Pollution Monitoring, Strategic Behavior, and Dynamic Representativeness

Lin Yang, Yatang Lin, Jin Wang, and Fangyuan Peng†

PRELIMINARY DRAFT. PLEASE DO NOT CIRCULATE

In major countries around the world, air quality evaluation is mainly based on stationary in-situ monitors that aim to provide representative measurements of the ambient air pollutant concentrations. Based on the staggered roll-out of an automated monitoring system in China, we use high-resolution remotely sensed data to examine local governments' strategic behavior in pollution reduction and its implications on dynamic spatial representativeness of ground monitors. We find strong evidence that despite the use of automation that pre-cludes the falsifying of data, local governments take targeted pollution reduction actions near the monitors, leaving pollution elsewhere unchanged or even increased. We also detect heterogeneity in the strategic measures taken by local officials facing different career concerns and in cities with varying degree of pre-automation data tampering conduct. Our findings underscore a new agency problem emerged: as the central government adopts more advanced monitoring technology to reduce local officials' capacity for misreporting, the latter might employ more concealed but costly means of obfuscation. One immediate and significant policy implication is that air quality evaluations should closely integrate ground monitoring data with remote sensing information to account for local agents' strategic behavior within decentralized environmental governance.

Keywords: Air Pollution Monitoring; Environmental Regulations; Strategic Response; China

---

†Yang: Dyson School of Applied Economics and Management, Cornell University (email: ly285@cornell.edu); Lin: Department of Economics, Division of Environment and Division of Social Science, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (email: linyt@ust.hk); Wang: Division of Social Science, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (email: sojinwang@ust.hk); Peng: Society Hub, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (email: fpengac@ust.hk). We are very grateful to the helpful comments and suggestions from Shanjun Li, Cathy Kling, Ariel Ortiz-Bobea, Ivan Rudik, Guojun He, Jimmy Fung, Todd Gerarden, Binglin Wang, Deyu Rao, Avralt-Od Purevjav, and seminar participants at OSWEET, the Applied Economics & Policy Seminar and SEERE at Cornell Dyson, and the workshops at HKUST for helpful comments. Yang acknowledges the financial support from the C.V. Starr Fellowship at Cornell University. Lin acknowledges the financial support of research grant ECS26504419 from the University Grants Committee of Hong Kong. This article subsumes and replaces our previous work, Yang (2020) and Lin et al. (2020). All errors remain our own.

1 Introduction

Enforcement of and compliance with regulations critically hinge on accurate measurements of implementation and outcomes.[1] Given the ubiquitous information asymmetry between central and local governments, imperfect monitoring of national regulations can lead to strategic compliance at the local level, which will further bias performance measurements and cause policy failures. Implementation of national policies at local levels under fiscal and political incentives is a principal-agent problem inherent in the delegation of authority by governments to bureaucratic officials (Aghion and Tirole 1997). Recent advancements in technology have enabled new modes of monitoring, substantially lowering the principal's costs of verifying and observing information, which are so deeply embedded in any agency theory (Hölmstrom 1979). These applications ultimately aim to mitigate the moral hazard problem (hidden actions) associated therewith (Greenstone et al. forthcoming; Hubbard 2000). Through the lens of a technology-aided environmental regulatory program in China, the goal of this paper is to analyze how local governments employ more concealed strategic responses to influence environmental quality measure, rendering the existing outcome-based incentive scheme inefficient.In a context of decentralized governance, it begs the question as to how political economy considerations of local officials can break air pollution regulations despite the use of improved monitoring technology.

In most countries, air quality standards are enforced according to ambient air pollution measurements, which are obtained from a ground-based air-quality monitoring network. Monitoring stations are located in limited geographic areas because of budget constraints. The degree to which the monitoring data are reliable and spatially representative of the ambient air pollutant concentrations, therefore, is of paramount importance within the context of enforcing environmental regulations. In 2012, amidst the public outcry on the lack of transparency in pollution data, the Ministry of Environment and Ecology of China (MEE)[2] launched a nation-wide, real-time air quality monitoring and disclosure program in a staggered fashion across various cities. The MEE centralized the planning, establishment, construction and maintenance of all state-level air quality monitoring stations, essentially

[1] For instance, crime reduction relies on correct detections of crime activities; tax reform requires precise estimation of population income distribution; transportation and environmental regulations need accurate monitoring of traffic and pollutants.

[2] In 2018, the Ministry of Environmental Protection (MEP) was renamed as the Ministry of Environment and Ecology of China (MEE), the name we use to refer to the ministry throughout this paper.

minimizing the room for data manipulation at the local level.(Greenstone et al. forthcoming).[3] It also put new pressures on local officials, who face a delicate trade-off between economic growth and environmental quality, and find that their careers may be affected by the ability to demonstrate progress in reducing air pollution levels. As a result, local governments might seek alternative ways to improve their air pollution readings. A handful of studies for the United States have shown that local regulators could strategically avoid local pollution hotspots when deciding on the location of monitoring stations (Grainger et al. (2019)); In response to intermittent monitoring, local governments issued more public air quality advisories for short-term suppression of polluting activities on monitored days (Zou (2021)). Evidence assessing dynamic strategic responses of local regulators to spatial gaps in monitored areas is, however, thin. The present paper, which exploits the quasi-natural experiment of China, seeks to address this research gap.

Leveraging high-resolution satellite-based air pollution measures, this article examines local officials' strategic behaviors in pollution reduction and the implications thereof on dynamic spatial representativeness of ground monitors. To do so, we employ a distance-based difference-in-differences approach with variable treatment intensity. The staggered roll-out of the new monitoring system allows cities that joined in different waves to serve as treated and control groups for each other. The pollution changes for areas near the monitors (classified as "monitored") are then compared to those "unmonitored" areas far away.

To shed light on whether strategic cleaning efforts of local governments would undermine regulatory effectiveness, we compare population-weighted average pollution levels of an entire city to the count based on the monitored locations. While most of the monitors represent the city's average air quality well in the initial stage of automation, the spatial representativeness has been changing over years, reflecting varying degree of pollution changes within the city. The environmental performance measure is a critical pillar of "merit pay" systems the central government, as principal, used to determine political reward for the local governments, the agents. For an increasing number of cities, we detect decreasing spatial representativeness of the monitoring data, implying a comprised performance evaluation system.

For the analyses, we draw primarily upon remote sensing data to detect pollution changes from space. We obtain a nationwide sample of annual $PM_{2.5}$ (fine inhalable particles, with

[3]Serious concerns about the data manipulation issues in China's air quality data have been well documented by the literature (Andrews 2008; Chen et al. 2012; Ghanem and Zhang 2014).

diameters that are generally 2.5 micrometers and smaller) grids at the 1km by 1km resolution for over nine million grids from 2000 to 2017.[4] In remarkable detail and accuracy, fine-scale satellite-derived $PM_{2.5}$ data fills the spatial gaps in ground monitoring networks and validates the data quality at the ground level (Fowlie et al. 2019; Sullivan and Krupnick 2018). With the use of annual data, we minimize measurement error caused by missing values, which are prevalent in the monthly and daily series. To gain insights into the mechanisms at work and the underlying political incentives, we also assemble data on thermal anomalies, a novel satellite-based measure proxying for local economic activities, along with city socio-economic characteristics and personnel information on local officials from the China Political Elite data. These data offer a unique opportunity to identify the causal impact of the automatic monitoring program on local governments' strategic responses.

Our first main finding is that areas adjacent to monitors experience a 6.1% reduction in $PM_{2.5}$ concentrations than those farther away. Prior to the automation, monitored and unmonitored areas have very similar trends in pollution. The gap between the two groups emerged since the monitor began operation, and grew even larger as the final assessment deadline set by the central government approaches.[5] These results are robust to a number of sensitivity checks and tests of omitted variable bias, including addressing mean-reversion of pollution, the use of raw satellite Aerosol Optical Depth (AOD) readings as an alternative outcome measure, along with others. Overall, we offer compelling evidence that targeted measures had been taken by local regulators to reduce air pollution close to the monitors.

We then proceed to explore the underlying mechanisms that give rise to the varying changes of pollution between monitored and unmonitored areas: strategic cleaning. Local governments may have ordered polluters near the monitors to decrease or even shut down their production to lower pollution readings. To test this important channel, we start with the premise that targeted measures should decay with the geographic distance from the monitor. An examination of pollution changes in inner-ring areas relative to outer-ring areas confirms that significant pollution reductions were only detected in areas within a 30 kilometer radius of the monitor, while the pollution level in those far away had not changed or even increased.

[4]By combining satellite-based measures of AOD with chemical-transport modeling and land characteristics, van Donkelaar et al. (2019) derive ground-level concentrations of $PM_{2.5}$ at high levels of spatial disaggregation.

[5]According to the Air Pollution Prevention and Control Action Plan announced in 2013, the central government conducted a final assessment of overall pollution reduction at the end of this action plan in 2017.

More importantly, by utilizing the novel thermal anomalies measure, both the extensive and intensive margin of local industrial production activities within the monitored areas were found to have significantly reduced, a most likely response to automation. These two sets of results provide corroborating evidence of our proposed mechanism.

We further examine the role of career concern incentives faced by local officials in the conducting of more or less strategic cleaning. To do so, both temporal and spatial heterogeneity are explored. We show that regional governments with newly automated monitors are more likely to prioritize improving pollution control, possibly due to the imminent pressure of central assessment. Cities which more aggressively engaged in data manipulation prior to the automation are also more inclined to take targeted pollution reduction actions near the monitor area in the post-automation era. Overall, the analysis underscores the misaligned incentives and interests of the central regulator and local bureaucrats within China's target-based performance evaluation system.

In light of more concealed strategic responses to the tightening monitoring system, our analyses offer important implications for how to devise effective regulatory methods that incentivize meaningful changes to limit pollution. One should be modest and cautious with the application of technology aided program. There is no easy solution to the salient moral hazard problem due to the existence of multiple dimensions of private actions. In designing monitoring policies, the central regulator should rigorously account for local agents' possible responses. While conducting performance evaluations for local officials, the central government should closely integrate ground monitoring data with other measures such as remote sensing information, the use of mobile monitors and increasing citizen participation in supervision of environmental quality. This would dampen at least some undesirable efforts of local agents to strategically clean merely the immediate areas surrounding the monitors rather than engaging in systematic reductions that benefit wider regions.

This paper makes the following contributions. First, we build on and contribute to a nascent literature on the environmental monitoring regulation and enforcement (Duflo et al. 2013; Gray and Shimshack 2011; Shimshack 2014). Our paper extends two predecessors (Grainger et al. 2019; Zou 2021) by providing among the first analysis that links strategic cleaning actions of local officials with the dynamic change in monitor representativeness in China. We provide new evidence that monitoring network and design that are ex-ante efficient will not necessarily remain efficient with the evolving responses of local agents. Our findings are potentially relevant for other developed and developing countries which built monitoring networks decades ago and thus faced increasing need for a new or upgraded system. Our paper also relates closely to two of the concurrent studies that share a similar regulatory context and highlight the immediate improvement of data quality associated with the automated monitoring system ((Greenstone et al. forthcoming)); and analyze the effect of monitor-based information disclosure on individuals' avoidance behavior (Barwick et al. (2020)). Our results complement these studies by uncovering more concealed actions that surface, when fabricating data is no longer a viable option. With the strategic responses emerged, the accuracy of information disclosed to the public will be undermined. Consequently, the avoidance behavior taken may be sub-optimal especially for the rural households.

Second, our paper adds to the growing literature on the political economy of environmental regulation in the framework of the principal-agent relationship with moral hazard. Studies have found that firms and local governments respond to stringent regulations in ways that result in unintended consequences such as pollution spillover (Chen et al. 2018b; Kahn 2004; Kahn and Mansur 2013; Kahn et al. 2015; Karplus et al. 2018). In particular, a recent study by He et al. (2020) examines how imperfect performance monitoring of water pollution led Chinese local officials to enforce regulations on polluters immediately upstream of monitoring station. In our paper, we document another unintended deviation of national regulations at the local level and offer insights into the underlying political incentives. We show that the local agents respond to the spatial gaps of the ground monitoring network by conducting strategic cleaning efforts near the automated monitor.

Third, this paper relates to a recent literature with applications of remote sensing data in studying environmental regulations, climate change, wildfire surveillance and biodiversity (Hansen et al. 2013; Ruminski et al. 2007; Turner et al. 2015). In particular, we unearth novel satellite measures at fine spatial units to evaluate the representativeness of the ground monitoring system, and to proxy for economic activities in investigating local strategic responses to the automation program.

Lastly, this paper provides evidence about uneven pollution control, adding to the economic literature on environmental justice (Banzhaf et al. 2019; Bento et al. 2015; Currie et al. 2020; Grainger and Schreiber 2019). Specifically, strategic cleaning in the monitored regions and undertaking air pollution mitigation in ways that only target these specific areas (urban areas) raise the prospect that residents living far away from the monitoring stations (for instance, rural households) may not benefit after all. They could even be harmed if pollution grows elsewhere as a result Such localized improvement can exacerbate geographic environmental inequality with profound distributional consequences.

The remainder of the paper is organized as follows. Section 2 provides a brief background on environmental regulations and the monitoring system in China. Section 3 describes the main data sources. Section 4 presents the empirical strategy. Section 5 elucidates the mechanisms underlying the strategic behavior to the automation program. Section 6 discusses policy implications for improving the air pollution monitoring system. Section 7 concludes.

2 Institutional Background

The benefits of China's unprecedented economic growth in the past decades are built upon the huge cost of a stained environment. China's unprecedented economic growth relies heavily on industrialization and fossil fuels, and lax environmental regulations. Over the last 40 years, China has experienced the fastest economic growth and became the largest consumer of energy and coal while also having many of the most polluted cities in the world.[6] Severe air pollution (known as "smog") in major cities attracted the attention of the international community, putting pressure on the central government of China. In the past decades, public awareness of air pollution rises, and more research has revealed the negative impact of air pollution on human health, both physical and mental. The Chinese government began to shift its policy priority from the long-lasting economic growth to environmental concerns and introduced stringent regulations on air pollution. This section introduces the political system and environmental regulations in China and discusses the underlying nature of local officials' strategic behaviors.

[6]"Helping China Fight Air Pollution", The World Bank. https://www.worldbank.org/en/news/feature/2018/06/11/helping-china-fight-air-pollution

2.1 Political System in China

Political incentives are one of the internal mechanisms of both economic development and environmental protection, especially in China. A salient feature in China's political system is that the central government sets targets and links the local officials' promotion to their performance in these targets. Local officials, in turn, are highly incentivized and are given great flexibility in local regulatory plans to meet the national targets. Studies in political economics have examined the principal-agent problem lies in China's economic development. The incentive-based strategic responses by local governments have led to many unintended consequences such as inequality, collusion, corruption, and cheating, which may undermine the policy goals. (Fisman and Wang 2015; Jia 2017; Jia and Nie 2017; Li and Zhou 2005; Oliva 2015)

The Target Responsibility System launched in the 11$^{\text{th}}$ Five-Year-Plans (FYPs) in 2005 marked an important transformation in China's national policy, where environmental targets were incorporated into the evaluation criteria of local officials.[7] In this system, local leaders who fail to attain environmental performance targets, no matter how successfully they accomplished all other tasks, would receive an unqualified evaluation in their year-end comprehensive assessment, and would not be eligible for any annual bonuses or career advancement. However, such a motivation system has also motivated strategic responses. More recent literature has placed the spotlight on the firms and local governments' behaviors under various water and air pollution regulations. The strategic responses to environmental regulations have led to issues like data manipulations (Chen et al. 2012; Ghanem and Zhang 2014; Karplus et al. 2018) and pollution spillovers (Chen et al. 2018b; Kahn 2004; Kahn et al. 2015).

2.2 Environmental Regulations in China

Air pollution regulation has been a top priority of the central government of China in the past decade. It declared "war on air pollution," implementing a series of mitigation actions, such as the "Air Ten" action plan that was announced in 2013, (the Air Pollution Prevention and Control Action Plan). The action plans add detailed pollution control requirements to the 12$^{\text{th}}$ FYPs in terms of targets, standards, measures, and technologies. In addition

[7] China's five-year planning process defines overarching principles to guide national policy and broadly sets forth regulatory objectives for both economic growth and environmental protection.

to the plans, a raft of new environmental protection laws and guidance are enacted, which are claimed to be the "strictest ever" environmental policies regulations to show the central government's determination to win this "war".

Under the set of stringent regulations that closely correlate with local official's own incentives, it is not a surprise to see that China has made significant progress in pollution reduction and prevention over the past decade. For example, the "Air Ten" evaluates local officials' performance in pollution reductions on an annual basis. In addition, the central government conducted a final assessment of overall performances at the end of this action plan in 2017. The promotion of local officials is not the only aspect linked with their performance in pollution control. The government budgets and new projects related to air pollution are linked to the local officials' performance as well.

Stringent central regulations have helped improving air quality in China, according to the ground monitor readings. For example, Greenstone et al. (2020) estimate the air pollution trend since 2013 ("Air Ten") and show that all of the air pollutant concentrations dropped sharply, except for O_3, which saw a modest increase. $PM_{2.5}$ levels dropped by 27.7 $\mu g/m^3$, or about 41 percent from the 2013 level. However, the sharp reduction in air pollution is based on the ground monitor readings, which may be subject to bias due to gaps in spatial coverages. Our paper aims to dig deeper into this pollution reduction trend and study the local governments' strategic pollution reduction behaviors using the newly disclosed monitoring system.

2.3 Monitoring Systems for Ambient Pollutants

In recent years, large cities and more developed areas have successfully mitigated coal-smoke pollution. In the wake of such mitigation efforts, haze, photochemical smog and acid rain have emerged as more severe pollution issues in China. $PM_{2.5}$ concentrations are increasing in key regions including the Pearl River Delta, the Yangtze River Delta and the Beijing-Tianjin-Hebei area. The surge has caused a reduction in visibility in urban areas, and has adversely affected human health and safety. Yet, China's earlier version of air quality standard did not include $PM_{2.5}$. Against this background, the State Council of China introduced Ambient Air Quality Standards (AAQS) in 2012, which included six principal pollutants: PM_{10}, NO_2, SO_2, O_3, $PM_{2.5}$, and CO. Three of these (O_3, $PM_{2.5}$, and CO) were added for the first time.

The enforcement of environmental regulations relies on accurate and comprehensive monitoring of compliance. As a critical measure to achieve this, the new standards stipulated measures required to further expand the national air quality monitoring network, and, more importantly, to establish a real-time reporting system. The data quality in China has been criticized a lot, especially for air pollution data before 2013: only 74 major cities had monitors, the data was reported by local governments as a daily air pollution index, and not available to the public. Obviously, local governments have great power to manipulate the reported air pollution data. As shown in Ghanem and Zhang (2014), when the policy goal is the number of "blue sky days" in a year, that is when the air pollution index is less than 100, the air pollution data reported by local governments is bunching at the cut-off. [8]

To win the "war against pollution" after 2013, China launched a nation-wide, real-time air quality monitoring and disclosure program, which quickly built-out over 1500 monitors. Several major improvements have been made in this new monitoring program. Firstly, $PM_{2.5}$ is listed as a major pollutant. Secondly, after the monitoring stations are built or upgraded, raw monitoring data from the monitors is directly transmitted to the central system, which significantly eliminates the data manipulation issue in the pre-automation self-reported pollution data. A recent study by Greenstone et al. (2020) shows the improvement in data quality with the new monitoring system, and the increased public awareness of pollution prevention. Lastly, the introduction of new air quality system has been accompanied by corresponding changes in the arrangement of responsibilities for various environmental authorities.

As illustrated in Figure 1, prior to 2012, the central-level MEE was responsible for setting out technical specifications and for publishing the monitoring data; local environmental authorities were responsible for managing and operating monitors, and for collecting and managing the environmental surveillance data submitted to the MEE. This procedure created non-trivial room for local data manipulation. Under the reformed system, the MEE is in charge of the design of national environmental monitoring stations. China National Environmental Monitoring Centre (CNEMC) is responsible for assessing the performance of instruments and equipment, such as surveillance cameras and the real-time reporting system within the stations, as well as for checking the accuracy of the data collected. The

[8] "Blue sky day" is a term introduced by the central government in 1998 when Beijing was bidding to host the Olympics, at which the city's Air Pollution Index is less than 100. The number of "Blue sky days" is a critical basis to evaluate a city's air quality condition.

local governments may purchase and install equipment only upon inspection and approval of CNEMC. Surveillance cameras are set up around the monitoring stations to prevent data tampering. Furthermore, CNEMC entrusted a third party to operate and maintain these monitors. This third party provides technical support, and conducts the quality assessment for the overall monitoring process of the new standards.

To address the possibility of selective siting, the placement of these new stations was decided centrally by the Ministry of Ecology and Environment. Locations for all the stations rolled out from 2012 to 2014 were determined in 2012. [9] The central government states that only central monitors will be counted into the evaluation of cities' average air quality conditions and local official's performance in pollution reduction. Local monitors, although built under the same guidelines, will only be helping local officials in detecting polluting sources and designing for local regulatory plans.

Three waves of prefectural cities entered the monitoring system successively in each year between 2012 and 2014. Major development regions such as the Jing-Jin-Ji region, the Yangtze River Delta region, and the Pearl River Delta region, as well as a few large cities such as provincial capitals, are the first wave to enter the new monitoring network. In these cities, many of the monitors were built and operated long before the new monitoring system was introduced. Entering the program means upgrading the existing monitors to automation, as well as adding new monitors. In the first wave, 74 cities upgraded 496 stations before January 1st, 2013. In the second wave, 116 cities installed the new monitoring system with 449 stations located within their jurisdictions by January 1st, 2014. The remaining 177 cities built 552 stations before January 1st, 2015.[10] Figure 2 shows the three roll-out waves of monitors in China. The national monitoring network with 1606 central monitors is designed to serve for urban areas of 336 cities. The number of monitors in each city is based on the population density and a city's pollution level in the past three years.

[9] There are three types of monitors in China: 1. Monitors controlled by the central government; 2. Monitors controlled by local government; 3. Micro Monitors for specific polluting sources. The central government control monitors are the first group of monitors set up before air pollution becomes a society-wide concern. As of 2016, there were more than 2000 monitors in China, including both central and local monitors.

[10] Satellite data for some stations were not available. Data for 542 stations were recorded for the third wave. The 10 stations for which data were not unavailable are located in Kizilsu Kirghiz (one station), Bortala (two stations), Kashgar (three stations), Wenshan (two stations) and Ngari (two stations).For our empirical analysis, we also include an additional 119 stations built after the automation deadline of each city; we label these as the same wave as the city they belong.

Since local officials do not have much control over the location choices of central monitors, ideally, as long as the central monitors well-represent local air quality conditions, the monitoring network should be efficient. Moreover, the central government encourages third party companies to gradually take over the operation and maintenance of these monitors, which greatly eliminates the possibility of direct data falsification, shutting down or destroying the monitoring devices. Data accuracy has been significantly improved after the involvement of third-party organizations (Niu et al. 2020). However, manipulations and strategic responses by local officials never ended. Medias covered several stories of constantly watering the monitored areas with fog cannon trucks, shutting down small-scale workshops, and food trucks near monitors, which burnt coal. Using the remote-sensing data to fill the spatial gaps in ground air pollution monitoring system, we find evidence for local officials spatially differentiated pollution control strategies.

3 Data

We draw upon several primary data sources including: 1) satellite remote sensing data which report air quality and industrial activities at a fine-granularity level; and 2) comprehensive monitoring station data which offer key information on a station's automation status.

3.1 Remote Sensing Data

Air Quality In order to examine the spatial difference in air pollution regulations, this paper fills the gap in the ground monitoring system using high-resolution images of the major air pollutant, $PM_{2.5}$, which are derived from the original satellite measures of Aerosol Optical Depth (AOD). The satellite AOD data comes from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) algorithm. AOD measures the total vertical distribution of particles and gases within a grid according to the light extinction coefficient. It indicates how much direct sunlight is prevented from reaching the ground by aerosol particles and can be used to infer ground-level pollution, particularly for fine particles such as $PM_{2.5}$ and PM_{10}. Atmospheric science literature has shown a strong correlation between satellite measure and ground-level pollution data.[11] Since the satellite measures are largely affected by cloud coverages over an area, missing data is a big issue when using remote sensing data

[11] Lee et al. (2012); Liu et al. (2007); Zhang and Li (2015). Previous economic research using the satellite measure as the proxy for ground-level pollution includes Bombardini and Li (2016); Chen et al. (2013); Foster et al. (2009); Fowlie et al. (2019); Sullivan and Krupnick (2018).

with fine spatial and temporal resolutions. Studies of the remote sensing techniques find better correlations between AOD and ground-level PM with coarser spatial and temporal resolutions by month or year (Hoff and Christopher 2009).

The satellite images this paper uses include annual $PM_{2.5}$ grids (1km by 1km resolution, nine million grids for whole China) from 2000 to 2017. By combining satellite-based measures of AOD with chemical-transport modeling and land characteristics, van Donkelaar et al. (2019) derive ground-level concentrations of $PM_{2.5}$ at high levels of spatial disaggregation. One concern with the satellite-derived ground-level pollution measure is the measurement errors caused by the calibration of the satellite data using ground monitoring data. Even though that van Donkelaar et al. (2019) use geographical weighting method to give smaller weights to cells further away from ground monitors, and larger weights to cells closed to ground monitors, one may be worried about different measurement errors may occur at cells with different distances to monitors. To address this concern, the authors conducted cross-validation tests, where they remove part to all of the ground monitors from the calibration. The derived $PM_{2.5}$ data still performs well. [12]

Thermal Anomalies To develop a measure of industrial activities at a fine spatial scale, we leverage a novel dataset on satellite-based thermal anomalies. A variety of industrial activities, such as power generation and cement production, is associated with the local release of an enormous amount of heat. This motivates the use of thermal anomalies tracked by remote sensing as a real-time and high-resolution measure of local industrial activities.[13]

We draw on the MODIS Version 6 Global Monthly Fire Location Product, MCD14ML, which traces active fires and other thermal anomalies such as industrial plants and volcanoes.[14] On a monthly basis, this product documents primarily the geographic location and date for each fire pixel detected by the Terra and Aqua MODIS sensors. It also provides information on the type of heat spot, including whether the source is presumed to be vegetation fire, active volcano, static land source, and offshore source. Figure 3 presents the high correlations between detected fire spots and the geo-locations of major air polluting firms in

[12]We have also used raw daily AOD data downloaded from NASA's MODIS system to check the robustness of our analysis to potential measurement errors that correlate with locations of ground monitors.

[13]Some previous studies that use thermal anomalies to identify industrial activities include (Huang et al. (2018)), (Wei et al. (2019)) and (Xia et al. (2018)).

[14]Data source: https://firms.modaps.eosdis.nasa.gov/.

2016.

For our analysis, we identify fire spots under the category of "static land source" as active industrial activities. We use the Band 21 brightness temperature to measure their intensity. Moreover, we exploit the richness of the data and conduct placebo tests on the agricultural fires (fire spots under "presumed vegetation fire").

3.2 Ground Monitoring Network

Monitoring Station Data The MEE has been publishing $PM_{2.5}$ data since 2012. Our access to this particular data source has been granted through the Hong Kong University of Science and Technology Atmospheric & Environmental Database, which covered 1,606 stations from 2012 to 2017. The data recorded each station's name, geographical coordinates and hourly pollution readings.

Spatial Representativeness of Ground Monitors With the fine-scale pollution data and spatial information of the new ground monitoring network, we examine the spatial representativeness of these monitors.

One thing to notice is that almost all monitors are located in urban areas, and the sparse central monitors are the only base in evaluating the air pollution condition of a city. The gaps in the ground monitoring network might cause the regulation focus to bias toward urban citizens. Instead, the less-monitored places, i.e., the rural areas' pollution, will not be considered in evaluating the local officials' environmental performance. Contrarily, the satellite-based measurements give a highly spatial resolved coverage of the air pollution in the entire city area. To examine the difference between monitor-based and satellite-based city average $PM_{2.5}$, we use the 1km by 1km gridded population count from 2015 Census to weigh each cell and calculate the weighted average $PM_{2.5}$ for each city. Taking this as the "true" city-level $PM_{2.5}$, we then compare it with the monitor-based population-weighted average $PM_{2.5}$. The map in Figure 4 shows the monitors representation errors in the years that cities joined the system. We regard the cities with errors within $\pm 15\%$ as having well-representative monitors. The warm colors are cities where monitors over-represent the "true" city-level $PM_{2.5}$, and the cool colors are cities with under-representative monitors. The representation errors exhibit large spatial variations, where two-thirds of cities have over-representative monitors.

The representation errors in Figure 4 are static at the moment of their openings. If the pollution reduction patterns are even across space, then the representativeness of monitors would not change as long as the monitors' locations do not change. However, though monitors are unlikely to move for a long period, local regulators' strategic responses to the static monitor locations would change the monitor's spatial representativeness overtime. By plotting the representation error maps in each year (Appendix A1), this is indeed the case. Monitors' spatial representativeness exhibits dynamic changes in years after cities joining the program, which greatly motivates our study of local regulators' strategic pollution reduction behaviors.

3.3 Other Data and Summary Statistics

To check if other factors would affect the spatial representativeness of ground monitors and the strategic environmental regulating behaviors, we collect data on city characteristics such as population, GDP, etc., and weather variables, such as temperature, humidity, wind directions, wind speed, etc. We have also collected information about local officials from the China Political Elite data, which includes local officials' career path, age, and education. The descriptive statistics of the variables are summarized in Table 1.

4 Strategic Pollution Reduction After Monitoring

4.1 Empirical Framework

We examine the strategic pollution reductions in monitored areas after monitoring using a Difference in Differences method with a staggered roll-out schedule. Joining the new monitoring program by either having new monitors or automation of existing monitoring data could change local officials' incentives and strategies to meet environmental targets. Thus, once a city joins the program, it will be considered as in the treated group. Within each treated city, there will be different treatment effects by distances away from monitors. We use the following empirical framework to examine the impact of monitoring on overall air quality and the heterogeneous treatment effects by treatment intensity:

$$\ln(PM2.5_{iwt}) = \alpha Open_{wt} + \beta Near_i \times Open_{wt} + Cell_i + Year_t + Trend_{wt} + \varepsilon_{it} \quad (1)$$

The outcome variable, $\ln(PM2.5_{iwt})$, is the logarithm of annual $PM_{2.5}$ concentration. To reduce computational burden, we aggregate the 1km×1km cells into 3km×3km grid cells.

i is the index for grid cells within cities opened in wave w at year t. The 3km-resolution satellite data we use has over a million cells' annual $PM_{2.5}$ from 2000 to 2017. $Open_{wt}$ is the treatment indicator that takes the value of 1 if cell i is in a wave w city after joining the new monitoring program. The treatment intensity is defined by $Near_i$, which equals 1 if the grid cell i is in an area adjacent to a ground monitor (monitored area), and 0 if the cell i is in areas far away from monitors (unmonitored area). In most cases, we are less interested in the causal effect of the monitoring program per se (α), but rather more in the difference in the causal effect in monitored vs. unmonitored areas (β) after monitoring. Due to the large spatial and temporal variations in air pollution, there may be confounders that would bias β from identifying the difference in pollution reductions across space. Especially, cells in monitored and unmonitored areas could have different location attributes that affect air quality. To address these concerns, we report results of estimations with a rich array of controls, including cell fixed effect and year fixed effect. We also include a wave-specific time trend to allow the unobserved time trend in pollution to vary across waves. The identification variation is then from comparing cells in monitored vs. unmonitored areas before vs. after new waves of monitor roll-out. Since pollution observed at a cell is likely driven by emissions elsewhere that also affect nearby cells, we cluster standard errors at the city level.

Cities selected into the program in different waves may be due to wave-specific unobservables that are time-variant. Cities in earlier waves tend to be larger cities with more population, higher GDP per capita, higher levels of air pollution and industrial emissions, etc. We include wave by year fixed effects in Equation (2). Although the fixed effects absorb the baseline impact of monitoring on overall pollution (α from Equation (1)), Equation (2) provides a clearer identification of changes in treatment effect by treatment intensity (β). It also has more flexible controls than the wave-specific time trend. The identification variation now is from comparing monitored vs. unmonitored cells in the same wave cities, before and after monitoring. The key explanatory variable is $Near_{it}$ which is an interaction of $Near_i$ and $Open_{wt}$.

$$\ln(PM2.5_{it}) = \beta Near_{it} + Cell_i + Wave_w \times Year_t + \varepsilon_{it} \tag{2}$$

4.2 Baseline Results

In the baseline results, we estimate equation (1) and (2) using our preferred sample from 2008 to 2017. The monitored area is defined as grid cells within 3km of a monitor. The results are robust to alternative definitions of unmonitored areas such as outside 15km, 30km, 60km and 90km where we drop the cells in between to address the concern of misclassifying monitoring status. The DID with treatment intensity provides estimates of local effects within the choice of the treatment intensity groups, where results using different monitored areas could represent different pollution control strategies that local officials adopt. We will discuss more in the next section.

Table 2 presents the baseline DID result by estimating Equation (1) and adding controls sequentially. In the first two columns, *Open* captures the baseline impact of joining the monitoring program on air pollution, comparing to control cities. [15] We then include the treatment intensity indicator $\mathbf{1}$(0-3km) in column (3) to (5) to capture the heterogeneous treatment effects of the monitoring program on pollution in monitored (cells within 3km) vs. unmonitored (cells outside 3km) areas.[16] In column (5), we use wave specific year fixed effects to absorb the baseline causal effect of monitoring and show the relative changes between treatment intensity groups (Equation (2)), which is the main finding of our paper. The results show that pollution in monitored areas is 6.1% less than that in unmonitored areas after monitors roll-out. Table 3 shows that our main DID results are robust to alternative definitions of unmonitored (control) groups.

4.3 Identification

The key assumption is that in the absence of a monitor opening or switching to automation, air quality in the monitored and unmonitored areas follow parallel trends. In other words, we assume that the only reason that ambient air quality might show a significant difference between areas nearby monitors and areas far away from monitors is because that local officials strategically put more efforts into reducing "local" air pollution. As directed by the central government, most monitoring stations are placed in urban centers to cover populated areas.

[15] Goodman-Bacon (2018) points out the concern of DID with heterogeneity in treatment timing, which could be a valid concern for our baseline DID estimation of the causal effect of monitoring (α). Thus, an event study is preferred than an average treatment effect. In our paper, the three waves of cities entered the program consecutively within three years. The potential impact of wave-specific factors affecting the pollution in different years has been controlled by the Wave by Year FE. The estimated key parameter of interest (β) is the different pollution changes among treatment intensity groups within a wave of cities after monitoring.

[16] Without controlling for cell fixed effect, the raw difference between two treatment intensity groups is positive. This result is likely driven by the fact that the urban centers, where most monitors are placed, tend to have higher pollution levels than other areas of a city. Once cell fixed effect is included, the results show that areas near monitors experiences larger pollution reductions after monitors opened.

One may be concerned that cells in the unmonitored areas are too far away from the city center and thus would have different pollution trends from those in the monitored areas. While the parallel trend assumption is not directly testable, we conduct a "placebo" test and an event study analysis to support the assumption. To address the identification concern of endogenous monitor locations, we conduct another "placebo" test with random monitor placements.

4.3.1 Placebo Tests

First, we conducted a "placebo" test using only pre-program periods and randomly assign opening years for all monitors at the same locations. The rationale behind the placebo test is that cells in "monitored" and "unmonitored" areas should not be significantly different over a false-opening year in the absence of the monitoring program. For each monitor, we randomly assign an opening year between 2007 to 2011 for 500 times. We then conduct 500 estimations of equation (2) and plot the distribution of the coefficients in Figure 5. Comparing with the observed coefficient, we find that the observed coefficient lies outside of the 99% confidence interval of the coefficients from 500 placebo tests, which center around 0.016. This result shows that before the monitoring program, a false opening would not lead to larger pollution reductions in monitored areas than unmonitored areas.

Second, in order to show that our findings indeed a result of local pollution reductions in monitored areas, we conduct a placebo test with random monitor locations. Keeping the number of monitors and the year of joining the program unchanged, we randomly relocate all the monitors within each city 500 times. The underlying idea is that if local officials only conduct strategic reductions in areas very closed to monitors, then no significantly different pollution reduction should be observed in areas with a false monitor opening compare to other areas in the city. After matching the 500 groups of placebo monitors with the satellite grid cells, we estimate equation (2) and plot the distribution of the coefficients in Figure 6. The observed coefficient lies outside of the 99% confidence interval, suggesting that local pollution reductions happened only at the observed monitored areas.

4.3.2 Event Study

We use an event study analysis to show the parallel trends between monitored and unmonitored groups hold for pre-opening periods in general. We divide the years around opening dates into four pre-opening periods $n = -4, ..., -1$, and four post-opening periods

$n = 0, 1, 2, 3$ and run the following regression:

$$\ln(PM2.5_{it}) = \sum_{n=-4}^{3} \beta_n \phi(n) \times Near_{it} + Cell_i + Wave_w \times Year_t \ + \varepsilon_{it} \tag{3}$$

where $\phi(n) = \mathbf{1}[n \leq t \leq n+1]$, indicating interval n. The base interval is the year before the opening year (i.e., $n = -1$). We use a balanced sample which include four years before and three years after monitoring for all monitored cities.

Figure 7 presents the coefficient estimates of $\phi(n)$. The results support the parallel trends assumption in general: compared with the base interval (1-year before opening years), the subsequent changes in air pollution between the monitored and unmonitored areas are not significantly different for the four pre-opening intervals in the specification. In contrast, we find statistically significant different air pollution reduction between the monitored and unmonitored groups in the post-opening intervals for the same specification. The fifth year prior to monitoring exhibits a significant difference, which could be due to more unobserved policy changes in years further before monitoring.

4.3.3 Eliminate Alternative Explanations

In this subsection, we discuss a few alternative explanations which may generate similar patterns, including pollution mean reversion, and the measurement error in the satellite-derived pollution measures. First, an identification concern may arise from the fact that monitors are in urban centers, which happen to be more polluted area. The difference in pollution reduction patterns between the monitored and unmonitored areas exists due to the nature of pollution transporting from dirty areas to clean areas. If this is the case, then one should expect to see larger differences in pollution changes after monitoring for monitored cells located at dirtier areas than monitored cells located in cleaner areas of a city.

A similar concern lies in the political interpretation of local officials' strategic behaviors. One may argue that local officials choosing to reduce more pollution in monitored areas is not a strategy that they play to gaming the performance evaluation. Instead, they choose a more cost-effective way to reduce pollution in a relatively more polluted area, which happens to be the area adjacent to a monitor. To address this type of concerns of monitored areas being coincident with polluted areas, we examine the treatment effects where we allow the impact to differ based on the relative pollution levels of cells within the vicinity of monitors as in Equation (4),

$$\ln(PM2.5_{it}) = \beta Near_{it} + \eta Near_{it} \times PopPM2.5_i + Cell_i + Wave_w \times Year_t + \varepsilon_{it} \qquad (4)$$

where $PopPM2.5_i$ is the population weighted $PM_{2.5}$ in cell i for the base year 2008. We use this measurement to indicate the baseline pollution exposures without monitoring. This specification examines the potential concern of monitors locating in the dirty area of a city. The results are reported in Table 4. We show that mean reversion should not be an issue and monitors being in an dirtier area does not lead to large pollution reductions as concerned. In fact, the magnitude of the triple interaction terms with *Pop_PM2.5* is almost zero comparing to the strategic pollution reductions in monitored areas.

Another possibility that may generate similar results is the measurement errors from satellite-derived pollution measures. The $PM_{2.5}$ data we use is derived from the raw satellite images, which require information from monitor-based sources. The Geographical Weighted Regression method used when deriving $PM_{2.5}$ from satellite images assigns larger weights to areas closer to ground monitors, and smaller weights to farther areas. One may be concerned that the resulted measurement errors from the data generating process will be correlated with the distances to monitors and also varied over time when more ground monitors are opened. If this is the case, then the spatially different pollution patterns could simply because of the spatially differentiated measurement errors. Although van Donkelaar et al. (2016) have conducted several out-of-sample cross-validation tests to justify their satellite-derived $PM_{2.5}$ data, we conduct a robustness check using the raw satellite images to further eliminate this possible explanation. Using the raw AOD data from the NASA MODIS product, we manually aggregate the daily AOD images at 3km by 3km resolution into annual AOD, and match with the ground monitors. The grid cells containing monitors are monitored cell, and those do not contain any monitors are unmonitored cells. Estimating Equation (2) using the AOD data shows a similar result. After monitoring, pollution in monitored cells decreases comparing to unmonitored cells. (Table 5)

After eliminating alternative explanations, the empirical results shown in this section suggest that after monitoring, the area adjacent to monitors experience larger pollution reductions relative to areas farther away. So far, we have not claimed that the spatial gaps in pollution changes are due to local officials' strategic responses to central environmental regulations.

5 Mechanisms and Heterogeneous Effects

In this section, we discuss the potential channels through which the heterogeneous effect by treatment intensity may occur and provide empirical evidence for one major mechanisms. We then conduct multiple heterogeneity analyses to support the political interpretation of the results, and show how the effect size varies in various circumstances, including a cities compliance level, economic development, leader characteristics, and information transparency.

5.1 Channels for Strategic Reduction

We present the spatial distribution of the impact of monitoring and discuss abundant qualitative evidence of the local officials' pollution control strategies to support the political interpretation of our findings. By replacing the binary indicator of one monitored group and one unmonitored group used in Eq (1) with fifteen treatment intensity groups, we show the spatial distribution of the treatment effect by distances from monitors in Table **??**. The changes in the impact of monitoring over space also indicate the potential channels of strategic pollution reductions. The coefficient estimates of *Open* represents the impact of monitoring on air pollution in the base group, which includes cells more than 300km away from the closest monitors. Combining with the interaction terms, the strategic pollution reductions exist within 70km ranges of monitors and are robust in magnitudes. Beyond 70km, the overall impact of monitoring turns positive and continues to increase for cells further away. With more distance bins in the unmonitored groups, Table **??** represents the potential pollution migration patterns across space after monitoring. Note that most of these central monitors are placed in population-dense (urban) areas. Column (2) in Table **??** summarizes the population in each distance bin. Although the monitoring enforcement seems to divert air pollution away from areas near monitors, this does not necessarily lead to policy failure when considering the population exposed to air pollution. However, this could exacerbate inequality issues if pollutions are leaking to rural areas. We provide more discussion in Section 6 on the dynamic changes of monitors' representativeness in population-weighted pollution exposure.

We reviewed numerous policy documents from both the central and local governments in China, collecting reports by national inspections teams, and media newsletters. They show that local governments have strong political incentives in improving air quality readings to meet the centrally designated air quality targets. As we introduced in Section 2, the most direct ways to falsify monitor readings from the devices are difficult to implement with the new monitoring system. Such direct manipulation methods include shutting down monitors during polluted days, blocking up the sensors inside monitoring devices, and deliberately damaging monitors. With the real-time data collecting monitoring system, any of these data manipulations would result in abnormal data patterns and trigger alarms. However, the advanced new system cannot eliminate all possible channels of "manipulating" the monitor readings. As the famous saying in China points out, "when the central government has a policy, the local governments have countermeasures". There are several major strategies that local regulators commonly adopt to "manipulate" the monitor readings.

The first type of strategy directly cleans up the air near ground monitors. Since the monitor locations are known to local regulators, many of them choose to clean up the adjacent areas by spraying water or using fog canon towards either monitors (higher risk of being caught, most effective), or towards trees near monitors (lower risk, less effective). A recent scandal was exposed by the media that in Jan 2018, the building of the Environmental Protection Agency in Shizhuishan, Ningxia Province, where a central monitor is located, was turned into an ice sculpture when the staff tried to reduce monitor readings with fog cannons.

The next set of strategies is the ones causing the largest pollution leakages into unmonitored areas. Short term strategies may include traffic controls in monitored areas, divert food trucks and other mobile polluting sources away from monitors, or restrict operation durations for certain polluters. An inspection report of Tianjin's environmental regulation states that the inspection team found strategic pollution reduction behaviors such as traffic controls and increased water spraying frequency in the monitored areas. Media also revealed temporarily shutting down of gas stations near monitors in Pingdingshan, China. [17]

A more effective strategy in the longer term would be relocating polluting sources from small-scale workshops, restaurants to large industrial plants to suburban or rural areas that are commonly unmonitored. This type of strategy would be preferred considering either economic development or environmental performance (improving monitor readings). However, it would impose the largest environmental damages and bias of central regulations. Based on the baseline DID results in Table **??**, relocation of pollution either by relocating

[17] Example of news and media coverages of the existing manipulation strategies: Yuqing, People.cn; Bloomberg Law; Guancha; People.cn

industrial activities or shifting operation sites seems to be the most common strategy given that unmonitored areas become more polluted after monitoring.

5.2 Mechanism: Changes in Industrial Activities

Considerable anecdotal evidence suggests that local governments targeted spots near monitors for pollution reduction by taking actions such as temporarily shutting down gas stations, and suspending the operation of coal-fired power plants. In this subsection, we employ a novel dataset of satellite-based industrial activities and empirically explore possible channels of strategic cleaning.

Specifically, we examine how the monitoring program leads to changes in industrial activities of the monitored and unmonitored areas. To obtain a measure of the *intensity* of industrial activities at a high degree of spatial resolution, we exploit a crucial fact that most industrial plants emit high-temperature waste gas, and make air above them thermal anomalies. Therefore, thermal anomalies in urban areas can serve as good proxies for industrial activities. To identify these hotspots, we draw on the MODIS Version 6 Global Monthly Fire Location Product, MCD14ML, which traces active fires and other thermal anomalies such as industrial plants and volcanoes.[18]

For our analysis, we identify fire spots under the category of "static land source" as active industrial activities, and measure the count and strength of thermal anomalies in this category. We select Band 21 brightness temperature notated as BR as a measure of thermal anomalies intensity. The fire radiative power (FRP) retrieval is based on the difference of brightness temperature between the emitted pixel and background, which is a measure of the rate of radiant heat output. FRP is related to the rate at which fuel is being consumed (Wooster et al., 2005) and smoke emissions released (Freeborn et al., 2008). We use FRP as another measure to indicate the intensity of industrial activities. The raw data contain multiple fire activities within one pixel one month, so we sum up the number of fire activities within each cell and month to generate "frequency" variable as a measure of monthly industrial activities frequency. Moreover, we generate a dummy variable "count" to indicate if a cell has fire activities (operating plants).

Many cells do not have industrial activities which result in a larger proportion of zero observations for our thermal anomalies count at the cell-month level. To address this particular

[18] Data source: https://firms.modaps.eosdis.nasa.gov/.

feature embedded in the data, we estimate the analysis using a Poisson Regression Model. [19] In a similar vein to the baseline model, the effect of monitoring on monthly industrial activities is evaluated using the following specification:

$$y_{it} = \alpha Open_{it} + \sum_{n=1}^{10} \beta_n Open_{it} \times Bin_n + Cell_i + YearMonth_t + Month \times CityFE + \epsilon_{st} \quad (5)$$

where y_{it} is the number or strength of hotspots in monitored cell i at time t. We include the interactions of monitor opening with ten distance bins in order to see the spatial pattern of the effect. The controls are again cell fixed effect, yearmonth fixed effect, and city-specific time fixed effect to address the endogeneity concerns. As a placebo test, we also perform parallel analysis on vegetation-based thermal anomalies such as agricultural fires. If local strategic cleaning is at work, agricultural fires should be less likely to respond to monitoring than industrial activities.

Table 7 reports the effects of monitoring program on the thermal anomalies (industrial activities) within different treatment intensity bins using the Poisson Regression Model. The coefficient estimates of Open represents the impact of monitoring on air pollution in the base group, which includes cells outside of 120km of monitors. The first two columns report the estimates for the effect of automation on industrial activities' strength (intensive margin) within each distance bin. Column (3) and (4) report the estimates on the operating frequency and operation counts. Consistent with the pollution migration pattern found in Table 6, Table 7 shows that the closer to monitors, the larger reduction of industrial activities is observed. Also, we see that for regions beyond 60km, the industrial activities start to increase after monitoring. Although industrial pollution is only one part of the air pollution captured by satellites, these results provide a direct evidence for an important strategic cleaning mechanism which diverts air pollution away from monitors.

5.3 Heterogeneity in Strategic Pollution Reductions

The complete strategies that local officials use to achieve better monitor readings are hard to test empirically due to data limitations. For example, traffic controls and water spraying in monitored areas are short-term actions which may only be caught by constantly observing

[19] We assume that industrial activities follow a Poisson process with parameter $\lambda : E(Y_{it}) = \lambda_{it} = exp(X_{it}\beta)$, the probability function of Y_{it} is given by $Prob(Y_{it} = y_{it}) = \frac{e^{-\lambda}\lambda_{it}^n}{y_{it}!}$.

the abnormal phenomenon near monitors. Instead, we use several heterogeneity analyses to indirectly support the findings of local regulators' strategic responses. In this subsection, we present evidence from heterogeneous analysis to show that the political incentives of local politicians are indeed the driving forces behind our main findings.

a) Leader Characteristics

City secretaries and mayors play essential roles in policy regulation and implementations. We investigate whether a city secretary/mayor's characteristics have an impact on the strategic pollution reductions after monitoring. Figure 8 shows the heterogeneity analysis by city leader's age, where we separate the sample into two sub-samples by city leaders' age. A mayor has better chances to be promoted to a higher position at an age younger than 57. Thus, a younger mayor may have larger incentives to perform well in the environmental evaluation and adopt more strategic pollution reduction methods in monitored areas. For mayors older than 57, which means they have little to no promotion opportunities, they would be less incentivised to achieve policy targets. Figure 8 shows such results that cities with mayors younger than 57 tend to have larger strategic reductions in monitored areas. On the other hand, we do not find any significant impact of a mayor's educational background on their strategic behaviors. This may suggest the strategic reduction methods are common knowledge for leaders across education levels and do not require elite training. Similar results are seen for city secretary's age, which suggest both types of leaders face similar evaluation pressure and promotion incentives.

b) Roll-out Waves of Entering the Program

As introduced in section 2, we may divide monitors into new monitoring stations, and upgraded existing stations. Cities in the first two waves mostly had built their monitoring systems years before 2012, and upgraded to automation under the new monitoring program, while cities in the third wave are building new monitoring systems. Following Greenstone et al. (forthcoming), we define the pre-automation data manipulation status to further explore the heterogeneous effect in cities which have or have not had data manipulation issues, i.e. under reporting the true pollution levels.

Figure 9 presents the heterogeneous strategic cleaning in cities with upgraded monitors with/without pre-automation data manipulation, and cities with new monitors. We find the cities with new monitors show the largest strategic pollution reductions in monitored areas. This could be a result of the approaching final assessment deadline since these cities joined the program the last. [20] For cities with existing monitors automated, Figure 9 shows that cities engaged in data manipulation prior to automation take more aggressive pollution control actions near monitors. This result indicates that local governments' various levels of pollution abatement pressures and leaders' preference in gaming the system still play important roles in their cleaning strategies.

c) Economic Growth

In general, there are trade-offs between economic development and pollution abatement for local regulators. Chen et al. (2018a) find evidence of trade-offs between SO_2 emissions and GDP growth rate when an emissions quota was included in the performance evaluation system for top local bureaucrats (city mayors and Party secretaries) for the first time in 2005. To examine the role of economic growth pressure, we generate a dummy variable indicating the growth or recession of a city's GDP in the previous year and interact with the DID treatment intensity term. Column (3) in Table 8 shows the results. We find that when a local official faces downward pressure on economic growth, they tend to reduce strategic measures that improve monitor readings. This set of heterogeneity results suggest that local regulators are balancing both their efforts and performance in economic growth and pollution control. The gap in pollution changes between monitored and unmonitored areas is indeed a result of local regulators' strategic pollution reductions.

6 Policy Implications and Suggestions

6.1 Dynamic Monitors Representativeness

From the representation error map in Figure 4, the current monitoring system in most cities shows good representativeness when the cities first joined the program. However, similar to the monitoring systems in developed countries, monitor locations are unlikely to change once the monitors were placed. For example, the current air quality monitors in the U.S. were built two decades ago, and covered populated areas following federal guidelines. Other than adding new monitors to nonattainment counties, the existing monitor locations have not changed ever since. Thus, even though monitors were sited to be well-representing

[20] In addition to the annual assessment, the local officials face a final assessment of air pollution reductions at the end of 2017. They may use more aggressive strategies to reduce monitor readings when the final assessment approaches.

counties' overall air quality in the 90s, the representativeness can be dynamic due to human interventions in monitored areas. Using our estimates for the relative pollution reductions in monitored areas (cells within 3km have 6.1% more pollution reductions), and the last observed year of pollution in our data in 2017, we calculate the projected pollution levels for five years from 2018 to 2022. We do not conduct simulations for a longer period into the future because there could be large uncertainty and new regulations. We find that in the near future, the over-representative monitors seem to become more representative of a city's overall air quality. However, by the end of 2022, monitors in around 180 cities will be negatively representing cities' true air pollution. Moreover, the potential pollution leakages to unmonitored areas, mostly rural regions, could cause adverse health impacts and biased evaluation of policy goals.

6.2 The Remote Sensing Data and Other Pollution Information

The key to eliminating or preventing local official's strategic responses to the ground monitoring system is to add referencing data sources into the evaluation. In an ideal world with ground monitors everywhere, local officials are impossible to predict which sets of monitors would be used to evaluate their environmental performance. Thus, the only strategy left is to improve air quality city-wide. This seems unrealistic because ground monitors are large in size and costly to build and maintain. The satellite-based pollution measures can be a good source to fill the gap in ground monitor coverages. As shown in (Sullivan and Krupnick 2018) and Fowlie et al. (2019), remote-sensing data has helped the authors to assess the extent to which the existing U.S. ground monitor-based measurements over- or under-estimate true exposure to $PM_{2.5}$ pollution. In our context, we have used the satellite-based data to re-evaluate the policy goals set by the "Air Ten" action plan for the end of 2017. Unlike the monitor-based pollution patterns estimated in Greenstone et al. (2020), $PM_{2.5}$ decreases by 40% from 2013 to 2018, our estimates find an overall increase in the city-wide pollution level. This suggests that monitor-based evaluation would overstate the environmental performance and distort future policy design.

However, it is important to recognize the limitations of completely relying on satellite images. Satellite-based data is not direct measures of ground pollution levels and is subject to missing data issues that are strongly correlated with cloud coverages. Ground monitors, on the other hand can provide more detailed hourly observations and better accommodate various weather conditions. Additionally, advanced monitoring technologies have provided broader coverages with mobile monitors and micro-monitors that local regulators have less control. Hence, the central government should use this information as supplementary evidence for city-wide pollution evaluation. This is true for any country relying on stationary, in situ monitors in environmental regulations. Overall, a better policy design of monitoring regulation and enforcement would need a mixed contribution from the ground monitoring system, remote-sensing technologies, mobile monitors, as well as public awareness, and third-party auditors.

7 Conclusion

Environmental regulations are often associated with strategic responses, and effective regulation relies on accurate monitoring and measurements. In major countries around the world, local governments face stringent pollution abatement targets, which often link local governments' federal funding or regulators' promotions with their success in achieving these targets. A growing literature has highlighted the unintended consequences of these policies, such as pollution spillover in China's water quality regulation, which undermines policy goals and bias evaluations. This paper adds to these studies by demonstrating strategic responses to central regulations at local levels and extending the literature to air pollution monitoring regulations. Using high-resolution satellite measures of pollution, we have shown that local officials have incentives to improve monitor readings by strategically reducing pollution in monitored areas. Such strategic behaviors will change the spatial representativeness of the current monitoring system and lead to biased policy evaluations.

We find that there exists a significant difference between pollution changes in areas adjacent to monitors and areas far away from monitors after monitoring. Although the new ground monitoring network has improved data quality significantly, the gaps in monitor coverage could still incentivize local officials' strategic cleaning behavior targeting monitored areas. The spatial radius results show that pollution in unmonitored areas either does not change or even increases after monitors roll-out. By studying the temporal and spatial heterogeneity, we provide evidence supporting the political interpretation of the strategic pollution reductions. Overall, our results are consistent with the expectation that strategic pollution reductions are more likely to arise with larger incentives to improve monitor readings, such as in cities with younger mayors and cities with approaching assessment deadlines.

Our results emphasize the importance of accurate and representative measurements in regulations and are widely applicable to any regulations with in situ monitoring systems globally. Our paper contributes to the growing literature on environmental monitoring regulation and enforcement by expanding the study to China's air quality monitoring system. We highlight the importance of a monitoring regulation that accounts for local regulators' strategic responses and considers the monitoring network from a dynamic point of view. The results are also widely applicable for building or improving monitoring systems in other countries, both in the developed and developing world. We provide policy suggestions for efficient regulations that require a mixed source of pollution information from ground-level monitors, advanced monitoring techniques, and the public to accurately evaluate local officials' environmental performance and improve air quality city-wide.

References

Aghion, Philippe and Jean Tirole, "Formal and real authority in organizations," *Journal of political economy*, 1997, *105* (1), 1–29.

Andrews, Steven Q, "Inconsistencies in air quality metrics: 'Blue Sky' days and PM10 concentrations in Beijing," *Environmental Research Letters*, July 2008, *3* (3), 034009.

Banzhaf, Spencer, Lala Ma, and Christopher Timmins, "Environmental justice: The economics of race, place, and pollution," *Journal of Economic Perspectives*, 2019, *33* (1), 185–208.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou, "From Fog to Smog: the Value of Pollution Information," *NBER Working Paper*, 2020, *w26541.*

Bento, Antonio, Matthew Freedman, and Corey Lang, "Who benefits from environmental regulation? Evidence from the Clean Air Act Amendments," *Review of Economics and Statistics*, 2015, *97* (3), 610–622.

Bombardini, Matilde and Bingjing Li, "Trade, Pollution and Mortality in China," *NBER Working Paper*, 2016, *w22804.*

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi, "Gaming in Air Pollution Data? Lessons from China," *The B.E. Journal of Economic Analysis & Policy*, December 2012, *12* (3).

_ , _ , _ , and _ , "The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality," *Journal of Environmental Economics and Management*, November 2013, *66* (3), 424–443.

Chen, Yvonne Jie, Pei Li, and Yi Lu, "Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in China," *Journal of Development Economics*, 2018, *133*, 84–101.

Chen, Zhao, Matthew E. Kahn, Yu Liu, and Zhi Wang, "The consequences of spatially differentiated water pollution regulation in China," *Journal of Environmental Economics and Management*, March 2018, *88*, 468–485.

Currie, Janet, John Voorheis, and Reed Walker, "What caused racial disparities in particulate exposure to fall? New evidence from the Clean Air Act and satellite-based measures of air quality," *Technical Report, National Bureau of Economic Research*, 2020.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*," *The Quarterly Journal of Economics*, November 2013, *128* (4), 1499–1545.

Fisman, Raymond and Yongxiang Wang, "The Mortality Cost of Political Connections," *The Review of Economic Studies*, October 2015, *82* (4), 1346–1382.

Foster, Andrew, Emilio Gutierrez, and Naresh Kumar, "Voluntary Compliance, Pollution Levels, and Infant Mortality in Mexico," *American Economic Review*, April 2009, *99* (2), 191–197.

Fowlie, Meredith, Edward Rubin, and Reed Walker, "Bringing Satellite-Based Air Quality Estimates Down to Earth," 2019, p. 12.

Ghanem, Dalia and Junjie Zhang, "'Effortless Perfection:' Do Chinese cities manipulate air pollution data?," *Journal of Environmental Economics and Management*, September 2014, *68* (2), 203–225.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," Technical Report w25018, National Bureau of Economic Research, Cambridge, MA September 2018.

Grainger, Corbett and Andrew Schreiber, "Discrimination in Ambient Air Pollution Monitoring?," *AEA Papers and Proceedings*, 2019, *109*, 277–82.

_ , _ , and Wonjun Chang, "Do Regulators Strategically Avoid Pollution Hotspots when Siting Monitors? Evidence from Remote Sensing of Air Pollution," 2019, p. 51.

Gray, W. B. and J. P. Shimshack, "The Effectiveness of Environmental Monitoring and Enforcement: A Review of the Empirical Evidence," *Review of Environmental Economics and Policy*, December 2011, *5* (1), 3–24.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu, "Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution," *American Economic Review: Insights*, forthcoming.

_ , _ , Shanjun Li, and Eric Zou, "China's War on Pollution: Evidence from the First Five Years," *Review of Environmental Economics and Policy*, 2020.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend, "High-Resolution Global Maps of 21st-Century Forest Cover Change," *Science*, November 2013, *342* (6160), 850–853.

He, Guojun, Shaoda Wang, and Bing Zhang, "Watering Down Environmental Regulation in China*," *The Quarterly Journal of Economics*, November 2020, *135* (4), 2135–2185.

Hoff, Raymond M. and Sundar A. Christopher, "Remote Sensing of Particulate Pollution from Space: Have We Reached the Promised Land?," *Journal of the Air & Waste Management Association*, June 2009, *59* (6), 645–675.

Hölmstrom, Bengt, "Moral hazard and observability," *The Bell journal of economics*, 1979, pp. 74–91.

Huang, Keyong, Qingyang Xiao, Xia Meng, Guannan Geng, Yujie Wang, Alexei Lyapustin, Dongfeng Gu, and Yang Liu, "Predicting monthly high-resolution PM2.5 concentrations with random forest model in the North China Plain," *Environmental pollution*, 2018, *242*, 675–683.

Hubbard, Thomas N, "The demand for monitoring technologies: the case of trucking," *The Quarterly Journal of Economics*, 2000, *115* (2), 533–560.

Jia, Ruixue, "Pollution for Promotion," *21st Century China Center Research Paper*, 2017, pp. No. 2017–05.

_ and Huihua Nie, "Decentralization, Collusion, and Coal Mine Deaths," *Review of Economics and Statistics*, March 2017, *99* (1), 105–118.

Kahn, Matthew E, "Domestic pollution havens: evidence from cancer deaths in border counties," *Journal of Urban Economics*, July 2004, *56* (1), 51–69.

Kahn, Matthew E. and Erin T. Mansur, "Do local energy prices and regulation affect the geographic concentration of employment?," *Journal of Public Economics*, May 2013, *101*, 105–114.

_ , Pei Li, and Daxuan Zhao, "Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives," *American Economic Journal: Economic Policy*, November 2015, *7* (4), 223–242.

Karplus, Valerie J., Shuang Zhang, and Douglas Almond, "Quantifying coal power plant responses to tighter SO_2 emissions standards in China," *Proceedings of the National Academy of Sciences*, July 2018, *115* (27), 7004–7009.

Lee, Hyung Joo, Brent A. Coull, Michelle L. Bell, and Petros Koutrakis, "Use of satellite-based aerosol optical depth and spatial clustering to predict ambient PM2.5 concentrations," *Environmental Research*, October 2012, *118*, 8–15.

Li, Hongbin and Li-An Zhou, "Political turnover and economic performance: the incentive role of personnel control in China," *Journal of Public Economics*, September 2005, *89* (9-10), 1743–1762.

Lin, Yatang, Jin Wang, and Fangyuan Peng, "Selective Siting or Strategic Cleaning: Comparing Chinese AmbientPollution Monitoring Data to Remote Sensing of Air Pollution," *Working Paper*, 2020.

Liu, Yang, Meredith Franklin, Ralph Kahn, and Petros Koutrakis, "Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: A comparison between MISR and MODIS," *Remote Sensing of Environment*, March 2007, *107* (1-2), 33–44.

Niu, XueJiao, XiaoHu Wang, Jie Gao, and XueJun Wang, "Has third-party monitoring improved environmental data quality? An analysis of air pollution data in China," *Journal of Environmental Management*, January 2020, *253*, 109698.

Oliva, Paulina, "Environmental Regulations and Corruption: Automobile Emissions in Mexico City," *Journal of Political Economy*, June 2015, *123* (3), 686–724.

Ruminski, Mark, Shobha Kondragunta, Roland Draxler, and Glenn Rolph, "Use of environmental satellite imagery for smoke depiction and transport model initialization," 01 2007.

Shimshack, Jay P., "The Economics of Environmental Monitoring and Enforcement," *Annual Review of Resource Economics*, November 2014, *6* (1), 339–360.

Sullivan, Daniel M and Alan Krupnick, "Using Satellite Data to Fill the Gaps in the US Air Pollution Monitoring Network," 2018, p. 33.

Turner, W., C. Rondinini, N. Pettorelli, B. Mora, A.K. Leidner, Z. Szantoi, G. Buchanan, S. Dech, J. Dwyer, M. Herold, L.P. Koh, P. Leimgruber, H. Taubenboeck, M. Wegmann, M. Wikelski, and C. Woodcock, "Free and open-access satellite data are key to biodiversity conservation," *Biological Conservation*, February 2015, *182*, 173–176.

van Donkelaar, Aaron, Randall V. Martin, Chi Li, and Richard T. Burnett, "Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors," *Environmental Science & Technology*, March 2019, *53* (5), 2595–2611.

_ , _ , Michael Brauer, N. Christina Hsu, Ralph A. Kahn, Robert C. Levy, Alexei Lyapustin, Andrew M. Sayer, and David M. Winker, "Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors," *Environmental Science & Technology*, 2016, *50* (7), 3762–3772.

Wei, Jing, Wei Huang, Zhanqing Li, Wenhao Xue, Yiran Peng, Lin Sun, and Maureen Cribb, "Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach," *Remote Sensing of Environment*, 2019, *231*, 111221.

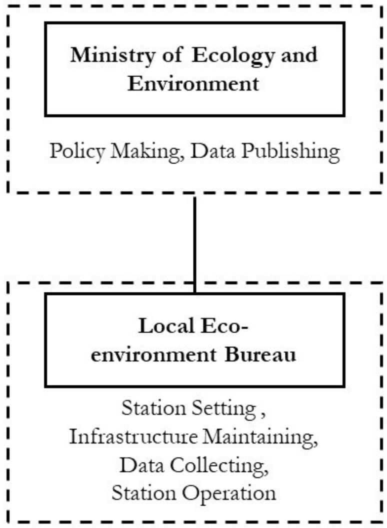
Xia, Haiping, Yunhao Chen, and Jinling Quan, "A simple method based on the thermal anomaly index to detect industrial heat sources," *International journal of applied earth observation and geoinformation*, 2018, *73*, 627–637.

Yang, Lin, "Pollution Monitoring, Strategic Behavior, and Dynamic Representativeness," *Working Paper*, 2020.

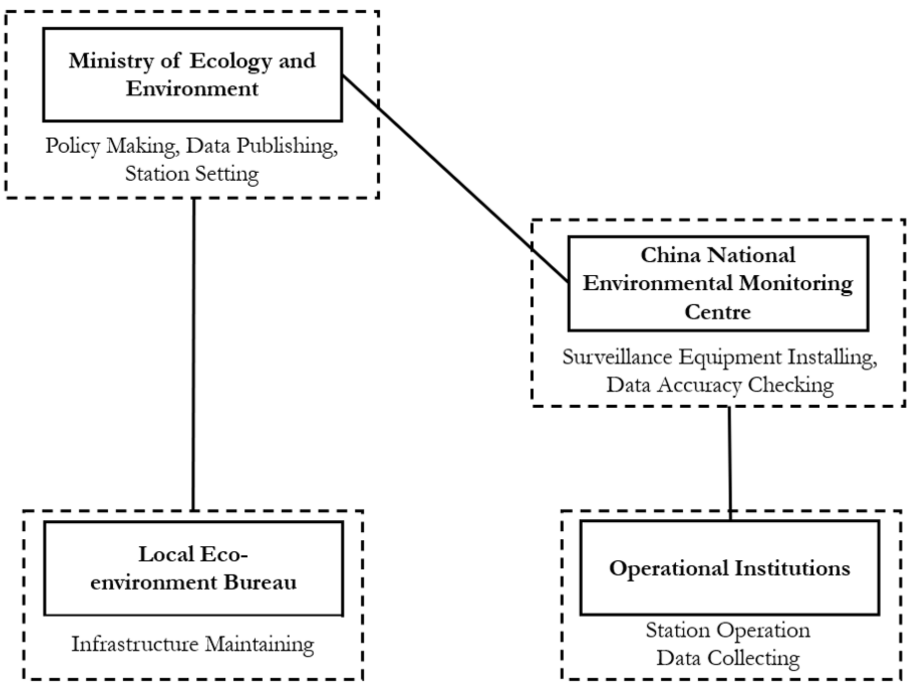
Zhang, Ying and Zhengqiang Li, "Remote sensing of atmospheric fine particulate matter (PM2.5) mass concentration near the ground from satellite observation," *Remote Sensing of Environment*, April 2015, *160*, 252–262.

Zou, Eric, "Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality," *American Economic Review*, 2021.

Figure 1: Change in Authorities' Responsibilities



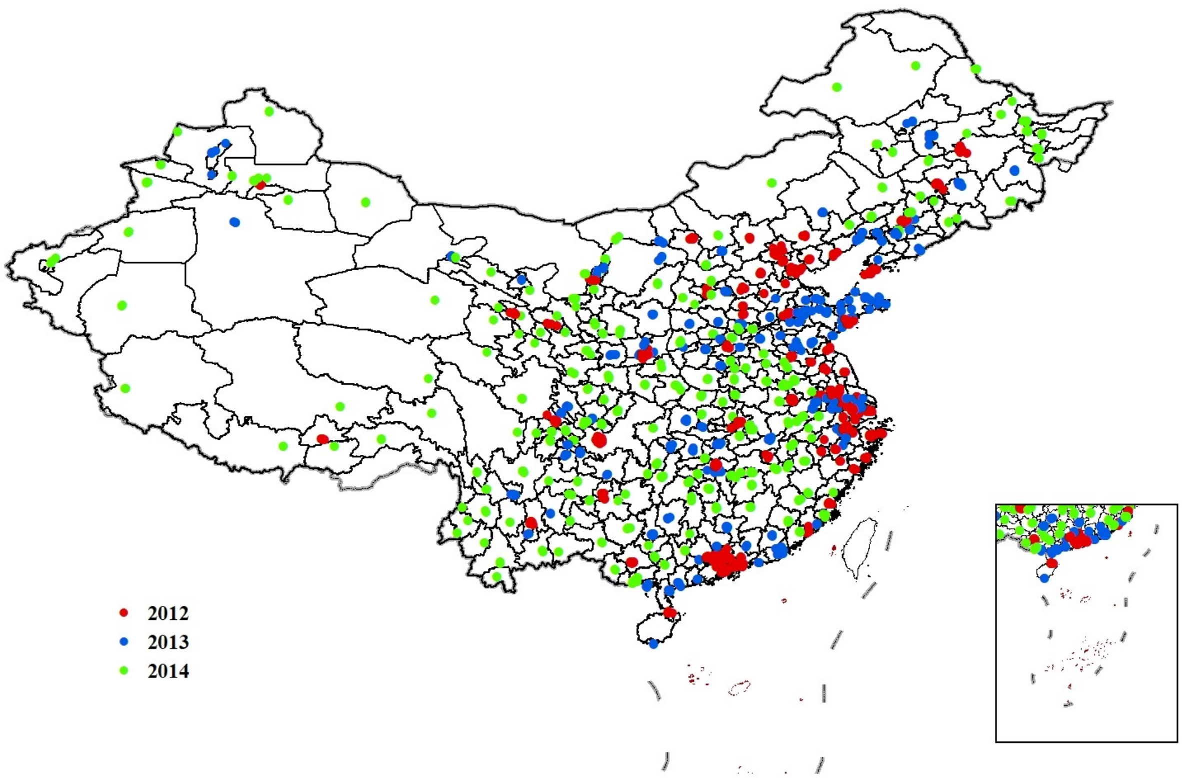
(a) Before



(b) After

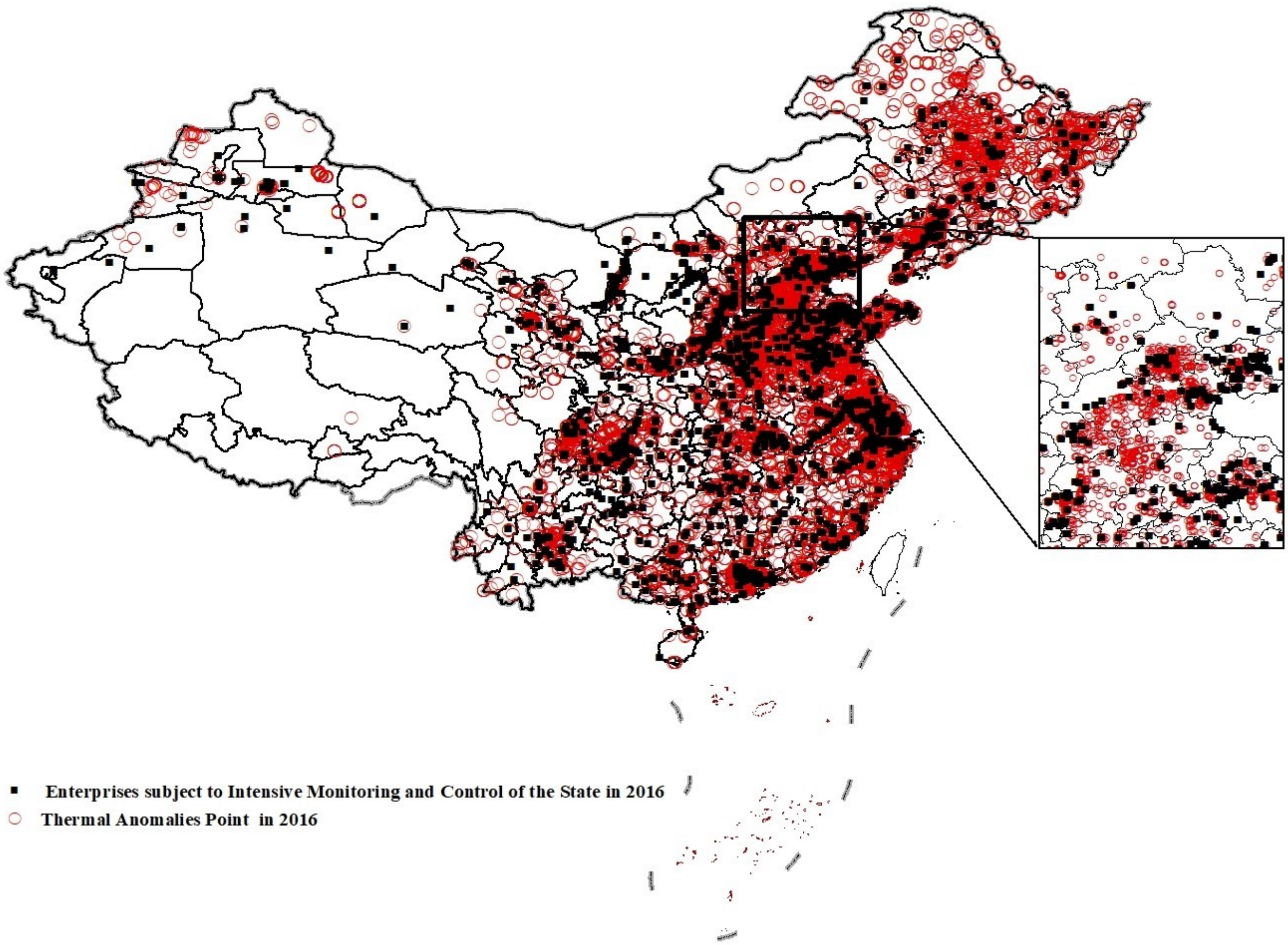
Notes: These two figures illustrate the roles and responsibilities of different environmental authorities, before (Panel A) and after (Panel B) the introduction of new standards. China National Environmental Monitoring Centre (CNEMC) is a newly established institution directly under the management of the Ministry of Environment and Ecology (MEE). It entrusts and oversees several third-party operational institutes to operate and maintain the monitoring stations. Among various responsibilities, $InfrastructureMaintenance$ refers to ensuring the supply of electricity and communications, and $DataAccuracyChecking$ denotes checking the anomaly data.

Figure 2: The Timeline of Monitoring Station Automation



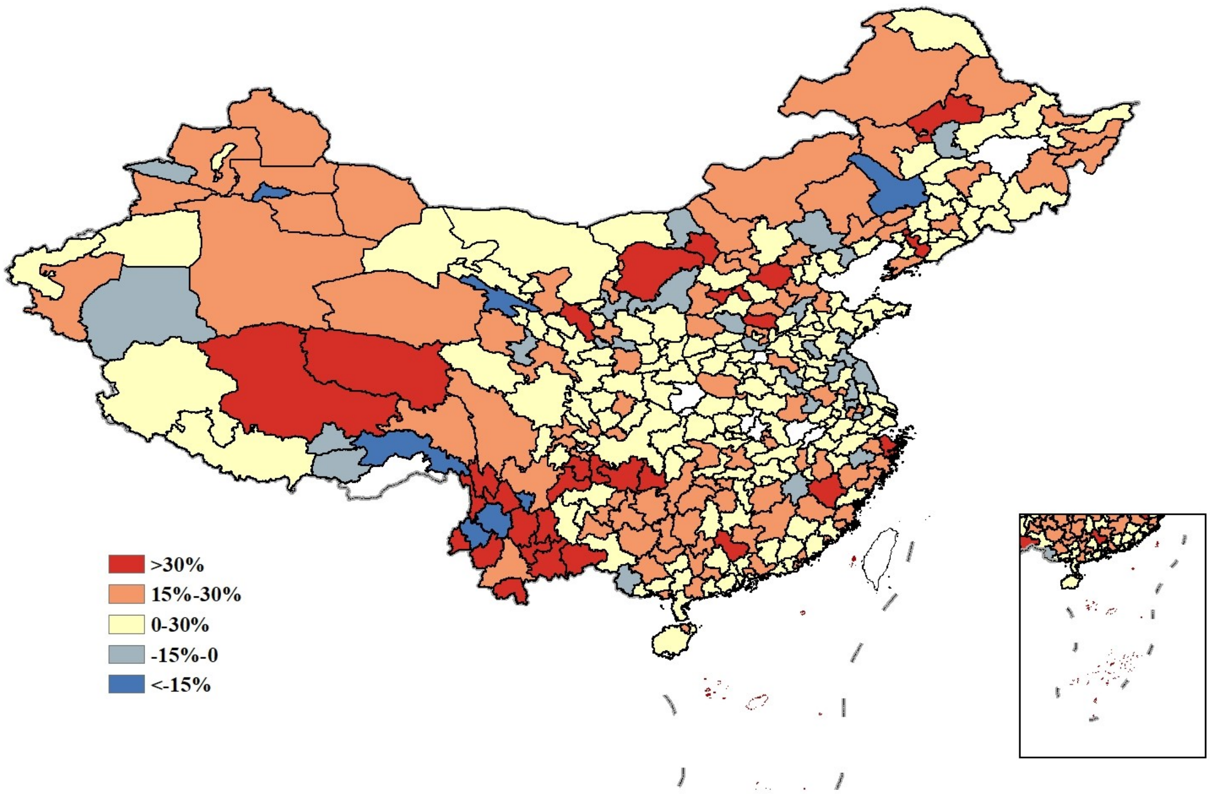
Notes: This figure displays the spatial distribution of stations automated in different waves. Wave 1, wave 2 and wave 3 denote the stations which were required to be automated before in 2012, 2013 and 2014 respectively.

Figure 3: Thermal Anomalies Hotpots and Industrial Plant location



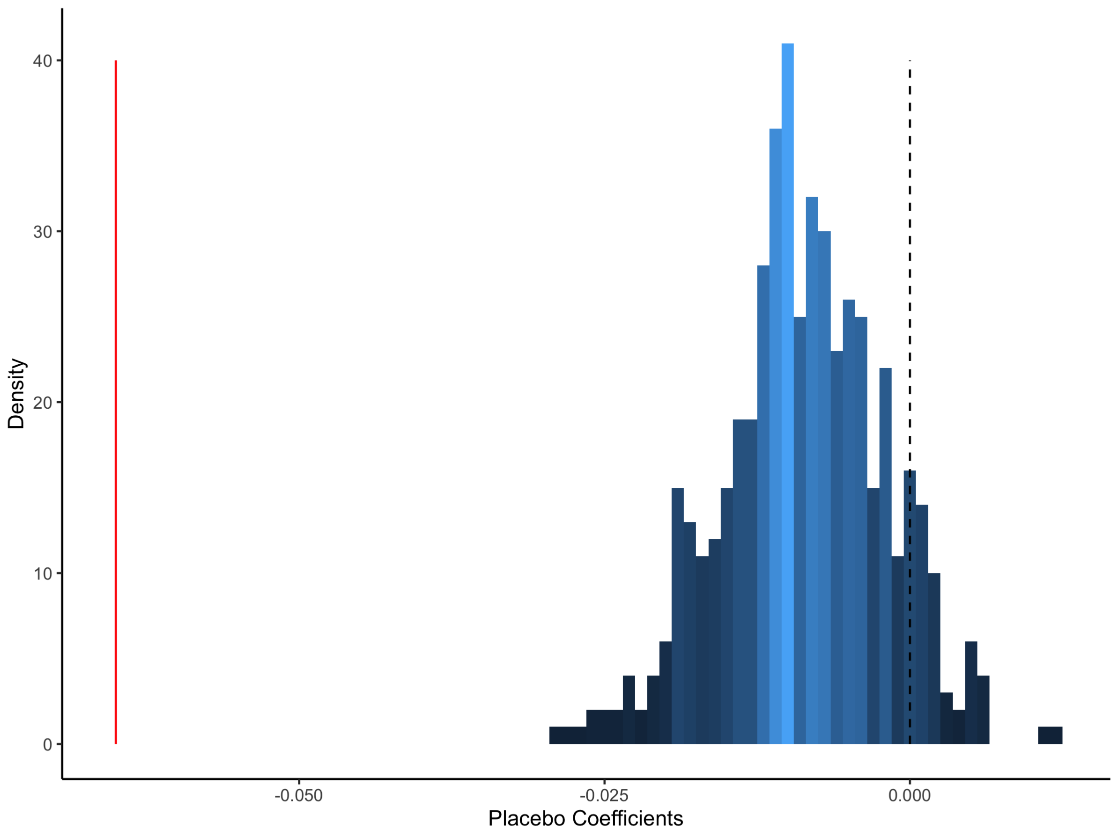
Notes: This figure shows the location of the enterprises subject to intensive monitoring and control of the State in 2016 and thermal anomalies point in 2016. Thermal anomalies point are restricted in static hot spot which could indicate industrial plants. There are 1829 enterprises subject to intensive monitoring and control of the State in 2016 and thermal anomalies static hot spot have 20134 points in 2016. The zoom-in area is Beijing-Tianjin-Hebei Urban Agglometates.

Figure 4: Monitor Representation Errors in openyear: All Cells vs. Monitored Cells



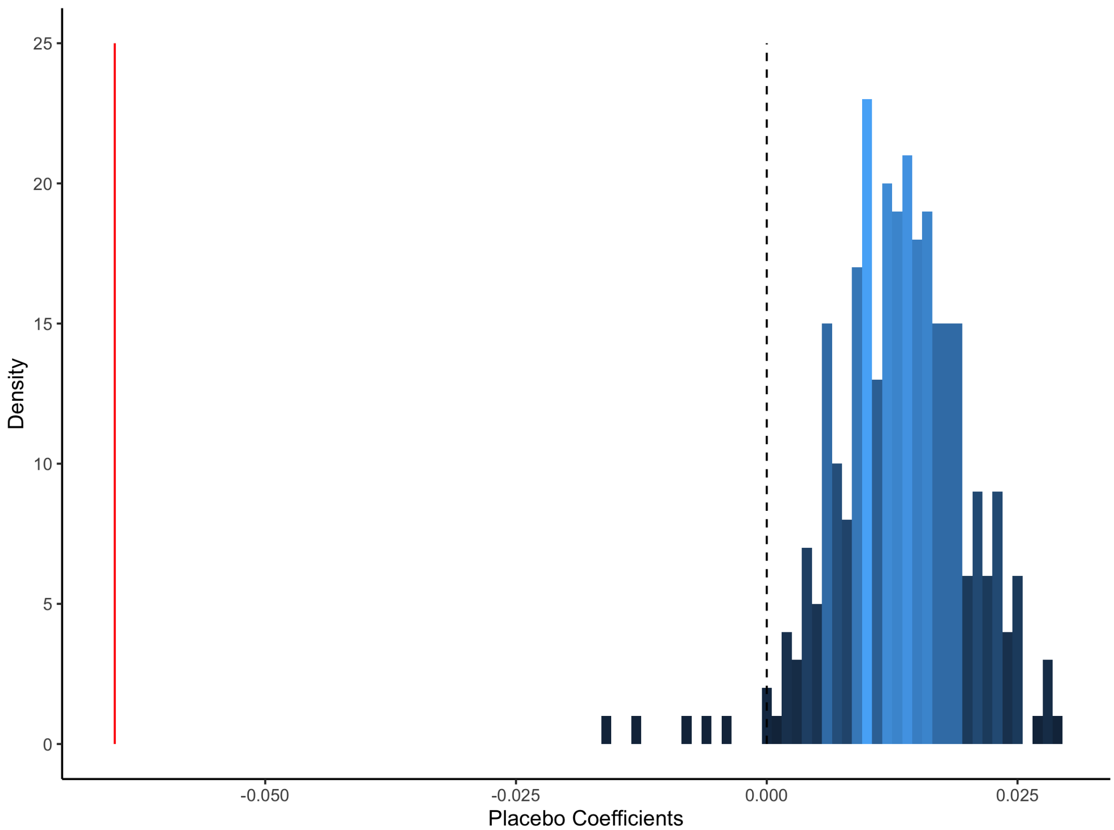
Notes: This figure shows the monitors representation errors in the year when city join monitoring program. The representation error is defined as the percentage difference between city average pollution level calculated based on only monitored cells and city average pollution based on all cells.

Figure 5: Placebo Test with Random Opening Years in Pre-Monitoring Periods



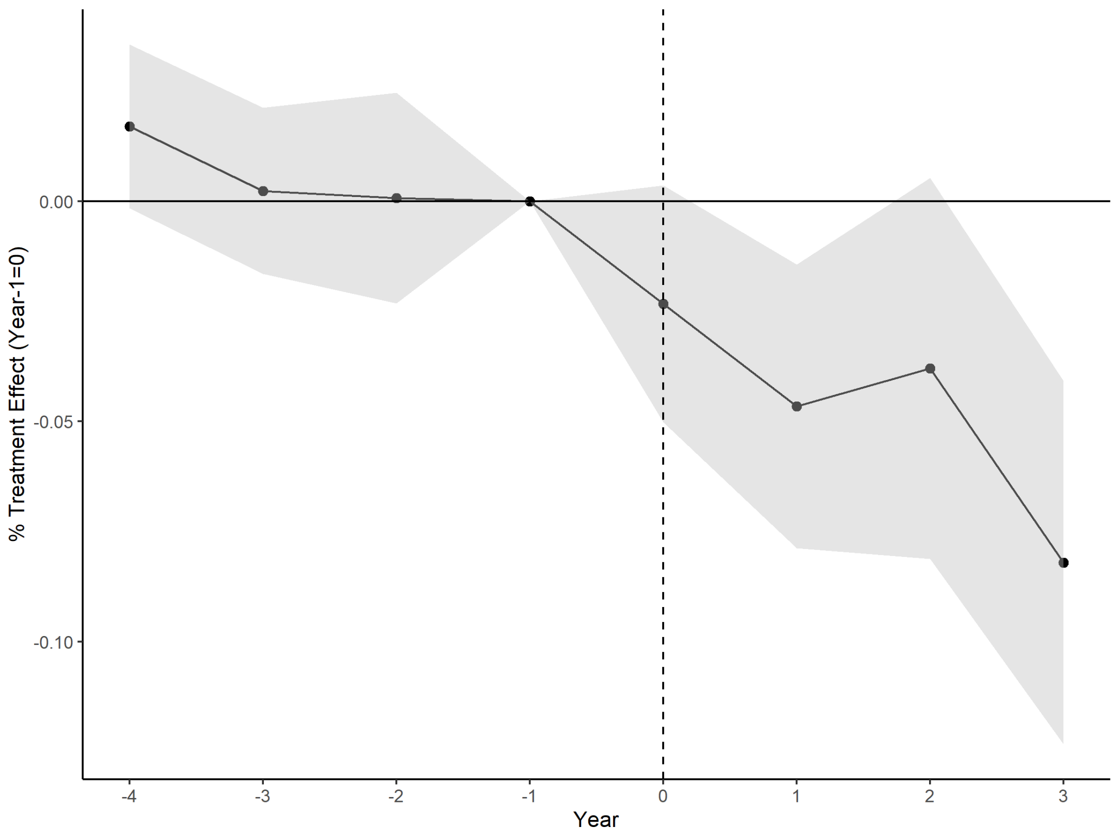
Note: This figure shows the results of a "placebo" test using only pre-program periods and randomly assign each monitor an opening year. We conduct 500 estimations of the treatment intensity analysis and plot the distribution of the 500 placebo coefficients and compare them with the observed effect size using the real sample (red line).

Figure 6: Placebo Test with Random Monitor Locations



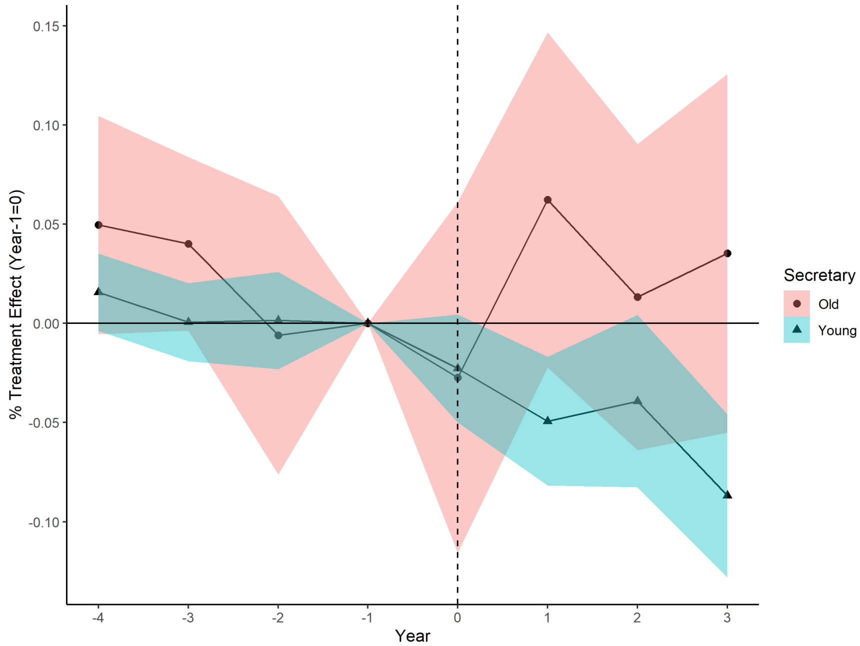
Note: This figure shows the results of a "placebo" test that conducts 500 randomly relocations of all monitors within a city and keep the opening year unchanged. We conduct 500 estimations of equation (2) and plot the distribution of the 500 placebo coefficients, and compare them with the observed effect size using the real sample (red line).

Figure 7: An Event Study of New Monitoring Program on the 3km-Around-Monitor Pollution Level

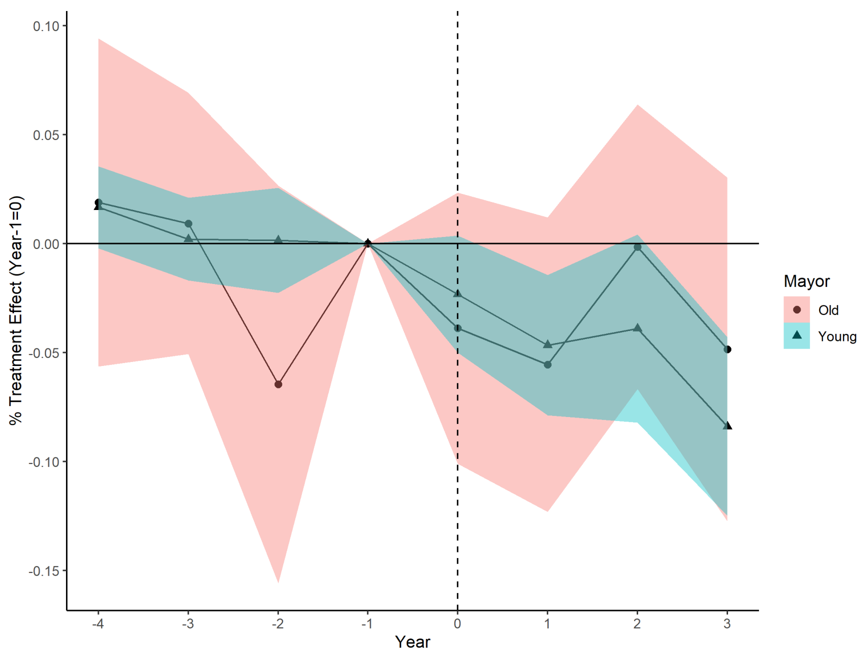


Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for each year within the event window (see Equation (3)), which shows the event study results of monitor opening with treatment intensity on air pollution. The omitted time category is one year before a city join the new monitoring program. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Figure 8: Heterogeneity Analysis by City Leader Age: An Event Study of New Monitoring Program on the 3km-Around-Monitor Pollution Level



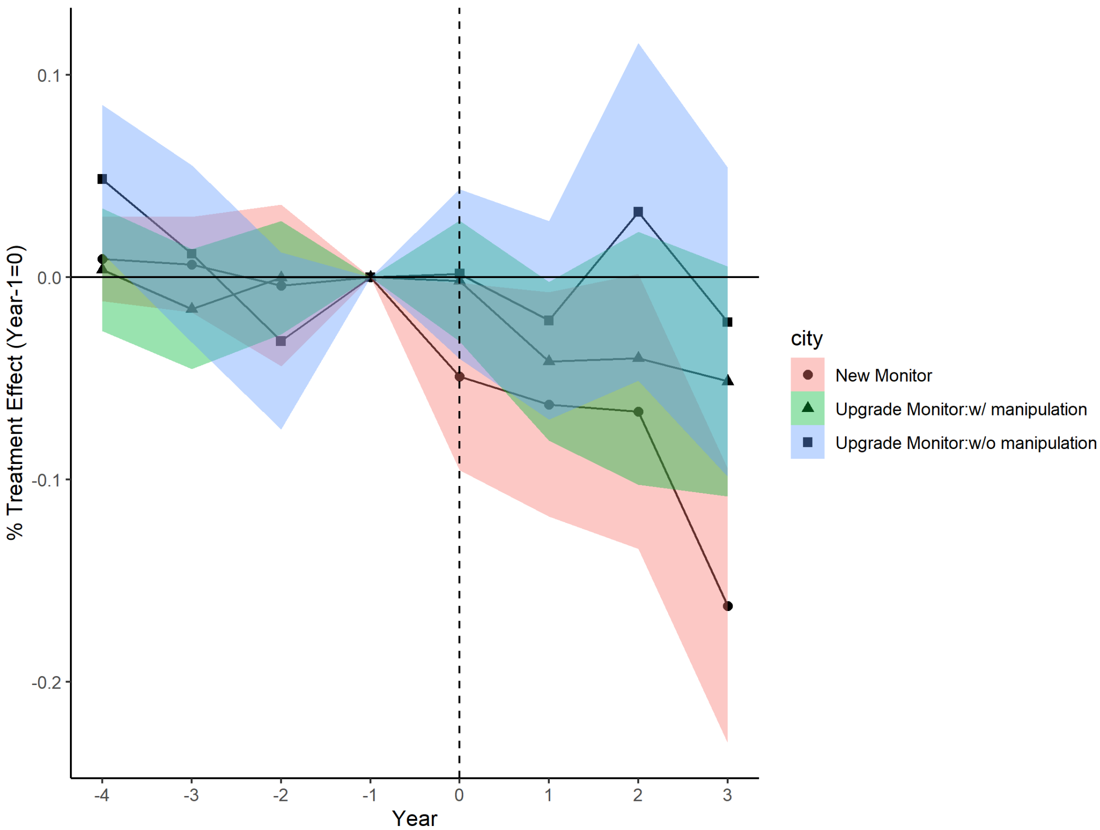
(a) City Secretary Age



(b) City Mayor Age

Notes: These two figure plots the estimated coefficients and their 95% level confidence intervals for each year within the event window, for two subsamples of cities classified by city leader's age. Regression include cell fixed effects, year fixed effect and wave by year fixed effects. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Figure(a) and (b) are heterogeneity analysis by city secretaries' age and mayors' age, respectively. Young is a dummy variable that equal to one if a city leader younger than 57. The omitted time category is one year before a city join the new monitoring program.

Figure 9: Heterogeneity Analysis by City Type: An Event Study of New Monitoring Program on the 3km-Around-Monitor Pollution Level



Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for each year within the event window, for three subsamples of cities classified by whether type of cities. Regression include cell fixed effects, year fixed effect and wave by year fixed effects. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. New is for cities that do not have monitors before the program. For cities that have monitors before program, manipulation status are defined by whether the local linear RD estimate is positive and calculated by using algorithm in Greenstone et al. (forthcoming). Upagrade Monitor:w/ manipulation refers to cities that manipulate data before the program and Upagrade Monitor:w/o manipulation means their RD estimate is negative. We also include missing data cities in the regression which do not show in this graph. Missing data city refers to cities that have stations before and their most of pollution monitoring data are missing so that manipulation status could not be estimated. The omitted time category is one year before a city join the new monitoring program.

Table 1: Variable Definitions and Descriptive Statistics.

Variables	N	Mean	Sd	Min	Max
	(1)	(2)	(3)	(4)	(5)
Panel A: Cells and Pollutants					
$PM_{2.5}$	10,414,296	35.217	27.234	0.000	223.889
Distance	10,427,410	160.373	149.849	0.312	1,380.487
Open	10,427,410	0.437	0.496	0	1
(0-3km)	10,427,410	0.003	0.059	0	1
Panel B: Satellite-based Economic Activity Measures					
Fire Radiative Power(FRP)	11,470,151	0.069	1.401	0	794
Brightness(BR)	11,470,151	1.878	24.316	0	440
Freq	11,470,151	0.018	0.977	0	554
Count	11,470,151	0.006	0.077	0	1
Panel C: City Characteristics					
$\text{Startage}_{\text{secretary}}$	10,137,302	49.829	3.835	40.000	65.000
$\text{Startage}_{\text{mayor}}$	10,225,959	48.462	3.613	35.000	60.000
New monitor indicator	10,427,410	0.837	0.369	0	1
Manipulation indicator	1,339,760	66.967	122.432	-75.744	749.508
Population in 2010	10,425,560	100.789	603.908	0.000	74,052.430
GDP	10,408,135	111.594	207.588	1.266	3,292.501

Notes: Observations are at the city-cell-year level. In Panel A, $PM_{2.5}$ is satellite-estimated $PM_{2.5}$. Distance is the distance between cell and its nearest monitoring stations. Open is a dummy variable which equals one after the city where cell located becomes joined the air monitoring program, and zero otherwise.(0-3km) is a dummy variable that equals one if cells are located around city monitor stations within 3km. In Panel B, fire radiative power(FRP) is defined as the rate of radiant heat output and related to the rate at which fuel is being consumed and smoke emissions released. Brightness(BR) is defined as the Band 21 brightness temperature of pixel and measures the intensity of thermal-related economic activities. Freq measures the times of operating industrial plants in each cell. Counts is a dummy variable that equal to one if have thermal-related economic activities in each cell yearly. In Panel C, $\text{Startage}_{\text{secretary}}$ and $\text{Startage}_{\text{mayor}}$ are ages of cities' secretary and mayors at the start of their office term. New monitor indicator is a dummy variable that equals one if city do not have monitors before monitoring station automation program. Manipulation indicators are the local linear RD estimate calculated by using algorithm in (). GDP (in 1 billion¥) are city-level GDP and population (*person*) and Population in 2010 are cell-level population.

Table 2: Strategic Cleaning Response to Monitoring Program Automation

	Dependent variable: $\ln(PM_{2.5})$				
	(1)	(2)	(3)	(4)	(5)
open	0.052**	0.054**	0.052**	0.055**	
	(0.023)	(0.024)	(0.023)	(0.024)	
(0-3km)×open			-0.091***	-0.061***	-0.061***
			(0.022)	(0.015)	(0.013)
CellFE	X	X	X	X	X
Year FE	X	X	X	X	X
Wave×Year Trend		X		X	
Wave×Year FE					X
Observations	10,413,707	10,413,707	10,413,707	10,413,707	10,413,707
R^2	0.966	0.966	0.966	0.966	0.967
Adjusted R^2	0.962	0.963	0.962	0.963	0.963

Note: This table reports the effects of new monitoring program on the satellite-based $\ln PM_{2.5}$. The data is full sample contain data from 2008 to 2017. $\ln PM_{2.5}$ is natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Open is a dummy variable that equals one after a city join new monitoring program. (0-3km) is a dummy variable that equals one if cells are located around city monitor stations within 3km.Standard errors are clustered at the city level. *** , ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

Table 3: Strategic Cleaning Response to Monitoring Program With Different Control Group

	Dependent variable: ln($PM_{2.5}$)			
	>15km	>30km	>60km	>90km
	(1)	(2)	(3)	(4)
(0-3km)×open	-0.063***	-0.069***	-0.083***	-0.097***
	(0.014)	(0.015)	(0.019)	(0.026)
CellFE	X	X	X	X
Year FE	X	X	X	X
Wave×Year FE	X	X	X	X
Observations	10,036,814	9,309,540	7,579,589	6,009,758
R^2	0.967	0.967	0.967	0.969
Adjusted R^2	0.963	0.963	0.964	0.965

Note: This table reports the effects of new monitoring program on the satellite-based ln$PM_{2.5}$ with different control group.The data is full sample contain data from 2008 to 2017. Column(1)-(4) use cells within 3km to the monitor as the monitored group and compare different unmonitored groups: cells outside 15km, 30km ,60km and 90km of the monitors, respectively.ln$PM_{2.5}$ is natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Open is a dummy variable that equals one after a city join new monitoring program. (0-3km) is a dummy variable that equals one if cells are located around city monitor stations within 3km.Standard errors are clustered at the city level. *** , ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

Table 4: Difference in Differences Estimation Robustness check: Mean Reversion

	Dependent variable: $\ln(PM_{2.5})$		
	(1)	(2)	(3)
(0-3km)*Open	-0.043***	-0.047***	-0.042***
	(0.013)	(0.014)	(0.013)
Pop_PM*Open		-0.00000	
		(0.00000)	
Pop_PM*(0-3km)*Open		0.00001	
		(0.00000)	
CellFE	X	X	X
Year FE	X	X	X
Wave×Year FE	X	X	X
Observations	6,817,131	6,817,131	6,817,131
R^2	0.966	0.966	0.966
Adjusted R^2	0.962	0.962	0.962

Note: This table reports the effects of new monitoring program on the satellite-based $\ln PM_{2.5}$ with different control group.The data is full sample contain data from 2008 to 2017. Column(1)-(4) use cells within 3km to the monitor as the monitored group and compare different unmonitored groups: cells outside 15km, 30km ,60km and 90km of the monitors, respectively. $\ln PM_{2.5}$ is natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Open is a dummy variable that equals one after a city join new monitoring program. (0-3km) is a dummy variable that equals one if cells are located around city monitor stations within 3km. Pop_PM is the population weighted $PM_{2.5}$ level in base year 2008. Standard errors are clustered at the city level. *** , ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

Table 5: Difference in Differences Estimation Robustness Check Using Raw AOD Data

	Dependent variable: AOD			
Unmonitored Areas:	>3km	>12km	>50km	>50km
	(1)	(2)	(3)	(4)
(0-3km)*Open	-0.021***	-0.022***	-0.033***	-0.034***
	(0.005)	(0.005)	(0.006)	(0.006)
(3-6km)*Open		-0.025***		-0.037***
		(0.005)		(0.006)
(6-9km)*Open		-0.030***		-0.042***
		(0.005)		(0.006)
(9-12km)*Open		-0.034***		-0.046***
		(0.005)		(0.006)
Observations	10,136,285	10,136,285	6,992,163	7,330,347
R^2	0.876	0.876	0.849	0.859

Note: Column (1)-(4) show DID estimation results with treatment intensity defined by distances to monitors. *Open* is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. The dependent variable is the annual AOD at 3km by 3km grid cells. Column (1) & (3) use cells within 3km to the monitor as the monitored group and compare different unmonitored groups: cells outside 3km, or 50km of the monitors. Column (2) & (4) add three distance bins to the monitored group and compare two unmonitored groups. All columns include both the cell FE and Wave×Year FE. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 6: Strategic Cleaning Response to Monitoring Program with Different Treatment Intensity Bins

	Dependent variable: $\ln(\text{PM}_{2.5})$
	(1)
open	0.137***
	(0.038)
(0-3km)	-0.167***
	(0.032)
(3-6km)	-0.163***
	(0.031)
(6-9km)	-0.162***
	(0.031)
(9-15km)	-0.160***
	(0.030)
(15-21km)	-0.158***
	(0.030)
(21-30km)	-0.156***
	(0.029)
(30-45km)	-0.149***
	(0.028)
(45-60km)	-0.140***
	(0.027)
(60-90km)	-0.118***
	(0.026)
(90-120km)	-0.085***
	(0.025)
(120-150km)	-0.048*
	(0.028)
YearFE	X
CellFE	X
Observations	10,413,707
R^2	0.967
Adjusted R^2	0.963

Note: This table reports the effects of new monitoring program on the satellite-based $\ln\text{PM}_{2.5}$ with treatment intensity bins.The data is full sample contain data from 2008 to 2017. $\ln\text{PM}_{2.5}$ is natural logarithm of the cell-level yearly satellite-based $\text{PM}_{2.5}$. Open is a dummy variable that equals one after a city join new monitoring program, which includes cells outside of 150km of monitors. The interactions represent the effect in each treatment intensity group. (0-3km) is a dummy variable that equals one if cells are located around city monitor stations within 3km. Standard errors are clustered at the city level. *** , ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

Table 7: Mechanism of Strategic Cleaning: Thermal Anomalies with Treatment Intensity Bins

	(1)	(2)	(3)	(4)
Dependent variable:	lnFRP	lnBR	lnfreq	count
open	0.322***	0.324***	0.373***	0.325***
	(0.0933)	(0.0870)	(0.0859)	(0.0871)
(0-3km)×Open	-0.581***	-0.584***	-0.667***	-0.585***
	(0.0916)	(0.0853)	(0.0839)	(0.0854)
(3-6km)×Open	-0.503***	-0.493***	-0.562***	-0.493***
	(0.0908)	(0.0852)	(0.0839)	(0.0853)
(6-9km)×Open	-0.485***	-0.477***	-0.532***	-0.477***
	(0.0969)	(0.0919)	(0.0915)	(0.0920)
(9-15km)×Open	-0.487***	-0.481***	-0.503***	-0.481***
	(0.0994)	(0.0932)	(0.0980)	(0.0933)
(15-21km)×Open	-0.428***	-0.419***	-0.451***	-0.419***
	(0.105)	(0.0974)	(0.0965)	(0.0975)
(21-30km)×Open	-0.387***	-0.371***	-0.401***	-0.371***
	(0.0995)	(0.0913)	(0.0942)	(0.0914)
(30-45km)×Open	-0.353***	-0.342***	-0.368***	-0.342***
	(0.0931)	(0.0858)	(0.0840)	(0.0859)
(45-60km)×Open	-0.297***	-0.283***	-0.329***	-0.283***
	(0.0920)	(0.0863)	(0.0864)	(0.0864)
(60-90km)×Open	-0.244***	-0.219**	-0.257***	-0.218**
	(0.0944)	(0.0870)	(0.0896)	(0.0870)
(90-120km)×Open	-0.0284	-0.0196	-0.0176	-0.0191
	(0.105)	(0.0922)	(0.0935)	(0.0922)
(120-150km)×Open	-0.0407	-0.0110	-0.186	-0.0105
	(0.149)	(0.126)	(0.176)	(0.126)
Observations	233,740	233,750	233,750	233,750
Cell FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: This table reports the effects of new monitoring program on the thermal anomalies with intensive bins using a Poisson Regression Model.The data is full sample contain data from 2008 to 2017. Column (1) reports the results on the lnFRP of thermal anomalies and FRP is defined as the rate of radiant heat output and related to the rate at which fuel is being consumed and smoke emissions released. Column (2) reports the results on the lnBR of thermal anomalies, and BR is defined as the Band 21 brightness temperature of pixel and measures the intensity of thermal-related economic activities. Column (3) reports results on the times of thermal anomalies, which measures the times of operating industrial plants in each cell. Column (4) reports results on the count of thermal anomalies, and count is a dummy variable that equal to one if have thermal-related economic activities in each cell yearly. ln$PM_{2.5}$ is natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Open is a dummy variable that equals one after a city join new monitoring program, which includes cells outside of 150km of monitors. The interactions represent the effect in each treatment intensity group. (0-3km) is a dummy variable that equals one if cells are located around city monitor stations within 3km. Standard errors are clustered at the city level. *** , ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

Table 8: Heterogeneity Analysis:

	Dependent variable: $\ln(PM_{2.5})$			
	(1)	(2)	(3)	(4)
(0-3km)×Open	-0.033*	0.040	-0.067***	0.031
	(0.020)	(0.039)	(0.015)	(0.047)
Open×New	-0.009			-0.009
	(0.034)			(0.033)
(0-3km)×New	0.063			0.046
	(0.043)			(0.039)
(0-3km)×Open×New	-0.062**			-0.067**
	(0.030)			(0.032)
Open×Young		0.070**		0.068**
		(0.027)		(0.028)
(0-3km)×Young		0.027		0.006
		(0.022)		(0.024)
(0-3km)×Open×Young		-0.105**		-0.071*
		(0.043)		(0.042)
Open×1(GDP Decline)			0.021	0.021
			(0.036)	(0.036)
(0-3km)×1(GDP Decline))			-0.036	-0.038*
			(0.023)	(0.023)
(0-3km)×Open×1(GDP Decline)			0.102***	0.111***
			(0.033)	(0.033)
YearFE	X	X	X	X
CellFE	X	X	X	X
Wave×Year FE	X	X	X	X
Observations	10,413,707	10,413,707	10,380,564	10,380,564
R^2	0.967	0.967	0.967	0.967
Adjusted R^2	0.963	0.963	0.963	0.963

Notes: This table reports the heterogeneous effects of new monitoring program on the satellite-based $\ln PM_{2.5}$. The data is full sample contain data from 2008 to 2017. $\ln PM_{2.5}$ is natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Open is a dummy variable that equals one after a city join new monitoring program. (0-3km) is a dummy variable that equals one if cells are located around city monitor stations within 3km.1(GDP Decline) is a dummy variable indicating cells inside a city that experienced an economic recession in the previous year (decreased GDP). Column (1) report the heterogeneous effects of new monitoring program on pollution control based on whether cities have air pollution monitors before. Cities which have monitors before the monitoring program will upgrade their monitor after joining while cities without monitors will establish new monitors. New is a dummy variable that equals one if city do not have monitors before monitoring station automation program. Column (2) report the heterogeneous effects of new monitoring program on pollution control based on city mayors' age. Young is a dummy variable that equal to one if city mayors younger than 57. Column (3) include both city and leader's characteristics. Standard errors are clustered at the city level. *** , ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.