Pollution Monitoring, Strategic Behavior, and Dynamic Representativeness

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

In major countries around the world, air quality evaluation is mainly based on stationary in-situ monitors that aim to provide representative measurements of 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 for the dynamic spatial representativeness of ground monitors. We find strong evidence that, despite the use of automation that precludes 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 in cities with varying degrees of pre-automation data tampering conduct, and by those who face different career concerns and public pressure. Our findings underscore a new agency problem: 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, mobile monitors, and public participation.
JEL Codes: K32, Q53, Q58, O13, R11.

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

---

[†]Yang (corresponding author): Urban Governance and Design, Society Hub, Hong Kong University of Science and Technology, Guangzhou, China (email: yangl@ust.hk); Lin: Department of Economics, Division of Public Policy 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@connect.ust.hk). We are very grateful for 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. 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, Wang and Peng (2020). All errors remain our own.

1 Introduction

Enforcement of regulations hinges on accurate measurements of implementation and outcomes. 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. Recent advances in information and communications technology have enabled new modes of monitoring, substantially lowering the principal's costs of observing and verifying actions (Hölmstrom 1979; Hubbard 2000; Greenstone et al. forthcoming). However, local governments might employ more concealed strategic responses, rendering existing outcome-based incentive schemes inefficient. In a context of decentralized governance, this paper asks how the high-powered incentives offered to local officials could undermine the effectiveness of improved monitoring technologies in solving agency problems.

Through the lens of a technology-aided environmental regulatory program in China, we analyze how local governments respond strategically to the *de facto* greater stringency in air pollution regulation resulting from the upgraded monitoring system. In 2012, amidst the public outcry on the lack of transparency in pollution data, the Ministry of Environment and Ecology of China (MEE)[1] launched a nation-wide, real-time air quality monitoring and disclosure program in a staggered fashion across three waves of cities. Throughout 2012-2014, the national environmental air monitoring network expanded from 113 cities to 335 cities. Meanwhile, the MEE centralized the planning, establishment, construction and maintenance of all state-level air quality monitoring stations in order to minimize data manipulation, which was previously rampant at the local level (Greenstone et al. forthcoming).[2] The changes placed new pressures on local officials, who face a delicate trade-off between economic growth and environmental quality, and find their career advancement increasingly tied to environmental performance.

In this paper, we provide the first thorough empirical investigation of local officials' strategic responses to this technology-aided monitoring program. We examine whether automation has led to larger pollution reductions in the areas immediately around monitoring stations

[1]In 2018, the Ministry of Environmental Protection (MEP) was renamed the Ministry of Environment and Ecology of China (MEE).

[2]Serious concerns about air quality data manipulation in China have been well documented (Andrews 2008; Chen et al. 2012; Ghanem and Zhang 2014).

relative to the wider surrounding areas, thereby undermining spatial representativeness of these stations.

We draw primarily upon the fine-scale remote sensing data which fills the spatial gaps in ground monitoring networks to detect pollution changes from space. We obtain ten years' observations of the annual $PM_{2.5}$ concentration at a 1km by 1km resolution for over nine million grids covering the whole of China. Using Difference-in-Differences with treatment intensity, our empirical specification consists of two layers. First, we explore the impact of automation on city-wide pollution by taking advantage of the staggered roll-out of the new monitoring system across cities. The second layer constitutes the core of our analyses, focusing on identifying strategic behaviors associated with automation. We compare the pollution gap between monitored and unmonitored areas before and after automation to the gap in cities that have yet to join the automated monitoring system.

Perhaps not surprisingly, the monitoring roll-out is not randomly assigned. Cities with larger population sizes and higher-level administrative hierarchies were among the first to join the new program. However, the assumption of parallel pre-program trends in earlier and later adopters is verified in an event study, where we include a rich set of pre-treatment city characteristics[3] interacted with flexible functions of time. We also augment our main specifications with additive automation wave-year fixed effects. Although aggregate city-level responses to automation in the first layer are absorbed, we are able to evaluate in the second layer whether subsequent reductions in air pollution are concentrated near automated monitors rather than unmonitored areas within the treated city.

Our main finding is that areas adjacent to monitors experience a 6.1% greater reduction in $PM_{2.5}$ concentrations than those farther away. Prior to the automation, monitored and unmonitored areas had very similar trends in pollution. The gap between the two groups emerged after the monitoring began, and grew even larger as the final assessment deadline set by the central government approached.[4] More importantly, these results are robust to a number of sensitivity checks, such as the use of raw satellite Aerosol Optical Depth (AOD) readings as an alternative outcome measure and the application of entropy-balancing

[3]These pre-program city-level characteristics include GDP, population, pollution level, and indicators for environmental priority cities (which include provincial capitals).

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

weighting (Hainmueller 2012) to ensure balance in pre-treatment characteristics between monitored and unmonitored cells within each city. We view the observed pollution reductions near the monitors as resulting from strategic efforts. As a further test, we rule out an alternative explanation that the gap was driven by the mean reversion of local pollution. Targeting the monitoring sites for very localized improvement did not improve the overall air quality of the automated cities. The overall pollution level in treated cities increased slightly after they joined the automatic monitoring system, despite wider confidence intervals, indicating a possible migration of pollution activities away from the monitors.

Several mechanisms could give rise to the observed strategic clean-up efforts. A wide range of government strategic responses to automation have been well documented by news articles and environmental inspection reports, including spraying water around the monitor, blocking nearby traffic, and temporarily shutting down or relocating polluters. We tested the latter two channels by leveraging a novel thermal anomalies database. Satellite-based thermal anomalies offer high-resolution, high-frequency information on local industrial activities, on the premise that industrial activities often emit high-temperature gases. Based on this approach, we find that automation has led to a 10% reduction in the probability of having industrial activities within 3 km of the monitor and 9.1% reduction in the days of operation, but few changes in the intensity of operation. Consistent with the main results, the number of thermal anomalies farther away from the monitors increased in response to automation, indicating increased emissions farther away.

We also find substantial heterogeneity in responses to automation across cities. We show that cities that had no monitors in the past are more likely to concentrate more pollution control efforts near the monitors in response to newly installed monitors, possibly due to the imminent pressure of central assessment. Cities with an apparent history of data manipulation in the old monitoring system, and cities headed by younger mayors who are competing to demonstrate air quality improvements for career advancement, are also more likely to take targeted pollution reduction actions near monitors. Finally, cities with more active civic participation in pollution monitoring engage less in strategic behavior.

Overall, the analysis underscores the misaligned incentives of the central regulator and local bureaucrats within China's target-based performance evaluation system. Our study highlights the complexity of the moral hazard problem with the existence of concealed actions. The key implication for the design of environmental regulations is that the central authority needs to rigorously account for local agents' possible strategic responses. Performance monitoring and evaluation should closely integrate ground-level data with other measures such as remote sensing information, 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 up just the immediate areas surrounding the monitors rather than engaging in systematic reductions that benefit wider regions.

Our paper makes the following contributions. First, we add to the literature on environmental monitoring, regulation and enforcement (Duflo et al. 2013; Shimshack 2014; Gray and Shimshack 2020). Our paper extends two predecessors (Grainger, Schreiber and Chang 2019; Zou 2021) by providing one of the first analyses that links strategic clean-up actions by local officials with the dynamic change in monitoring representativeness in China. We provide new evidence that monitoring designs that were efficient ex-ante 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 face increasing need for a new or upgraded system.

Second, our paper relates closely to two concurrent studies that utilized the same regulatory context to investigate the role of monitor-based pollution information in shaping avoidance behaviors (Barwick et al. 2020) and in detecting pre-automation pollution data manipulation (Greenstone et al. forthcoming). Our results complement these studies by uncovering actions that surface when fabricating data is no longer a viable option. With the strategic responses, the accuracy of information disclosed to the public will be undermined. Consequently, the avoidance behavior taken may be sub-optimal.

Third, our paper adds to the growing literature on the political economy of environmental regulation in the framework of the principal-agent relationship. Studies have found that firms and local governments respond to stringent regulations in ways that result in unintended consequences such as pollution spillover (Kahn 2004; Kahn and Mansur 2013; Kahn, Li and Zhao 2015; Chen et al. 2018; Karplus, Zhang and Almond 2018). In particular, a recent study by He, Wang and Zhang (2020) examines how imperfect performance monitoring of water pollution led Chinese local officials to enforce regulations on polluters immediately upstream of monitoring stations. In our paper, we document another unintended deviation of national regulations at the local level and offer insights into the underlying political incentives.

Fourth, this paper relates to a recent literature with applications of remote sensing data in studying environmental regulation (Sullivan and Krupnick 2018; Fowlie, Rubin and Walker 2019), climate change, wildfire surveillance (Ruminski et al. 2007), forest land cover (Hansen et al. 2013) and biodiversity (Turner et al. 2015). In particular, we unearth novel satellite measures of thermal anomalies, which provide a new source of high-resolution, high frequency measurement of local industrial activities. It allows us to dig deeper into the mechanisms of strategic cleaning.

Lastly, this paper provides strong evidence on uneven pollution control, adding to the economic literature on environmental justice (Bento, Freedman and Lang 2015; Banzhaf, Ma and Timmins 2019; Grainger and Schreiber 2019; Currie, Voorheis and Walker 2020). Undertaking air pollution mitigation in ways that only target monitored areas raises the prospect that residents living far away from the monitoring stations may not benefit after all. They could even be harmed when 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 and estimation results. Section 5 elucidates the underlying mechanisms and explores the role of regional heterogeneity. Section 6 turns to policy implications. Section 7 concludes.

2 Institutional Background

The benefits of China's unprecedented economic growth in the past few decades are built upon huge environmental costs due to heavy reliance on industrialization and fossil fuels, combined with lax environmental regulation. Major Chinese cities rank among the most polluted in the world.[5] Air pollution has adversely affected human health, and public environmental awareness has noticeably increased as a result.

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

Facing mounting pressure at home and abroad, the Chinese government began to shift its policy priorities from economic growth to environmental protection. To do so, it has introduced more stringent regulations to tackle air pollution. This section presents a brief overview of the Chinese institutional background, including its political system and air pollution regulations, highlighting the role played by local officials' career incentives in shaping their environmental protection efforts.

2.1 Political Hierarchy and Incentive Structure

China is characterized by a regional decentralized authoritarian regime (Xu 2011). Within this gigantic bureaucracy, orders are issued, implemented, and monitored in a top-down manner associated with a strict performance-based reward-and-punishment scheme, in which the higher-level principal sets performance targets for lower-level agents. Motivated by strong career concerns, the local bureaucrats devote more resources and efforts to those criteria that are highly valued by the upper-level government. In the past, city leaders competed fiercely with each other for promotions by increasing economic growth in their jurisdictions (Li and Zhou 2005). A number of unintended actions and consequences followed, including inequality, collusion, corruption and manipulation, deviating from the central government's stated policy goals (Fisman and Wang 2015; Oliva 2015; Jia and Nie 2017; Jia 2017) .

The Target Responsibility System launched in the eleventh Five-Year-Plans (FYPs) (2006–2010) marked an important transformation in China's development priorities, in which the central government has begun to put more emphasis on "green" development. Beyond economic targets, more environmental goals have been included in local officials' performance metrics.[6] In the reformed incentive system, for local leaders, failing to attain related environmental targets would cancel out all other work performance, however successful, in the comprehensive evaluation at the end of the year. They would not be eligible for any annual bonuses or career advancement.

2.2 Air Pollution Regulations

To mitigate worsening air pollution, the central government of China declared a "war on air pollution," and implemented a series of regulatory policies, which were claimed to be the "strictest ever".

[6] 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.

Air Ten Action Plan As part of the regulatory change, the "Air Ten" action plan, also known as the "Air Pollution Prevention and Control Action Plan", was announced in 2013. It specified concrete pollution control requirements pertaining to targets, standards, measures and technologies for inclusion in the 12$^{\text{th}}$ FYPs. In particular, a pollution reduction target has been set explicitly as one of the hard targets in governmental officials' evaluation and promotion. Under the plan, the Ministry of Ecology and Environment (MEE), with support from the Development and Reform Commission, the Economic and Information Technology Commission, and other authorities, conducted an annual evaluation of local officials' performance in pollution control over the five-year period. The central government conducted a final assessment of overall performance in 2017.

National Ambient Air Quality Standards (NAAQS) The earlier version of air quality standards was introduced at the end of the last century, predominantly to tackle coal-smoke pollution. However, it failed to keep pace with the evolution of air pollution concerns in China. By design, the determinants of pollutants under the old standard did not include $PM_{2.5}$, and, therefore, the measures were insufficient to assess and manage the emerging challenge of multi-source air pollution.[7] Against this backdrop, 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 air quality and key pollutants were closely tracked by the nationwide air quality monitoring network.

Monitoring Systems for Ambient Pollutants Accurate and comprehensive monitoring of compliance is a crucial ingredient of air pollution regulation. Yet, the quality of China's official air quality data prior to 2012 has been heavily criticized. Monitoring stations operated in only 113 environmental priority cities, hindering the implementation of air pollution control policy in other regions. While the central-level MEE was responsible for setting out technical specifications and for publishing the monitoring data, the local environmental authorities were tasked with managing and operating monitors, and with collecting and

[7] Coal smoke was previously the main source of air pollution in China. The principal pollutants from coal smoke are SO_2, NO_2 and PM_{10}. Current principal pollutants include those pollutants, but also include O_3, $PM_{2.5}$, CO and VOCs. At present, air pollutants come from multiple sources, including road vehicles, vessels, and coal plants. Details about GB 3095-1996 are available in the *National Ambient Air Quality Standards (Consultation Paper)* (Ministry of Environment and Ecology of China 2010), which describes why a revision of NAAQS was needed and seeks public comments on the new standards.

managing the environmental surveillance data submitted to the MEE. This procedure created non-trivial room for local data manipulation (Ghanem and Zhang 2014), undermining the very foundations of effective regulation. Typical practices included shutting down the monitors during heavily polluted days, deliberately blocking the sensors, and sabotaging the monitors.

In order to tackle the problem, China expanded the national air quality monitoring network to 335 cities. More importantly, it launched a real-time reporting system and disclosure program. Several major improvements have been made to this new monitoring program. First, $PM_{2.5}$ is listed as a major pollutant. Second, after the monitoring stations are upgraded or built, raw monitoring data from the monitors is directly transmitted to the central system. Lastly, the responsibilities of various environmental authorities have undergone corresponding changes, as illustrated in Figure A.1.

Under the reformed system, the MEE is in charge of the design and placement of national environmental monitoring stations. To address the possibility of selective siting, locations for the monitors to be rolled out from 2012 to 2014 were centrally determined in 2012.[8] Only air quality data obtained from the central monitors are counted towards the evaluation of cities' environmental performance.

The 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 checking the accuracy of the data collected. The 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.

The automation program was rolled out across the country in three waves between 2012 and 2014. Large cities and more developed areas such as the Beijing-Tianjin-Hebei region, the Yangtze River Delta region, the Pearl River Delta region and a few provincial capitals

[8] There are three types of monitors in China: monitors controlled by the central government, monitors controlled by local governments, and micro monitors for specific polluting sources. The centrally-controlled monitors are the first group of monitors, set up before air pollution became a serious concern. As of 2016, there were more than 2000 monitors in China, including both central and local monitors.

were among the first wave. In these cities, the majority of the monitors were built and operated long before the new real-time monitoring system was established. Upon joining the new system, these cities upgraded their existing monitors; they also established some new monitoring stations, all of which were automated. Specifically, 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. Figure 1 displays the three waves of the automation program in China. The national monitoring network with 1497 central monitors is designed to serve the urban areas of 367 cities.The number of the monitors designated is based on a city's population density and pollution level in the past three years.

Local governments were not involved in the location choices of central monitors, and the introduction of automatic air pollution monitoring precludes the falsifying of data. Consequently, data accuracy has significantly improved (Niu et al. 2020; Greenstone et al. forthcoming). Ideally, as long as these monitors accurately represent local air quality conditions ex-ante, the monitoring network should be efficient. However, the popular Chinese phrase "when the central government has a policy, the local governments have countermeasures" highlights the difficulty in solving the agency problem. With a complex and dynamic principal-agent relationship that underpins the execution of environmental protection, the questions are whether these stringent regulatory measures prevent local officials from taking any hidden actions and whether they incentivize meaningful changes to limit pollution. We aim to examine these important questions.

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, 2) comprehensive monitoring station data, which offer key information on a station's automation status, and 3) city socio-economic characteristics and personnel information on local officials. Here, we present these data sources in detail and describe the key variables and definitions used in the article. Table 1 reports the summary statistics for these variables.

3.1 Remote Sensing Data

Air Quality In order to examine the spatial difference in local governments' response to air pollution regulations, this paper fills the gap in the ground monitoring system by using high-resolution images of a major air pollutant, $PM_{2.5}$, which are derived from the original satellite measures of Aerosol Optical Depth (AOD). The satellite AOD data were generated based on 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}. Overall, the atmospheric science literature has documented a strong correlation between remotely sensed data and ground-level pollution data.[9] However, clouds or shadows over certain areas could contaminate satellite images, in turn causing the problem of missing data. This is more notable for data with high spatial and temporal resolution. For instance, studies concerning remote sensing techniques have detected a weaker association between AOD and ground-level PM measures when utilizing daily counts than when employing monthly or yearly coarse spatial resolution satellite observations (Hoff and Christopher 2009). With this caveat in mind, this paper uses the satellite images that include annual $PM_{2.5}$ grids at 1km by 1km resolution of nine million grids covering the whole of China from 2008 to 2017.

An important issue pertaining to satellite-derived pollution data is that calibration with ground-based monitoring data could cause measurement error that correlates with the locations of ground monitors. To address such potential threats, we begin by assessing the quality of the data. The satellite-derived $PM_{2.5}$ concentration data come from a variety of sources (Van Donkelaar et al. 2015, 2016; van Donkelaar et al. 2019), where these studies have conducted rigorous exercises to validate the satellite-derived measurement of pollution.[10] Second, we have also used raw daily AOD data obtained from NASA's MODIS

[9] See Liu et al. (2007), Lee et al. (2012) and Zhang and Li (2015) for more details. Previous economic research using the satellite measure as the proxy for ground-level pollution includes Foster, Gutierrez and Kumar (2009), Chen et al. (2013), Sullivan and Krupnick (2018), Fowlie, Rubin and Walker (2019), and Bombardini and Li (2020).

[10] By combining satellite-based measures of AOD with chemical-transport modeling and land characteristics, that study used a geographical weighting method to derive ground-level concentrations of $PM_{2.5}$ at high levels of spatial disaggregation. They also conducted cross-validation tests by removing part or all of the ground monitors from the calibration to tackle the concern that varying degrees of measurement errors may

system to verify the robustness of our analysis to potential measurement errors arising from calibration, thus additionally offering external validation of the data used in the current study.

Thermal Anomalies To develop a measure of industrial activities at a fine spatial scale, we leverage a novel dataset on satellite-based thermal anomalies. Various industrial activities, such as power generation and cement production, are 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.[11]

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.[12] 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 2 presents the high correlations between detected fire spots and the geolocations of major air polluting firms and power plants. We explore the validity of the thermal anomaly measure in measuring local pollution in greater detail later in the paper.

3.2 Ground Monitoring Network

Monitoring Station Data The MEE has been publishing $PM_{2.5}$ data since 2012. Our access to this data source has been granted through the Hong Kong University of Science and Technology Atmospheric & Environmental Database, which covered 1,497 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.

Several features of the monitor siting decision are noteworthy. Almost all monitors are

occur at cells with different distances to the monitors. Reassuringly, the derived $PM_{2.5}$ data still performed well.

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

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

located in urban areas. These monitors (centrally-controlled) are sparsely deployed and serve as the basis for the upper-level government to evaluate the air quality of a city. With the spatial gaps in the ground monitoring network, the regulatory focus is likely to bias towards urban citizens who reside in these monitored areas. On the other hand, the pollution level of less-monitored areas is not included in the assessment of local officials' environmental performance. By offering more comprehensive coverage of air pollution in the entire city at much higher spatial resolutions, the satellite remote sensing can show pollution levels in these less-monitored areas.

To examine the difference between the monitor-based and satellite-based city average $PM_{2.5}$, we proceed in two steps. First, we use the 3km by 3km gridded population count from the 2015 National Population Census as the weight for each cell and calculate the weighted average $PM_{2.5}$ for each city. Taking this estimate 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 3 displays the representation errors of various monitors in the years when their hosting cities joined the automatic pollution monitoring program. We regard the cities with errors within $\pm 15\%$ as siting well-representative monitors. As shown by the red shaded areas, these are cities where monitors over-represent the "true" city-level pollution $PM_{2.5}$. The blue shaded areas denote cities with under-representative monitors. The representation errors exhibit large spatial variations, where two-thirds of cities have over-representative monitors.

By definition, the representation errors in Figure 3 are static at the moment of the three roll-out waves (including both automated and newly opened monitors). Assuming that the pollution reduction patterns are even across space, a one-period analysis would simply conclude that the representativeness of monitors tends to persist as long as the monitors' locations do not change. However, in a multiple-period setting, local regulators' dynamic choices over how to respond to the automated system would potentially shape the monitor's spatial representativeness over time. By plotting the representation error of each year (2012-2017) on the digital maps (results available on request), we find that this is indeed the case. Monitors' spatial representativeness exhibits dynamic changes after various cities joined the automated program, motivating us to study of local regulators' strategic pollution reduction behaviors.

3.3 Socioeconomic and Personnel Data

To address the concern that other factors could affect the spatial representativeness of ground monitors and possibly confound our identification of strategic clean-up behavior, we construct a rich set of control variables. First, we assemble city-level demographic and socioeconomic data, including GDP and population statistics. Second, we extract personnel information from the China Political Elite data, which provides local officials' career path, age and education.

4 Strategic Response to the Automatic Pollution Monitoring

In this section, we examine how local governments respond to the use of the advanced monitoring technology. As discussed, the real-time data reporting system leaves the local governments with virtually no ability to directly tamper with data. For these local officials, a cost-effective way to improve the reading of pollution monitors may be to target the monitoring sites for very localized clean-up. This could become an unintended and undesirable consequence of the upgraded regulatory system.

4.1 Empirical Framework

Using a distance-based difference-in-differences (DiD) method with variable treatment intensity, we test whether the pollution reduction between monitored areas and the un-monitored surrounding areas differs after the rollout of the automatic pollution monitoring program in three staggered waves. The estimation specification is the following:

$$\ln(PM_{2.5icwt}) = \alpha Auto_{wt} + \beta Near_i \times Auto_{wt} + \gamma X_{ct} + Cell_i + Year_t + \varepsilon_{it} \quad (1)$$

where i, c, w and t denote grid cell, city, rollout wave and year, respectively. The outcome variable, $\ln(PM_{2.5icwt})$, is the logarithm of the satellite-based $PM_{2.5}$ concentration. We aggregate nine million 1km×1km cells to one million 3km×3km grids prior to estimation to reduce computational demands. $Auto_{wt}$ is the treatment indicator, which takes the value of 1 after a city joined the automatic monitoring program in wave w, and zero otherwise.[13] The

[13]Here we use *Auto* to represent the automatic monitoring program, which includes both new monitors and existing monitors upgraded to automation.

treatment intensity is defined by $Near_i$, which equals 1 if the grid cell i is in an area adjacent to a ground monitor (or equivalently, a monitored area), and zero if it is located far away from the monitor (or equivalently, an unmonitored area). We add a rich array of covariates to the specification, including cell fixed effects and year fixed effects, to control for time-invariant spatial characteristics and the macro shocks common to all cell units. Since air pollution observed in a cell and the nearby cells are likely influenced by emissions elsewhere, we cluster standard errors at the city level to safeguard against spatial correlation.

Throughout the paper, the parameter of primary interest is the coefficient (β), which identifies the difference in the causal effects of the technology-aided monitoring program on monitored and unmonitored areas. It represents the change in the pollution gap between areas near air monitors (or new monitors) and those farther away, in response to the technical upgrade (or new installment) of these monitors. The coefficient α represents the impact of automation on the overall pollution for the whole city. Although this was not the central focus of our paper, such aggregate-level responses could be important when considering policy counterfactuals and mechanisms.

As with any DiD design, the identification of α relies on the parallel trend assumption. We assess the validity of this assumption by examining the pretrends of treatment and comparison cities. We exploit the fact that the staggered sequence of the automation program is largely dependent on the priority of cities; specifically, larger cities and those with higher-level administrative hierarchy tended to receive the treatment earlier. In order to correct for this potential source of endogeneity bias, we include in X_{ct} interactions between a fourth order polynomial in time with city determinants of automation priority, such as baseline GDP, population, pollution level and indicators for environmental priority cities (which include provincial capitals). In doing so, we allow the temporal variation in pollution to flexibly differ across factors most relevant for the timing of automation. Figure 4 shows the results of the event study on city-level pollution. There were no clear pre-existing trends in the years prior to automation, lending support to our identification assumption.

The identification of β further uses variation across subareas with varying propensities of being affected by the automatic monitoring system. To fully exploit the triple difference design, we next use a most demanding specification, which includes wave by year fixed effects to control for different yearly variations in outcomes across cities automated in three waves.

We estimate regressions of the following form:

$$\ln(PM_{2.5iwt}) = \beta Near_i \times Auto_{wt} + Cell_i + Wave_w \times Year_t \ + \varepsilon_{it} \tag{2}$$

where the key explanatory variable is denoted by the interaction term. With this augmented specification, the baseline impact of automated monitoring on air pollution in the preceding Equation (1), which is denoted by the coefficient α, is absorbed. In comparison, Equation (2) isolates a more credible identifying variation by comparing monitored and unmonitored cells amongst cities of the same wave, prior to and following the operation of the automatic monitoring system.

The identification assumption for β is that the monitored-unmonitored pollution gap would have moved in parallel across cities and over time, had the automation not been implemented. Potential violations of the assumption would require confounders not only to be specific to the automated cities at the time of automation, but also to be associated with the monitored cells only within the treated cities. Given the centrally designated rollout sequence and pre-determined monitor locations, the probability that confounding factors can simultaneously meet all these conditions is very small.

We nevertheless formally conduct a parallel trends test of the triple difference setup. The cell-level event study is implemented on a sample covering four years prior to and three years following the automation of the monitoring system. The specification is as follows:

$$\ln(PM_{2.5iwt}) = \sum_{n=-4}^{3} \beta_n \phi(n)_{it} + Cell_i + Wave_w \times Year_t \ + \varepsilon_{it} \tag{3}$$

where n defines the period relative to the automation year, $n = -4, -3, -2, 0, 1, 2, 3$. The dummy variables $\phi(n)_{it} = \mathbf{1}\left[n \leq t \leq n+1\right]$ jointly represent the automation event. We omit the dummy variable for $n = -1$, the year right before automation. The control variables are as previously defined.

Figure 5 plots the estimated coefficients β_n for the event study analysis. $PM_{2.5}$ air pollution concentrations for the monitored and unmonitored areas were trending similarly prior to automation. The coefficients associated with event times $n = -4, -3, -2$ are small and not statistically significant. In contrast, we find a large and significant $PM_{2.5}$ decrease in the monitored areas relative to unmonitored areas during the three years following the automation. Taken together, the event study analysis produces strong evidence on the causal impact of the policy change on varying pollution outcomes in different areas, lending further support to the validity of our empirical approach.[14]

4.2 Baseline Results

For the baseline analysis, we use a sample from 2008 to 2017. Table 2 reports the estimation results by adding controls sequentially to the regression equation. In the first column, the coefficient of *Auto* captures the average impact of joining the automatic monitoring program on a city's air pollution, compared to those cities whose monitors have not yet been opened or automated.[15] The treatment intensity indicator $Near_i$ is denoted by a dummy variable **1**(0-3km), assigning grid cells within the 3km radius of a monitor as monitored (the indicator thus equals one) and others as unmonitored (the indicator equals zero).

In column (2), the coefficient on the key independent variable in our analyses–the interaction term of **1**(0-3km) and *Auto*–captures the differential effects of the automatic monitoring program on air pollution in different areas, which is estimated from Equation (1).[16] Column (3) further controls for the wave-specific year fixed effects and estimates β in the spirit of Equation (2). This is our preferred specification. The estimates imply that air pollution in the monitored areas is 6.1% lower than in the unmonitored areas after the automation roll-out. When using cells outside a 30km and 60km radius of a monitor as control groups, columns (4) and (5) show that the coefficients remain similar in size and statistically significant, which confirm the robustness of our main results to alternative definitions of unmonitored areas. The coefficient on *Auto*, which captures the impact of monitor automation on pollution, is positive and marginally significant, showing that strategic pollution reduction may lead to pollution leakages to unmonitored areas. We will further explore the spatial pattern of pollution relocation in a ring analysis in section 4.4.

[14] We also demonstrate in Figure A.4 that our event study results are robust to controlling for existing pre-trends, building on recent methods by Rambachan and Roth (2019).

[15] Goodman-Bacon (2021) 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). In our paper, the potential impact of wave-specific factors affecting 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 the 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 the cell fixed effect is included, the results show that areas near monitors experience larger pollution reductions after the automatic monitoring program.

4.3 Robustness Checks

In the previous subsections, we have demonstrated that the introduction of a technology-aided upgrade in pollution monitoring widens the gap in satellite-sensed pollution between monitored and unmonitored areas, which indicates the existence of strategic behavior in relocating pollution away from in-situ monitors.

In this subsection, we summarize the additional tests we conducted to address other identification concerns, which include (1) falsification tests with randomized automation timing and location, (2) tests of alternative explanations, (3) re-weighting of monitored and unmonitored cells with entropy balancing, and (4) replacing the outcomes with raw AOD data.

Randomly Generated Automation Timing and Monitor Locations First, we use only the pre-program periods and randomly assign automation years for monitors located in various cities. The underlying premise is that pollution levels of cells in "monitored" and "unmonitored" areas should not differ significantly over falsely generated automation timing. To do so, for each monitor, we randomly assign the timing of automation within the sample period from 2008 to 2017. Using this false *Auto* variable, a placebo DiD estimation is conducted using the specification in Equation (2). To increase the identification power of this placebo test, it is repeated 1000 times. Figure 6 (a) plots the distribution of the estimates from the 1000 runs along with the benchmark estimate. As shown in the figure, the benchmark coefficient, which centers around -0.061, lies outside the 99% confidence interval of those estimated from the placebo test. This indicates that, before the monitoring program was implemented, a counterfactual rollout schedule of the automatic monitoring program did not lead to larger pollution reductions in monitored areas than unmonitored areas.

Similarly, we conduct another placebo test by randomly assigning monitor sites to various locations. Keeping the number of monitors and the year of automation unchanged, areas are selected at random. Again, this is repeated 1000 times. By matching the groups of placebo monitors with the gridded satellite data, we re-estimate equation (2) and plot the distribution of the coefficients in Figure 6 (b). The comparison between placebo estimates and the benchmark estimate suggests that pollution reductions occurred only at the observed monitored areas, not at the neighborhood of the counterfactual sites. Overall, these findings suggest that local governments might have engaged in strategic clean-up as a response to the automatic monitoring.

Alternative Explanations Our causal interpretation of the estimated coefficient on the interaction term $Near_i \times Auto_{wt}$ as strategic clean-up hinges on the assumption that, had there been no automation, the gaps in pollution level between monitored and unmonitored cells should have been largely similar across early and later automated cities. In this section, we evaluate a number of possible explanations which might undermine this causal interpretation of our main findings. For example, a plausible alternative scenario is that monitors are often sited in urban centers, which are also likely to be heavily polluted. The dynamics of local air pollution could display a certain degree of mean reversion, in part because pollutants could disperse away through the air from the source to other places in the city. If this is indeed the case, part of the differences in pollution reduction between the monitored and unmonitored areas could be explained by some autoregressive process interacted with selective monitor siting. Consequently, one should expect to witness a larger reduction in pollution following the automation for monitored cells located in dirtier areas than those located in cleaner areas of a city.

A similar concern relates to the interpretation of local officials' strategic behaviors. One could argue that local officials who choose to reduce more pollution in monitored areas are not necessarily gaming the system. Instead, they are adopting a more cost-effective way to mitigate pollution in more polluted areas, which coincide with urban localities adjacent to a monitor. To check against these alternative views, we provide a version of the regression that accounts for the pre-existing pollution condition in various areas:

$$\ln(PM2.5_{iwt}) = \beta Near_i \times Auto_{wt} + \eta Near_i \times Auto_{wt} \times Dirtier_i + Cell_i + Wave_w \times Year_t + \varepsilon_{it} \tag{4}$$

where $Dirtier_i$ is an indicator which equals one if cell i is more polluted than the city average in the base year 2008. This measure of initial conditions controls for either a pre-existing trend or an autoregressive process. The results are reported in Table 3. Columns (1) to (2) address the possible role of initial condition heterogeneity in accounting for our results. Furthermore, column (3) controls flexibly for any difference in the time trends of outcome across locations depending on their pre-existing pollution. The estimate on the main parameter of interest remains relatively stable throughout the specifications. This helps rule out alternative explanations, as illustrated above.

Entropy Balancing between Monitored and Unmonitored Cells As a more direct way to address the concern that the selective siting of monitors might lead to divergent trends in pollution, we apply the entropy balancing technique following Hainmueller (2012) and Hainmueller and Xu (2013) to further ensure the comparability of monitored and unmonitored cells. It is a generalized weighting approach which allows us to select a group of control (unmonitored) gridcells similar to the monitored cells with respect to a battery of cell-level confounding factors. The technique is similar to propensity matching, but it has an attractive feature that dissimilar observations are not dropped from the analysis but are assigned smaller weights instead. This flexibility allows us to achieve a higher level of balance between the treated and control groups, ensuring not just the balance in means but also in higher-order moments. For our analysis, we include proximity to polluters, cell-level population, and GDP in 2008 as matching variables and ensure the balance in the mean and variance of these variables between monitored and unmonitored cells within each city.

Column (4) of Table 3 reports the results. The point estimates are smaller than those reported in the baseline, but the signs and significance are consistent. After ensuring strict balance between monitored and unmonitored cells, automation leads to a 2.9% increase in their pollution gap.

Measurement Errors In recent research that utilized satellite-derived pollution indicators, measurement errors have been highlighted as a problem of significant concern. The $PM_{2.5}$ data employed in the current study were derived from the raw satellite images, and the calibration procedure also required information from land-based monitoring stations. Specifically, the Geographical Weighted Regression method assigns larger weights to areas closer to ground monitors, and smaller weights to farther areas. One might worry that the resulting measurement errors are correlated with the distance to monitors and could also be systematically linked to the establishment of new ground monitors. Beyond the validation evidence in section 3.1, we conduct our own analysis with raw AOD measurements as an alternative outcome indicator. The pertinent results are reported in Table 4.[17] Reassuringly, they are largely consistent with the baseline.

[17] We also conduct an event study by using AOD data, as shown in Figure A.2.

4.4 Ring analysis

In order to establish the key point that the spatial gaps in pollution changes are attributable to strategic responses of local governments to more advanced central monitoring technology, we now present the geographic patterns of the automation program impact. To do so, we deploy a concentric ring approach under an implicit assumption that targeted pollution control measures should decay with the geographic distance from the monitor. Specifically, we replace the binary indicator of monitored areas in Eq (1) with a set of bin indicators, which denote cells within a 3km radius of a monitor, then those within 6 km, continuing step-wise out to 150 km.

The results are reported in Figure 7. Note that the reference category includes grid cells more than 150km away from the closest monitor and experienced a 10.4% increase in the pollution level, as denoted by the estimated coefficient of β. The coefficient estimates for a set of bin indicators represent the impact of automatic monitoring on air pollution in each distance bin relative to the reference group. Based on these estimates, the strategic pollution reductions were detected within a 30km range of monitors. Beyond that distance, the overall impact of automatic monitoring on pollution turns positive and continues to rise for cells farther away, implying potential pollution migration across space in light of the automation.

5 Mechanisms and Heterogeneity in Strategic Responses

So far, we have established a number of new empirical results. First, improved monitoring technology leads to significant reductions in air pollution within a 3 kilometer radius around the monitor. Second, the impact of monitor automation on local pollution is reduced with distance from the automated monitor, and pollution increases beyond 30km, suggesting possible relocation of polluting activities away from the monitors.

In this section, we analyze the potential channels through which differential spatial patterns of air pollution may have occurred. We offer compelling evidence that targeted measures have been taken by local regulators to reduce air pollution close to the monitors. Moreover, to emphasize the importance of political economy factors in driving the results, we explore multiple sources of heterogeneity, including a city's pre-automation compliance level, leader characteristics, and public pressure.

5.1 Anecdotal Evidence on Strategic Reduction

As noted in section 2, manipulation of environmental data was previously common among local governments. However, the new monitoring system has made it virtually impossible to directly tamper with air quality monitoring equipment or to falsify data. With the real-time data collection and release of information to the public, any attempt to manipulate the readings would result in abnormal data patterns and alerts would be triggered. However, such sophisticated monitoring technology cannot eliminate all possibilities of manipulation. To learn about the targeted clean-up measures taken by the local authorities, we extensively reviewed a large number of government policy documents, work reports prepared by discipline inspection teams, and newspaper articles. The policy documents reveal that local governments have strong political incentives to exert more effort toward improving air quality readings in order to meet the centrally designated targets. Below we summarize their strategies into three types.

First, local governments could directly engage in cleaning up air pollution near the ground monitors. Since the monitor locations are fully known to the local governments, many of them spray water in the adjacent areas or target fog cannons at monitors (a high-risk yet effective approach) or toward other subjects near the monitors (associated with a lower risk, albeit less effective). As a case in point, a scandal recently publicized by the press revealed that, in January 2018, the Environmental Protection Agency's office building in Shizhuishan City of Ningxia Province, where a central monitor is located, was turned into an ice sculpture when targeted by fog cannons.

The next set of strategies involves various traffic restriction policies which focus on the monitored areas. An official report of Tianjin's environmental inspection team documented the strategic use of temporary traffic control plans by the local agency.[18] Media also reported incidents of temporarily shutting down the gas stations near the monitors in Pingdingshan City, again a carefully calibrated approach taken by the local government to improve air quality in the immediate area around the monitors.[19]

[18]See "The Central Environmental Protection Supervision Team: A short-cut plan to guarantee good air quality is set up around the Tianjin Monitoring Station" (dated on July 29th, 2017) for an example. Source: http://news.sina.com.cn/o/2017-07-29/doc-ifyinwmp0696032.shtml.

[19]See news and media coverages of the existing manipulation strategies Yuqing(People.cn), Bloomberg Law, Guancha and Changcheng(People.cn) for more details.

A long-run strategy would be relocating major polluting sources ranging from barbecue restaurants to large industrial plants to unmonitored suburban areas.[20] We find empirical support for this notion based on the ring analyses in Figure 7, where the impact of monitor automation on surrounding areas' pollution turns positive beyond 30 km.

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 suspending the operation of coal-fired power plants and other heavily polluting activities. In this section, we test this channel empirically. To obtain a measure of the *intensity* of industrial activities at a high degree of spatial and temporal resolution, we exploit the crucial fact that most industrial plants emit high-temperature waste gas, which turns the air above them into thermal hotspots. To identify these hotspots, we draw upon the MODIS Version 6 Global Monthly Fire Location Product, which offers the count and strength of thermal anomalies on static land.[21] We utilize this novel data set to measure the number of operating plants and the intensity of their operation.

5.2.1 Validation of Thermal Anomalies Data

Before proceeding to the empirical specification, we conduct a battery of external validation exercises. We start by assessing the correlation between thermal anomalies and polluting activities. To do so, we obtain two lists of major polluting plants: the first is 1829 heavily polluting industrial firms, drawn from the MEE's Key Centrally Monitored Polluting Enterprises database. These firms have been placed under more intensive monitoring by the central authority. The other is composed of 10491 power plants, sourced from the China Emissions Accounts for Power Plants (CEAP) database. Figure 2 maps the locations of thermal anomalies along with the geographic distribution of those polluting plants. It is clear from the figure that key centrally monitored industrial firms (Panel A) and power plants (Panel B) are always located in spots with observed thermal anomalies, albeit more spatially dispersed. On that basis, we argue that thermal anomalies provide rather comprehensive coverage of major polluting sources. Table A.3 shows that, for each 10km-by-10km cell, the

[20]See "Linfen Data Falsification Case" One Year Later, Part of Shanxi's Environmental Information Still Undisclosed" (dated on Jun13, 2019) for an example. Source: http://sxhjjcw.com/index.php?p=news_show&id=468&lanmu=4.

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

presence of any thermal anomalies increases the probability of the existence of a polluting plant by 99%. Beyond the extensive margins described above, Table A.4 further compares the pollution intensity of industrial plants with the thermal anomaly intensity measured at the plant site. At the intensive margins, Column (3) shows that a one percent increase in the radiant heat output around each power plant (capturing the rate at which fuel is consumed and smoke emissions released) is associated with a 0.14 percent growth in the satellite-derived $PM_{2.5}$ measures from the plant, confirming the quality of these thermal anomaly data.

We then examine whether short-run variations in thermal anomalies respond to drastic government measures such as temporarily closing some industrial plants. As a case in point, a series of strict emission control policies were adopted in Beijing and surrounding regions to ensure blue skies during the 2014 Asia-Pacific Economic Cooperation (APEC) summit. Figure A.3 presents the time series of two measures of thermal anomalies one month prior to and one month after the summit for the affected region of Beijing, Tianjin and Hebei. Both indicators dropped sharply preceding the event, and picked up immediately after the summit ended. The observed pattern synchronizes with the temporary suspension and resumption of industrial activities around Beijing, again highlighting the validity of the thermal anomaly measure in measuring temporal variation of local pollution.

5.2.2 Changes in Industrial Activities: Results

In a similar vein to the baseline model, the effect of automation on industrial activities is evaluated using the following specification:

$$y_{it} = \alpha Auto_{wt} + \beta Near_i \times Auto_{wt} + \gamma X_{ct} + Cell_i + Year_t + \varepsilon_{it} \tag{5}$$

where y_{it} denote measures of thermal anomalies' presence or intensity at cell i in year t. For our analysis, we identify fire spots under the category of "static land source" as active industrial activities. A gridcell might experience multiple thermal anomalies during a year. So, we consider the treatment effects of monitor automation at both the extensive margin and the intensive margin. In particular, the former denotes whether or not thermal anomalies are observed in a certain gridcell, while the latter counts the days with active hotspots and the average intensity throughout a year. In an alternative specification, we also include the interactions of monitor automation with ten distance bins around monitors to check the spatial pattern. Throughout our analysis, we control for cell and year fixed effects.

Many gridcells have never had any thermal anomalies, resulting in a large proportion of observations with zero values. To address this issue, we estimate our baseline model using a Poisson count data model for a set of outcomes. The results are reported in Table 5. Column (1) presents estimates of a logit model where the dependent variable switches to one with thermal anomalies present (TAP) in gridcell i in year t. The marginal effect estimates suggest that automation leads to a 10% reduction in the probability of observing industrial activities within a 3km radius of the monitors, relative to the unmonitored areas. A possible explanation is that activities closer to the monitors were relocated away from the original area after monitor automation. Columns (2) and (3) report results from the Poisson estimation, where the dependent variables are the logarithm of the number of days with active thermal anomalies and the average intensity of thermal anomalies, respectively.[22] The coefficients of interest are on the interaction of $Near_i \times Auto_{wt}$, which are negative and significant for both outcomes. In Columns (4) and (5), we focus on the effects of the automatic monitoring system at the intensive margin, and therefore restrict the sample to only those gridcell-year observations with active thermal anomalies. When the logarithm of the number of days with active thermal anomalies is the dependent variable, the coefficient on the interaction, β, is negative and significant. But the estimated magnitude is small and the coefficient is statistically insignificant with the average intensity of thermal anomalies as the dependent variable. These results suggest that monitor automation is likely to reduce the operating days of plants nearby but not the intensity of operation during these days.

Table 6 further reports the impact of monitor automation on the thermal anomalies (proxy for industrial activities) across different distance bins. Moving away from the monitors, there is a decrease in the estimated magnitudes of the monitors' impacts in reducing pollution. These findings are consistent with the pollution migration pattern documented in Section 4.4. Overall, the evidence indicates the relocation of polluters away from the monitors and a decrease in days of operation for existing polluters surrounding the monitors. However, we do not find evidence of reducing operating intensity for existing plants during their working days. The results on thermal anomalies offer new insight on the mechanisms through which local governments implement strategic clean-up.

[22]FRP stands for Fire Radiative Power. It is defined as the rate of radiant heat output and is related to the rate at which fuel is being consumed and smoke emissions are released.

5.3 Heterogeneity in Strategic Pollution Reductions

a) Underreporting before Automation

To convincingly connect the observed strategic behavior to local governments' responses to the tightening of environmental regulations, we divide monitors into three types: newly installed monitors, upgraded monitors with a high degree of pre-automation data manipulation, and upgraded monitors without prior manipulation. The conjecture is that, for cities that had not previously engaged in data tampering conduct, the introduction of the automatic pollution monitoring system would have very little impact on de facto regulatory stringency. In contrast, automation shut down the channel of "effortless perfection"–namely, meeting environmental standards by misreporting pollution readings–for those cities with detected data manipulation misbehavior in the past. Similarly, the installation of new monitors in previously unmonitored areas is expected to impose extra regulatory pressure on local officials in the newly-monitored areas.

Out of the 335 cities with active ground-based air monitoring stations by the end of the sample period (2017), only 113 had installed monitors before 2013. For this sub-sample, we build upon the work of Greenstone et al. (forthcoming) to classify whether or not their environmental data had previously been manipulated. That measure was obtained from a regression discountinuity design that captures the sharp increase in *reported* pollutants immediately after the monitoring system was automated.[23]

Figure 8 presents the heterogeneity in strategic clean-up across city groups. The first group comprises cities which were included recently in the monitoring network (denoted by "New Monitor"). The other two groups are those with or without pre-automation data manipulation (denoted by "Upgraded Monitor w manipulation" and "Upgraded Monitor w/o manipulation", respectively). As one can see in Figure 8, the air pollution gap between monitored and non-monitored areas has experienced the largest increase in cities recently incorporated into the monitoring network. For cities with upgraded monitors, we observe such a gap only in those that previously manipulated data, but not in the others. This pattern is consistent with strategic clean-up being a direct response to the improved monitoring technology. Most notably, cities with no track record of data manipulation prior to the

[23] For cities that had monitors before automation, manipulation status is defined by whether the local linear RD estimate is positive, as calculated by using the algorithm in Greenstone et al. (forthcoming).

monitor automation did not face additional regulatory burdens following the occurrence of automation. As a result, there was very little incentive for them to undertake strategic clean-up actions.

b) Political Incentives of Leaders

To delve deeper into the political incentives behind the documented strategic pollution reduction, we explore regional heterogeneity in the career concerns of government officials. As discussed in Section 2.1, the Chinese central government sets targets, and local officials' promotion is linked to their achievement of the targets. In addition, city leaders, including party secretaries and mayors, routinely experience political turnover, which arguably is institutionalized. Their incentives to reach the environmental targets critically rely on other determinants of their promotion prospects, such as political loyalty, educational qualification, and age. For example, party secretaries and mayors of the prefecture-level cities have greater chances of political advancement at an age younger than 57. Thus, in the current setting, city leaders below that critical age threshold would have stronger incentives to deliver lower pollution readings and devote strategic pollution reduction efforts towards monitored areas. Figure 9 shows the heterogeneity across cities with differential leader age in their responses to automatic monitoring. Specifically, we divide the sample into two sub-samples according to whether the party secretary is over 57 years old, as shown in the upper panel. The results in the bottom panel correspond to the effects on mayors. Several key insights emerge from the figure. Leaders' age is a critical factor at play here. Strategic clean-up behavior is exclusively observed in cities headed by younger mayor or party secretaries, highlighting the role of career concerns in shaping their strategic reactions to environmental regulation. Furthermore, the observed patterns do not appear to differ by the concrete position of city leaders, indicating that the underlying incentives and capacities of both party secretaries and mayors could drive them to pursue strategic clean-up.

c) Public Pressure

Environmental regulation in China largely depends on top-down supervision and executive orders. Such approaches are subject to implementation gaps and fraudulent reporting. As shown above, although the improved monitoring technology alleviates the principal-agent problem inherent in a top-down regulatory system, it leads to new strategic reactions from the local agents. Some argue that civic engagement and bottom-up supervision could be complementary to the current regulatory regime. In this part, we seek to relate regional heterogeneity in the strategic clean-up responses to local public pressure, as measured by the number of complaints from local residents.

The environmental authorities in China are paying more attention to public complaints. The 2003 Environmental Impact Assessment Law and the 2004 Administrative Licensing Law included concepts of disclosure and public participation (Wang 2017). In 2009, the MEE officially launched the 12369 hotline for whistleblowing related to environmental issues. Subsequently, online and social media platforms have been established with similar functions. As of 2017, more than 0.6 million complaints had been lodged through various channels described above. The MEE requires each reported case to be resolved within 60 days after the complaint is filed, and 99% of the documented cases met the deadline. For our analysis, we construct an indicator of local public pressure based on the per capita number of environmental complaints in 2017. This measure could capture both supply and demand side factors that determine the level of public oversight of environmental problems.[24] Figure 10 presents differential responses to monitor automation across provinces with varying levels of environmental complaints. Evidently, a stronger strategic response, represented by wider gaps in the effects of automation between monitored and unmonitored areas, coincides with provinces that had low levels of civic engagement in environmental issues. In provinces with more "participatory" environmental monitoring, industrial pollution emitted by plants in both monitored and unmonitored areas is likely to arouse public concern and therefore put pressure on local authorities to respond quickly. These results offer some suggestive evidence concerning the complementarity between top-down and bottom-up approaches in environmental monitoring.

6 Policy Implications

6.1 Dynamic Monitors Representativeness

According to the map of representativeness errors in Figure 3, the monitoring system exhibited good representativeness in most Chinese cities at the start of the program. Although

[24] On the supply side, the functioning and responsiveness of the hotline or online complaint platform plays an important role in influencing citizen engagement. On the demand side, local residents may pay varying degrees of attention to the surrounding pollution, which in turn drives their willingness to lodge a complaint.

monitor locations are unlikely to change once they are placed,[25] and the monitors were well-sited to represent overall air quality in the 2010s, the representativeness is dynamic and evolutionary and can be driven by local interventions in monitored areas. Recall the estimate for the strategic pollution reductions in monitored areas: gridcells within a 3km radius of monitors experience a 6.1% greater reduction in $PM_{2.5}$ concentrations than those farther away. Using this coefficient and the last year of our sample (2017), we calculate the projected pollution levels over the five-year period from 2018 to 2022.[26] As shown in Figure 11, the forecast suggests that, in the near future, the previously over-representative monitors seem to become more representative of a city's air quality. However, by the end of 2022, monitors in approximately 182 cities are predicted to under-represent overall air pollution, indicating the power of local strategic conduct in a dynamic setting. Moreover, the potential pollution migration to unmonitored areas could cause adverse health impacts and biased evaluation of policy goals. Considering the fact that the demographic characteristics of the population in monitored and unmonitored areas are likely to be different, such targeted pollution control measures can also exacerbate geographic environmental inequality, which has received scant attention in China.

6.2 The Remote Sensing Data and Other Pollution Information

A key insight of our study is that the current coverage of the monitoring network in China is not sufficient to preclude local officials' strategic responses. In an ideal scenario with monitoring network coverage almost everywhere, local officials would find it impossible to predict the exact sets of monitors used by the upper-level authority to evaluate their environmental performance. Consequently, only one choice would be left open for them: work harder to improve city-wide air quality. Yet, the real world does not function in this way and governments face budget constraints. Given its large size, the ground monitoring system is costly to build and maintain. Satellite-based pollution measures can be a good data source to fill the gap in the spatial coverage of monitoring networks.[27] In the current context, we

[25] 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.

[26] We did not attempt to perform an extended-period simulation because of large uncertainty and the possibility of new regulations.

[27] As shown in Sullivan and Krupnick (2018) and Fowlie, Rubin and Walker (2019), remote-sensing data helped these authors assess the extent to which the existing U.S. ground monitor-based measurements over- or under-estimate the true exposure to $PM_{2.5}$ pollution.

have used the remote sensing data to re-evaluate the policy goals set by the "Air Ten" action plan by the end of 2017. Unlike the monitor-based pollution estimated in Greenstone et al. (2020), who found that $PM_{2.5}$ decreased by 40% from 2013 to 2018, our study finds a mild increase in city-wide pollution levels. This suggests that monitor-based evaluation might overstate the environmental performance and hence distort future policy design.

However, we caution in taking the utilization of satellite images too far. After all, remote sensing data are not direct measures of ground pollution levels and are subject to missing data problems resulting from cloud coverage. In contrast, ground monitors can provide more detailed hourly observations and better accommodate changing weather conditions. It is also noteworthy that advanced monitoring technologies such as mobile monitors and micro-monitors have enabled broader network coverage over which local regulators can exert little control. The central government could make use of this recent innovation as supplementary information for city-wide pollution evaluation. Overall, the incentive and information issues apply to the design of any monitoring regulation and enforcement. Importantly, contributions from multiple sources, including the ground monitoring system, remote-sensing technologies, mobile monitors, public awareness, and third-party auditors, are needed to achieve a better regulatory outcome.

6.3 Public Participation

China's environmental governance has long been dominated by two major players – the government and firms – while public participation has largely been kept out. Amidst the growing level of environmental awareness and increasing complexity of pollution monitoring, it appears to be the appropriate time for the country's environmental authority to more actively engage local citizens in the management of pollution.

The call for these governance changes has been made possible through information and communication technology (ICT). It is highly cost-effective for the central and local environmental authorities to improve public communication and engagement using e-governance. Some examples include the real-time disclosure of monitoring data to the public and mobile applications that encourage citizens to report pollution incidents.[28] Regulators could in turn gather more accurate information and take corresponding actions. As a case in point, local

[28]Some e-platforms request the public to submit evidence such as geo-coded images of pollution occurring from an industry or business.

governments in Qingdao and Linyi of Shandong Province provided speedy follow-ups to the public claims on their social media (Weibo) accounts.[29] Moreover, an increasing number of provinces are launching their own online platforms for air pollution disclosure, on which detailed data obtained from various monitors are accessible to the public. As a promising tool, environmental information disclosure not only facilitates individual avoidance behavior, but also allows local residents to supervise monitors and check data consistency.

7 Conclusion

In major countries around the world, local governments are responsible for the enforcement of national environmental regulations, typically under incentive contracts that tie rewards to performance. However, inadequate monitoring and misaligned incentives often lead to strategic compliance at the local level. Recent technological developments, such as the increasing use of automated systems, have greatly improved the central regulator's information-monitoring capability and helped to mitigate the long-standing problem of moral hazard.

In the Chinese context of an automated air quality monitoring system, our study has shown that areas near the monitors experienced a larger drop in pollution than the surrounding areas after the monitors were technologically upgraded. Although the data quality of reported pollution has been substantially improved at the monitoring sites, the spatial gaps in monitoring coverage appear to incentivize local officials to take strategic pollution control measures targeting monitored areas rather than engaging in systematic reductions that benefit wider regions. Such unintended responses will change the spatial representativeness of the monitoring system, in that the monitoring stations are becoming less representative of the surrounding environmental quality, leading to biased policy evaluations in the long run. Our results emphasize the central role of accurate and representative measurements in enforcing environmental regulations and can be generalized to in-situ monitoring systems in other contexts.

Our study also reveals that incentive scheme design, of which career concerns are a key component, affects the degree of strategic pollution reductions in the Chinese context. This effect is shown to be stronger for cities with approaching assessment deadlines, those that

[29] See the news report http://www.gov.cn/xinwen/2014-06/20/content_2705303.htm. for details.

had engaged in pre-automation data manipulation, those led by younger mayors, and those with less public pressure.

One of the main contributions of the paper is to uncover more concealed local strategic responses to improved monitoring capacity in the presence of high-powered political incentives. A technology-aided program is not necessarily an easy solution to a moral hazard problem due to a highly complex set of potential responses from the agent (Holmstrom and Milgrom 1991).

In light of the 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. Central regulators should corroborate the existing ambient pollution monitoring system with a combination of information from ground-level monitors, advanced monitoring techniques, and public input, in order to accurately evaluate local officials' environmental performance and improve air quality city-wide. In particular, it may help when pressures occur outside the system. When pollution is publicly known and draws a fair amount of attention, the state must respond effectively in an effort to retain its legitimacy.

References

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

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

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou. 2020. "From Fog to Smog: the Value of Pollution Information." NBER Working Paper 26541.

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

Bombardini, Matilde, and Bingjing Li. 2020. "Trade, pollution and mortality in China." *Journal of International Economics*, 125: 103321.

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

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. 2013. "The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality." *Journal of Environmental Economics and Management*, 66(3): 424–443.

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

Currie, Janet, John Voorheis, and Reed Walker. 2020. "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.*

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

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

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

Fowlie, Meredith, Edward Rubin, and Reed Walker. 2019. "Bringing Satellite-Based Air Quality Estimates Down to Earth." *AEA Papers and Proceedings*, 109: 283–88.

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

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*.

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

Grainger, Corbett, Andrew Schreiber, and Wonjun Chang. 2019. "Do regulators strategically avoid pollution hotspots when siting monitors? Evidence from remote sensing of air pollution." *Unpublished*.

Gray, Wayne B, and Jay P Shimshack. 2020. "The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence." *Review of Environmental Economics and Policy*.

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

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

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political analysis*, 25–46.

Hainmueller, Jens, and Yiqing Xu. 2013. "Ebalance: A Stata package for entropy balancing." *Journal of Statistical Software*, 54(7).

Hansen, Matthew C, Peter V Potapov, Rebecca Moore, Matt Hancher, Svetlana A Turubanova, Alexandra Tyukavina, David Thau, SV Stehman, Scott J Goetz, Thomas R Loveland, et al. 2013. "High-resolution global maps of 21st-century forest cover change." *science*, 342(6160): 850–853.

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

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

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

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *JL Econ. & Org.*, 7: 24.

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

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

Jia, Ruixue. 2017. "Pollution for Promotion." *21st Century China Center Research Paper*, 2017–05.

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

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

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

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

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

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

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

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

Liu, Yang, Meredith Franklin, Ralph Kahn, and Petros Koutrakis. 2007. "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*, 107(1-2): 33–44.

Ministry of Environment and Ecology of China. 2010. "National Ambient Air Quality Standards (Consultation Paper)."

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

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

Rambachan, Ashesh, and Jonathan Roth. 2019. "An honest approach to parallel trends." *Unpublished.*

Ruminski, Mark, Shobha Kondragunta, Roland Draxler, and Glenn Rolph. 2007. "Use of environmental satellite imagery for smoke depiction and transport model initialization." *16th Annual International Emission Inventory Conference: Emission Inventories—Integration, Analysis, and Communications.*

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

Sullivan, Daniel M, and Alan Krupnick. 2018. "Using satellite data to fill the gaps in the US air pollution monitoring network." *Resources for the Future Working Paper*, 18–21.

Turner, Woody, Carlo Rondinini, Nathalie Pettorelli, Brice Mora, Allison K Leidner, Zoltan Szantoi, Graeme Buchanan, Stefan Dech, John Dwyer, Martin Herold, et al. 2015. "Free and open-access satellite data are key to biodiversity conservation." *Biological Conservation*, 182: 173–176.

van Donkelaar, Aaron, Randall V. Martin, Chi Li, and Richard T. Burnett. 2019. "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*, 53(5): 2595–2611.

Van Donkelaar, Aaron, Randall V Martin, Michael Brauer, N Christina Hsu, Ralph A Kahn, Robert C Levy, Alexei Lyapustin, Andrew M Sayer, and David M Winker. 2016. "Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors." *Environmental science & technology*, 50(7): 3762–3772.

Van Donkelaar, Aaron, Randall V Martin, Robert JD Spurr, and Richard T Burnett. 2015. "High-Resolution Satellite-Derived $PM_{2.5}$ from Optimal Estimation and Geographically Weighted Regression over North America." *Environmental Science & Technology*, 49(17): 10482–10491.

Wang, Alex L. 2017. "Explaining environmental information disclosure in China." *Ecology LQ*, 44: 865.

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

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

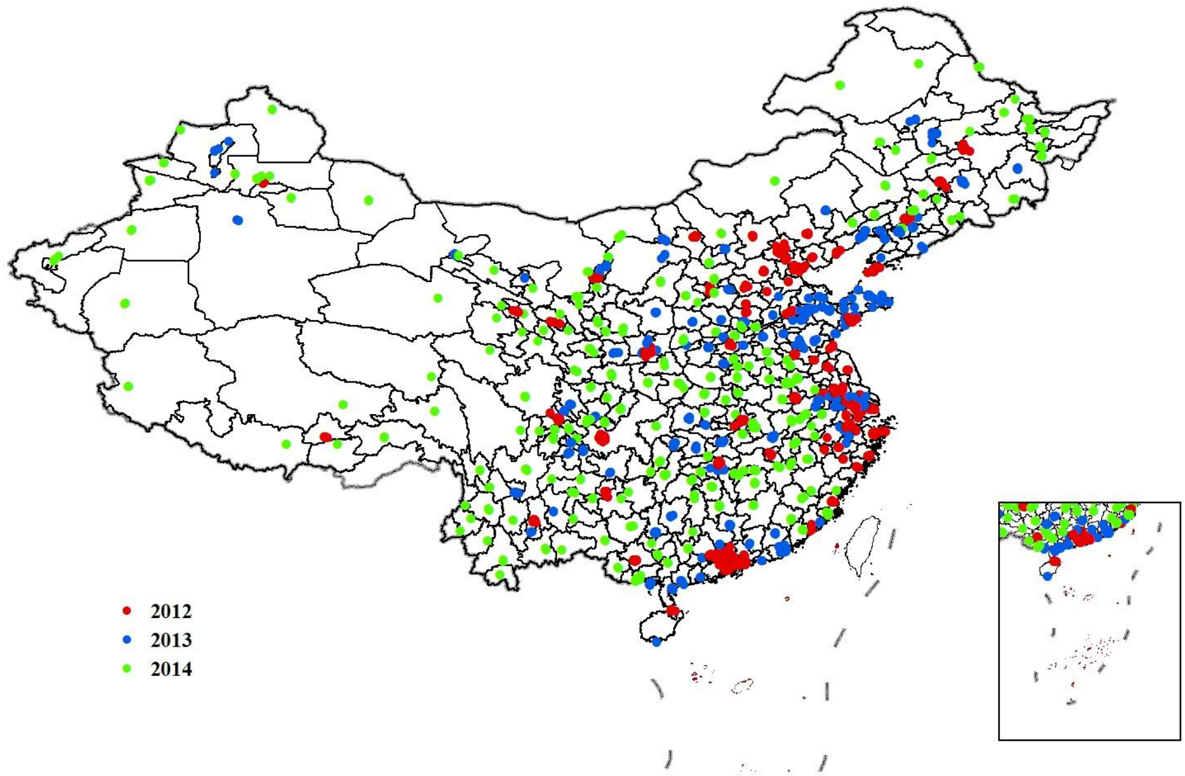
Xu, Chenggang. 2011. "The fundamental institutions of China's reforms and development." *Journal of economic literature*, 49(4): 1076–1151.

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

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

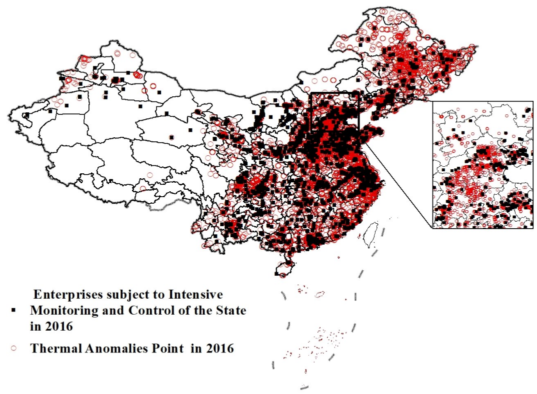
Zou, Eric Yongchen. 2021. "Unwatched pollution: The effect of intermittent monitoring on air quality." *American Economic Review*, 111(7): 2101–26.

Figure 1: The Timeline of Monitoring Station Automation

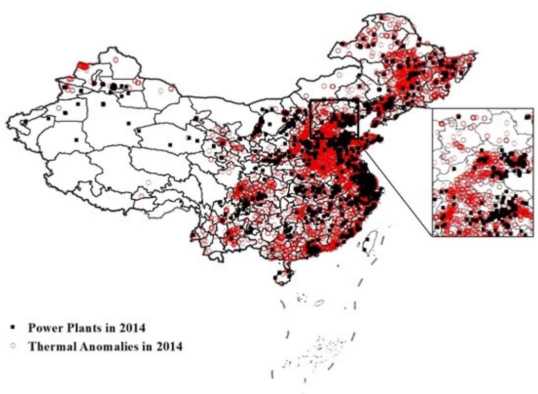


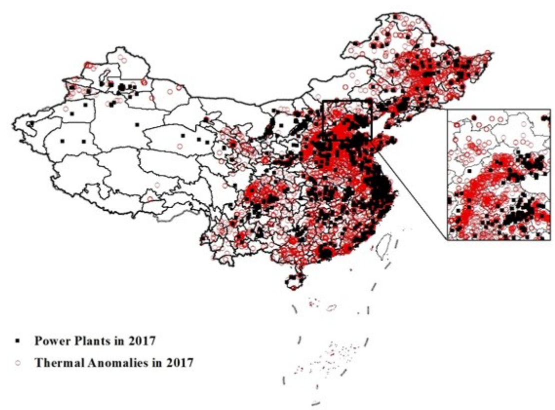
Notes: This figure displays the spatial distribution of monitoring stations automated in three waves. 2012, 2013 and 2014 denote the stations which were required to be automated in 2012, 2013 and 2014 respectively.

Figure 2: The Location of Thermal Anomalies Hotspots and Plants



(a) Industrial Plants

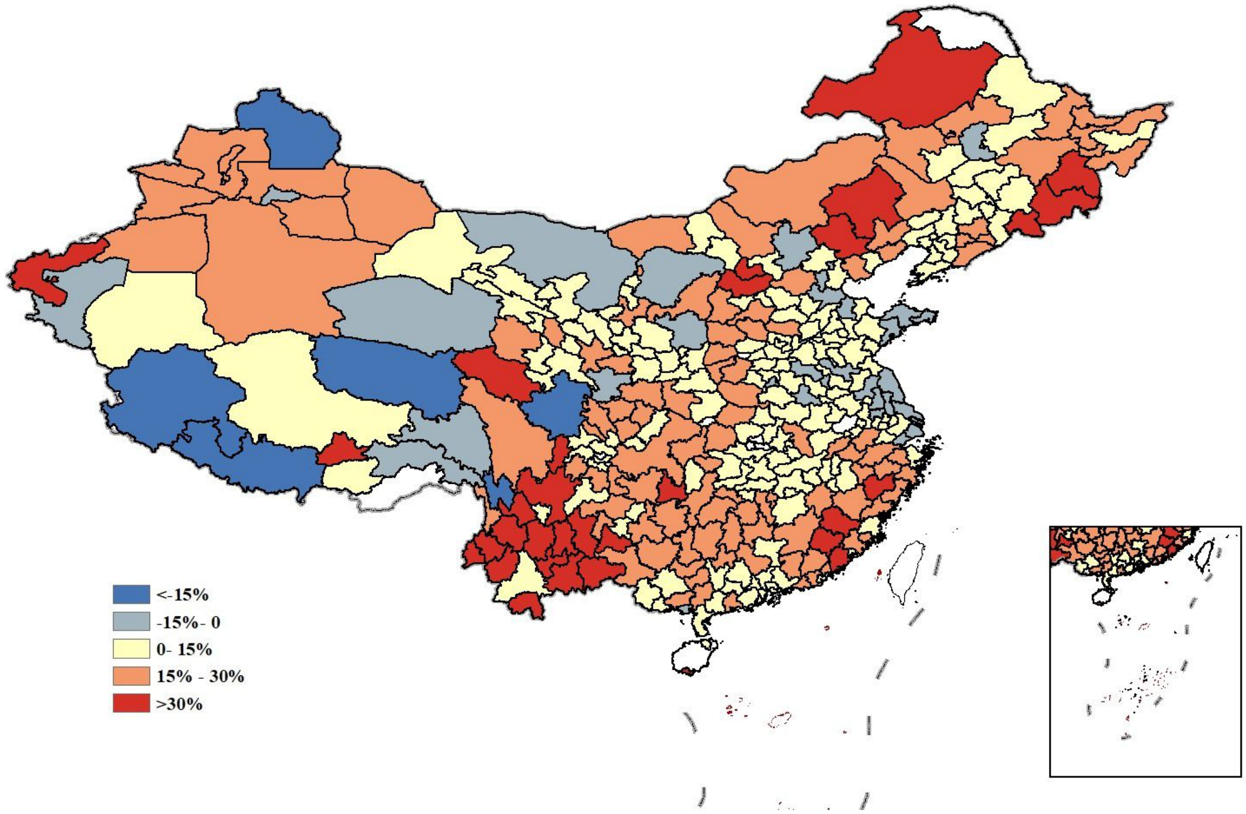




(b) Power Plants

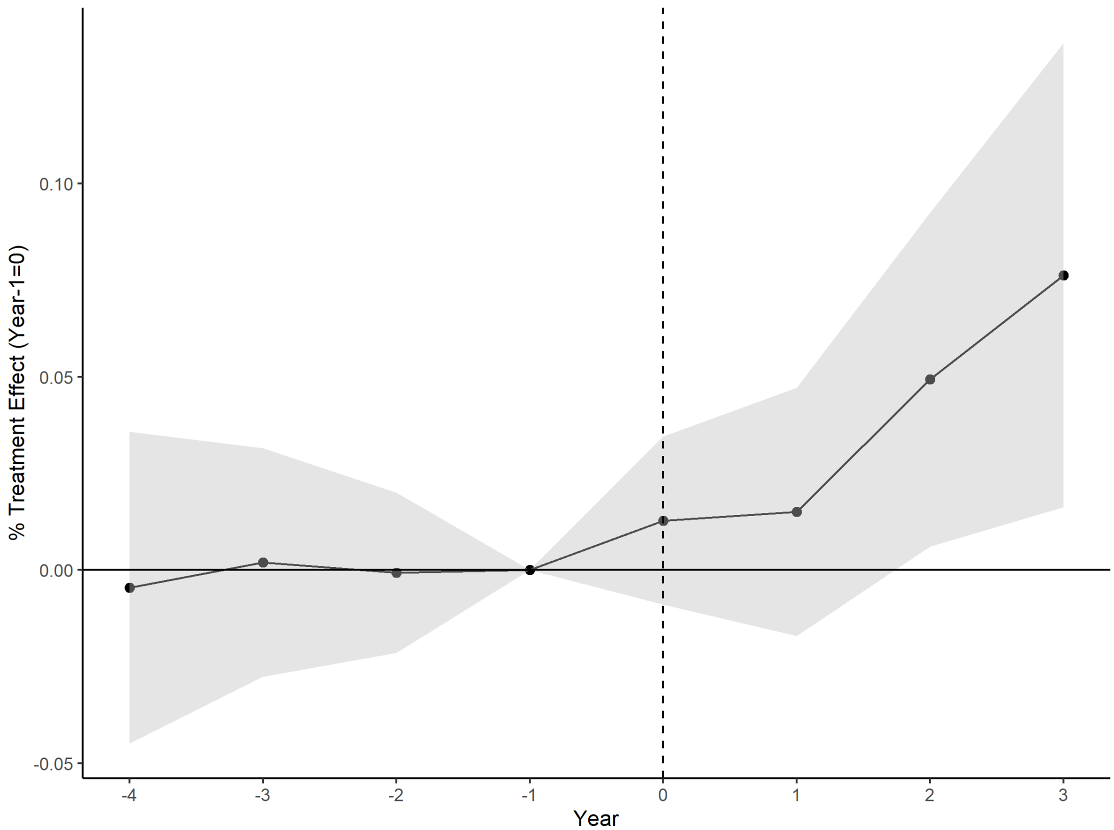
Notes: These figures compare the location of thermal anomalies hotspots and major polluting plants. Panel (a) compares the location of 1,829 heavily polluting industrial plants to the presence of 20,134 satellite-based thermal anomalies points (static hot spots) in 2016. These plants were drawn from the MEE's Key Centrally Monitored Polluting Enterprises database. Panel (b) shows the location of power plants and thermal anomalies points in 2014 and 2017, respectively. The power plants were obtained from the China Emissions Accounts for Power Plants (CEAP) database. There are 10,491 power plants and 82,082 thermal anomalies static hot spots from 2014 to 2017. The zoom-in area represents the Beijing-Tianjin-Hebei Urban Agglomerates.

Figure 3: Monitor Representation Errors in the Year of Automation: All Cells vs. Monitored Cells



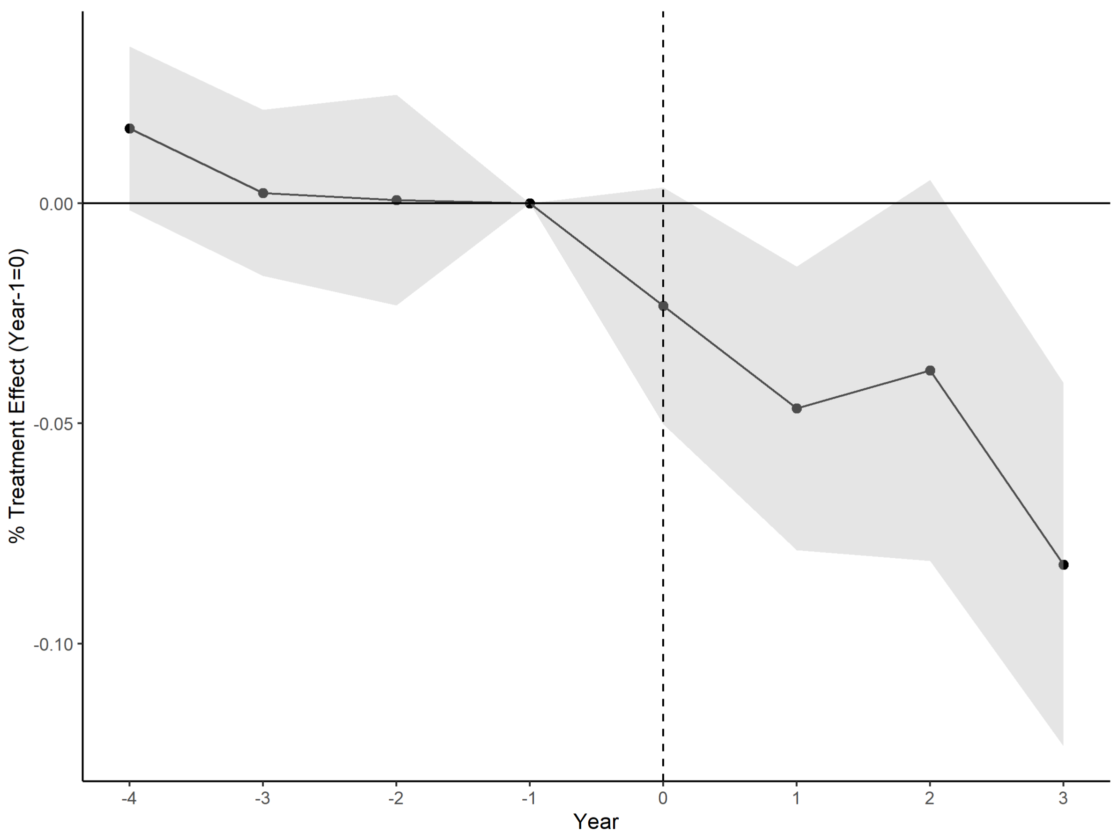
Notes: This figure presents the monitor representation errors in the year when the monitors were automated. The representation error is defined as the percentage difference between the population-weighted, satellite-based average pollution at monitored cells and the average pollution across the cells within the city boundary. Positive measures indicate over-representation by monitors.

Figure 4: Event Study: the Effect of Monitor Automation on the City-level Pollution



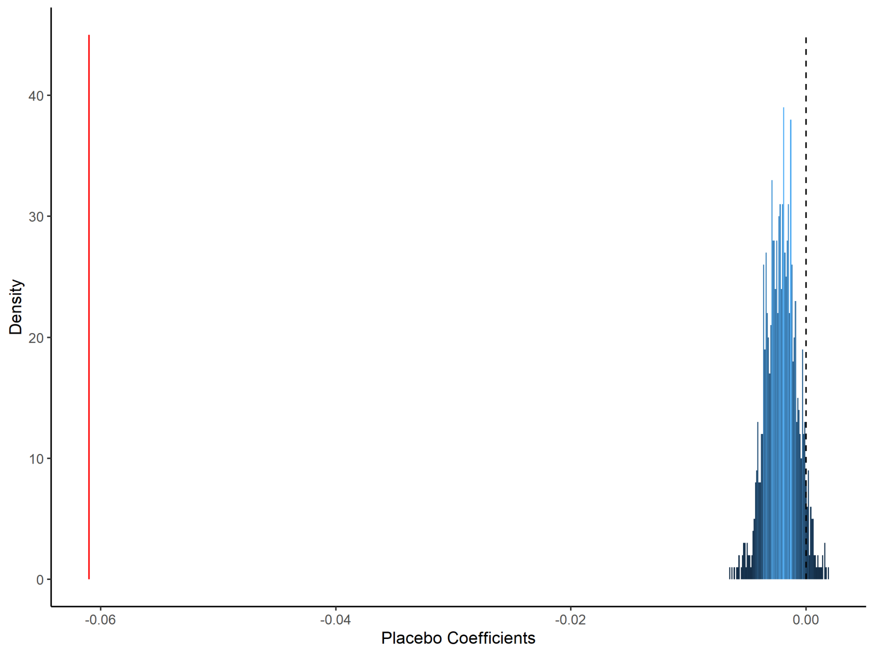
Notes: This figure presents regression coefficients and their 95% confidence intervals from an event study of the monitor automation's effect on the city-level pollution. The omitted time category is the year before a city joined the automatic monitoring program. The regression includes cell fixed effects and year fixed effects, along with the time polynomial interactions. Time polynomial interactions are a fourth-order polynomial in time interacted with average city level GDP in 2008-2010, average city level population in 2008-2010, average city level $PM_{2.5}$ in 2008, and a dummy variable which indicates whether a city is an environmental priority city. Standard errors are clustered at the city level.

Figure 5: Event Study: the Effect of Monitor Automation on Air Pollution within 3km Around the Monitor

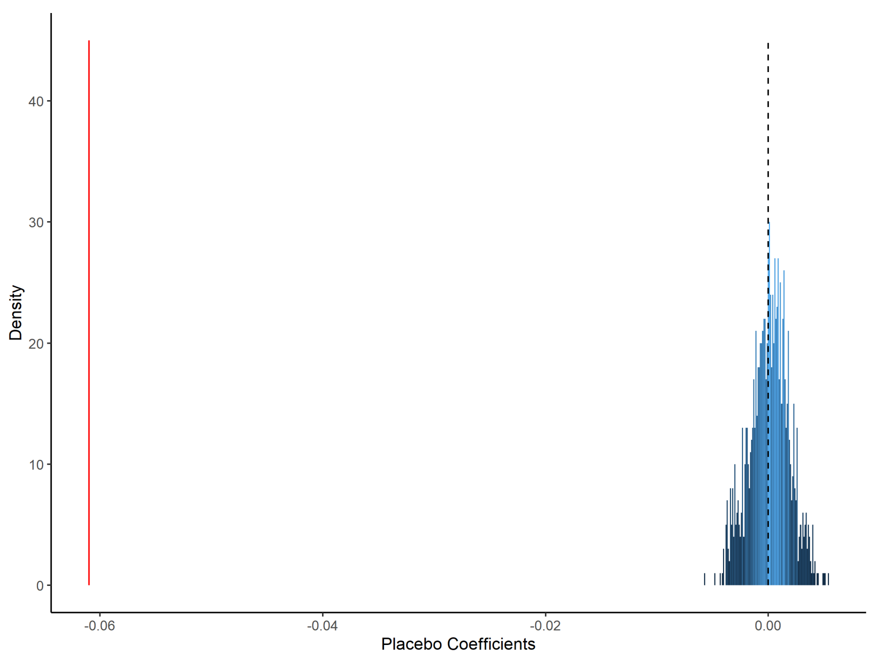


Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for β_n from Equation (3). Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period (also reported in Table A.2). The omitted time category is the year before a city joined the automatic monitoring program. The regression includes cell fixed effects, wave$\times$year fixed effects, and year fixed effects. Standard errors are clustered at the city level.

Figure 6: Placebo Tests: Randomizing Treatment Timing and Locations



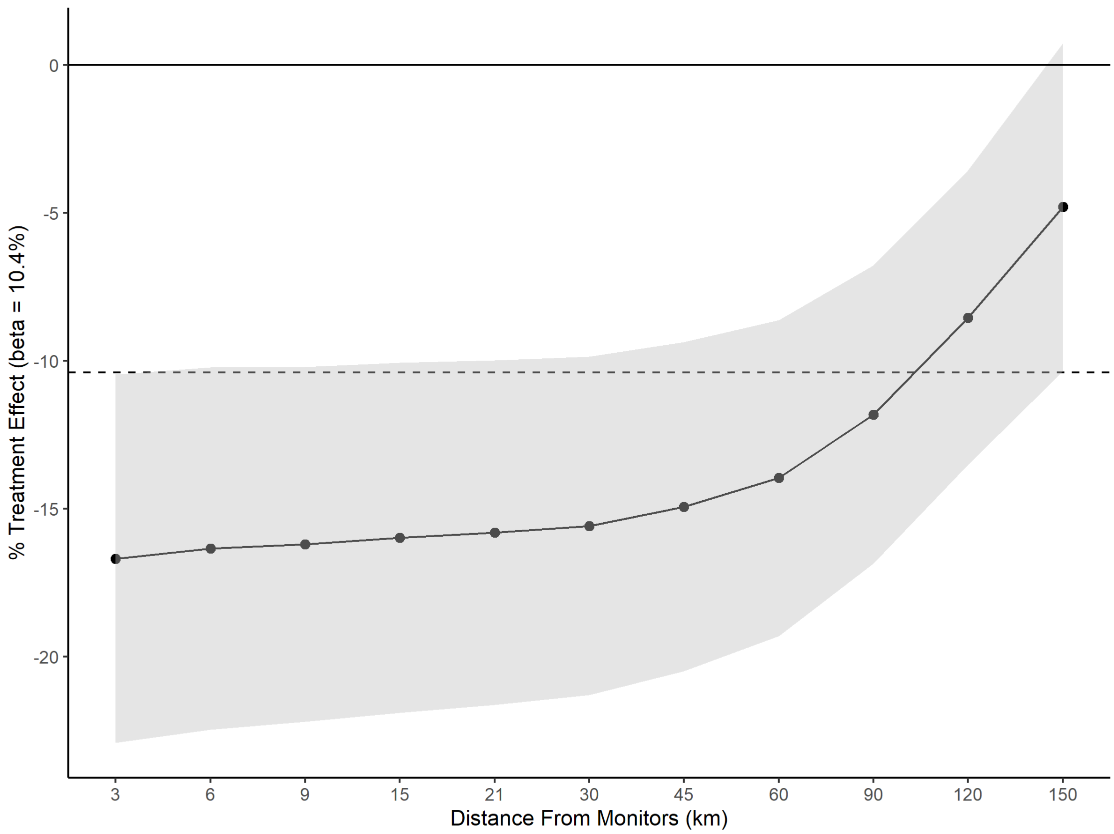
(a) Random Automation Years



(b) Random Monitor Locations

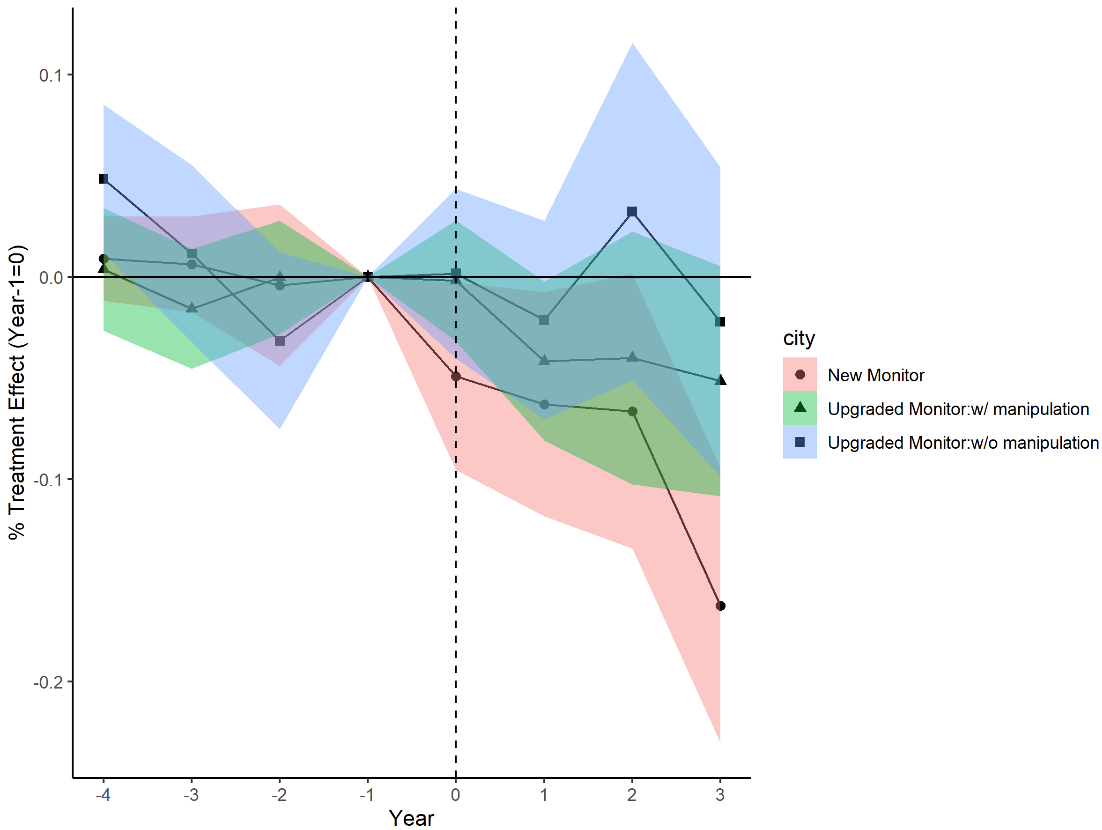
Notes: This figure presents the results of two placebo tests (See Equation (2)). Figure (a) plots a "placebo" test that randomly assigns each monitor an automation year within the sample period from 2008 to 2017. Figure (b) plots a "placebo" test that randomly assigns monitor sites to various locations while keeping the number of monitors and the year of automation unchanged. For each placebo test, the DiD estimation is repeated 1000 times. The distribution of the estimates from the 1000 runs (blue lines) is then plotted along with the benchmark estimate (red line).

Figure 7: Strategic Cleaning Response to Monitoring Program with Different Treatment Intensity Bins



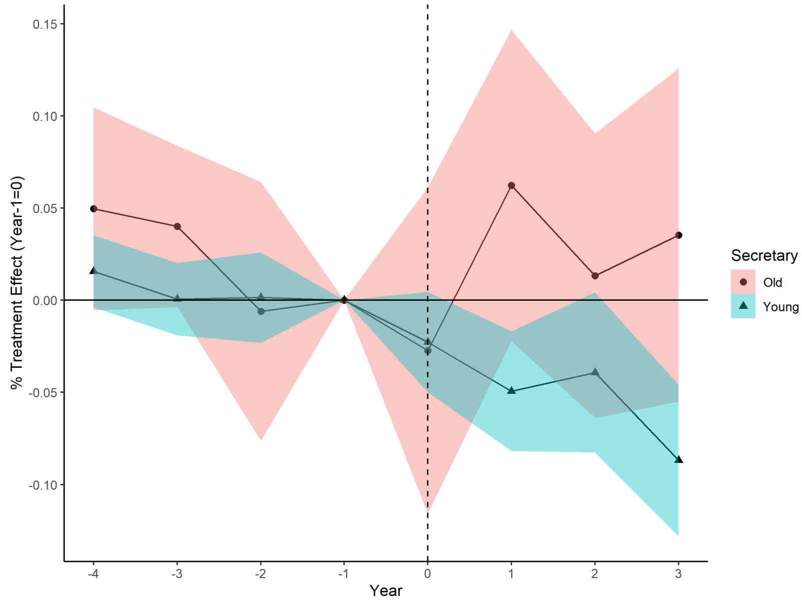
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals on the effects of monitor automation on the satellite-based ln$PM_{2.5}$ at different distance bins from the monitor. Each point estimate represents the pollution change in each distance bin relative to the baseline group at the outer range (distance to monitor >150 km), which on average experiences a 10.4% pollution increase. The absolute effect becomes positive above the dotted line. The regression includes cell fixed effects and year fixed effects, along with the time polynomial interactions. The time polynomial interactions are a fourth-order polynomial in time interacted with average city level GDP in 2008-2010, average city level population in 2008-2010, average city level $PM_{2.5}$ in 2008, and a dummy variable which indicates whether a city is an environmental priority city. Standard errors are clustered at the city level.

Figure 8: Heterogeneity Analysis by Pre-Automation Status: Event Study Results

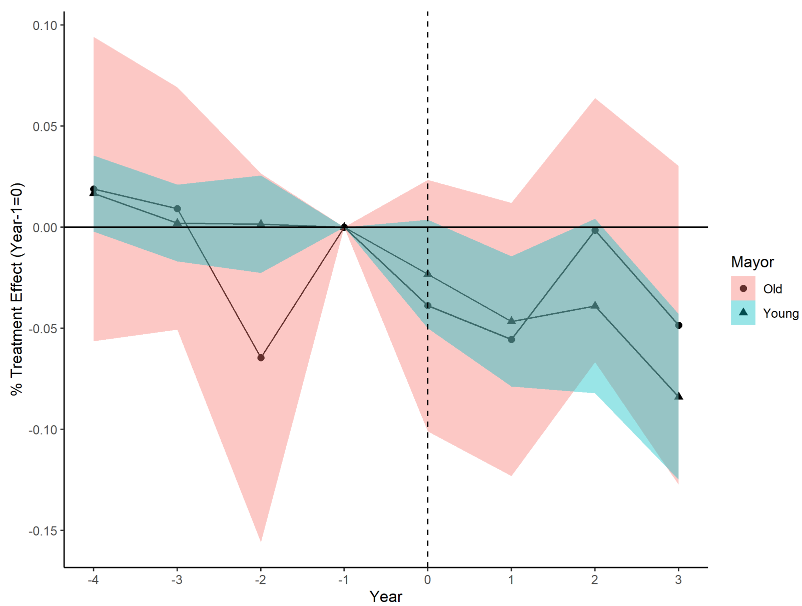


Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the interaction terms between year-specific automation treatment effects and the indicators for three subsamples of cities. Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within a 3km radius of the monitors) and unmonitored areas (cells more than 3km away from the monitors) at a given period for each subgroup. "New Monitor" denotes cities that do not have monitors before the program. Cities with monitors before the program are divided according to their pre-automation pollution reading manipulation status, obtained using the local linear RD estimates in Greenstone et al. (forthcoming). "Upgraded Monitor: w/ manipulation" refers to cities that experience sudden drops in pollution levels directly following monitor automation, which is indicative of pre-automation data manipulation. "Upgraded Monitor: w/o manipulation" refers to those cities with negative RD estimates, indicative of no pre-automation data manipulation. The omitted time category is one year before a city joined the automatic monitoring program. Regressions include cell fixed effects; year fixed effect and wave by year fixed effects.

Figure 9: Heterogeneity Analysis by City Leader Age: Event Study Results



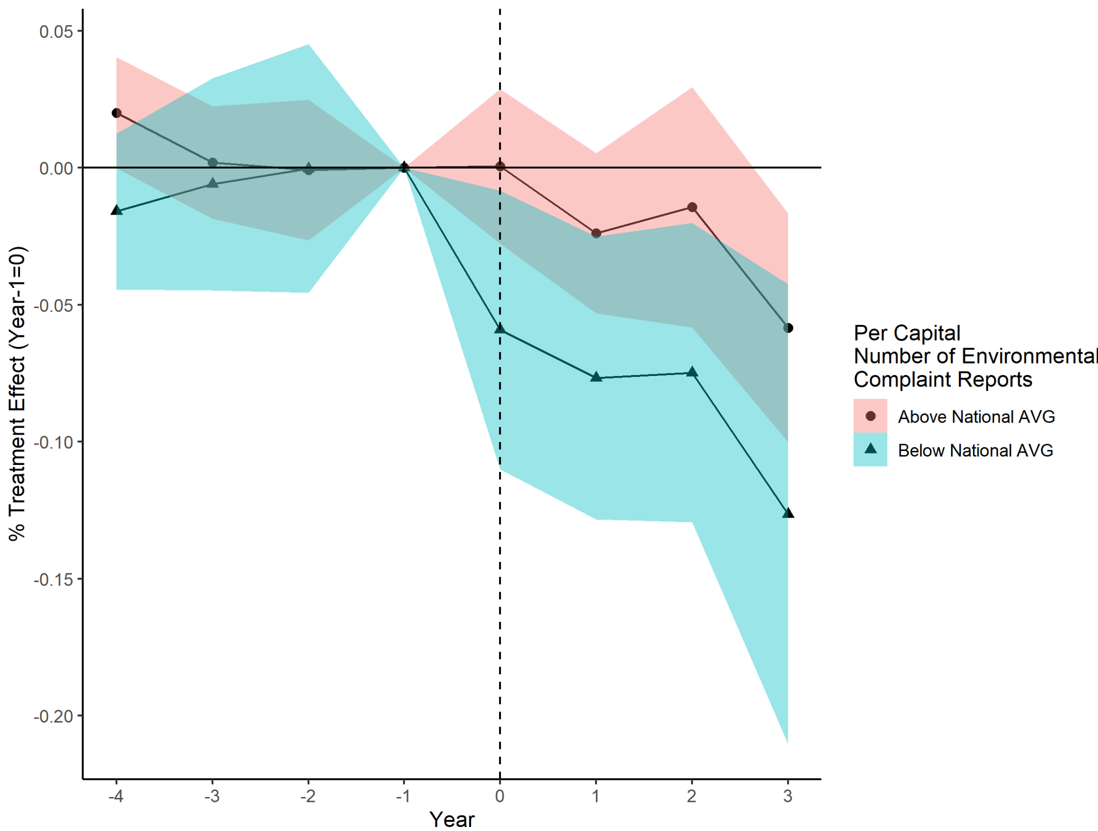
(a) City Secretary Age



(b) City Mayor Age

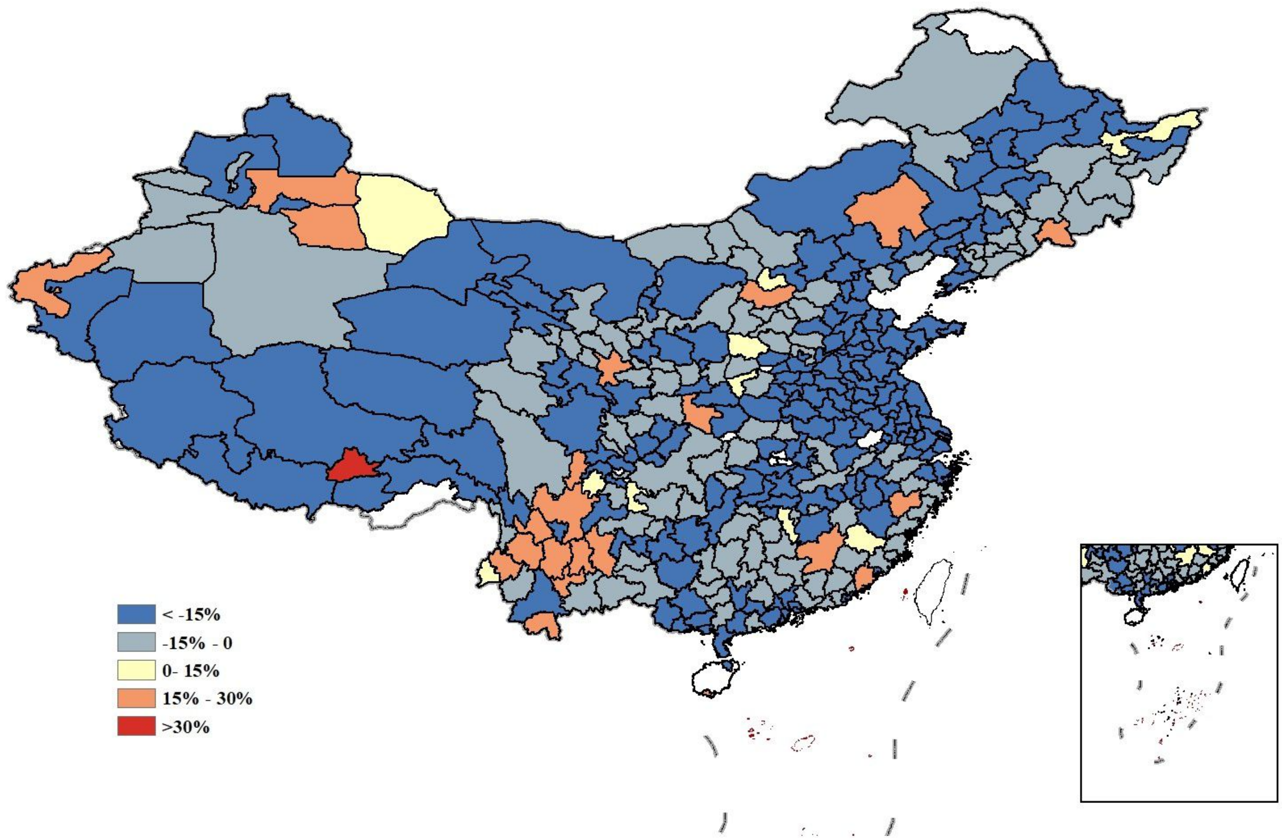
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the interaction terms between year-specific automation treatment effects and indicators for two subsamples of cities classified by city leader's age. Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within a 3km radius of monitors) and unmonitored areas (cells located more than 3km away from the monitors) at a given period for each subgroup. Figures (a) and (b) are heterogeneity analyses by city secretaries' age and mayors' age, respectively. Young is a dummy variable that is equal to one if a city leader is younger than 57, and zero otherwise. The omitted time category is one year before a city joined the automatic monitoring program. Regressions include cell fixed effects, year fixed effects and wave by year fixed effects.

Figure 10: Heterogeneity Analysis by Local Public Pressure: Event Study Results



Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the interaction terms between year-specific automation treatment effects and indicators for two subsamples of cities, classified by local residents' attention to environmental issues. Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within a 3km radius of the monitors) and unmonitored areas (cells beyond a 3km distance from the monitors) at a given period for each subgroup. Above National AVG indicates provinces which have above-median per capita complaints on environmental issues in 2017. The omitted time category is one year before a city joined the automatic monitoring program. Regressions include cell fixed effects, year fixed effects and wave by year fixed effects.

Figure 11: Prediction of Monitor Representation Errors in 2022: All Cells vs. Monitored Cells



Notes: This figure shows the predicted representation errors of monitors in 2022, which are calculated based on the estimated pollution reductions in monitored areas (cells within a 3km radius of monitors have experienced a 6.1% greater reduction in air pollution relative to unmonitored areas), and are projected beyond the last observed sample year of pollution (2017). The representation error is defined as the percentage difference between the population-weighted average pollution level calculated based only on the monitored cells and the city's average pollution based on all cells within the city boundary. Positive measures indicate over-representation by monitors.

Table 1: Descriptive Statistics

Variables	N	Mean	Sd	Min	Max
	(1)	(2)	(3)	(4)	(5)
Panel A: Cells and Satellite-based Pollutants					
$PM_{2.5}$	10,414,296	35.217	27.234	0.000	223.889
Distance to monitor	10,427,410	160.373	149.849	0.312	1,380.487
Auto	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
Days	11,470,151	0.018	0.977	0	554
1(TAP)	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	10,427,200	872.341	4,206.690	0.000	409,222.400
GDP in 2008	10,416,010	49.360	81.102	1.443	1,097.079

Notes: Observations are at the cell-year level. Each cell is at 3km×3km resolution. In Panel A, $PM_{2.5}$ is satellite-based $PM_{2.5}$ measured for each cell. Distance to monitor is the distance between a cell and its nearest monitoring station. Auto is the treatment indicator that switches to one after a city joined the automatic monitoring program. (0-3km) is a dummy variable that switches on for cells located within a 3km radius around monitoring stations. In Panel B, Fire Radiative Power (FRP) is a measure of the average intensity of thermal anomalies, defined as the rate of radiant heat output and is related to the rate at which fuel is being consumed, and smoke emissions are released. Days measures the number of days with active thermal anomalies in each cell annually. 1(TAP) denotes 1(Thermal Anomalies Presence), which is a dummy variable that equals one if there are thermal-related economic activities in a cell in a year. In Panel C, $\text{Startage}_{\text{secretary}}$ and $\text{Startage}_{\text{mayor}}$ are ages of cities' secretaries and mayors at the start of their term. New monitor indicator is a dummy variable that equals one if a city does not have any monitor before the automatic monitoring program. Manipulation indicators are the local linear RD estimates calculated using the algorithm in Greenstone et al. (forthcoming). Population in 2010 (*person*) are cell-level population, and GDP in 2008 (in 1 billion¥) are city-level GDP.

Table 2: Strategic Cleaning Response to Monitoring Program Automation

	Dependent variable: $\ln(PM_{2.5})$				
Unmonitored Area	>3 km	>3 km	> 3 km	>30km	>60km
	(1)	(2)	(3)	(4)	(5)
Auto	0.041*	0.042*			
	(0.025)	(0.025)			
(0-3km)×Auto		-0.042***	-0.061***	-0.069***	-0.083***
		(0.012)	(0.013)	(0.015)	(0.019)
CellFE	X	X	X	X	X
Year FE	X	X	X	X	X
Wave×Year FE			X	X	X
Time polynomial interactions	X	X			
Observations	9,357,471	9,357,471	10,413,707	9,309,540	7,579,589
R-squared	0.965	0.965	0.967	0.967	0.967

Notes: This table reports the effects of the monitor automation program on the satellite-based $\ln PM_{2.5}$. $\ln PM_{2.5}$ is the natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Auto is the treatment indicator that switches on after a city joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if the cells are located within the 3km radius of a city's monitoring stations. Columns (3)-(5) use cells within 3km of the monitor as the monitored group and compare it to different unmonitored groups: cells outside 3km, 30km, 60km of the monitors, respectively. Time polynomial interactions are a fourth-order polynomial in time interacted with average city level GDP in 2008-2010, average city level population in 2008-2010, average city level $PM_{2.5}$ in 2008, and a dummy variable which indicates whether or not a city is an environmental priority city. Standard errors are clustered at the city level. Significance: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 3: Robustness Check: Test for Mean Reversion and Entropy Balancing

	Dependent variable: $\ln(PM_{2.5})$			
	(1)	(2)	(3)	(4)
Unmonitored Areas:	>3km	>3km	>3km	>60km
(0-3km)×Auto	-0.044***	-0.053**	-0.046***	-0.029**
	(0.013)	(0.022)	(0.013)	(0.0144)
Auto×Dirtier	-0.048***	-0.048***		
	(0.018)	(0.018)		
(0-3km)×Auto×Dirtier		0.011		
		(0.022)		
CellFE	X	X	X	X
Year FE	X	X	X	X
Wave ×Year FE	X	X	X	X
Dirtier × Year FE			X	
Entropy Balancing				X
Observations	10,411,191	10,411,191	10,411,191	4,972,049
R-squared	0.967	0.967	0.968	0.958

Notes: This table reports the effects of the monitor automation program on the satellite-based $\ln PM_{2.5}$ in alternative specifications. $\ln PM_{2.5}$ is the natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Auto is the treatment indicator that takes the value of 1 after a city joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if cells are located within 3km around a city's monitor. Dirtier is a dummy variable which equals one if the $\ln PM_{2.5}$ of a cell is higher than the city's average $\ln PM_{2.5}$ in 2008 (base year). Columns (1) and (2) include the (0-3km)×Auto×Dirtier and Auto×Dirtier to test for mean reversion as an alternative explanation of the observed automation treatment effects. Column (3) includes Dirtier × Year fixed effects. Column (4) reports the estimation result when observations are reweighted with entropy balance weights. All regressions control for cell fixed effects, year fixed effects and wave × year fixed effects. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table 4: Robustness Check Using Raw AOD Data

	Dependent variable: AOD		
Unmonitored Areas:	>3km	>30km	>60km
	(1)	(2)	(3)
(0-3km)*Auto	-0.009*	-0.022***	-0.033***
	(0.005)	(0.006)	(0.007)
Cell FE	X	X	X
Year FE	X	X	X
Wave ×Year FE	X	X	X
Observations	8,954,502	9,213,068	6,820,362
R-squared	0.885	0.862	0.836

Notes: This table shows the estimation results replacing outcome variables with the raw AOD data. *Auto* is the treatment indicator that equals one if a cell is in a city that has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if cells are located within 3km around a city's monitor. The dependent variable is the annual AOD measures at 3km by 3km grid cells from 2008 to 2017. Columns (1) - (3) use cells within 3km of the monitor as the monitored group and compare it with different unmonitored groups: cells more than 3km, 30km and 60km away from the monitors. All columns include cell fixed effects, year fixed effects and wave × year fixed effects. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 5: Mechanism of Strategic Cleaning: Industrial Activities (Thermal Anomalies)

VARIABLES	(1) 1(TAP)	(2) ln(Days+1)	(3) ln(FRP+1)	(4) ln(Days+1)	(5) ln(FRP+1)
Auto	0.368***	-0.0173	-0.0421	0.0121	0.00634
	(0.0111)	(0.0372)	(0.0410)	(0.0139)	(0.0113)
Marginal Effect	*0.0904****	*-0.00483*	*-0.0268*		
	(0.00264)	*(0.0104)*	*(0.0262)*		
(0-3km)×Auto	-0.411***	-0.296***	-0.236***	-0.0910***	-0.0164
	(0.0436)	(0.0391)	(0.0386)	(0.0216)	(0.0159)
Marginal Effect	*-0.101****	*-0.0826****	*-0.150****		
	(0.0107)	*(0.0109)*	*(0.0246)*		
Cell FE	X	X	X	X	X
Year FE	X	X	X	X	X
Time polynomial interactions		X	X	X	X
Model	Logit	Poisson	Poisson	OLS	OLS
Sample	All	All	All	1(TAP)	1(TAP)
Observations	227,740	232,150	232,140	51,859	51,859
R-squared				0.719	0.552

Notes: This table reports the effects of the monitor automation program on the thermal anomalies. Column (1) uses a logit regression model. Columns (2) and (3) use a Poisson regression model. Columns (4) and (5) use an OLS model. For the logit regression model and the Poisson regression model, the marginal effects are also reported. Column (1) reports the results on a dummy indicator of thermal anomalies presence (TAP), denoted by 1(TAP), which is equal to one if thermal-related economic activities are present in a cell in that year. Column (2) reports the results on the number of days with active thermal anomalies by using the full sample, which measures the operating time of industrial plants in each cell. Column (3) reports the results on the average intensity of thermal anomalies, denoted by ln(FRP+1). FRP is defined as the rate of radiant heat output and is related to the rate at which fuel is being consumed and smoke emissions are released. We use the natural logarithm of (FRP+1) and (Days+1) to tackle zero observations. Column (4) reports the effect of automation on the logarithm of the number of days with active thermal anomalies by restricting the sample to only those gridcell-year observations when 1(Thermal Anomalies Presence) is equal to one. Column (5) reports the effect of automation on the average intensity of thermal anomalies per day (denoted by ln(FRP+1)) when 1(Thermal Anomalies Presence) is equal to one. $\ln PM_{2.5}$ is the natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Auto is the treatment indicator that takes the value of 1 after a city joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if cells are located within 3km around a city's monitoring stations. Time polynomial interactions are a fourth-order polynomial in time interacted with average city level GDP in 2008-2010, average city level population in 2008-2010, average city level $PM_{2.5}$ in 2008, and a dummy variable which indicates whether a city is an environmental priority city. Standard errors are clustered at the city level. Significance: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

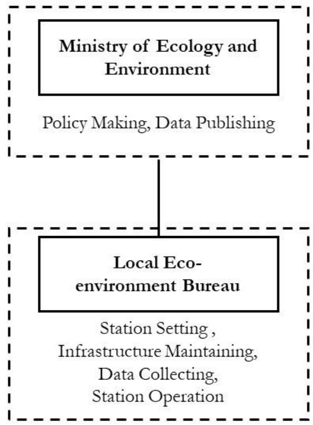
Table 6: Mechanism of Strategic Cleaning: Thermal Anomalies with Treatment Intensity Bins

VARIABLES	(1) 1(TAP)	(2) ln(Days+1)	(3) ln(FRP+1)	(4) ln(Days+1)	(5) ln(FRP+1)	(6) Pop (Million)
Auto	0.245***	0.0832***	0.171***	0.0563	-0.0138	22.119
	(0.0168)	(0.0227)	(0.0562)	(0.0402)	(0.0214)	
(0-3km)×Auto	-0.256***	-0.173***	-0.354***	-0.137***	0.00325	111.842
	(0.0223)	(0.0554)	(0.0432)	(0.0246)	(0.0230)	
(3-6km)×Auto	-0.208***	-0.139***	-0.293***	-0.0963**	-0.0120	86.577
	(0.0224)	(0.0552)	(0.0424)	(0.0228)	(0.0205)	
(6-9km)×Auto	-0.202***	-0.120***	-0.262***	-0.0912**	0.000674	44.529
	(0.0191)	(0.0239)	(0.0582)	(0.0435)	(0.0255)	
(9-15km)×Auto	-0.201***	-0.102***	-0.248***	-0.0530	0.0214	68.324
	(0.0184)	(0.0252)	(0.0604)	(0.0439)	(0.0256)	
(15-21km)×Auto	-0.178***	-0.0911***	-0.218***	-0.0312	0.0446*	61.013
	(0.0188)	(0.0250)	(0.0634)	(0.0432)	(0.0235)	
(21-30km)×Auto	-0.159***	-0.0841***	-0.202***	-0.0329	0.0123	84.522
	(0.0185)	(0.0249)	(0.0606)	(0.0428)	(0.0238)	
(30-45km)×Auto	-0.155***	-0.0898***	-0.206***	-0.0321	0.0375	124.066
	(0.0182)	(0.0222)	(0.0567)	(0.0417)	(0.0230)	
(45-60km)×Auto	-0.135***	-0.0890***	-0.189***	-0.0539	0.0169	126.786
	(0.017)	(0.0186)	(0.0233)	(0.0568)	(0.0431)	
(60-90km)×Auto	-0.119***	-0.0700***	-0.160***	-0.00853	0.0290	54.028
	(0.0186)	(0.0241)	(0.0578)	(0.0479)	(0.0235)	
(90-120km)×Auto	-0.0496**	-0.0145	-0.0295	-0.00396	0.0346	50.600
	(0.0210)	(0.0257)	(0.0644)	(0.0523)	(0.0297)	
(120-150km)×Auto	-0.0379	-0.0515	-0.0298	-0.00440	0.0703**	20.875
	(0.0252)	(0.0488)	(0.0928)	(0.0526)	(0.0347)	
Cell FE	X	X	X	X	X	
Year FE	X	X	X	X	X	
Time polynomial interactions		X	X	X	X	
Model	Logit	Poisson	Poisson	OLS	OLS	
Marginal Effect	X	X	X			
Sample	All	All	All	1(TAP)	1(TAP)	
Observations	227,740	233,750	233,740	52,097	52,097	
R-squared				0.551	0.718	

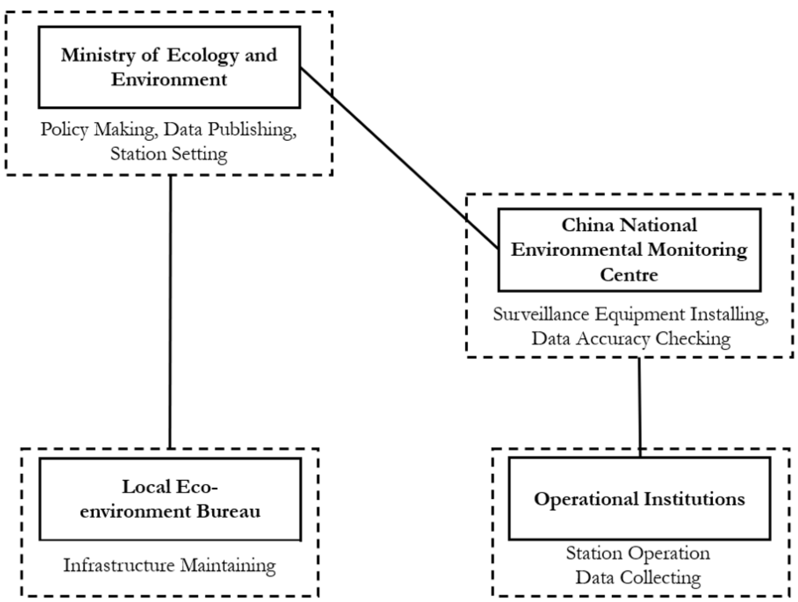
Notes: This table reports the effects of the monitor automation program on various distance bins. Column (1) uses a logit regression model, Column (2) and (3) use a Poisson regression model, and (4) and (5) use an OLS model. For the logit regression model and the Poisson regression model, the marginal effects are reported. Column (1) reports the estimation results on a dummy indicator of thermal anomalies, 1(TAP), which is equal to one if thermal-related economic activities are present in a cell in that year. Column (2) reports the estimation results on the number of days with active thermal anomalies using the full sample, which measures the operating time of industrial plants in each cell. Column (3) reports the results on the average intensity of thermal anomalies, denoted by ln(FRP+1). FRP is defined as the rate of radiant heat output and is related to the rate at which fuel is being consumed and smoke emissions are released. We use the natural logarithm of of (FRP+1) and (Days+1) to tackle zero observations. Column (4) reports the effect of automation on the total active days of thermal anomalies by restricting the sample to only those gridcell-year observations with thermal anomalies present, i.e., when 1(Thermal Anomalies Presence) is equal to one. Column (5) reports the effect of automation on the average intensity of thermal anomalies per day (denoted by ln(FRP+1)) when 1(Thermal Anomalies Presence) is equal to one. Column (6) shows the total population (million) in each distance bin using the 2015 population data. ln$PM_{2.5}$ is the natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. Auto is the treatment indicator that takes the value of 1 after a city joined the automatic monitoring program. Its coefficient denotes the effect of automation on cells beyond 150km radial distance from a monitor. The coefficients of the interaction terms represent the effect of automation on the pollution change in each distance bin relative to the baseline group at the outer range (distance to monitor>150 km). (0-3km) is a dummy variable that equals one if the cells are located within 3km around a city's monitoring stations. Time polynomial interactions are fourth-order polynomials in time interacted with average city-level GDP in 2008-2010, average city-level population in 2008-2010, average city-level $PM_{2.5}$ in 2008, and a dummy variable which indicates whether a city is an environmental priority city. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

A Appendix

Figure A.1: Change in Environmental Authorities' Responsibilities



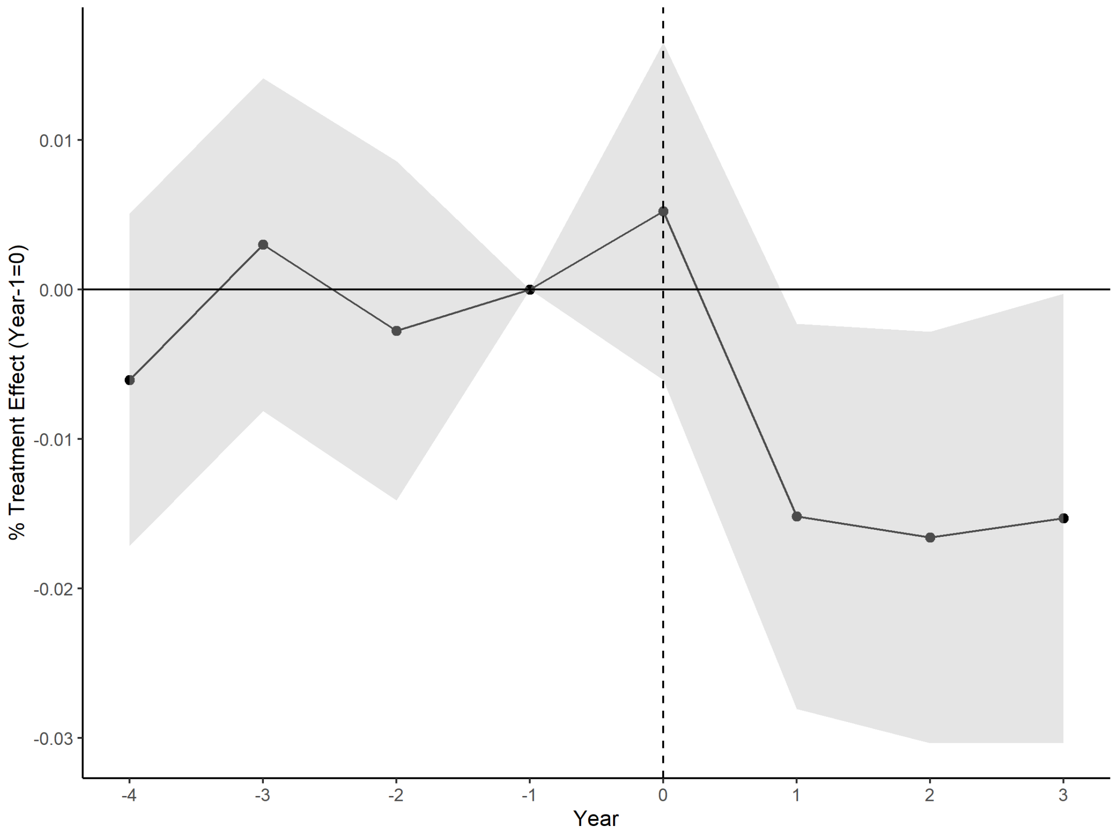
(a) Before



(b) After

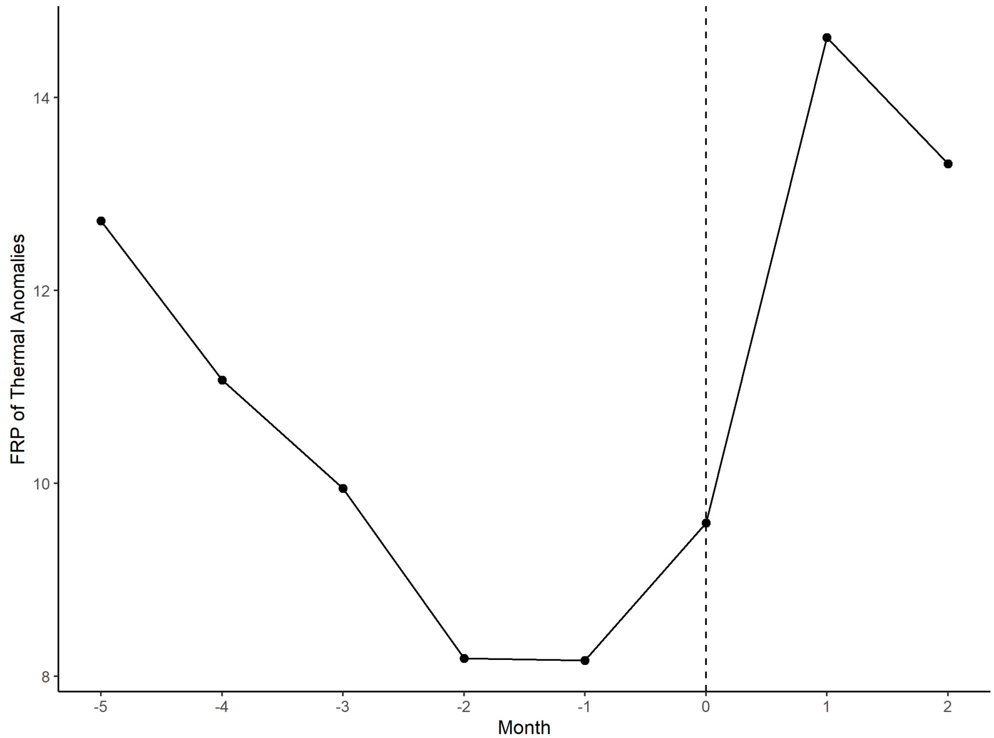
Notes: These two figures illustrate the roles and responsibilities of different environmental authorities, before (Panel A) and after (Panel B) the introduction of new standards. China National Environmental Monitoring Centre (CNEMC) is a newly established institution directly under the management of 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, Infrastructure Maintaining refers to ensuring the supply of electricity and communications, and Data Accuracy Checking denotes checking the anomaly data.

Figure A.2: An Event Study of New Monitoring Program: Robustness Check Using AOD Raw Data



Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for β_n from Equation (3), replacing outcome variables with raw AOD data. Detailed estimates are reported in Table A.2. The omitted time category is one year before a city joined the automatic monitoring program. Each estimate represents the difference in AOD between monitored areas (cells within a 3km radius of a monitor) and unmonitored areas (cells outside the 3km range of a monitor) at a given period. The regression includes cell fixed effects, year fixed effects and wave×year fixed effects. Standard errors are clustered at the city level.

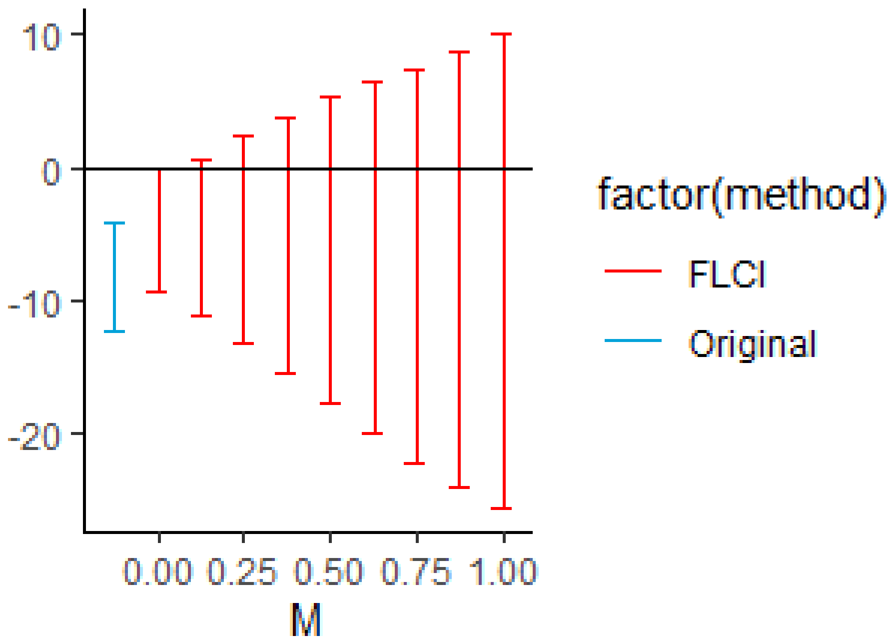
Figure A.3: Validation of Thermal Anomalies Using the APEC Event



(a) FRP

Notes: This figure shows the time-series of the thermal anomalies measure shortly before and after the APEC event. Month 0 denotes the month during which APEC was held (November, 2014). FRP is defined as the rate of radiant heat output and is related to the rate at which fuel is being consumed and smoke emissions are released.

Figure A.4: Robust Analysis of Event Study on $PM_{2.5}$



Notes: This figure shows the sensitivity analysis of estimated effects on $PM_{2.5}$ to potential violations of the parallel trends assumption following the methods proposed by Rambachan and Roth (2019). The blue bar represents the 95% confidence interval of the DiD estimate for the last period ($\tau = 3$) from estimation of Equation (3). The red bars represent corresponding 95% confidence intervals when allowing for per-period violations of parallel trends of up to M, which is the largest allowable change in the slope of an underlying linear trend between two consecutive periods.

Table A.1: The Summary Statistics of the Automation Sequence

Wave	N	BasePM$_{2.5}$ ($\mu g/m^3$)	BaseGDP (billion¥$/km^2$)	POP2010 (person$/km^2$)
	(1)	(2)	(3)	(4)
Wave 1	99,632	46.00	270.84	6,061,255
(2012)		(19.49)	(244.99)	(5,028,577)
Wave 2	191,388	47.72	86.90	2,272,809
(2013)		(25.84)	(52.41)	(1,393,642)
Wave 3	751,721	30.96	27.85	1,300,608
(2014)		(25.59)	(24.69)	(1,195,594)

Notes: The table reports the average pollution, local population and GDP across stations in various automation waves, measured from baseline years. Wave 1, wave 2 and wave 3 denote the stations which were required to be automated in 2012, 2013, and 2014, respectively. BasePM$_{2.5}$ is yearly PM$_{2.5}$ in the base year 2008. BaseGDP (in 1 billion¥) are city-level GDP in 2008 and population ($person$) and POP2010 are cell-level population measured in 2010, the first year when such information was available.

Table A.2: Event Study: the Effect of Monitor Automation on Air Pollution within 3km Around the Monitor

	Dependent variable:	
	$\ln(PM_{2.5})$	AOD
	(1)	(2)
(0-3km)×before4	0.017*	-0.006
	(0.009)	(0.006)
(0-3km)×before3	0.002	0.003
	(0.010)	(0.006)
(0-3km)×before2	0.001	-0.003
	(0.012)	(0.006)
(0-3km)×after0	-0.023*	0.005
	(0.014)	(0.006)
(0-3km)×after1	-0.047***	-0.015**
	(0.016)	(0.007)
(0-3km)×after2	-0.038*	-0.017**
	(0.022)	(0.007)
(0-3km)×after3	-0.082***	-0.015**
	(0.021)	(0.008)
CellFE	X	X
YearFE	X	X
Wave × YearFE	X	X
Time polynomial interactions	X	
Observations	10,413,707	8,954,502
R-squared	0.967	0.885

Notes: The table reports the event study results of monitor automation on air pollution with different dependent variables. Column (1) shows the effect the monitor automation program on $\ln(PM_{2.5})$ of 3km-Around-Monitor monitored areas (Figure 5) and column (2) shows the effect the monitor automation program on the annual AOD of 3km-Around-Monitor monitored areas (Figure A.2). Time polynomial interactions are fourth-order polynomials in timeinteracted with average city-level GDP in 2008-2010, average city-level population in 2008-2010, average city-levelPM2.5in 2008, anda dummy variable which indicates whether a city is an environmental priority city. Standard errors are clustered at the city level. Significance: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A.3: Validating Thermal Anomalies Data: Extensive Margins

Dependent variable:	Presence of any polluting firm		Presence of any power plant	
	(1)	(2)	(3)	(4)
Thermal Anomalies Presence	22.569***	40.627***	22.912***	41.703***
	(0.134)	(0.358)	(0.140)	(0.451)
Marginal Effect	*0.993****	*0.999****	*0.995****	*0.999****
	(0.001)	*(0.000)*	*(0.001)*	*(0.000)*
City FE		X		X
Observations	95,168	95,168	761,344	761,344

Notes: This table shows the association between thermal anomalies and polluting firms or power plants at the extensive margin, using the logit model. In columns (1) and (2), the dependent variable is a dummy variable that equals one should there exist any polluting plants within a 10km-by-10km cell. The polluting plants are defined as the MEE's Key Centrally Monitored Polluting Enterprises database. In columns (3) and (4), the dependent variable is a dummy indicator that equals one if there exists any power plant within the 10km-by-10km cell. The power plants sample is obtained from the China Emissions Accounts for Power Plants (CEAP). "Thermal Anomalies Presence" is a dummy variable which is equal to one if there are any thermal-related economic activities in a cell. Columns (2) and (4) include city fixed effects. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A.4: Validating Thermal Anomalies Data: Intensive Margins

	Dependent variable: $\ln(PM_{2.5})$	
Sample	Polluting firms	Power plants
	(1)	(2)
ln(FRP+1)	0.156***	0.137***
	(0.032)	(0.004)
Observations	1,329	10,491
R-squared	0.020	0.083

Notes: This table shows the relationship between thermal anomalies' intensity and satellite-derived pollution levels of Chinese polluting firms or power plants at the intensive margins. The samples are restricted to gridcells with polluting firms or power plants only. The polluting plants are defined as the MEE's Key Centrally Monitored Polluting Enterprises database, and the power plants sample is obtained from the China Emissions Accounts for Power Plants (CEAP). The dependent variable is $\ln PM_{2.5}$ defined as the natural logarithm of the cell-level yearly satellite-based $PM_{2.5}$. FRP measures the intensity of thermal-related economic activities, which is defined as the average rate of radiant heat output around a 10km radius of polluting firms and is related to the rate at which fuel is consumed and smoke emissions are released. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.