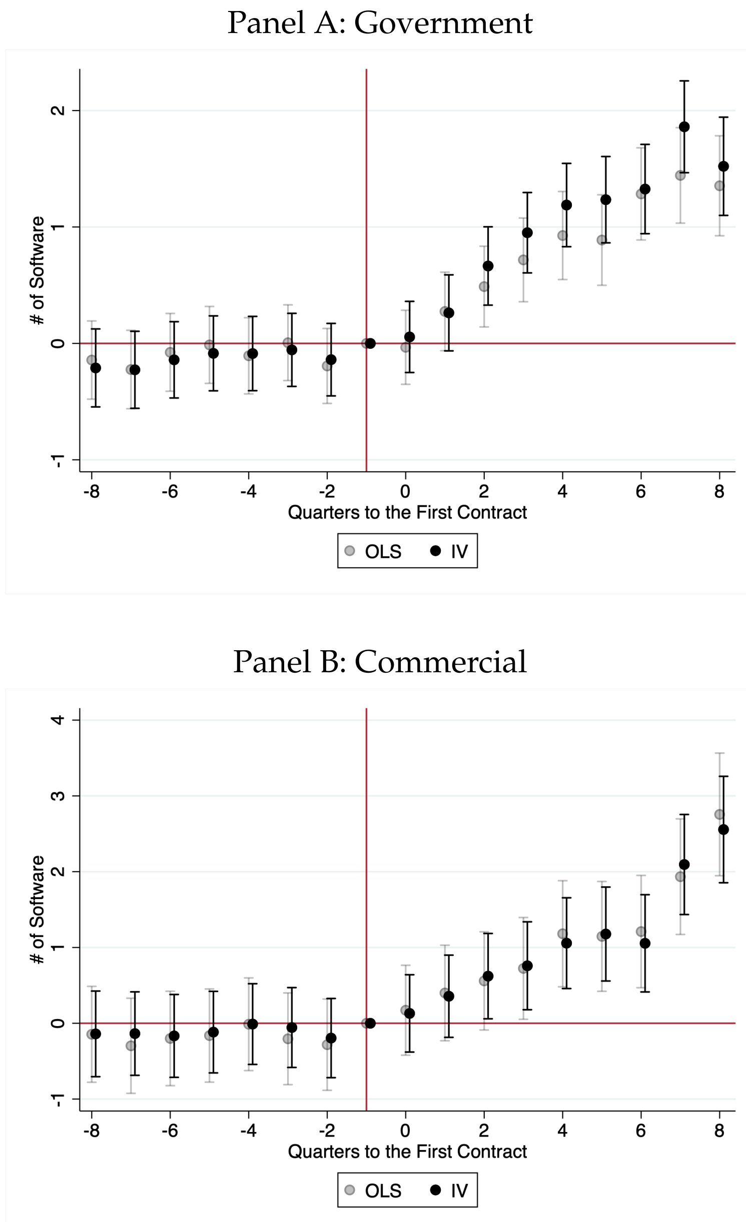
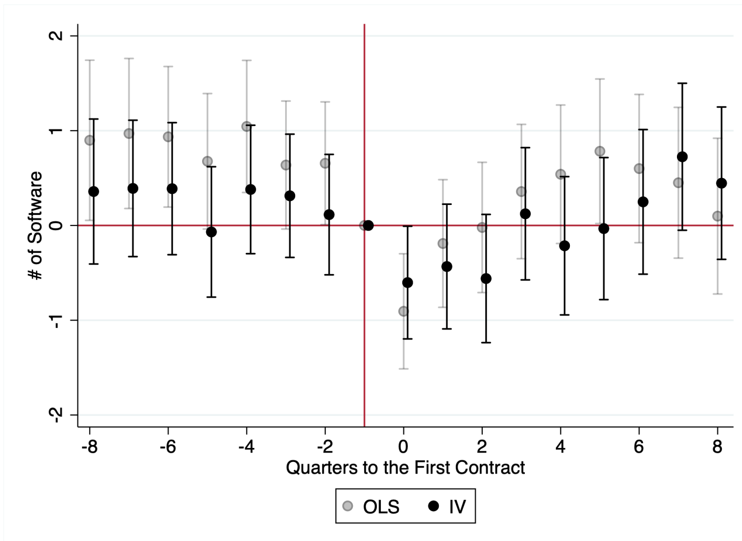


Figure 4: Differential software development by firms that receive public security contracts versus non-public security ones, relative to the time of receiving the initial contract. All panels restrict firms to those that receive above median unrest contracts, and control for firm and time period fixed effects. Panel A shows software intended for government uses, and Panel B for commercial. Translucent lines/markers use weather to instrument for unrest.

Panel A: Commercial



Panel B: Commercial surveillance

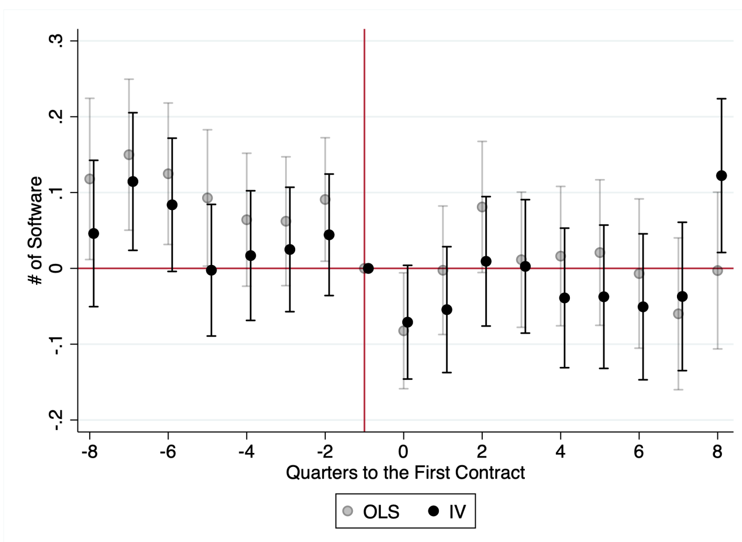


Figure 5: Differential software development by firms that receive politically motivated contracts (above median unrest) versus non-politically motivated ones, relative to the time of receiving the initial contract. All panels restrict firms to those that receive above median unrest contracts, and control for firm and time period fixed effects. Panel A shows software intended for commercial uses, and Panel B for commercial surveillance. Translucent lines/markers use weather to instrument for unrest.

Table 1: Summary statistics

	Mean	S.D.
	(1)	(2)
Panel A: Political unrest		
All events (per prefecture-quarter)	2.419	18.490
Protests	0.607	4.603
Demands	0.720	5.009
Threats	1.092	9.479
Panel B: Procurement of AI and the technology of political control		
All AI contracts (per prefecture-quarter)	3.976	7.818
Non-public security contracts	2.285	5.118
Public security contracts	1.691	3.476
First public security contracts	0.082	0.327
Surveillance cameras (per prefecture-quarter)	2,118	12,684
Police hires (per prefecture-year)	59.278	84.991
Panel C: Innovation of AI firms		
All software (per firm-quarter)	5.756	7.124
Government software	1.724	3.337
Commercial software	2.353	3.675
Data-complementary software	2.273	3.605
Surveillance software	0.588	2.126

Notes: This table presents summary statistics at the prefecture-quarter level (firm-quarter for Panel C) for variables of interest. Column 1 shows the sample mean and column 2 the standard deviation. Panel A presents counts of unrest events, Panel B presents counts of local government procured facial AI contracts and other technologies of political control, and Panel C presents counts of software produced by facial AI firms. For Panels A and B, $N = 8,167$ (Panel B police hires, $N = 2,672$). For Panel C, N = 23,697

Table 2: Effect of unrest events on facial AI procurement

	Public security AI procurement					*Non-public security AI procurement*				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: OLS										
Event	0.962***	0.971***	0.962***	0.969***	0.972***	0.033	0.034	0.031	0.034	0.034
	(0.244)	(0.239)	(0.243)	(0.240)	(0.238)	(0.029)	(0.028)	(0.028)	(0.029)	(0.028)
Panel B: IV										
Event	0.702**	0.777**	0.772**	0.766**	0.759**	-0.122	-0.105	-0.117	-0.111	-0.104
	(0.351)	(0.356)	(0.347)	(0.359)	(0.357)	(0.097)	(0.088)	(0.094)	(0.092)	(0.085)
D.V. mean	12.856	12.865	12.905	12.865	12.905	3.017	3.019	3.028	3.019	3.028
D.V. sd	60.906	60.926	61.016	60.926	61.016	17.887	17.893	17.920	17.893	17.920
N	8424	8418	8392	8418	8392	8424	8418	8392	8418	8392
GDP × time	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Population × time	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Gov. revenue × time	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Notes: This table presents regressions at the prefecture-quarter level. The outcome is the number of facial AI contracts procured by the local government per capita, scaled up by 1,000,000. In columns 1 - 4, these are public security contracts, while in columns 5 - 8, these are non-public security contracts. There is a one quarter lag between the quarter of unrest events occurring and the number of public security AI contracts procured by the local government. Columns 2 and 7 control for prefecture GDP × quarter effects, columns 3 and 8 controls for prefecture population × quarter effects, columns 4 and 9 controls for prefecture government tax revenue × quarter effects, and columns 5 and 10 includes all controls. Panel B uses weather variables as selected by LASSO to instrument for unrest events. These variables are: max. temperature over 95 dummy X hail, thunder X hail, hail X max. gust speed, thunder X max. gust speed, and snow depth X precipitation, each interacted with a dummy for whether an unrest event occurred on the day. All specifications include prefecture and quarter fixed effects. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table 3: Effect of AI procurement on suppressing unrest

	Standardized number of unrest events									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Fine weather	0.7618***	0.7548***	0.7599***	0.7556***	0.7581***	0.7922***	0.7850***	0.7917***	0.7860***	0.7890***
	(0.1377)	(0.1359)	(0.1368)	(0.1360)	(0.1363)	(0.1495)	(0.1478)	(0.1487)	(0.1478)	(0.1480)
Public security AI_{t-1}	0.0014	0.0024	0.0051	0.0023	0.0030					
	(0.0041)	(0.0076)	(0.0181)	(0.0068)	(0.0170)					
Fine weather × public security AI_{t-1}	-0.3660**	-0.3656**	-0.3822**	-0.3660**	-0.3739**					
	(0.1699)	(0.1691)	(0.1780)	(0.1693)	(0.1752)					
Non-public security AI_{t-1}						-0.0014	-0.0012	-0.0025	-0.0013	-0.0024
						(0.0014)	(0.0015)	(0.0019)	(0.0015)	(0.0018)
Fine weather × non-public security AI_{t-1}						-0.0447	-0.0447	-0.0474	-0.0451	-0.0427
						(0.0301)	(0.0290)	(0.0307)	(0.0296)	(0.0279)
GDP × time	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Population × time	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Gov. revenue × time	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Notes: Panels A presents regressions at the prefecture-quarter level. The dependent variable is the standardized number of events in the prefecture. Fine weather is the standardized number of predicted events from the good weather LASSO variables interacted with whether there was an event elsewhere with fixed effects. AI (public security AI contracts per capita in columns 1 - 5, non-public in columns 6 - 10) is also standardized. Prefecture and quarter fixed effects are included. Columns 1 and 6 present baseline results, columns 2 and 7 add controls for local GDP by quarter fixed effects, columns 3 and 8 add controls for local population by quarter fixed effects, columns 4 and 9 add controls for prefectural government tax revenue by quarter fixed effects, and columns 5 and 10 add all prior controls. Standard errors are two-way clustered by prefecture and quarter and allow for autocorrelation. * significant at 10% ** significant at 5% *** significant at 1%.

Table 4: Effect of public security contracts on software production in high unrest prefectures

	Government software				Commercial software			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
8 quarters before contract	0.042	0.006	-0.048	-0.065	0.030	0.026	0.038	0.055
	(0.114)	(0.170)	(0.122)	(0.179)	(0.214)	(0.278)	(0.207)	(0.272)
7 quarters before contract	0.018	-0.011	-0.065	-0.068	0.103	0.091	0.085	0.099
	(0.112)	(0.167)	(0.120)	(0.176)	(0.207)	(0.269)	(0.201)	(0.264)
6 quarters before contract	-0.016	-0.033	-0.050	-0.051	0.087	0.092	0.048	0.068
	(0.110)	(0.166)	(0.118)	(0.173)	(0.206)	(0.269)	(0.199)	(0.262)
5 quarters before contract	-0.016	-0.031	-0.076	-0.075	0.006	0.002	-0.007	0.008
	(0.109)	(0.164)	(0.117)	(0.171)	(0.202)	(0.263)	(0.194)	(0.256)
4 quarters before contract	0.004	-0.009	-0.073	-0.075	0.023	0.019	-0.017	-0.005
	(0.107)	(0.160)	(0.113)	(0.166)	(0.199)	(0.259)	(0.190)	(0.251)
3 quarters before contract	-0.021	-0.034	-0.050	-0.054	0.033	0.032	-0.004	0.006
	(0.106)	(0.160)	(0.112)	(0.165)	(0.197)	(0.257)	(0.188)	(0.248)
2 quarters before contract	0.018	0.003	-0.008	-0.018	0.074	0.051	0.042	0.024
	(0.105)	(0.157)	(0.111)	(0.163)	(0.195)	(0.255)	(0.186)	(0.245)
Receiving 1st contract	0.308***	0.270*	0.277**	0.236	0.315	0.305	0.286	0.286
	(0.104)	(0.156)	(0.109)	(0.161)	(0.194)	(0.253)	(0.184)	(0.243)
1 quarter after contract	0.499***	0.458***	0.545***	0.519***	0.632***	0.630**	0.652***	0.685***
	(0.109)	(0.164)	(0.116)	(0.170)	(0.204)	(0.265)	(0.195)	(0.256)
2 quarters after contract	0.764***	0.717***	0.783***	0.763***	1.026***	1.009***	1.009***	1.038***
	(0.113)	(0.170)	(0.121)	(0.177)	(0.213)	(0.276)	(0.203)	(0.267)
3 quarters after contract	1.014***	0.954***	1.001***	0.967***	1.204***	1.174***	1.252***	1.272***
	(0.118)	(0.176)	(0.125)	(0.182)	(0.219)	(0.284)	(0.209)	(0.274)
4 quarters after contract	1.329***	1.256***	1.279***	1.241***	1.538***	1.526***	1.463***	1.512***
	(0.124)	(0.185)	(0.130)	(0.189)	(0.227)	(0.294)	(0.216)	(0.283)
5 quarters after contract	1.695***	1.616***	1.538***	1.501***	1.919***	1.904***	1.864***	1.926***
	(0.126)	(0.188)	(0.133)	(0.194)	(0.234)	(0.303)	(0.221)	(0.290)
6 quarters after contract	2.015***	1.918***	1.782***	1.742***	2.548***	2.562***	2.469***	2.570***
	(0.132)	(0.197)	(0.139)	(0.203)	(0.245)	(0.317)	(0.232)	(0.304)
7 quarters after contract	2.622***	2.520***	2.440***	2.406***	2.770***	2.770***	2.692***	2.791***
	(0.140)	(0.208)	(0.146)	(0.213)	(0.261)	(0.339)	(0.246)	(0.323)
8 quarters after contract	3.159***	3.055***	2.964***	2.938***	3.526***	3.513***	3.414***	3.510***
	(0.147)	(0.220)	(0.158)	(0.230)	(0.273)	(0.354)	(0.263)	(0.345)
8 quarters before contract × public security	-0.137	-0.185	-0.207	-0.223	-0.147	-0.219	-0.135	-0.177
	(0.202)	(0.304)	(0.202)	(0.297)	(0.384)	(0.500)	(0.342)	(0.452)
7 quarters before contract × public security	-0.218	-0.260	-0.220	-0.241	-0.299	-0.342	-0.131	-0.153

	(0.203)	(0.305)	(0.200)	(0.293)	(0.381)	(0.497)	(0.334)	(0.441)
6 quarters before contract × public security	-0.072	-0.106	-0.135	-0.152	-0.200	-0.253	-0.160	-0.178
	(0.201)	(0.302)	(0.198)	(0.290)	(0.378)	(0.493)	(0.331)	(0.438)
5 quarters before contract × public security	-0.011	-0.042	-0.088	-0.107	-0.164	-0.201	-0.109	-0.121
	(0.199)	(0.299)	(0.194)	(0.285)	(0.373)	(0.487)	(0.326)	(0.431)
4 quarters before contract × public security	-0.100	-0.129	-0.087	-0.101	-0.016	-0.044	-0.011	-0.021
	(0.197)	(0.296)	(0.192)	(0.283)	(0.371)	(0.484)	(0.323)	(0.427)
3 quarters before contract × public security	0.006	-0.013	-0.059	-0.066	-0.206	-0.231	-0.057	-0.063
	(0.196)	(0.295)	(0.189)	(0.278)	(0.367)	(0.479)	(0.319)	(0.422)
2 quarters before contract × public security	-0.189	-0.193	-0.142	-0.134	-0.280	-0.270	-0.194	-0.175
	(0.194)	(0.292)	(0.188)	(0.276)	(0.365)	(0.477)	(0.317)	(0.419)
Receiving 1st contract × public security	0.044	0.056	0.043	0.059	0.077	0.050	0.196	0.168
	(0.191)	(0.288)	(0.184)	(0.270)	(0.360)	(0.469)	(0.309)	(0.408)
1 quarter after contract × public security	0.394*	0.361	0.347*	0.340	0.386	0.321	0.444	0.402
	(0.203)	(0.305)	(0.197)	(0.289)	(0.382)	(0.498)	(0.329)	(0.434)
2 quarters after contract × public security	0.554***	0.544*	0.714***	0.716**	0.508	0.488	0.695**	0.688
	(0.209)	(0.313)	(0.203)	(0.297)	(0.393)	(0.512)	(0.341)	(0.450)
3 quarters after contract × public security	0.707***	0.695**	1.054***	1.074***	0.724*	0.696	0.789**	0.789*
	(0.216)	(0.325)	(0.208)	(0.305)	(0.407)	(0.530)	(0.352)	(0.464)
4 quarters after contract × public security	1.059***	1.069***	1.262***	1.321***	1.153***	1.107**	1.155***	1.157**
	(0.228)	(0.341)	(0.216)	(0.316)	(0.424)	(0.553)	(0.363)	(0.478)
5 quarters after contract × public security	1.002***	0.975***	1.308***	1.321***	1.191***	1.128**	1.268***	1.243**
	(0.234)	(0.351)	(0.224)	(0.328)	(0.439)	(0.572)	(0.376)	(0.495)
6 quarters after contract × public security	1.245***	1.235***	1.416***	1.432***	1.163***	1.077*	1.166***	1.102**
	(0.238)	(0.357)	(0.231)	(0.339)	(0.449)	(0.585)	(0.389)	(0.512)
7 quarters after contract × public security	1.427***	1.417***	1.976***	1.978***	1.980***	1.896***	2.144***	2.068***
	(0.247)	(0.370)	(0.238)	(0.349)	(0.462)	(0.602)	(0.400)	(0.527)
8 quarters after contract × public security	1.353***	1.359***	1.626***	1.632***	2.695***	2.634***	2.683***	2.632***
	(0.258)	(0.388)	(0.254)	(0.373)	(0.490)	(0.639)	(0.426)	(0.561)
Regression	OLS	OLS	IV	IV	OLS	OLS	IV	IV
Event-study weighting	No	Yes	No	Yes	No	Yes	No	Yes

Notes: The table presents regression coefficients for facial AI firms that earn contracts from local governments when there is an above median amount of unrest in the quarter prior to the contract. The table shows the difference in software production between firms that earn politically motivated (public security) contracts versus non-politically motivated contracts. In the IV specification (columns 3-4, 7-8), local unrest is instrumented by weather. Columns 1-4 present results for amount of government software produced by the firm, while columns 5-8 present results for commercial software. All columns control for time period fixed effects and subsidiary firm fixed effects. Columns 2, 4, 6, and 8 weight the control group by 10 times more than the treatment, following Borusyak et al. (2017). Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

ONLINE APPENDIX

Appendix A Auxiliary data sources

In addition to the primary data sources described in Section 3, we also use a number of auxiliary data sources for the empirical analysis.

Uyghur minority share As an auxiliary measurement of (perceived risk of) political unrest, we use the Uyghur minority population present in each prefecture. Uyghurs have been one of the primary targets of Chinese state surveillance and are viewed as a security risk by the central government. See, for example, a recent *Reuter*'s report that states, "Beijing accuses separatists among the Muslim Uyghur ethnic minority there [in Xinjiang] of stirring up tensions with the ethnic Han Chinese majority and plotting attacks elsewhere in China" (source: `https://reut.rs/332IYs9`). Also, a recent article from the *Atlantic* notes that "Uyghurs can travel only a few blocks before encountering a checkpoint outfitted with one of Xinjiang's hundreds of thousands of surveillance cameras. Footage from the cameras is processed by algorithms that match faces with snapshots taken by police at 'health checks' " (source: `https://bit.ly/3aWC1gB`).

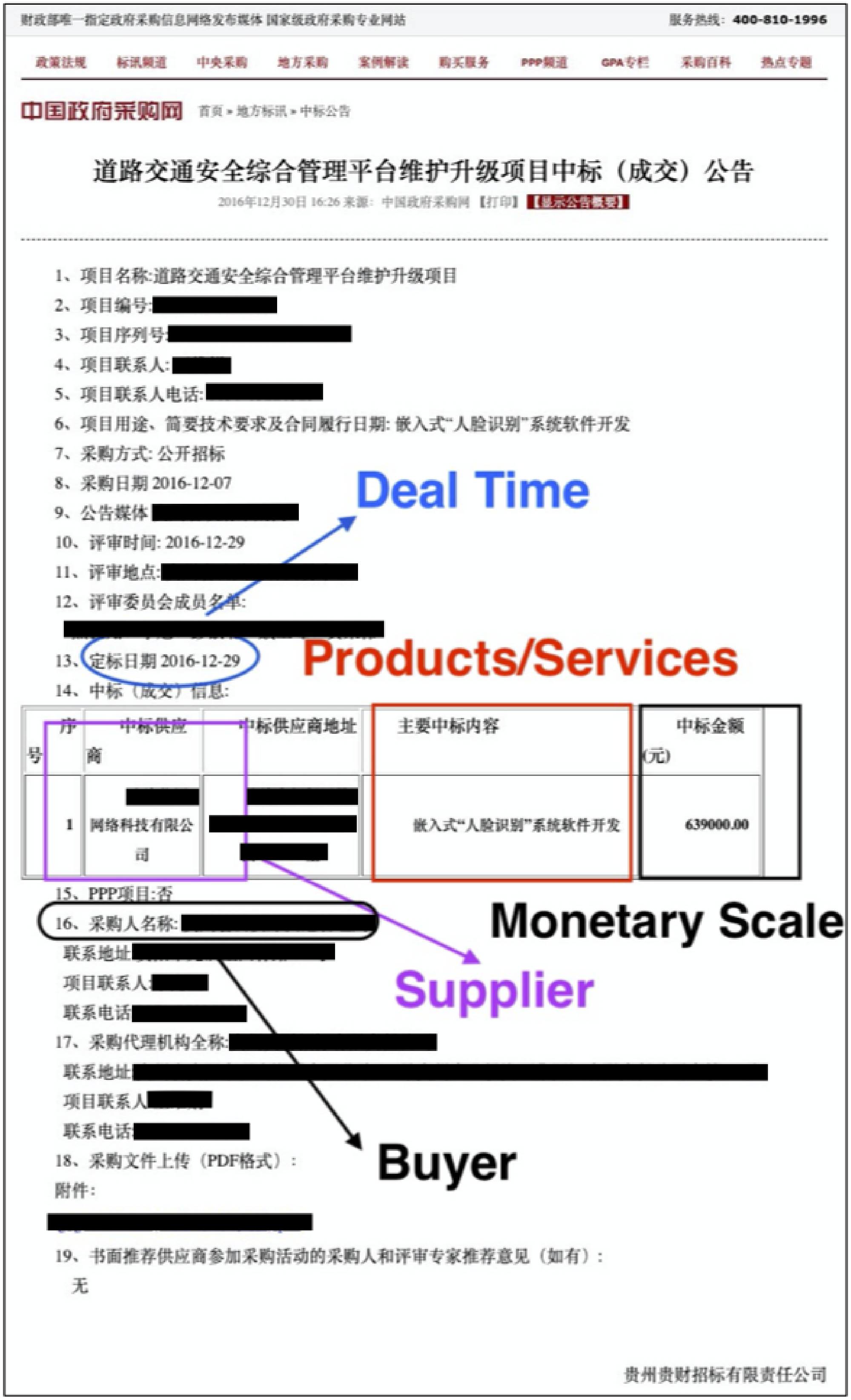
We collect data on the number of Uyghurs and Uyghur men in each prefecture from the Chinese Statistical Yearbooks in the year 2000, and use the fraction of the population that are Uyghur or Uyghur men to proxy for government concern for political unrest.

Local governments' procurement of surveillance cameras In addition to the public security procurement of AI technology, we also observe local government's investments in two complementary technologies for public security purposes. First, we identify local public security units' procurement of high-resolution surveillance cameras, which are capable of collecting data for any AI control systems that may be in place. We construct a panel of the number of surveillance cameras in a given prefecture at the monthly level; when the number of cameras purchased in a given contract is not disclosed, we use the monetary value of the contract to impute the number of cameras purchased. In total, we identify 17,306 public security procurement contracts for surveillance cameras; during the period between 2013 and 2019, the average prefecture purchased 60,437 surveillance cameras (median = 20,439 and standard deviation = 117,672).

Local governments' police hiring Second, we collect data on personnel hiring by local police departments. From the website of OffCN Education Technology, we collect comprehensive listings of the number of police officers' job openings posted and fulfilled by each department in a given year. OffCN Education Technology is a private firm providing labor market services specializing in the public sector; see `http://sd.offcn.com/` for details.

Using job-specific details, we are able to observe changes in police department hiring composition over time, by classifying police new hires into "field jobs" (e.g., police on the street) that require lower human capital, and "office jobs" (e.g., police working in the office) that require higher human capital. There are approximately 15,500 unique job positions to classify. We manually classify the 2,000 most common jobs as either field or office based on the job's title, description and requirements, and use keyword matching to classify the remainder. During the period of 2013 to 2019, the average local police department makes 32 hires in a year, of which 14 hires are for desk jobs.

Appendix B Additional figures and tables


财政部唯一指定政府采购信息网络发布媒体 国家级政府采购专业网站 服务热线：400-810-1996

政策法规 标讯频道 中央采购 地方采购 案例解读 购买服务 PPP频道 GPA专栏 采购百科 热点专题

中国政府采购网 首页 » 地方标讯 » 中标公告

道路交通安全综合管理平台维护升级项目中标（成交）公告

2016年12月30日 16:26 来源：中国政府采购网 【打印】【显示公告概要】

1、项目名称:道路交通安全综合管理平台维护升级项目
2、项目编号:[redacted]
3、项目序列号:[redacted]
4、项目联系人:[redacted]
5、项目联系人电话:[redacted]
6、项目用途、简要技术要求及合同履行日期: 嵌入式"人脸识别"系统软件开发
7、采购方式: 公开招标
8、采购日期 2016-12-07
9、公告媒体 [redacted]
10、评审时间: 2016-12-29
11、评审地点:[redacted]
12、评审委员会成员名单:
[redacted]
13、定标日期 2016-12-29
14、中标（成交）信息:

序号	中标供应商	中标供应商地址	主要中标内容	中标金额(元)
1	[redacted] 网络科技有限公司	[redacted]	嵌入式"人脸识别"系统软件开发	639000.00

15、PPP项目:否
16、采购人名称:[redacted]
联系地址 [redacted]
项目联系人 [redacted]
联系电话 [redacted]
17、采购代理机构全称:[redacted]
联系地址 [redacted]
项目联系人 [redacted]
联系电话 [redacted]
18、采购文件上传（PDF格式）：
附件:
[redacted]
19、书面推荐供应商参加采购活动的采购人和评审专家推荐意见（如有）：
无

贵州贵财招标有限责任公司

Figure A.1: Example of a procurement contract record; source: Chinese Government Procurement Database.

Figure A.2: Example of AI firm record from *Tianyancha* (excerpt).


Highlights

Employees
1,000
As of 24-Oct-2018

Last Deal Details
Undisclosed
Later Stage VC 06-May-2019

Total Raised to Date
$355.16M
As of 06-May-2019

Edit Highlights

Timeline

Round & Amount

General Information

Description
Provider and developer of artificial intelligence technology used in the fields of smart cities, smart medical, and smart commerce. The company is engaged in the research of computer vision, image and video intelligent understanding, distributed system and big data application, it offers traffic management software, medical diagnostic technology and intelligent hardware, enabling companies to apply AI technology in their products.

Most Recent Financing Status (as of 13-Feb-2020)
The company raised an undisclosed amount of venture funding from [redacted]
Previously, the company raised $300 million of Series C+ venture funding from [redacted]

Website	[redacted]		
Entity Types	Private Company Acquirer	Financing Status	Venture Capital-Backed
		Year Founded	2012
Legal Name	[redacted]	Universe	Venture Capital
Business Status	Generating Revenue	Employees	1,000
Ownership Status	Privately Held (backing)		View Employee History

Industries & Verticals

Primary Industry
Business/Productivity Software

Verticals
Artificial Intelligence & Machi...
Big Data
Digital Health
TMT

What PitchBook Analysts Say
View More Analyst Insights

"Both incumbents and startups are developing new hardware. While Google is putting their custom tensor processing units (TPUs) to use for many recent breakthroughs, independent leaders such as Cerebras and Graphcore have raised significant capital and developed other novel designs to cater to AI & ML applications."

| 10-Dec-2019 | Cameron Stanfill | Artificial Intelligence & Machine Learning +3

Contact Information

Primary Contact
[redacted]
Co-Founder & Chief Executive Officer
Phone: [redacted]

Primary Office
[redacted]
China
Phone: [redacted]

Alternate Offices (4)
Beijing
[redacted]
China
Phone: [redacted]

Figure A.3: Example of AI firm record from *Pitchbook* (excerpt).

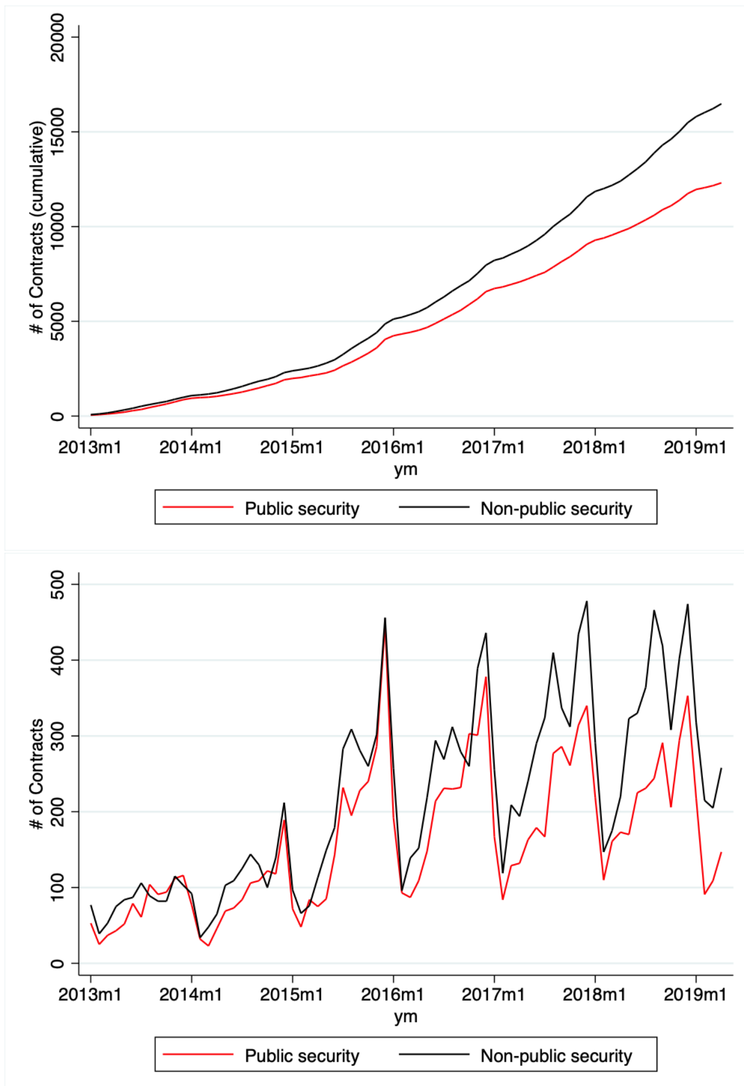


Figure A.4: Cumulative number of public security and non-public security contracts (top panel), and the flow of new contracts signed in each month (bottom panel).

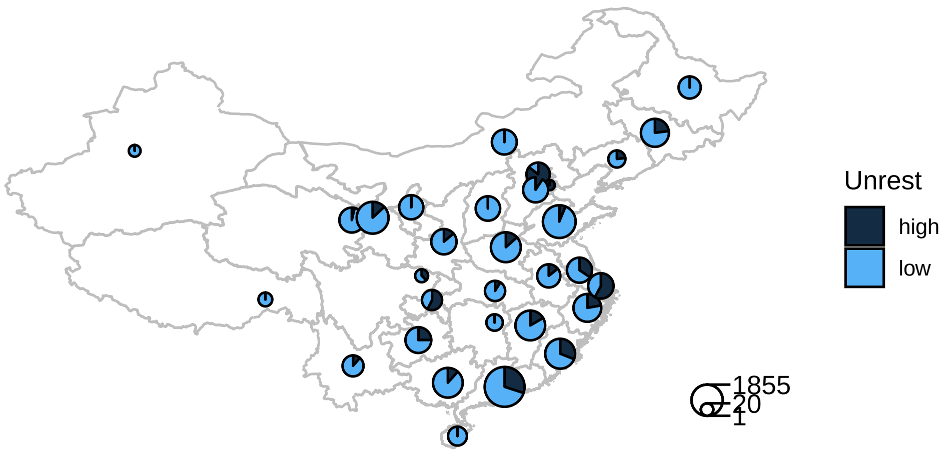


Figure A.5: Each circle represents a province in our dataset that has at least one public security AI contract that is some AI firm's first government contract. Circle size indicates the number of public security AI contracts awarded to a prefecture in the province (larger circles indicate more contracts, log scale), where prefecture-level contracts are weighted by the number of prefectures in the province. Circle shading indicates the fraction of first AI contracts that were procured during high or low unrest periods, where the within-prefecture variation comes from changes in the number of unrest events in a prefecture over time (a larger fraction of dark shading indicates a larger fraction of prefecture contracts procured during high unrest periods).

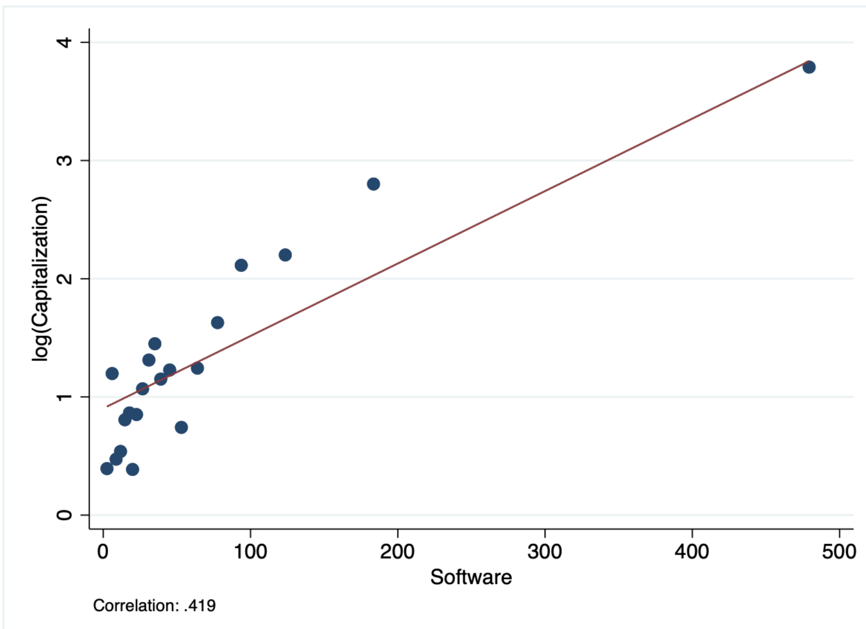
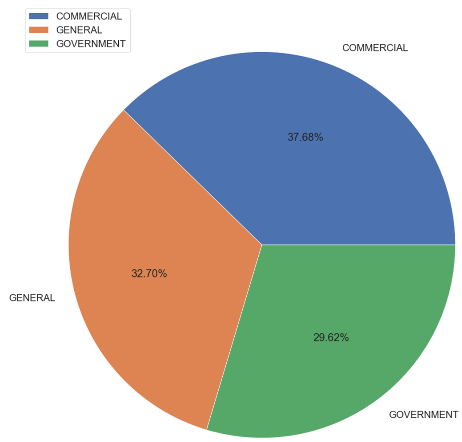
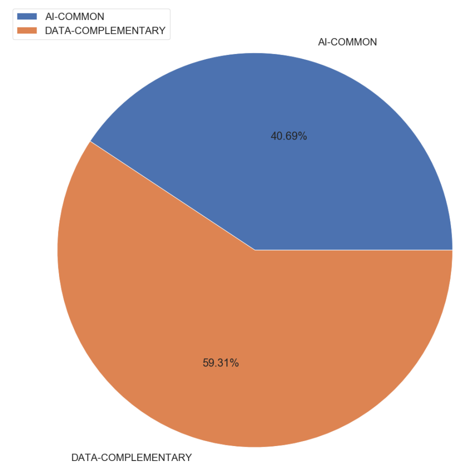


Figure A.6: Binscatter plot at the firm level of log(firm capitalization) and amount of software produced.

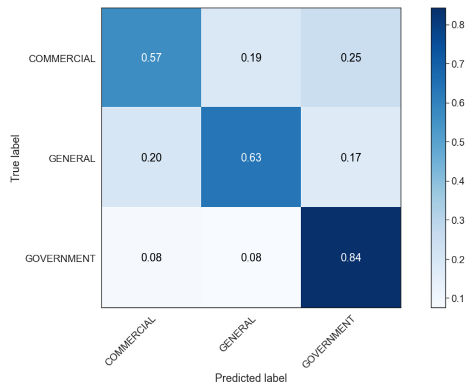


(a) Customers

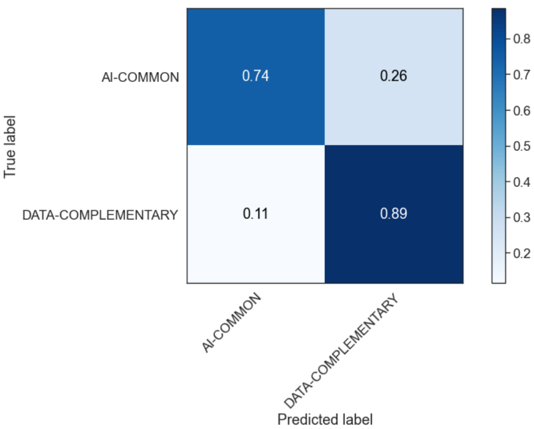


(b) Function

Figure A.7: Summary statistics of categorization outcomes for software categorizations based on Recurrent Neural Network with Long Short-Term Memory algorithm. Left panel shows categorization by customers; right panel shows categorization by function.

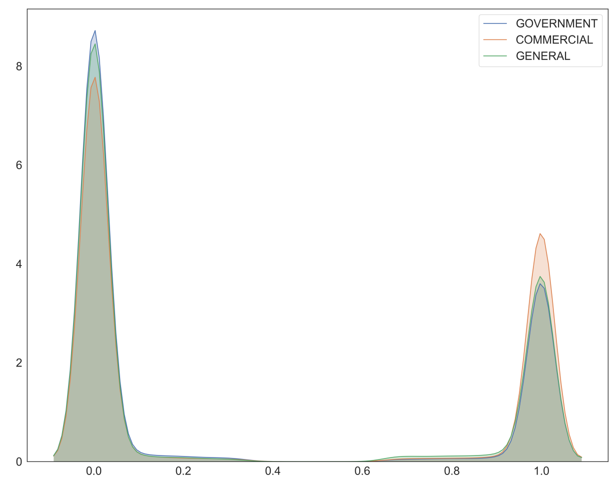


(a) Customers

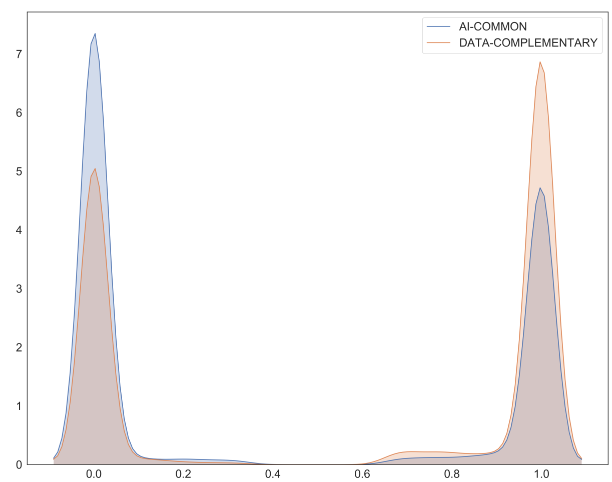


(b) Functions

Figure A.8: Confusion matrix of categorization outcomes for software categorizations. True labels are based on training set constructed by human categozations (performed by two individuals). Predicted labels are outputs based on Recurrent Neural Network with Long Short-Term Memory algorithm. Left panel shows categorization by customers; right panel shows categorization by function.



(a) Customers



(b) Function

Figure A.9: Probability density plots of software categorizations based on Recurrent Neural Network with Long Short-Term Memory algorithm. Top panel shows categorization by customers; bottom panel shows categorization by function.

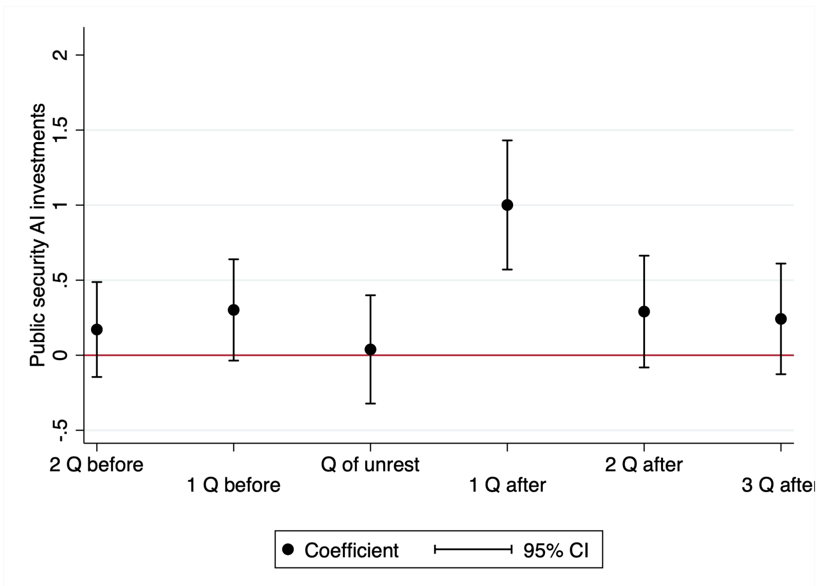


Figure A.10: Public security AI investments relative to the quarter of political unrest. Coefficients and confidence intervals displayed are from a regression following the specification in Table 2, Panel B (weather IV for unrest), column 5, but additionally stacking observations containing (residualized) unrest from multiple quarters. Public security AI investments are per million residents.

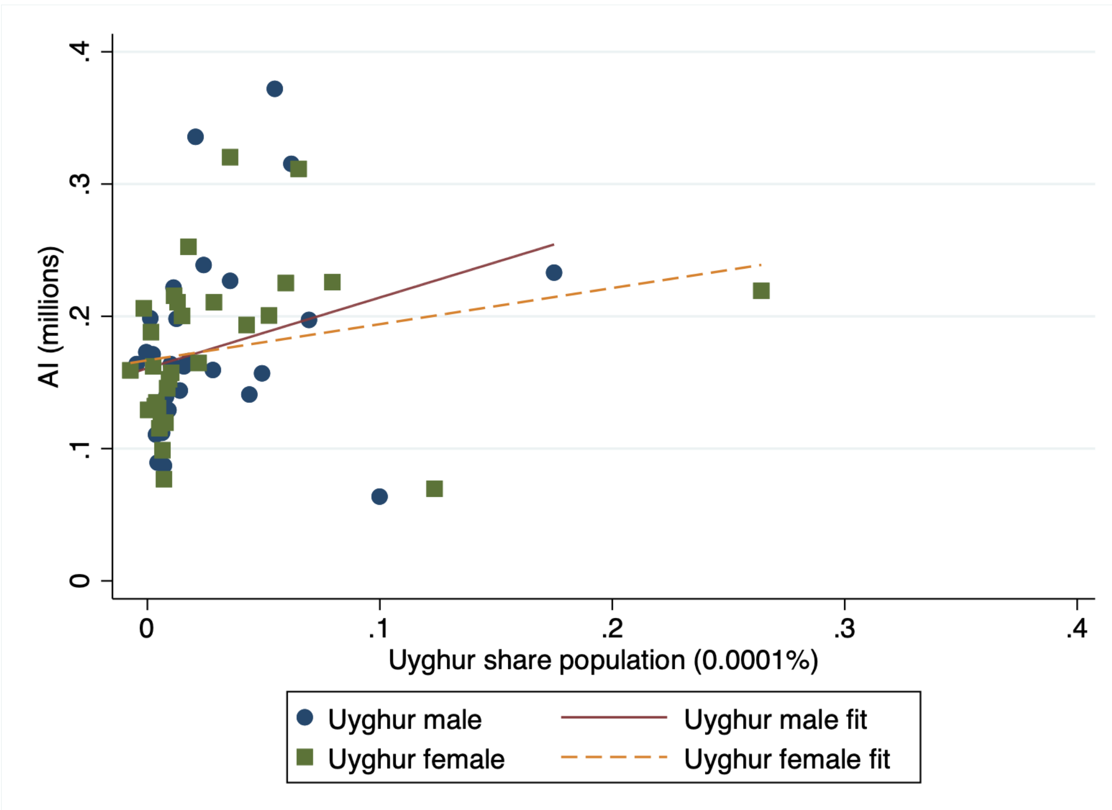


Figure A.11: Binscatter of Uyghur share of population (0.0001%) on millions of AI contracts procured at the prefecture level. Blue circles show the binscatter for Uyghur men and green squares show the binscatter for Uyghur women. The solid red line shows the linear best fit for Uyghur men and the dashed orange line shows the linear best fit for Uyghur women.

Panel A: Government surveillance

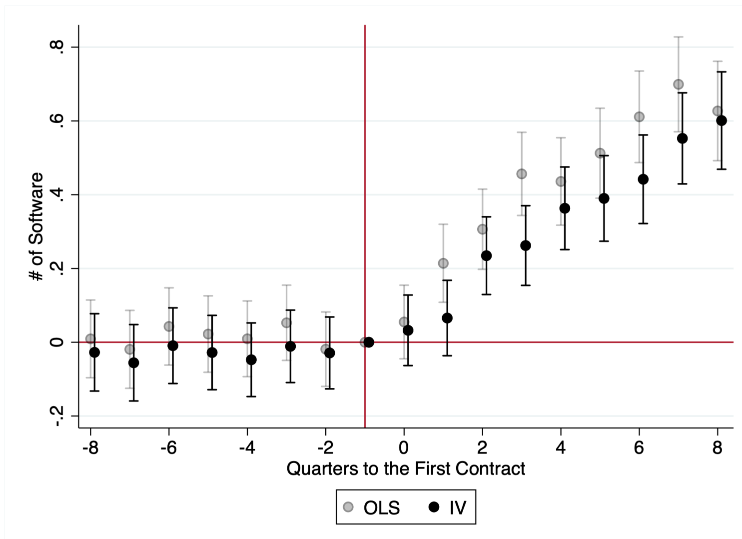


Figure A.12: Differential software development by firms that receive public security contracts versus non-public security ones, relative to the time of receiving the initial contract. All panels restrict firms to those that receive above median unrest contracts, and control for firm and time period fixed effects. Panel A shows software intended for government surveillance uses. Translucent lines/markers use weather to instrument for unrest.

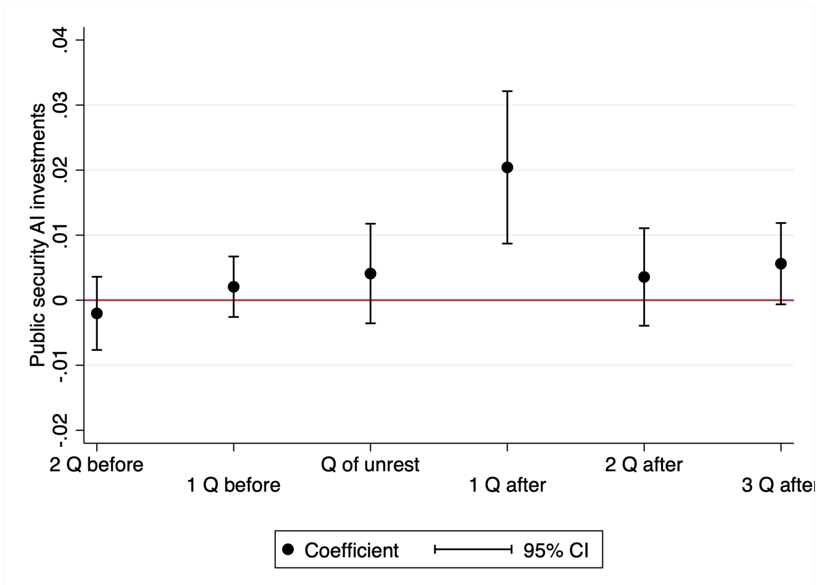


Figure A.13: Surveillance cameras per capita relative to the quarter of political unrest. Coefficients and confidence intervals displayed are from a regression following the specification in Table 2, Panel A, but additionally stacking observations containing (residualized) unrest from multiple quarters.

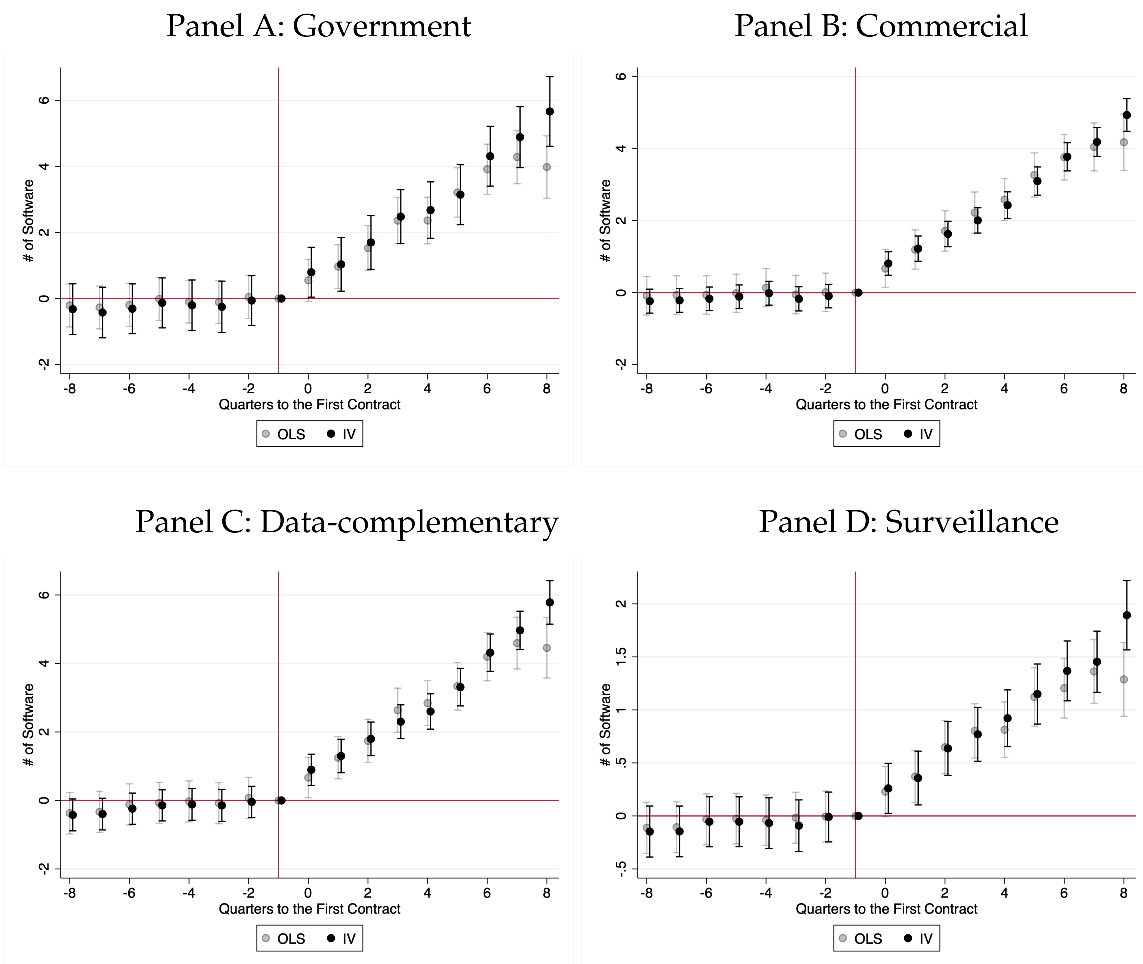


Figure A.14: Software development intended for government (Panel A), commercial (Panel B), data-complementary (Panel C), or surveillance (Panel D) uses relative to the time of receiving initial contract, controlling for firms and time period fixed effects. All subfigures display results for firms with first contracts that are politically motivated and have above median unrest in the year before the contract. Translucent lines/markers use weather to instrument for unrest.

Panel A: Data-complementary

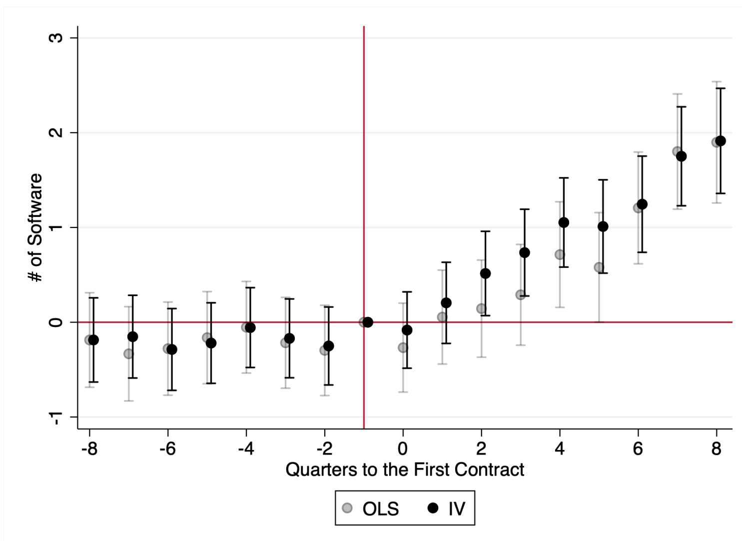


Figure A.15: Differential software development by firms that receive public security contracts versus non-public security ones, relative to the time of receiving the initial contract. All panels restrict firms to those that receive above median unrest contracts, and control for firm and time period fixed effects. Panel A shows software intended for data-complementary uses. Translucent lines/markers use weather to instrument for unrest.

#	Developer	VISA Photos FNMR@ FMR ≤ 0.000001	VISA Photos FNMR@ FMR ≤ 0.0001	MUGSHOT Photos FNMR@ FMR ≤ 0.0001	WILD Photos FNMR@ FMR ≤ 0.00001	CHILD EXP Photos FNMR@ FMR ≤ 0.01	Submission Date
1	yitu-002	0.004[1]	0.001[1]	0.013[7]	0.052[13]		2018_10_19
2	yitu-001	0.007[2]	0.003[7]	0.013[8]	0.058[26]	0.579[13]	2018_06_12
3	sensetime-001	0.009[3]	0.003[6]	0.013[11]	1.000[76]		2018_10_19
4	sensetime-002	0.010[4]	0.003[10]	0.015[29]	1.000[77]		2018_10_19
5	siat-002	0.013[5]	0.004[15]	0.014[15]	0.055[20]	0.428[3]	2018_06_13
6	ntechlab-004	0.013[6]	0.003[4]	0.013[12]	0.046[6]	0.420[2]	2018_06_14
7	ntechlab-005	0.014[7]	0.002[2]	0.013[10]	0.050[10]		2018_10_19
8	megvii-002	0.014[8]	0.004[12]	0.030[63]	0.071[35]		2018_10_19
9	vocord-005	0.016[9]	0.003[3]	0.015[32]	0.048[9]		2018_10_18
10	everai-001	0.016[10]	0.004[14]	0.013[2]	0.031[2]		2018_10_30

Figure A.16: Face Recognition Vendor Test (FRVT) ranking of top facial recogntion algorithms. Source: *National Institute of Standards and Technology (NIST)*.

Table A.1: Top predicted words from LSTM model — non-binary categorization of software

Panel A: Customer type

Government			Commercial			General		
Chinese	English	Freq. (%)	Chinese	English	Freq. (%)	Chinese	English	Freq. (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
交通	Traffic	.603	手机	Mobile Phone	.821	视觉	Vision	.474
威视	Prestige	.382	APP	App	.645	学习	Learning	.378
海康	Haikang	.369	IOS	IOS	.438	腾讯	Tencent	.340
平安	Safety	.351	iOS	iOS	.430	三维	3D	.312
海信	Hisense	.318	企业	Enterprise	.331	识别系统	Recognition System	.301
城市	City	.311	金蝶	Kingdee	.327	算法	Algorithm	.270
金融	Finance	.296	电子	Electronics	.307	计算	Computing	.252
安防	Safety	.281	健康	Health	.212	深度	Depth	.225
数字	Numbers	.272	自助	Self-Help	.209	无人机	Drone	.212
中心	Center	.269	手机游戏	Mobile Game	.201	实时	Real-time	.209
公交	Public Transport	.216	助手	Assistance	.196	认证	Certification	.207
社区	Community	.207	支付	Pay	.191	处理	Processing	.196
调度	Scheduling	.200	后台	Backstage	.189	引擎	Engine	.194
中控	Central Control	.191	门禁	Access Control	.176	技术	Technique	.187
人像	Portrait	.163	人工智能	AI	.174	分布式	Distributed	.183
指挥	Command	.161	车载	Vehicle	.174	仿真	Simulation	.179
辅助	Auxilary	.159	智能家居	Smart Appliance	.169	网易	Netease	.173
摄像机	Camera	.158	工业	Industry	.169	工具软件	Tool Software	.172
万达	Wanda	.148	DHC	DHC	.168	程序	Program	.170
高速公路	Highway	.148	营销	Marketing	.161	互动	Interactive	.166

Panel B: Function type

AI-Common			Data-Complementary			AI-Video		
Chinese	English	Freq. (%)	Chinese	English	Freq. (%)	Chinese	English	Freq. (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
指纹	Fingerprint	.342	存储	Storage	.206	人脸	Face	1.104
训练	Training	.203	可视化	Visualization	.167	深度	Depth	.321
管家	Housekeeper	.201	一体化	Integration	.164	抓拍	Snapshot	.310
文本	Text	.151	分布式	Distributed	.162	商汤	SenseTime	.287
高速公路	Highway	.150	仿真	Simulation	.157	考勤	Attendance	.258
虹膜	Iris	.147	医学影像	Medical Imaging	.148	科达	Kedacom	.258
汽车	Car	.143	通用	General	.144	跟踪	Track	.249
海尔	Haier	.137	集成	Integrated	.141	全景	Panoramic	.224
WPS	WPS	.134	数据管理	Data Management	.136	广电	Broadcastt	.209
翻译	Translate	.126	宇视	UTV	.136	目标	Target/Objective	.189
推荐	Recommend	.124	管控	Manage	.126	车牌	License Plate	.189
图片	Image	.119	高速	High Speed	.126	特征	Feature	.184
测量	Test	.116	媒体	Media/Medium	.125	铂亚	Platinum	.175
征信	Credit	.111	手机软件	Phone Software	.125	预警	Warning	.166
指纹识别	Fingerprint Recognition	.106	设计	Design	.117	运通	American Express	.163
作业	Operation	.106	接口	Interface	.117	指挥	Command	.158
微信	WeChat	.105	开发	Development	.116	统计	Statistics	.149
评估	Assessment	.105	服务器	Server	.116	安居	Safety	.146
灵云	AIcloud	.102	处理软件	Processing Software	.113	SDK	SDK	.141
活体	Living Body	.098	传输	Transmission	.111	布控	Deploymentt	.141

Table A.2: Effect of different kinds of unrest events on AI procurement

	Public security AI procurement					*Non-public security AI procurement*				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A.1: Protests — OLS										
Event	2.991***	2.996***	2.988***	2.995***	2.996***	0.002	0.002	0.002	0.002	0.002
	(0.732)	(0.733)	(0.729)	(0.733)	(0.731)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Panel A.2: Protests — IV										
Event	2.775*	3.103**	3.111**	3.086**	3.081**	0.032	0.019	0.030	0.024	0.016
	(1.497)	(1.490)	(1.475)	(1.502)	(1.502)	(0.042)	(0.035)	(0.040)	(0.039)	(0.031)
Panel B.1: Demands — OLS										
Event	2.121**	2.150**	2.118**	2.143**	2.150**	-0.002	-0.001	-0.001	-0.001	-0.001
	(0.884)	(0.875)	(0.882)	(0.877)	(0.875)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)
Panel B.2: Demands — IV										
Event	2.432	2.565	2.554*	2.541	2.493	0.007	-0.001	0.004	0.001	-0.003
	(1.511)	(1.591)	(1.547)	(1.589)	(1.569)	(0.024)	(0.021)	(0.023)	(0.022)	(0.019)
Panel C.1: Threats — OLS										
Event	1.852***	1.874***	1.853***	1.869***	1.876***	0.007	0.008	0.007	0.007	0.008
	(0.483)	(0.467)	(0.480)	(0.470)	(0.464)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Panel C.2: Threats — IV										
Event	1.400**	1.531**	1.530**	1.502**	1.520**	0.018	0.012	0.017	0.014	0.009
	(0.706)	(0.746)	(0.719)	(0.744)	(0.747)	(0.018)	(0.015)	(0.017)	(0.016)	(0.013)
GDP $\times$ time	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Population $\times$ time	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Gov. revenue $\times$ time	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Notes: This table follows Table 2 and presents regressions at the prefecture-quarter level. Panel A restricts unrest events to only protests, Panel B restricts unrest events to only demands, and Panel C restricts unrest events to only threats. The outcome is the number of facial AI contracts procured by the local government per capita, scaled up by 1,000,000. In columns 1 - 4, these are public security contracts, while in columns 5 - 8, these are non-public security contracts. There is a one quarter lag between the quarter of unrest events occurring and the number of public security AI contracts procured by the local government. Columns 2 and 7 control for prefecture GDP $\times$ quarter effects, columns 3 and 8 controls for prefecture population $\times$ quarter effects, columns 4 and 9 controls for prefectural government tax revenue $\times$ quarter effects, and columns 5 and 10 include all controls. Panels A.2, B.2, and C.2 use weather variables as selected by LASSO to instrument for unrest events. These variables are: max. temperature over 95 dummy X hail, thunder X hail, hail X max. gust speed, thunder X max. gust speed, and snow depth X precipitation, each interacted with a dummy for whether an unrest event occurred on the day. All specifications include prefecture and quarter fixed effects. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.3: Weather first stage, LASSO variables

	Number of events			
	(1)	(2)	(3)	(4)
Max temp over 95 X hail	-1.172	-0.283	0.314	-0.822
	(1.602)	(0.199)	(0.485)	(1.078)
Thunder X hail	2.614	0.602***	-0.259	2.137
	(2.198)	(0.231)	(0.647)	(1.706)
Hail X max gust speed	16.799***	2.921***	-1.635	0.807
	(6.502)	(0.803)	(4.383)	(5.996)
Thunder X max gust speed	-423.903	-136.943**	180.007	-244.705**
	(313.205)	(57.304)	(203.089)	(119.325)
Snow depth X precipitation	-122.237***	-8.508	-7.071	-18.547
	(37.538)	(7.309)	(8.121)	(17.109)
Max temp over 95 X hail X event elsewhere	1.189	0.528**	-0.020	-0.403
	(1.946)	(0.206)	(0.487)	(1.932)
Thunder X hail X event elsewhere	-7.022**	-3.464***	-1.500	-4.261*
	(2.998)	(1.189)	(1.083)	(2.192)
Hail X max gust speed X event elsewhere	-39.387***	-12.719***	-4.338	-19.028**
	(11.107)	(4.523)	(4.896)	(9.177)
Thunder X max gust speed X event elsewhere	1380.056***	585.926***	72.923	865.741**
	(522.983)	(200.472)	(239.437)	(402.239)
Snow depth X precipitation X event elsewhere	143.504**	48.557***	16.288	-15.957
	(62.979)	(14.807)	(43.907)	(46.580)
Event type	All	Protest	Demand	Threat

Notes: The table contains the first stage of the two sample two stage least squares regression specification. Regressions are at the prefecture-day level, where weather variables as selected by LASSO are interacted with a dummy for whether there was an unrest event elsewhere in China on the day to predict whether there was an unrest event in a given prefecture. Results from this first stage are aggregated to the prefecture-quarter level for Tables 2 - 4 and others using the weather IV. Coefficients are scaled up by 1000x so that the full coefficient can be seen. Prefecture and day fixed effects are used throughout. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.4: First stage - LASSO selected variables and weights

Variable	Weight
(1)	(2)
Max temperature $> 95\times$ hail	4.717
Thunder $\times$ hail	3.341
Hail $\times$ max gust speed	0.032
Thunder $\times$ max gust speed	0.001
Snow depth $\times$ precipitation	0.007

Notes: This table displays the weather variables selected by LASSO alongside the weights placed on each variable by the LASSO regression. Maximum temperature is measured in Celsius, hail and thunder are dummies for the presence of hail and thunder respectively, and snow depth and precipitation are measured in centimetres.

Table A.5: Alternate weather first stage

	Number of events			
	(1)	(2)	(3)	(4)
Fine temp dummy	-5.167***	-0.196	-0.536	-1.238*
	(1.778)	(0.263)	(0.399)	(0.741)
Fine temp dummy X event elsewhere	9.693***	2.086***	2.579***	3.549***
	(2.110)	(0.681)	(0.711)	(0.975)
Event	All	Protest	Demand	Threat

Notes: This table displays an alternative first stage at the prefecture-day level, where a fine temperature dummy (defined as the minimum temperature on the day being more than 0 degrees Celsius and maximum temperature less than 97 degrees) is interacted with a dummy for whether there was an unrest event elsewhere in China on the day to predict whether there was an unrest event in a given prefecture. Coefficients are scaled up by 1000x so that the full coefficient can be seen. Prefecture and day fixed effects are included. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.6: Effect of unrest events on surveillance camera procurement

	Surveillance cameras				
	(1)	(2)	(3)	(4)	(5)
Panel A: OLS					
Event	2.309***	2.346***	2.192***	2.335***	2.241***
	(0.536)	(0.516)	(0.548)	(0.522)	(0.517)
Panel B: IV					
Event	2.358**	2.960**	2.723**	2.793**	2.728**
	(1.096)	(1.171)	(1.118)	(1.155)	(1.160)
D.V. mean	61.453	61.497	61.688	61.497	61.688
D.V. sd	230.259	230.335	230.666	230.335	230.666
N	8424	8418	8392	8418	8392
GDP $\times$ time	No	Yes	No	No	Yes
Population $\times$ time	No	No	Yes	No	Yes
Gov. revenue $\times$ time	No	No	No	Yes	Yes

Notes: This table follows the specification in Table 2 and presents regressions at the prefecture-quarter level. There is a one quarter lag between the quarter of unrest events occurring and the number of surveillance cameras procured by the local government. Column 2 controls for local GDP by quarter fixed effects, column 3 controls for local population by quarter fixed effects, column 4 controls for local government revenue by quarter fixed effects, and column 5 adds all prior controls. Panel B uses weather variables as selected by LASSO to instrument for unrest events. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.7: Effect of public security AI on police hiring

	Police hires	
	(1)	(2)
Panel A: IHS(police hires)		
IHS(public security AI)	-0.069**	-0.069**
	(0.029)	(0.029)
Panel B: % office police		
IHS(public security AI)	0.025**	0.024**
	(0.011)	(0.011)
FE	Place, Year	Place, Year
Controls	None	Prefecture revenue

Notes: This table presents regressions at the prefecture-year level, with police hiring data one year after the unrest events. The outcome in Panel A is the inverse hyperbolic sine transformed number of new police hired, the outcome in Panel B is the share of desk jobs among new police hires. In both panels, the dependent variable of interest is the inverse hyperbolic sine transformed number of public security AI contracts. Column 2 additionally adds local prefecture government revenue as a control. All specifications include prefecture and year fixed effects. Standard errors are robust.

Table A.8: Effect of past unrest on current unrest

	Standardized number of events				
	(1)	(2)	(3)	(4)	(5)
Panel A: weather shock and local unrest at $t-1$ on local unrest					
Past unrest	-0.0051	-0.0078	-0.0081	-0.0038	-0.0062
	(0.0665)	(0.0607)	(0.0656)	(0.0624)	(0.0612)
Good weather (LASSO)	0.7972***	0.7902***	0.7977***	0.7911***	0.7947***
	(0.1683)	(0.1668)	(0.1659)	(0.1668)	(0.1676)
Good weather (LASSO) $\times$ past unrest	-0.0037	-0.0032	-0.0048	-0.0041	-0.0040
	(0.0176)	(0.0172)	(0.0179)	(0.0175)	(0.0176)
Panel B: weather shock and local unrest at $t-2$ on local unrest					
Past unrest	0.0269	0.0291	0.0242	0.0298	0.0336
	(0.0648)	(0.0605)	(0.0603)	(0.0613)	(0.0607)
Good weather (LASSO)	0.8191***	0.8114***	0.8198***	0.8126***	0.8174***
	(0.1653)	(0.1637)	(0.1634)	(0.1637)	(0.1646)
Good weather (LASSO) $\times$ past unrest	-0.0742	-0.0715	-0.0779	-0.0730	-0.0795
	(0.1119)	(0.1099)	(0.1073)	(0.1096)	(0.1084)
GDP $\times$ time	No	Yes	No	No	Yes
Population $\times$ time	No	No	Yes	No	Yes
Gov. revenue $\times$ time	No	No	No	Yes	Yes

Notes: This table follows the specification in Table 3, and presents regressions at the prefecture-quarter level. Good weather (LASSO) is the standardized number of predicted events from the good weather LASSO variables interacted with whether there was an event elsewhere with fixed effects. Local unrest in prior periods is also standardized; Panel A uses past local unrest in the quarter before and Panel B uses local unrest two quarters before. Prefecture and quarter fixed effects are included. Column 1 presents baseline results, column 2 adds controls for local GDP by quarter fixed effects, column 3 adds controls for local population by quarter fixed effects, column 4 adds controls for local government revenue by quarter fixed effects, and column 5 adds all prior controls. Standard errors are robust. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.9: Robustness and evaluating alternative hypotheses

	Government		Commercial	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Panel A.1: Baseline result				
8 quarters before contract	0.033	-0.053	0.024	0.033
	(0.115)	(0.123)	(0.214)	(0.207)
8 quarters after contract	3.249***	3.064***	3.536***	3.415***
	(0.149)	(0.159)	(0.273)	(0.263)
8 quarters before contract × public security	-0.132	-0.208	-0.142	-0.142
	(0.204)	(0.203)	(0.384)	(0.342)
8 quarters after contract × public security	1.477***	1.584***	2.774***	2.764***
	(0.261)	(0.256)	(0.490)	(0.425)
Panel A.2: Control for firm age by time fixed effects				
8 quarters before contract	0.083	-0.035	0.032	0.035
	(0.116)	(0.124)	(0.216)	(0.209)
8 quarters after contract	3.164***	2.902***	3.494***	3.421***
	(0.149)	(0.159)	(0.275)	(0.265)
8 quarters before contract × public security	-0.121	-0.202	-0.161	-0.112
	(0.205)	(0.204)	(0.387)	(0.345)
8 quarters after contract × public security	1.232***	1.612***	2.675***	2.588***
	(0.261)	(0.256)	(0.493)	(0.428)
Panel A.3: Control for pre-contract software production by time fixed effects				
8 quarters before contract	0.105	0.007	0.156	0.098
	(0.108)	(0.117)	(0.198)	(0.191)
8 quarters after contract	3.061***	2.865***	3.220***	3.237***
	(0.139)	(0.151)	(0.253)	(0.242)
8 quarters before contract × public security	-0.196	-0.228	-0.263	-0.166
	(0.192)	(0.193)	(0.355)	(0.316)
8 quarters after contract × public security	1.350***	1.684***	2.369***	2.476***
	(0.245)	(0.242)	(0.454)	(0.393)
Panel A.4: Control for pre-contract firm capitalization by time fixed effects				
8 quarters before contract	1.147**	0.622	1.376	1.369*
	(0.527)	(0.456)	(0.971)	(0.795)
8 quarters after contract	1.355**	1.567***	-1.737	0.783
	(0.631)	(0.547)	(1.160)	(0.945)
8 quarters before contract × public security	-0.144	-0.178	-0.188	-0.353
	(0.671)	(0.587)	(1.240)	(1.023)
8 quarters after contract × public security	1.464*	2.279***	4.005***	3.612***
	(0.826)	(0.714)	(1.528)	(1.230)
Panel A.5: Control for contract size by time fixed effects				
8 quarters before contract	0.564	0.354	0.612	0.632*
	(0.525)	(0.489)	(0.600)	(0.361)
8 quarters after contract	3.307***	1.613***	1.351**	1.386***
	(0.599)	(0.566)	(0.685)	(0.414)
8 quarters before contract × public security	0.114	0.084	-0.343	-0.326
	(0.655)	(0.626)	(0.748)	(0.461)

8 quarters after contract × public security	0.426	1.847**	2.063**	2.558***
	(0.784)	(0.741)	(0.903)	(0.539)
Panel B.1: Drop ambiguous public security agencies				
8 quarters before contract	0.019	-0.078	0.001	-0.002
	(0.112)	(0.124)	(0.216)	(0.210)
8 quarters after contract	3.268***	3.169***	3.661***	3.512***
	(0.145)	(0.160)	(0.275)	(0.267)
8 quarters before contract × public security	-0.130	-0.233	-0.044	-0.085
	(0.204)	(0.209)	(0.397)	(0.357)
8 quarters after contract × public security	0.742***	1.477***	2.277***	2.477***
	(0.263)	(0.265)	(0.511)	(0.447)
Panel C.1: LSTSM categorization model configuration (vary timestep = 10)				
8 quarters before contract	0.037	-0.054	0.109	0.087
	(0.118)	(0.120)	(0.292)	(0.321)
8 quarters after contract	2.638***	2.726***	3.346***	3.425***
	(0.153)	(0.155)	(0.380)	(0.420)
8 quarters before contract × public security	-0.222	-0.069	-0.335	-0.383
	(0.210)	(0.198)	(0.528)	(0.539)
8 quarters after contract × public security	0.909***	1.436***	4.425***	4.592***
	(0.268)	(0.249)	(0.681)	(0.681)
Panel C.2: LSTSM categorization model configuration (vary timestep = 30)				
8 quarters before contract	0.001	-0.035	0.095	0.062
	(0.101)	(0.106)	(0.285)	(0.307)
8 quarters after contract	2.627***	2.742***	3.723***	3.613***
	(0.130)	(0.136)	(0.367)	(0.396)
8 quarters before contract × public security	-0.283	-0.262	-0.129	-0.280
	(0.180)	(0.176)	(0.516)	(0.515)
8 quarters after contract × public security	1.264***	1.950***	3.532***	3.528***
	(0.230)	(0.221)	(0.671)	(0.648)
Panel C.3: LSTSM categorization model configuration (vary embeddings = 16)				
8 quarters before contract	0.018	0.023	0.044	-0.020
	(0.164)	(0.140)	(0.238)	(0.262)
8 quarters after contract	4.163***	3.782***	3.033***	3.171***
	(0.213)	(0.182)	(0.307)	(0.338)
8 quarters before contract × public security	-0.308	-0.268	-0.200	-0.242
	(0.293)	(0.231)	(0.427)	(0.436)
8 quarters after contract × public security	1.372***	2.183***	3.331***	3.433***
	(0.372)	(0.291)	(0.554)	(0.554)
Panel C.4: LSTSM categorization model configuration (vary nodes = 16)				
8 quarters before contract	0.059	-0.024	0.066	0.019
	(0.116)	(0.116)	(0.288)	(0.307)
8 quarters after contract	3.157***	3.329***	3.120***	2.986***
	(0.151)	(0.150)	(0.374)	(0.399)
8 quarters before contract × public security	-0.337	-0.180	-0.189	-0.301
	(0.207)	(0.193)	(0.516)	(0.511)
8 quarters after contract × public security	1.036***	2.062***	3.802***	3.638***
	(0.266)	(0.243)	(0.678)	(0.648)
Panel D.1: Time frame (full balanced panel)				

8 quarters before contract	0.034	-0.067	0.018	-0.014
	(0.125)	(0.138)	(0.235)	(0.233)
8 quarters after contract	3.208***	3.043***	3.513***	3.430***
	(0.161)	(0.177)	(0.300)	(0.296)
8 quarters before contract × public security	-0.137	-0.198	-0.138	-0.132
	(0.221)	(0.226)	(0.420)	(0.384)
8 quarters after contract × public security	1.490***	1.606***	2.835***	2.593***
	(0.282)	(0.283)	(0.535)	(0.476)
Panel D.2: Time frame (extended time frame)				
9 quarters before contract	0.005	-0.046	0.034	0.086
	(0.118)	(0.127)	(0.218)	(0.212)
18 quarters after contract	6.699***	3.905***	9.265***	6.252***
	(0.338)	(0.457)	(0.635)	(0.796)
9 quarters before contract × public security	-0.154	-0.179	-0.149	-0.187
	(0.207)	(0.206)	(0.389)	(0.346)
18 quarters after contract × public security	6.945***	14.964***	2.390**	4.813***
	(0.502)	(0.640)	(0.944)	(1.143)
Panel E.1: Access to commercial opportunities - control Beijing and Shanghai X time				
8 quarters before contract	0.174	0.042	0.220	0.293
	(0.113)	(0.121)	(0.212)	(0.199)
8 quarters after contract	3.000***	2.856***	3.215***	3.103***
	(0.146)	(0.157)	(0.271)	(0.253)
8 quarters before contract × public security	0.056	-0.142	0.110	0.052
	(0.201)	(0.200)	(0.380)	(0.330)
8 quarters after contract × public security	1.286***	1.629***	2.511***	2.546***
	(0.257)	(0.252)	(0.486)	(0.410)
Panel E.2: Access to commercial opportunities - contracts outside of Xinjiang				
8 quarters before contract	0.037	-0.051	0.026	0.033
	(0.115)	(0.123)	(0.214)	(0.207)
8 quarters after contract	3.304***	2.956***	3.520***	3.415***
	(0.149)	(0.158)	(0.273)	(0.263)
8 quarters before contract × public security	-0.138	-0.215	-0.143	-0.141
	(0.204)	(0.203)	(0.383)	(0.342)
8 quarters after contract × public security	1.348***	1.670***	2.772***	2.696***
	(0.261)	(0.255)	(0.490)	(0.425)
Panel E.3: Access to commercial opportunities - firm based outside contract prefecture				
8 quarters before contract	0.852	0.498	1.507	1.585
	(0.744)	(0.795)	(2.078)	(1.559)
8 quarters after contract	-0.194	2.030**	0.216	1.435
	(0.988)	(0.984)	(2.731)	(1.908)
8 quarters before contract × public security	-0.381	-0.322	-0.158	-0.201
	(0.956)	(1.024)	(2.674)	(2.008)
8 quarters after contract × public security	3.745***	2.403*	6.608*	5.038**
	(1.282)	(1.311)	(3.567)	(2.553)
Panel E.4: Access to commercial opportunities - firm based outside contract prefecture				
8 quarters before contract	0.023	-0.047	-0.023	-0.039
	(0.110)	(0.134)	(0.233)	(0.234)
8 quarters after contract	3.423***	3.340***	3.790***	4.038***
	(0.144)	(0.173)	(0.299)	(0.297)

8 quarters before contract × public security	-0.212	-0.400*	-0.212	-0.130
	(0.237)	(0.242)	(0.504)	(0.421)
8 quarters after contract × public security	2.220***	2.084***	5.181***	3.342***
	(0.302)	(0.301)	(0.644)	(0.516)
Panel F.1: Control for province by quarter fixed effects				
8 quarters before contract	1.017*	0.737*	1.252	1.556**
	(0.529)	(0.415)	(1.117)	(0.700)
8 quarters after contract	2.669***	1.249**	1.341	1.140
	(0.636)	(0.497)	(1.341)	(0.830)
8 quarters before contract × public security	-0.178	-0.205	-0.059	-0.484
	(0.670)	(0.531)	(1.420)	(0.896)
8 quarters after contract × public security	1.406*	2.752***	4.017**	4.277***
	(0.826)	(0.648)	(1.756)	(1.081)
Panel G.1: Non-government AI software production				
8 quarters before contract			0.034	0.009
			(0.315)	(0.322)
8 quarters after contract			6.266***	6.351***
			(0.404)	(0.412)
8 quarters before contract × public security			-0.310	-0.259
			(0.565)	(0.531)
8 quarters after contract × public security			3.869***	4.232***
			(0.723)	(0.666)

Notes: Specifications include full set of time indicators and interactions with politically motivated (public security) contracts; only selected coefficient estimates are presented. Standard errors clustered at mother firm level are reported in parentheses. Panel A.1 replicates the baseline specification in Table 4, Panel A.2 adds controls for firm age interacted by time, Panel A.3 adds controls for pre-contract firm software production interacted by time, Panel A.4 adds controls for pre-contract firm capitalization interacted by time, and Panel A.5 adds controls for contract monetary size interacted by time. Panel B.1 drops companies whose first contract is an ambiguous contract, or one that contains the keywords 'local government' ('人民政府') or 'government offices' ('政府办公室') which may be used for either public security or non-public security depending on interpretation. The baseline LSTM specification uses a timestep (phrase length) of 20, embedding size (number of dimensions in a vector to represent a phrase) of 32, and 32 nodes in the model. Panels C.1 and C.2 present results for the baseline model trained with a timestep of 10 and 30 respectively. Panel C.3 presents results for for the baseline model trained with an embedding size of 16, and Panel C.4 presents results for the baseline model trained with 16 nodes. Panel D.1 restricts the sample to firms that have non-missing observations during the entire time frame of 8 quarters before and 8 quarters after the initial contracts; Panel D.2 extends the time frame to 9 quarters before and 18 quarters after the initial contracts. Panel E.1 includes fixed effects for contracts from Beijing and Shanghai (the two highest capacity prefectures/provinces) interacted with quarter to contract, Panel E.2 omits contracts from Xinjiang, Panel E.3 restricts the analysis to firms that have their first contract outside of their home prefecture, and Panel E.4 restricts to firms with first contract outside their home prefecture. Panel F.1 adds fixed effects at the province by quarter level. Panel G.1 uses total non-government AI software production as the outcome with columns 3 and 4 continuing to show OLS and IV. * significant at 10% ** significant at 5% *** significant at 1%.