Achieving Air Pollution Control Targets with Technology-aided Monitoring: Better Enforcement or Localized Efforts?

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

Abstract

Weak enforcement of environmental regulations remains a global issue due to inadequate monitoring and misaligned incentives. This paper examines the effects of automated monitoring on achieving air pollution control targets amidst China's war on pollution. Utilizing the staggered roll-out process and remote-sensing data, we find that local governments respond to the advanced monitoring system by strategically targeting areas near monitors, resulting in a 3.2% decrease in pollution adjacent to automated monitors compared to areas farther away. Furthermore, we observe heterogeneity in response across cities with varying degrees of pre-existing data manipulation and among officials facing different incentives and public pressure.

Keywords: Air Pollution; Monitoring Technology; Environmental Regulations; Localized Clean-up; China

JEL Codes: K32, Q53, Q58, O13, R11.

[†]Yang: Urban Governance and Design Thrust, the Hong Kong University of Science and Technology (HKUST) (Guangzhou), and Department of Economics, HKUST (email: yangl@ust.hk); Lin: Department of Economics, Division of Public Policy and Division of Social Science, HKUST (email: linyt@ust.hk); Wang: Division of Social Science, HKUST (email: sojinwang@ust.hk); Peng: Society Hub, HKUST(email: fpengac@connect.ust.hk).

Acknowledgment

1 Introduction

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 that is not aligned with the intended policy goals (Zou 2021). Recent technological developments, such as the use of automated pollution monitoring systems, have greatly increased the central regulator's information capacity (Hölmstrom 1979; Greenstone et al. 2022).[1] Despite both scholarly and policy interest, evidence has been scant on whether technology-aided monitoring can lead to improved enforcement or perhaps only induce localized efforts in achieving pollution control targets.

This paper provides a comprehensive assessment of how the massive rollout of a nationwide, real-time air quality monitoring program has impacted air pollution in China, a highly polluted country, with a particular focus on the spatial dimension. In 2014, amidst public outcry over air quality, China declared an unprecedented 'war on pollution' featuring a staggered rollout of the monitoring program across three waves of cities. Led by the Ministry of Environment and Ecology (MEE), the national environmental air monitoring network expanded from 113 to 335 cities between 2012 and 2014. Additionally, the MEE centralized the planning, establishment, construction, and maintenance of all central monitors to minimize data manipulation, which was previously rampant at the local level (Andrews 2008; Chen et al. 2012; Ghanem and Zhang 2014).[2] Judging from ground monitor readings, the campaign has achieved remarkable success: the national-level $PM_{2.5}$ levels declined by about 40%

---

[1]Many countries have invested heavily in monitoring systems. Between 2011 and 2022, the number of cities with monitoring networks globally has risen from 1,100 in 91 countries to 6,000 in 117 countries, with a recent emphasis on real-time monitoring (World Health Organization 2021).

[2]Central monitors denote nationally-controlled (*guokong*) monitoring station sites.

from 2014 to 2019 (Greenstone et al. 2021). However, questions remain as to whether these achievements touted at the aggregate level are inflated by potentially a localized pollution control approach targeting monitored areas.

To analyze the role of automated monitoring in achieving pollution control targets, we draw primarily upon fine-scale remote sensing data, which can fill the spatial gaps in ground monitoring networks by detecting pollution changes from space. We obtain ten years of observations of the annual $PM_{2.5}$ concentration at a 1km by 1km resolution for the whole of China. Using a spatial difference-in-differences (DiD) design, our empirical specification consists of two layers. In the first layer, we take advantage of the staggered roll-out of the new automated monitoring system across cities to explore its impact on citywide pollution. The second layer constitutes the core of our analyses, wherein we compare areas located at different distances from monitoring stations within each city to assess how automation impacts the pollution gap between monitored and unmonitored regions.

Perhaps unsurprisingly, the roll-out of automated monitoring was not randomly assigned. Cities with larger populations and those higher in the administrative hierarchy were among the first to join the new program. We deal with the selection issue by including a rich set of pre-treatment city characteristics, including GDP, population, pollution level, indicators for environmental priority cities, number of monitors, and geographical size, interacting with time dummy variables in our regressions. Additionally, we control for potential confounding policies such as city-level concurrent pollution reduction targets. To verify parallel pre-program trends between earlier and later adopters, we utilize an event study with flexible controls.

Our main finding is that areas adjacent to automated monitors experienced a 3.2% decrease in $PM_{2.5}$ concentrations compared to those farther away. Prior to the automation, monitored and unmonitored areas exhibited very similar pollution trends. The gap between the two groups emerged after the monitoring began and grew even larger as the final assess-

2

ment deadline set by the central government approached.[3] Importantly, these results are robust to several sensitivity checks. We address potential measurement errors of satellite-based $PM_{2.5}$ data, use cities without monitoring stations as a placebo group, and apply entropy-balancing weighting (Hainmueller 2012) to ensure balance in pre-treatment characteristics between monitored and unmonitored cells within each city. Furthermore, building upon recent advances in the econometrics of staggered difference-in-differences (DiD) (Goodman-Bacon 2021; De Chaisemartin and d'Haultfoeuille 2020; Callaway and Sant'Anna 2021), we present estimates that correct for biases in traditional two-way fixed effects estimators.

To examine the city-wide impact of automation, we conduct an event study using cities' annual average $PM_{2.5}$ levels as the outcome and a concentric ring analysis that traces in detail the spatial scope of the treatment effects around the monitors. We find that as the distance from the monitors increases, pollution reduction in response to automation decays and eventually becomes insignificant. Moreover, in the outer rings, the pollution reduction shows a positive effect. These ring-based results are broadly consistent with the overall findings, as the overall pollution level in treated cities decreased slightly after they joined the automatic monitoring system, despite insignificant coefficients.

Several mechanisms could give rise to the observed localized clean-up efforts. Through a comprehensive review of local government policy documents, we have uncovered a wide range of action plans implemented in close proximity to the monitors. These plans explicitly state measures such as regulating coal use, suppressing dust through water spraying, restricting nearby traffic, prohibiting open burning and outdoor cooking, and shutting down major sources of pollution. We then empirically test for some of these channels. Leveraging a novel satellite-based thermal anomalies database that offers high-resolution, high-frequency information on local industrial activities, we find that in the vicinity of monitor sites, automation has led to a 10% reduction in industrial activities and an 8.6% reduction in the number of

---

[3]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.

days industrial plants operate. We observed little change in the intensity of industrial activities on the days when plants were operating, though. Additionally, we document an increase in relative humidity near monitors during high pollution episodes, potentially reflecting the use of water spraying.

We also find substantial heterogeneity in responses to automation across cities. We document larger pollution gaps in cities with a history of data manipulation under the old monitoring system, cities subject to more stringent pollution reduction targets, cities led by officials with stronger career incentives, and cities with less active civic participation in pollution monitoring.

Overall, the results demonstrate that enhanced monitoring technology indeed leads to significant improvements in air quality in monitored areas without a substantial reduction in city-wide pollution. This localized approach could arguably be cost-effective in a static sense, given that monitors are often located in dense areas. Nevertheless, it raises environmental justice concerns as a larger, lower-income population still resides in unmonitored areas with exposure to unhealthy air quality. Moreover, the localized cleaning and resulting monitor readings may greatly overstate the actual city-wide air quality improvement. Finally, pollution control efforts of limited scope are unlikely to be dynamically efficient once the monitored areas have been cleaned up. An important policy implication arising from our findings is that performance monitoring and evaluation should closely integrate ground-level data with other measures such as remote sensing information, the use of mobile monitors, and citizen participation in the supervision of environmental quality.

Our paper makes the following contributions. First, we contribute to the literature on environmental monitoring, regulation, and enforcement (Auffhammer, Bento and Lowe 2009; Duflo et al. 2013; Shimshack 2014; Gray and Shimshack 2020). Our paper extends the work of two previous studies (Grainger, Schreiber and Chang 2019; Zou 2021) by providing one of the first analyses that link strategic pollution reduction by local officials with the dynamic change in monitoring representativeness in China. We provide new evidence that monitoring

designs that were efficient ex-ante do not necessarily remain efficient as the responses of local agents evolve. Our findings are potentially relevant to other developed and developing countries that built monitoring networks decades ago and thus face an increasing need for a new or upgraded system.

Second, our paper is closely related 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. 2023) and in detecting pre-automation pollution data manipulation (Greenstone et al. 2022). Our results complement the former by uncovering actions that surface when fabricating data is no longer a viable option. With the localized clean-up, the accuracy of information disclosed to the public is undermined, potentially leading to sub-optimal avoidance behavior. Our study differs from the latter by focusing on the post-automation strategic behavior that arises.

Third, our paper contributes to the growing literature on the political economy of environmental regulation within the framework of the principal-agent relationship. Previous studies have found that firms and local governments respond to stringent regulations in ways that can have 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 from the goals of national regulations and offer insights into the underlying political incentives at the local level.

Lastly, this paper provides strong evidence of uneven pollution control, contributing 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). Conducting air pollution mitigation exclusively in monitored areas raises concerns that residents living far away from monitoring stations may not benefit from the improvements,

and they may even experience harm if pollution increases elsewhere as a result. Targeting improvement efforts to specific areas can worsen geographic, environmental inequality, with significant distributional consequences.

The remainder of the paper is organized as follows. Section 2 provides a brief background on environmental regulations and monitoring systems in China. Section 3 describes the main data sources. Section 4 presents the empirical strategy and estimation results. Section 5 elucidates the underlying mechanisms. Section 6 explores the role of regional heterogeneity. Section 7 discusses policy implications, and Section 8 concludes.

2   Institutional Background

The benefits of China's unprecedented economic growth in the past few decades have come at a significant environmental cost. In response to mounting pressure both domestically and internationally, the Chinese government has started to redirect its policy focus from economic growth to environmental protection. This section provides a brief overview of China's institutional background, including its political system and recent air pollution regulations. It highlights the influence of local officials' career incentives in shaping their efforts towards environmental protection.

2.1   Air Pollution Regulations and Political Incentives

China is characterized by a regional decentralized authoritarian regime (Xu 2011). Within this vast bureaucracy, orders are issued, implemented, and monitored in a top-down manner, accompanied by a strict performance-based reward-and-punishment system, where higher-level principals set performance targets for lower-level agents. Motivated by strong career concerns, local bureaucrats allocate more resources and efforts to criteria that are highly valued by the upper-level government. In the past, city leaders engaged in fierce competition for promotions by driving economic growth in their jurisdictions (Li and Zhou 2005). This

competition led to unintended actions and consequences, including inequality, collusion, corruption, and manipulation, which deviated from the central government's stated policy goals (Fisman and Wang 2015; Oliva 2015; Jia and Nie 2017; Jia 2017).

In 2014, the central government of China declared a "war on air pollution", implementing a series of regulatory policies described as the "strictest ever". This anti-pollution campaign effectively incorporated environmental targets into the evaluation of local officials along with traditional metrics such as GDP growth. The implementation of these regulatory policies was detailed through a number of official documents. The major policy document that sets the stage for China's "war on pollution" is the "Action Plan for Air Pollution Prevention and Control" (referred to as Air Ten hereafter), announced in September 2013. To better implement Air Ten at the local level, the Ministry of Ecology and Environment (MEE), the Chinese equivalence of EPA, signed Target Responsibility Agreements—essentially performance contracts ('Target')—with 31 provinces after the issuance of Air Ten. Additionally, it stipulated details on environmental performance assessment metrics with respect to Air Ten in another document ("Notice of the General Office of the State Council on Performance Assessment Measures for Air Pollution Prevention and Control Action Plan", referred to as Assessment hereafter) on April 30th, 2014. A more detailed description of these policy documents is included in Appendix A1.1.

In the agreements, provincial leaders pledge to attain certain air pollution reduction targets for the 2013–2017 period using the 2012 pollution level as the base. There are two major performance metrics: the air quality targets and the fulfillment of the ten tasks specified in Air Ten, such as vehicle or coal power plant pollution control. Notably, the air quality targets are based on the annual reduction rate in city-wide PM2.5 or PM10 concentrations. These targets are calculated by averaging the pollution readings across monitoring stations, and they vary by province. To ensure accountability and successful implementation, the provincial targets are further broken down and allocated to city governments through the target responsibility system. Regions that fail the annual assessment for reducing air pollution will

face penalties, including the summoning of local leaders for questioning, financial penalties, and the suspension of polluting project approvals.

2.2 Monitoring Systems for Ambient Pollutants

The upgrading of the monitoring system was part of the infrastructure support to enforce the Air Ten environmental regulations. 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 policies in other regions. While the central-level MEE was responsible for setting technical specifications and for publishing the data collected from monitors, local environmental authorities were tasked with managing and operating monitors and with collecting and managing the environmental surveillance data submitted to the MEE. This procedure created a non-trivial room for data manipulation at the local level (Ghanem and Zhang 2014). Typical practices included shutting down the monitors during heavily polluted days, deliberately blocking the sensors, and sabotaging the monitors.

In order to address the problem, China expanded its national air quality monitoring network to cover 335 cities. Additionally, it introduced a real-time reporting system and disclosure program. This new monitoring program brought several significant improvements. Firstly, $PM_{2.5}$ was recognized as a major pollutant. Secondly, after upgrading or constructing monitoring stations, the raw data collected by these monitors would be transmitted directly to the central system. Lastly, there were corresponding changes in the responsibilities of various environmental authorities, as depicted in Figure A11.

Concretely, the automation program was implemented in three waves from 2012 to 2014. It began with larger cities and more developed regions, such as the Beijing-Tianjin-Hebei region, the Yangtze River Delta region, and the Pearl River Delta region. The first wave, which concluded on January 1st, 2013, involved 74 cities upgrading a total of 496 monitoring stations. In the second wave, which concluded on January 1st, 2014, 116 cities installed new monitoring systems in 449 stations. The third wave concluded on January 1st, 2015, with

177 cities establishing 552 stations. Figure 1 provides a visual representation of the three phases of the program. The national air quality monitoring network consists of 1,497 central monitors and is designed to serve the urban areas of 367 cities.

To address the possibility of selective siting, locations where the monitors would be rolled out between 2012 and 2014 were centrally determined in 2012.[4] Only air quality data obtained from the central monitors count towards the cities' environmental performance.[5] Monitoring stations are located in a way that is well intended to represent the air quality of a city. According to the 2013 "Technical Regulation for Selection of Ambient Air Quality Monitoring Stations" (Ministry of Ecology and Environment 2013), urban monitors should ensure that the mean air quality readings of stations represent the city's average quality within a 10% margin of error.[6] The location choice also considers natural geography, meteorology, and socio-economic factors, notably with the number of stations hinged on a city's population size and land area. We discuss the representativeness of the monitoring system

_____

[4]There are three types of monitors in China: monitors controlled by the central government, monitors controlled by local governments, and micro-monitors to address specific polluting sources. The centrally-controlled monitors were 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.

[5]The China National Environmental Monitoring Centre (CNEMC) evaluates equipment performance and data accuracy, including surveillance cameras and the real-time reporting system in stations. Local governments must obtain CNEMC approval before purchasing and installing equipment. Surveillance cameras are used to prevent data tampering. A third party operates and maintains the monitors, providing technical support and conducting quality assessments.

[6]Stations were chosen based on data simulation from 15-day monitoring of air quality for each 2km*2km grid in a city. Each station can be representative of a circular area with a radius between 0.5km and 4km.

in Appendix Section A3, where we show it is indeed fairly representative of most Chinese cities in the baseline.

Local governments were not involved in choosing the locations of the central monitors, and the introduction of automatic air pollution monitoring has prevented the falsification of data. As a result, there has been a significant improvement in data accuracy (Niu et al. 2020; Greenstone et al. 2022). However, despite these improvements, local officials, driven by the dynamics of the principal-agent relationship, often find ways to undermine the representativeness of the monitoring system. This phenomenon is captured by the popular Chinese adage, "when the central government has a policy, the local governments have countermeasures". In our empirical analysis, we aim to uncover these hidden actions taken by local officials.

3   Data

We draw upon several primary data sources, including 1) satellite remote sensing data, which report air quality and industrial activities with fine granularity; 2) monitoring station data, which offer key information on a station's location and automation status; and 3) city socio-economic characteristics. 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.

A. Remote Sensing Data

Air Quality–We fill spatial gaps in the ground monitoring system using high-resolution images of the major air pollutant $PM_{2.5}$, which are derived from original satellite measures of Aerosol Optical Depth (AOD). 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.[7] However, clouds or shadows over certain areas can contaminate satellite images, primarily for data at high spatial and temporal resolutions. With this caveat in mind, this paper uses satellite images that include annual satellite-derived $PM_{2.5}$ concentration data from a variety of sources (Van Donkelaar et al. 2015, 2016a, 2019) in grids with a 1km by 1km resolution for a total of nine million grid cells covering the whole of China from 2008 to 2017. In Appendices A2.1 and A2.2, we discuss caveats with satellite-based measures of pollution concentrations and conduct a battery of robustness checks to address related concerns.

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.[8]

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.[9] It provides information on the type of heat spot, including whether the source is presumed to be a vegetation fire, active volcano, static land source, or offshore source. We discuss the data set in more detail in Section 5.2.

---

[7]See Liu et al. (2007), Lee et al. (2012) and Zhang and Li (2015) for more details. Previous economic research using satellite measures as a 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).

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

[9]Source: NASA Fire Information for Resource Management System.

B. Ground Monitoring and Socioeconomic Data

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, and it covers 1,497 stations from 2012 to 2017. The data include each station's name, geographical coordinates and hourly pollution readings. This allows us to measure cities' automation status precisely and gauge monitored grids, which are the key explanatory variables in our analysis.

Socioeconomic Data–To address the concern that other factors could affect the automation roll-out sequence and possibly confound the identification, we construct a rich set of control variables. First, we assemble city-level demographic and socioeconomic data from the National Bureau of Statistics (NBS), including GDP and population statistics. Second, we calculate the number of monitors for each city and the maximum distance between cells and monitors within a city based on the aforementioned geo-location of each monitor and a city-level GIS map of China. These two variables account for city heterogeneity in terms of treatments received and geographical size. Third, we obtain an environmental priority city list from the official website of the MEE.[10] Furthermore, we manually collect concurrent policies on city-level $PM_{2.5}$ and PM_{10} pollution control targets from the Air Pollution Prevention and Control Action Plans of various provinces.

To explore the heterogeneous impacts of automation, we consider three important dimensions: pre-automation data manipulation, local officials' political incentives, and local public pressure for air quality improvements. For the former, we use a regression discontinuity (RD) design that captures the sharp increase in reported pollutants immediately after the monitoring system was automated and the algorithm introduced by Greenstone et al. (2022). Based on the resulting local linear RD estimates, we construct a binary indicator of a city's pre-existing manipulation status, defined as whether its RD estimate is positive. For

---

[10]Source: Official website of the State Council of the People's Republic of China.

the measures of political incentives, we first leverage the Chinese Political Elite Database (Jiang 2018) and estimate the promotion likelihood of city leaders following the method of Wang, Zhang and Zhou (2020). Second, we create a dummy variable of a city's target status, indicating whether it has any assigned pollution reduction targets or not. Lastly, we also extract the provincial per capita number of environmental complaints in 2017, again from the MEE,[11] and use it as a proxy for public pressure.

4 The Effect of Automated Monitoring on Pollution

In this section, we examine whether the use of advanced monitoring technology helps reduce air pollution. As discussed, the real-time data reporting system leaves local governments with virtually no ability to tamper with data directly. Therefore, a cost-effective way to improve the reading of pollution monitors may be to target the monitoring sites for very localized clean-up.

4.1 Empirical Framework

Using a distance-based DiD method, we test whether the difference between the pollution reduction in monitored areas versus the un-monitored surrounding areas differs after the rollout of the automatic pollution monitoring program, which occurred in three staggered waves. The estimation specification is the following:

$$\ln(PM_{2.5ict}) = \alpha Auto_{ct} + \beta Near_i \times Auto_{ct} + \gamma X_{ct} + Cell_i + Year_t + \varepsilon_{ict} \qquad (1)$$

where i, c and t denote grid cell, city, and year, respectively. The outcome variable, $\ln(PM_{2.5ict})$, is the logarithm of the satellite-based $PM_{2.5}$ concentration. As noted in Section 2.1, pollution control targets and the associated rewards/punishments for local officials are tied to the average annual PM changes from their cities' monitoring stations. Therefore,

---

[11]Source: '12369' Environmental Reports in 2017.

we use the annual data for the baseline specification. We aggregated the nine million 1km by 1km cells into one million 3km by 3km grid cells prior to the estimation to reduce the computational burdens. $Auto_{ct}$ is the treatment indicator for the first layer of our analysis, which takes the value of one after a city c joins the automatic monitoring program, and zero otherwise.[12] By comparing early automated cities to late ones over time, the coefficient α represents the impact of automation on the overall city-level pollution[13]. For the second layer of analysis, we use a distance indicator denoted by $Near_i$, which equals one 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 from the monitor (or equivalently, an unmonitored area). The coefficient (β) is the parameter of primary interest in our paper. By further comparing areas at different distances from monitors within cities, it identifies the effects of automation on the pollution gap between monitored and unmonitored areas. 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 the air pollution observed in each cell is likely influenced by emissions elsewhere, we cluster standard errors (denoted by ε_{ict}) at the city level to safeguard against spatial correlation.

As with any DiD design, the identification of α relies on the parallel trend assumption. We exploit the fact that the sequence of the automation rollout was largely dependent on how cities were prioritized; specifically, larger cities and those higher in the administrative hierarchy tended to receive the treatment earlier. Arguably, these cities might have also been targeted by other concurrent environmental policies. In order to address these concerns, we control for the interactions between year dummies and an array of pre-treatment city characteristics, such as average GDP, population, average pollution between 2008 and 2011,

---

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

[13]Although this is not the central focus of our paper, such aggregate-level responses can be important when considering policy counterfactuals and mechanisms.

the number of monitors in each city, the maximum distance between cells and monitors within a city and indicators for environmental priority cities (which include provincial capitals), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets, denoted by X_{ct}. In doing so, we allow the temporal variation in pollution to differ flexibly across factors most relevant to the timing of automation. We further assess the validity of the parallel trend assumption by examining the pre-trends in average pollution levels between treatment and comparison cities in an event study following a variant of Equation (1):

$$\ln(PM_{2.5ict}) = \sum_{n=-4}^{3} \alpha_n \phi(n)_{ct} + \gamma X_{ct} + Cell_i + Year_t + \varepsilon_{ict} \tag{2}$$

where n defines the period relative to the automation year, $n = -4, -3, -2, 0, 1, 2, 3$. The dummy variables $\phi(n)_{ct} = \mathbf{1}\left[n \leq t \leq n+1\right]$ jointly represent the automation event that varies by city. We omit the dummy variable for $n = -1$, the year right before automation. The set of control variables is the same as in the baseline specification. It is clear from Figure 2 that there are no clear pre-existing trends in pollution levels prior to automation, lending support to our identification assumption. The estimated coefficients for $t > 0$ become negative, though marginally insignificant, which suggests that the average pollution levels for the whole automated cities slightly drop after automation.

We also conduct a similar event study analysis in the second layer of our DiD setup. The cell-level event study is implemented using a sample that covers four years prior to and three years following the automation of the monitoring system. The specification is as follows:

$$\ln(PM_{2.5ict}) = \sum_{n=-4}^{3} \alpha_n \phi(n)_{ct} + \sum_{n=-4}^{3} \beta_n Near_i \times \phi(n)_{ct} + \gamma X_{ct} + Cell_i + Year_t + \varepsilon_{ict} \tag{3}$$

where n defines the period relative to the automation year, $n = -4, -3, -2, 0, 1, 2, 3$. The dummy variables $\phi(n)_{ct} = \mathbf{1}\left[n \leq t \leq n+1\right]$ 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 3 plots the estimated coefficients β_n for the event study analysis. $PM_{2.5}$ concentrations in the monitored and unmonitored areas have similar trends prior to automation, with small and insignificant pre-automation coefficients. 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.[14]

4.2   Baseline Results

We perform our baseline analysis on a large sample of grid cells over the 2008–2017 period. 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. In column (2), after controlling for the possibility of confounding trends in the pollution that vary with pre-treatment city-level characteristics, the effects of automation on city-wide pollution changes are statistically insignificant. The distance indicator $Near_i$ is denoted by a dummy variable $\mathbf{1}$(0-3km). Grid cells within the 3km radius of a monitor are considered monitored (the indicator thus equals one), and others are considered unmonitored (the indicator equals zero).

In columns (3)-(6), we include the interaction terms of $\mathbf{1}$(0-3km) and *Auto*, capturing the differential effects of the automatic monitoring program on air pollution between monitored and unmonitored areas. The coefficient of *Auto* then represents the average effect of monitor automation on unmonitored areas. Column (3) presents results estimated from Equation (1) with only the cell and year fixed effects. The specification in column (4) includes additional controls by interacting year dummies with the pre-existing city-level population and the average level of pollution. The estimated coefficient on *Auto* drops from 0.052 to -0.001

---

[14]We also demonstrate in Figure A12 that our event study results are robust to controlling for existing pre-trends, building on recent methods by Rambachan and Roth (2019).

and is statistically insignificant, and the coefficient on $Near \times Auto$ changes from -0.091 to -0.032 but remains highly significant. In terms of magnitude, the change in the coefficient estimates suggests that earlier-automated cities tended to be more populous and polluted and were also more likely to experience an upward trend in pollution, especially in central urban areas. In our preferred specification, column (5) further controls for the interactions of year dummies with a rich set of city-level covariates (such as pre-treatment average GDP, the number of monitors for each city, the maximum distance between cells and monitors within a city and a dummy variable that indicates environmental priority city status), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. The estimated coefficients on both $Auto$ and $Near \times Auto$ remain almost unchanged from the previous column, so pre-treatment population and pollution levels appear to be the most important endogenous factors that determine the rolling sequence of automation. The estimates in column (5) suggest that air pollution in the monitored areas is 3.2% lower than in the unmonitored areas after the automation roll-out. Column (6) uses cells outside a 60km, as opposed to a 3km radius of a monitor, as control units. The coefficients on the interaction term are statistically significant and roughly similar in magnitude, confirming that our main results are robust to alternative ways of defining unmonitored areas. The coefficient on $Auto$, which captures the impact of monitor automation on city-level pollution, is negative but small and statistically insignificant. Section 4.3 will explore in further detail the spatial pattern of possible pollution relocation using a concentric ring analysis.

4.3   Uneven Pollution Control: A Concentric Ring Analysis

To better showcase the geographical scale of the clean-up efforts, we deploy a concentric ring approach under the implicit assumption that targeted pollution control measures should decay with greater geographic distance from the monitor. Specifically, we replace the binary indicator for monitored areas in Equation (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.[15]

The results are reported in Figure 4. The reference category includes grid cells more than 150km away from the closest monitor. These experienced a 4.5% increase in pollution level, as denoted by the estimated coefficient of β. The coefficient estimates for the set of bin indicators then represent the impact of automatic monitoring on air pollution in each distance bin relative to the reference group. Based on these estimates, the reduction in $PM_{2.5}$ within the 3km ring around monitors is about 4.7% (i.e., $9.2\% - 4.5\%$), and the effect reduces monotonically with distance to the monitor, slowly approaching zero at the distance bin beyond 120km. As a robustness check, we replace the outcome variable with log(AOD). A similar pattern is documented in Figure A14. In general, the spatial scale of pollution reduction suggests uneven pollution control. Nevertheless, there does not appear to be any clear pattern of pollution migration across space. Areas close to the monitors reap a much larger proportion of pollution reduction from the automatic monitoring technology than those far away (which show almost negligible change), thereby undermining the dynamic representativeness of the monitoring network and generating environmental justice concerns.

4.4 Robustness Checks

In this subsection, we conduct a series of additional robustness checks to address various empirical challenges and rule out other possible confounds discussed below.

Measurement Errors in Satellite-based $PM_{2.5}$ Data   Our primary dependent variable, the satellite-based $PM_{2.5}$ data obtained from Van Donkelaar et al. (2016b) and Van Donkelaar et al. (2016c), relies on a combination of remotely sensed Aerosol Optical Depth (AOD) measurements, ground station data, and a complex chemical transport model. While these satellite-derived pollution indicators have been widely used, recent studies have highlighted

---

[15]Figure A13 shows the distribution of the maximum distance between cells and monitors across cities. The average maximum distance in our sample is 152.29km.

concerns regarding potential measurement errors (Jain 2020; Proctor, Carleton and Sum 2023). In Appendix Section A2.1, we conduct a thorough evaluation of the quality of the satellite-derived pollution indicators through various analyses. Additionally, we address three primary concerns related to these indicators in a series of robustness checks, as presented in Appendix Section A2.2. First, we use raw AOD data instead of the imputed $PM_{2.5}$ data as the outcome variable. This adjustment allows us to account for potential measurement errors that may be correlated with the distance to monitoring stations, as the ground monitor data are utilized in the imputation process. Second, we investigate the potential aggregation bias, commonly referred to as the "ecological fallacy," associated with using annual data (Banzhaf, Ma and Timmins 2019). To address this concern, we replicate our analyses using monthly AOD-based $PM_{2.5}$ data, as well as other monthly moments such as the maximum and minimum values of daily observations from raw AOD data. Lastly, we examine the possibility of measurement errors resulting from remote sensing by applying the bias-correction method developed by Proctor, Carleton and Sum (2023). This approach helps mitigate potential biases in the satellite-derived data. By conducting these robustness checks, we enhance the validity and reliability of our findings regarding the impact of automation on pollution levels.

Heterogeneous Treatment with Staggered DiD A growing applied econometrics literature has identified potential biases in the traditional two-way fixed effects (TWFE) estimator in a staggered difference-in-differences design. The biases may emerge when the early treated units serve as control groups for the later treated ones, especially in the presence of heterogeneous dynamic treatment effects (Goodman-Bacon 2021; De Chaisemartin and d'Haultfoeuille 2020). These concerns are relevant to our estimation of the effects of automatic monitoring on pollution. Although the main specification is a triple difference design, we essentially leverage the staggered program roll-out. To address this issue, we perform a Goodman-Bacon decomposition of the DiD estimation, where we focus on the pollution gap, defined as the difference in pollution levels within 3km of a monitor in a given city and that

city's average pollution level. As a robustness check, we exclude the monitored treated in wave three from the treated group, which is identified as the main contributor to the bias associated with the TWFE estimator. We also employ an alternative estimation approach proposed by Callaway and Sant'Anna (2021). The details of the additional analyses are reported in Appendix Section A2.3.

Placebo Tests We perform three placebo tests to verify our findings. The first and most important test involves the use of nine cities without monitoring automation. We identify "placebo" monitor spots within these cities, adhering to existing siting rules, and assign their automation timings to either Wave 2 or 3.[16] A regression analysis with the same source of satellite data shows that air quality near the "likely" monitor spots in these non-treated cities does not experience a trend break compared to those farther away before and after station "automation" (see Appendix Table A11).[17] In two additional placebo tests, we randomly 1) generate automation timing within the sample period for each monitor and 2) assign monitor sites to locations (see Appendix Figure A15). The distributions of placebo estimates reinforce our principal finding, with the benchmark coefficient (-0.032) lying outside the 99% confidence intervals.

Alternative Explanation We evaluate an alternative explanation other than strategic cleaning that could yield differences in pollution reduction between monitored and unmonitored areas. It is possible that monitors are often located in highly polluted areas, with pollution dynamics exhibiting mean reversion due to air dispersion. If so, one should ex-

---

[16]Specifically, we assign counterfactual monitors to the location of either the municipal Environmental Protection Bureau, the municipal government building, or both.

[17]In another exercise, we use the other type of monitoring stations for cross-validation. The results are reported in Appendix Figure A16, again lending support to our main finding.

pect that post-automation monitored cells located in dirtier areas of a city may show larger pollution reductions than those located in cleaner parts. We check our results against these alternative scenarios with a regression accounting for pre-existing pollution conditions:

$$\ln(PM2.5_{ict}) = \alpha Auto_{ct} + \beta Near_i \times Auto_{ct} + \eta Near_i \times Auto_{ct} \times Dirtier_i + \gamma X_{ct} + Cell_i + Year_t + \varepsilon_{ict}$$
(4)

where $Dirtier_i$ is an indicator that equals one if cell i is more polluted than the city average in the base year 2008. The results are reported in Columns (1)–(3) of Table 3, where we find consistent estimates on the main parameter of interest and no significant difference between clean and dirty monitored areas, ruling out competing explanations of mean reversion in pollution.

Entropy Balancing between Monitored and Unmonitored Cells   To directly address the potential bias due to selective monitor siting, we employ the entropy balancing technique (Hainmueller 2012; Hainmueller and Xu 2013)–a generalized weighting method–to enhance the comparability of monitored and unmonitored cells. By selecting control grid cells that closely match monitored ones based on confounding factors such as proximity to polluters, cell-level population, and 2008 GDP, this approach ensures a balance in the mean and variance of these variables between the monitored and unmonitored cells within each city.[18] Column 4 in Table 3 reports the result: automation leads to a 3.1% increase in the pollution gap, a consistently significant impact similar to our baseline estimate.

[18]The technique is similar to propensity matching, but it has an attractive feature in 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.

5   Mechanisms of Localized Clean-up

As noted in Section 2.2, manipulation of environmental data was previously prevalent among local governments. Although the new air quality monitoring system has made direct tampering with equipment virtually impossible, manipulation cannot be completely eliminated. In this section, we analyze the potential channels by which differential spatial patterns of air pollution control may have emerged.

5.1   Evidence of Strategically Targeted Reductions

To gain first-hand evidence on the channels of strategic cleaning, we extensively reviewed a large number of government policy documents. To this end, we query the China Law Journal Database,[19] which hosted the entire corpus of more than three million government documents, including policies and legal regulations. We adopt a two-step methodology where we begin by searching for keywords such as "near monitoring stations" and their synonyms. Following this, we manually validate if the search results correspond to pollution control action plans that were issued by local authorities.

The analysis of our search outcome reveals a striking finding. It is not uncommon for local governments to put down strategic cleaning and the targeted measures explicitly in their official documents. We managed to locate 121 documents from 72 cities that directly mentioned "atmospheric pollution control around monitoring station". Among these documents, 42 of them even specified pollution control measures within a specific distance range of 0.05–10km from the monitoring station. Table A12 provides an overview of the targeted measures mentioned in these 121 documents, which include regulations on coal use near the monitors, dust suppression through water spraying, traffic restrictions in the vicinity, bans on open burning and outdoor cooking, and the closure of major polluting plants. Fur-

---

[19]Source: PKULAW (Beidafabao, in Chinese) maintained by the Legal Information Center of Peking University.

thermore, our findings are supported by several newspaper articles that corroborate these measures. Appendix A1.2 presents a more detailed analysis of the document contents, along with anecdotal evidence from newspapers. In Appendix A1.3, we provide examples of official documents, including the original text and our translations, highlighting the sections related to strategic cleaning.

5.2 Empirical Test of Strategic Cleaning Mechanisms

Change in Industrial Activities In this subsection, we examine how monitoring automation affects the spatial patterns of industrial activities, utilizing a novel satellite-based measure of the *intensity* of industrial activities at high spatial and temporal resolutions. We draw upon the MODIS Version 6 Global Monthly Fire Location Product, which offers the number and strength of thermal anomalies on static land, leveraging the fact that most industrial plants emit high-temperature waste gas, creating thermal hotspots above them.

We create three measures of thermal-based industrial activities. *Fire Radiative Power (FRP)* is a measure of the average intensity of thermal anomalies, defined as the rate of radiant heat output, the rate at which fuel is consumed and smoke emissions are released. *Days* measures the annual number of days with active thermal anomalies in each cell. *1(Thermal Anomalies Presence (TAP))* is a dummy variable that equals one if there are thermal-related economic activities in a given cell in a given year. Figure 5 presents the high correlations between detected thermal anomalies and the locations of major air-polluting firms and power plants. We further explore the validity of the thermal anomaly measure for the measurement of local pollution in greater detail in Appendix A2.4.

We adopt a similar specification as in the baseline model (Equation 1), replacing the outcome variables with measures of the presence of thermal anomalies or their intensity at cell i in year t. To address the issue that a large proportion of grid cells have no thermal anomalies, we estimate a Poisson count data model for a set of outcomes. The results are reported in Table 4. Column (1) presents estimates of a logit model where the dependent variable switches to one with thermal anomalies present (TAP) in grid cell i in year t. The marginal effect

23

estimates suggest that automation leads to a 10.1% reduction in the probability of observing industrial activities within a 3km radius of the monitors relative to unmonitored areas. A possible explanation is that industrial sources closer to the monitors were more likely to be closed down following automation. Columns (2) and (3) report results from the Poisson estimation, where the dependent variables are the log number of days with active thermal anomalies and the average intensity of thermal anomalies, respectively. The coefficients of interest are on the interaction of $Near_i \times Auto_{ct}$, which are negative and significant for both outcomes. Columns (4) and (5) examine the effects of the automatic monitoring system on the intensive margin by restricting the sample to grid cell-year observations with active thermal anomalies. The coefficient on the interaction, β, is negative and significant when the dependent variable is the log number of days with active thermal anomalies. However, when the dependent variable is the average intensity of thermal anomalies, the estimated magnitude is small, and the coefficient is statistically insignificant. These findings indicate that monitoring automation is likely to reduce the operating days of nearby plants but not the intensity of activities on the days plants are operating.

Water-Spraying   Next, we look for corroborating evidence on the "water spraying" channel frequently cited in government documents. Table A13 evaluates the effects of the monitor automation program on satellite-based relative humidity using meteorological data from He et al. (2020). Clearly, automation increases the relative humidity near monitors, especially during the winter when the particulate matter pollution tends to be more severe, and water spraying is more likely to be deployed.

6   Heterogeneity in Localized Clean-up

By far, we offer compelling evidence that targeted measures have been taken by local regulators to reduce air pollution, specifically in areas 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 level of compliance with air pollution reduction policies, leader characteristics and designated targets for pollution reduction that proxy for their political incentives, and public pressure for cleaner air.

a) Underreporting before Automation

To convincingly connect the observed strategic behavior to local governments' responses to the introduction of stricter environmental regulations, we focus on cities that already had monitors before the automation rollout and divide those cities into two types: cities with a high degree of pre-automation data manipulation, and cities without prior manipulation. The conjecture is that, for cities that did not previously engage in data tampering, the introduction of automated pollution monitoring would have little impact on their de facto regulatory stringency. In contrast, for cities detected to have engaged in data manipulation in the past, automation shut down their channel for achieving "effortless perfection"—meeting environmental standards by misreporting pollution reading.

Out of the 335 cities that had 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. (2022) to classify whether or not their environmental data had previously been manipulated. That measure was obtained from a regression discontinuity design that captured the sharp increase in *reported* pollutants immediately after the monitoring system was automated.[20]

Figure 6 presents the heterogeneity in the pollution gap between cities that previously engaged in data manipulation and those that did not. As shown, we observe a pollution gap only in those cities that previously manipulated data and not in the others. This pattern is consistent with strategic localized targeting of clean-up as a direct response to the automated

---

[20]For cities that had monitors before automation, manipulation status is defined by whether the local linear RD estimate is positive, as calculated using the algorithm in Greenstone et al. (2022).

monitoring technology.

b) Local Officials' Political Incentives

As we have noted in Section 2.1, the salient points in the regulatory changes all help establish a closer tie between monitored pollution levels and the officials' performance assessment. Under a target-based responsibility system, local officials are faced with stronger incentives to achieve a reduction in monitored PM concentrations.

Motivated by these institutional details, we conduct two empirical exercises to shed light on local governments' incentives for strategic cleaning. First, we evaluate the heterogeneity in strategic cleaning across city leaders with different likelihoods of promotion–a proxy for their career incentives. Aided by the Chinese Political Elite Database (Jiang 2018), we estimate the promotion likelihood for leaders (specifically, party secretaries) of each city by following the lead of Wang, Zhang and Zhou (2020).[21] The hypothesis is that local officials are more likely to prioritize strategic cleaning efforts if they have a higher probability of promotion compared to their "effective" competitors within the same province (Li et al. 2019). The heterogeneity analysis, displayed in Panel (a) of Figure 7, reveals that city leaders with a higher likelihood of promotion than the provincial median indeed demonstrate a greater inclination toward strategic cleaning after the implementation of automated monitoring.

Second, we further explore whether pollution reduction targets matter. Specifically, we examine the impact of pollution reduction targets on their strategic cleaning practices by comparing cities with pollution reduction targets to those without. With no pollution re-

---

[21]Specifically, we regress the promotion dummy (which equals one if the city leader was promoted to a higher-level position by the end of their term) on various characteristics of the leader (age, political hierarchy level when they entered the office and educational attainment). We then use the estimated coefficients, which are independent of the leaders' *ex post* performance, to predict the *ex ante* promotion likelihood of each city leader in our sample.

duction targets, cities would only be required to maintain their baseline levels of pollution and would therefore face less pressure. The pollution reduction target for each city is obtained from the provincial pollution control plan. In Panel (b) of Figure 7, cities are divided into two groups based on whether they "have any assigned pollution reduction targets during 2013-2017" or not. We observe greater strategic cleaning responses among cities in the group with pollution reduction targets.

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, while automated monitoring technology can mitigate the principal-agent problem, it can also induce new strategic reactions from local agents. Some have argued that civic engagement and bottom-up supervision could complement the current regulatory regime. In this section, we seek to relate regional heterogeneity in strategic clean-up responses to local public pressure, measured by the number of air quality complaints from local residents.

The environmental authorities in China have begun to pay 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. Online and social media platforms with similar functions have also been established. As of 2017, more than 600,000 complaints had been lodged. The MEE requires that each reported case be resolved within 60 days, and 99% of documented cases met this deadline. For our analysis, we constructed 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.[22] Figure 8 presents differential responses to monitor automation across provinces with varying

---

[22]On the supply side, the functioning and responsiveness of the hotline or online complaint platform play an important role in influencing citizen engagement. On the

levels of environmental complaints. We found that a stronger strategically targeted response, represented by wider gaps in the effects of automation between monitored and unmonitored areas, took place in provinces that had low levels of civic engagement in environmental issues. In provinces with more "participatory" environmental monitoring, any industrial pollution, whether emitted by plants in monitored or 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 when it comes to environmental monitoring.

7   Discussion and Policy Implications

7.1   Could Strategic Cleaning be Efficient?

Our main finding is that after the upgrade to automated monitoring systems, performance evaluation based solely on in-situ monitors leads to localized clean-up efforts. This localized clean-up can be rationalized due to the placement of monitoring stations in densely populated urban areas with high pollution levels, where the marginal social cost of pollution is high. We provide evidence of this in Table A14, which shows that areas within 3 kilometers of monitors contain 12.4% of China's population. However, the marginal benefit of local pollution (equivalently, the marginal cost of abatement) could also be high in these central locations due to higher agglomeration and industrial productivity, making it unclear how pollution should be optimally allocated across space. Furthermore, even if it were efficient to prioritize cleaning up near monitors in a static sense by considering the location-specific cost and benefit of pollution control, it is unlikely to remain dynamically efficient because the incentive gap between cleaning up in monitored and unmonitored areas persists even after the monitored areas have been cleaned up. As shown in Figure A17, a large group of

---

demand side, local residents may pay varying degrees of attention to the surrounding pollution, which in turn drives their willingness to lodge a complaint.

the population living far away from the monitors is exposed to unhealthy levels of pollution. From the perspective of aggregate impacts, localized clean-up fails to yield a positive externality that could benefit broader regions and populations, such as the unmonitored areas of the same city, neighboring cities, or even countries (Heo, Ito and Kotamarthi 2023).

In the following subsection, we proceed to explore two additional costs of strategic cleaning: the monitors' representativeness in terms of capturing citywide air quality and environmental justice. We then offer policy recommendations for the design of a more effective monitoring and evaluation scheme.

7.2 Consequences of Uneven Pollution Controls

A Re-examination of Environmental Performance Principals such as the central government or higher-level regulators often have to rely on monitor readings to reward or punish agents. Therefore, the localized clean-up could lead to biased policy evaluations and subsequent actions.[23]

As an illustrative example, we utilize remote sensing data to reassess the policy goals set by the "Air Ten" action plan in the present context. The original plan, published in 2013, aimed at a 10% reduction in the PM_{10} concentration nationwide and a 25% reduction in the $PM_{2.5}$ concentration in the Beijing-Tianjin-Hebei region by 2017. A progress report released in 2017 stated that both goals had been achieved with flying colors, documenting a 22% reduction in PM_{10} nationwide and a 40% reduction in Beijing-Tianjin-Hebei (hereafter referred to as BTH).

A cross-examination of the satellite-based pollution measures reveals a somewhat different story. As shown in Table 5, the city-wide recorded reductions in satellite-based $PM_{2.5}$ from 2013 to 2017 tend to be much lower: 2.592% for the whole country and 14.638% within the BTH region. Of course, the discrepancy could be driven by some other factors, such

---

[23]We provide a comprehensive analysis of the worsening spatial representativeness of the monitors in Appendix A3.

as inherent differences between satellite and ground-based measures. To highlight that the localized clean-up around the monitors is likely a key factor here, we leverage the full coverage of satellite data to compare reductions in satellite-based pollution directly at the monitored areas to reductions in rings at different distances from the monitored areas and for the whole city. Specifically, for the BTH region, the recorded reductions in satellite-based $PM_{2.5}$ drop monotonically from 21.82% at the monitored areas (i.e., those 3 by 3km grid cells where the monitors are situated) to 17.13%, 12.83% and 5.49% in cells that are within 3–30km, 30–60km, and 60–90km concentric rings to the monitors, and eventually down to 1.39% in cells beyond 90 km away from the monitors. Reassuringly, a similar pattern has been documented for the whole country, with the recorded satellite-based pollution reductions decreasing monotonically from 5.03% in the monitored cells to 0.29% in cells over 90km away from the monitors.

Overall, localized clean-up explains about 30% to 50% of the satellite-based reductions in PMs from 2013 to 2017. If we scale the ground-based monitor measures down by the same margin, then the "representative" reductions in PMs will be 11% for the whole country and 27% for the BTH region, barely meeting the initial goals.

Distributional Impacts The uneven pollution control could generate an uneven distribution of cleanup benefits among residents. Figure A18 shows the monitor share—defined as the percentage of counties with monitors—plotted against income (Panel a) and urbanization rate (Panel b) across 50 county groups ranked by income or urbanization rate, respectively. This figure conveys a clear message: more developed and urbanized counties tend to contain more monitors, and the slope of the curve is steeper at higher income ranges. As a result, strategic cleanup tends to direct a greater proportion of the health and amenity benefits from pollution reduction to high-income and urban residents.

The distributional consequences of uneven pollution control could be further amplified by large disparities in the utilization of health services, awareness of pollution-related health risks, and prevention measures between rural and urban areas. For instance, Li et al. (2018)

have found that urban adults in China are more than three times more likely to access health care than their rural counterparts. The increasing urbanization rate notwithstanding, a preponderant proportion of the population still lives in rural, largely unmonitored areas.

Appendix Table A14 reports results on a back-of-the-envelope calculation of the monetized health benefits (i.e., avoided mortality and morbidity costs) of automation for different distance bins from the monitors. To evaluate the impact of strategic cleaning on mortality, we obtain the estimate of the $PM_{2.5}$-attributable mortality rate from He, Liu and Zhou (2020), which found that a 10-μg/m3 increase in $PM_{2.5}$ increases monthly mortality by 3.25 percent. Regarding the morbidity impact, we follow a comprehensive study by Barwick et al. (2018), which estimated that a medium-run reduction of 10-μg/m3 in daily $PM_{2.5}$ would lead to $22.4 annual savings in healthcare spending per household. Using the preferred specification (i.e., Column (4) in Table 2), the annual mortality and morbidity benefits per capita are 126.22 and 2.10 USD for residents within a 3km radius of the automated monitors, more than double the benefits accrued to those living 60 km away from the monitors. The total monetized benefits from reduced pollution-related mortality and mobility are 67.53 and 1.12 billion USD annually, much higher than the annualized set-up and maintenance costs of the automated air monitoring system.[24] However, this comparison does not necessarily imply that the monitoring network passes the cost-benefit test. We caution that various forms of cleanup efforts could also incur substantial abatement costs, the estimation of which is beyond the scope of our paper.

7.3 Policy Suggestions

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 insufficient to preclude local officials' strategic targeting of monitor sites for localized cleanup. In an ideal scenario

---

[24]MEE spent 2.7 billion USD to build the new monitoring system and the annual maintenance cost of each monitor is around 22,000 USD. Source: Sohu.

with monitoring network coverage almost everywhere, local officials would find it impossible to predict the exact sets of monitors used by the upper-level authorities to evaluate their environmental performance. Consequently, they would be left with only one choice: to work harder to improve air quality citywide. Yet, the real world does not function in this way, and governments face budget constraints. The large size of the ground monitoring system makes it 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.[25]

We caution that the utilization of satellite images should not be taken 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. Meanwhile, ground monitors can provide more detailed hourly observations and better accommodate changing weather conditions. However, satellite data offers evidence of the dynamic spatial representativeness of the ground-monitor data, which could help evaluators better interpret the monitored data. 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 when evaluating city-wide pollution. Overall, incentive and information issues apply to the design, regulation, and enforcement of any monitoring system. Importantly, contributions from multiple sources, including the ground monitoring system, remote-sensing technologies, mobile monitors, observations by the public, and third-party auditors, are needed to achieve a better regulatory outcome.

Public Participation China's environmental governance has long been dominated by two major players—the government and firms—while the public has largely been shut out. Given

---

[25]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.

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 engage local citizens in the management of pollution more actively.

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.[26] Regulators could in turn, gather more accurate information and take corresponding actions. As a case in point, local governments in Qingdao and Linyi in Shandong Province have provided speedy follow-up to public claims on their social media (Weibo) accounts.[27] 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 the monitors and check the consistency of the data.

8   Conclusion

Weak enforcement of environmental laws remains a global issue. At the heart of it lies a classic principal-agent problem between central and local governments. Imperfect monitoring and misaligned incentives could result in strategic compliance at the local level. While advanced monitoring technology has been proposed as a solution to this problem, there has been very little direct empirical evidence of its efficacy.

We offer comprehensive evidence of the impact of a large-scale automated monitoring

---

[26]Some e-platforms request that the public submit evidence such as geo-coded images of pollution occurring produced by a specific industry or business.

[27]See the news report by Xinhua for details.

program on local air pollution in China. Our study finds that areas near automated monitors experienced a 3.2% decrease in pollution relative to areas farther away. Although the quality of the pollution data from monitoring sites has substantially improved, the spatial gaps in monitoring coverage appear to incentivize local officials to take localized pollution control measures that target only monitored areas rather than engaging in systematic reductions that would benefit broader regions and populations. Such localized efforts weaken the spatial representativeness of the monitoring system, leading to biased policy evaluations and affecting environmental justice in the long run. Our study also reveals that in the Chinese context, political institutions and incentive scheme designs, of which career concerns are a key component, appear to affect the degree of strategic conduct at the local level. The pollution gap between monitored and unmonitored areas is shown to be larger for cities with a history of data manipulation, in cities with officials facing stronger career incentives, and in cities with less civic participation in pollution monitoring.

Our analyses present important implications for the design of effective regulatory methods that incentivize meaningful changes to limit pollution. Central regulators should complement 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 citywide. In particular, it may help when the pressures come from outside the system. When pollution is publicly observed 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.

Auffhammer, Maximilian, Antonio M Bento, and Scott E Lowe. 2009. "Measuring the effects of the Clean Air Act Amendments on ambient PM10 concentrations: The critical importance of a spatially disaggregated analysis." *Journal of Environmental Economics and Management*, 58(1): 15–26.

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, Deyu Rao, and Nahim Bin Zahur. 2018. "The Healthcare Cost of Air Pollution: Evidence from the World's Largest Payment Network." NBER Working Paper 24688.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou. 2023. "From Fog to Smog: the Value of Pollution Information." Conditional acceptance at *American Economic Review*.

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.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

35

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." *NBER Working Paper 26659.*

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110(9): 2964–96.

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.

Fan, Maoyong, Guojun He, and Maigeng Zhou. 2020. "The winter choke: coal-fired heating, air pollution, and mortality in China." *Journal of Health Economics*, 71: 102316.

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*, 225(2): 254–277.

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*, 5(1): 3–24.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2022. "Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution." *American Economic Review: Insights*, 4(1): 54–70.

Greenstone, Michael, Guojun He, Shanjun Li, and Eric Yongchen Zou. 2021. "China's war on pollution: Evidence from the first 5 years." *Review of Environmental Economics and Policy*, 15(2): 281–299.

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political analysis*, 20(1): 25–46.

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

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

He, Guojun, Tong Liu, and Maigeng Zhou. 2020. "Straw burning, PM2. 5, and death: Evidence from China." *Journal of Development Economics*, 145: 102468.

He, Jie, Kun Yang, Wenjun Tang, Hui Lu, Jun Qin, Yingying Chen, and Xin Li. 2020. "The first high-resolution meteorological forcing dataset for land process studies over China." *Scientific data*, 7(1): 25.

Heo, Seonmin Will, Koichiro Ito, and Rao Kotamarthi. 2023. "International spillover effects of air pollution: evidence from mortality and health data." *NBER Working Paper 30830*.

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

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.

Jain, Meha. 2020. "The benefits and pitfalls of using satellite data for causal inference." *Review of Environmental Economics and Policy*.

Jiang, Junyan. 2018. "Making Bureaucracy Work: Patronage Networks, Performance Incentives, and Economic Development in China." *American Journal of Political Science*, 62(4): 982–999.

Jia, Ruixue. 2017. "Pollution for Promotion." *Unpublished manuscript.*

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.

Li, Jiajia, Leiyu Shi, Hailun Liang, Gan Ding, and Lingzhong Xu. 2018. "Urban-rural disparities in health care utilization among Chinese adults from 1993 to 2011." *BMC health services research*, 18(1): 1–9.

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

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.

Li, Xing, Chong Liu, Xi Weng, and Li-An Zhou. 2019. "Target Setting in Tournaments: Theory and Evidence from China." *The Economic Journal*, 129(623): 2888–2915.

Minitry of Ecology and Environment. 2013. "Technical regulation for selection of ambient air quality monitoring stations (on trial)."

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.

Proctor, Jonathan, Tamma Carleton, and Sandy Sum. 2023. "Parameter recovery using remotely sensed variables." National Bureau of Economic Research.

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

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.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Van Donkelaar, Aaron, Melanie S Hammer, Liam Bindle, Michael Brauer, Jeffery R Brook, Michael J Garay, N Christina Hsu, Olga V Kalashnikova, Ralph A Kahn, Colin Lee, et al. 2021. "Monthly global estimates of fine particulate matter and their uncertainty." *Environmental Science & Technology*, 55(22): 15287–15300.

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*a*. "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, Michael Brauer, N. Christina Hsu, Ralph A. Kahn, Robert C. Levy, Alexei Lyapustin, Andrew M. Sayer, and David M. Winker. 2016*b*. "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, Michael Brauer, N Christina Hsu, Ralph A Kahn, Robert C Levy, Alexei Lyapustin, Andrew M Sayer, and David M Winker. 2016*c*. "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.

Wang, Zhi, Qinghua Zhang, and Li-An Zhou. 2020. "Career incentives of city leaders and urban spatial expansion in China." *Review of Economics and Statistics*, 102(5): 897–911.

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.

World Health Organization. 2021. "WHO Global Ambient Air Quality Database."

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." *Unpublished manuscript.*

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
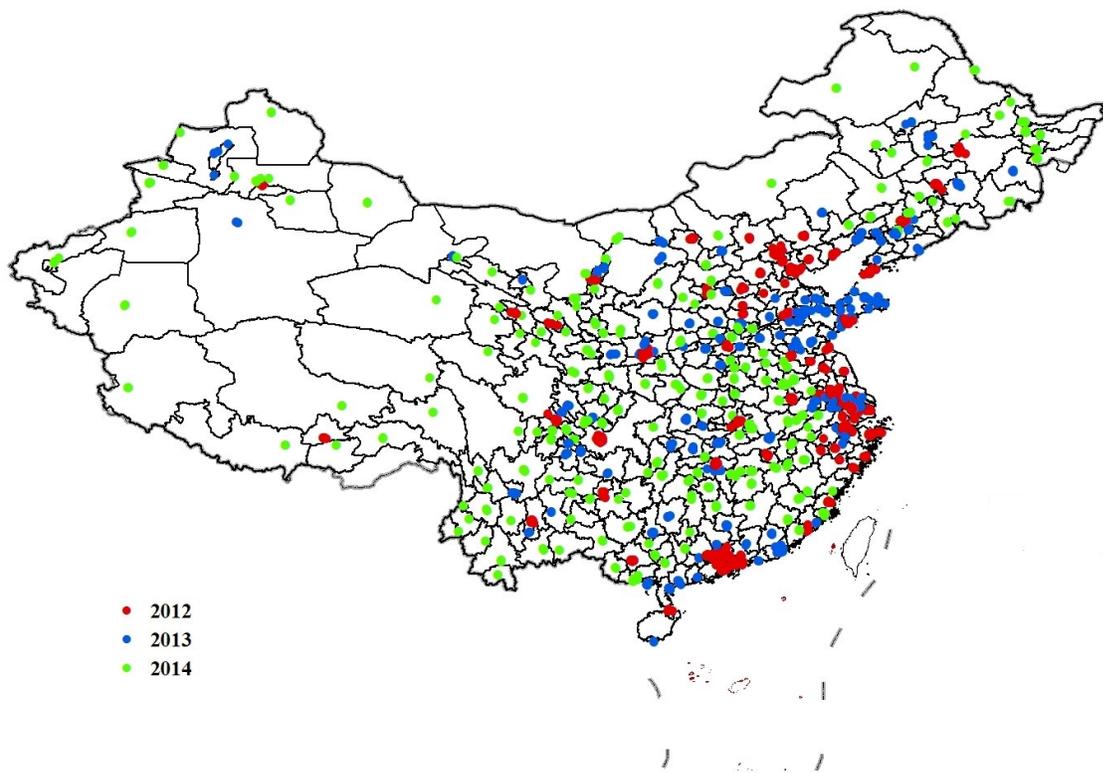
Notes: This figure displays the spatial distribution of monitoring stations that were auto-mated in three waves, which took place in 2012, 2013 and 2014, respectively.

Figure 2: Event Study: The Effect of Monitor Automation on City-Level Pollution
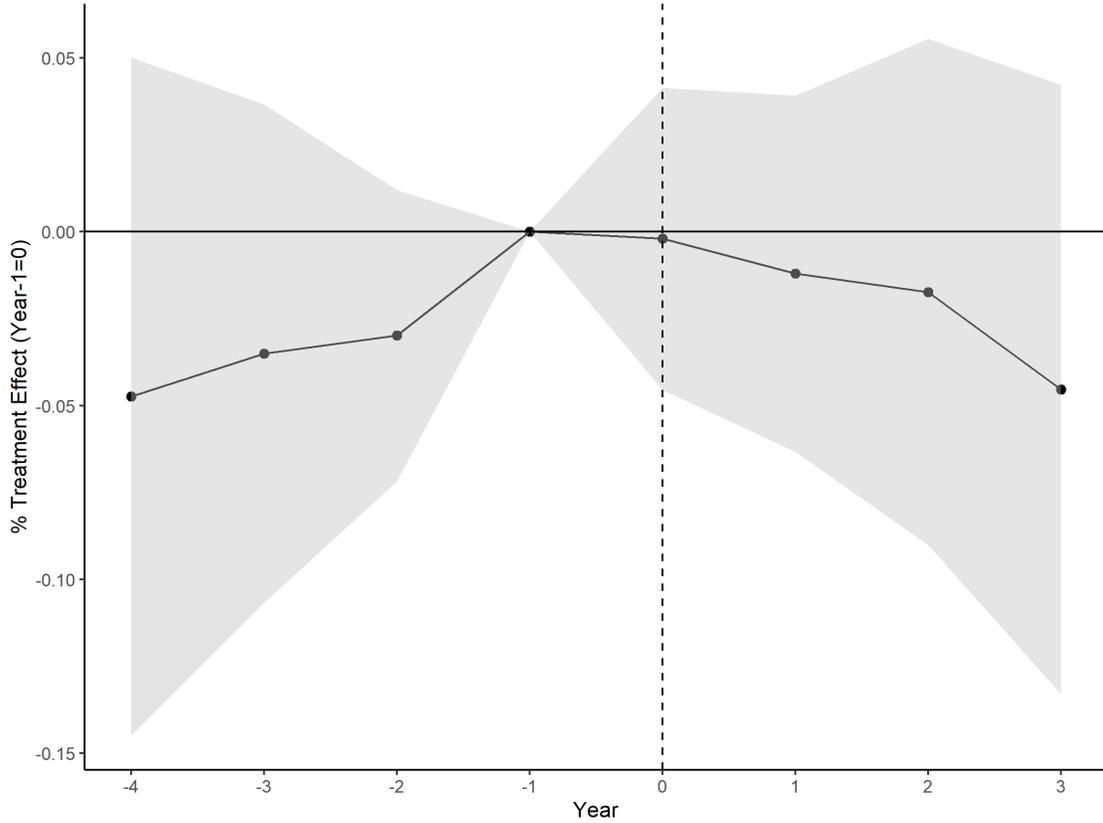
Notes: This figure presents the regression coefficients and their 95% confidence intervals from an event study of the monitor automation's effect on the pollution level of all grids within a city, following Equation 2. The omitted time category is the year before a city joined the automatic monitoring program. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population and $PM_{2.5}$ at the city level from 2008 to 2011 and the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure 3: Event Study: The Effect of Automation on Air Pollution within 3km of A Monitor
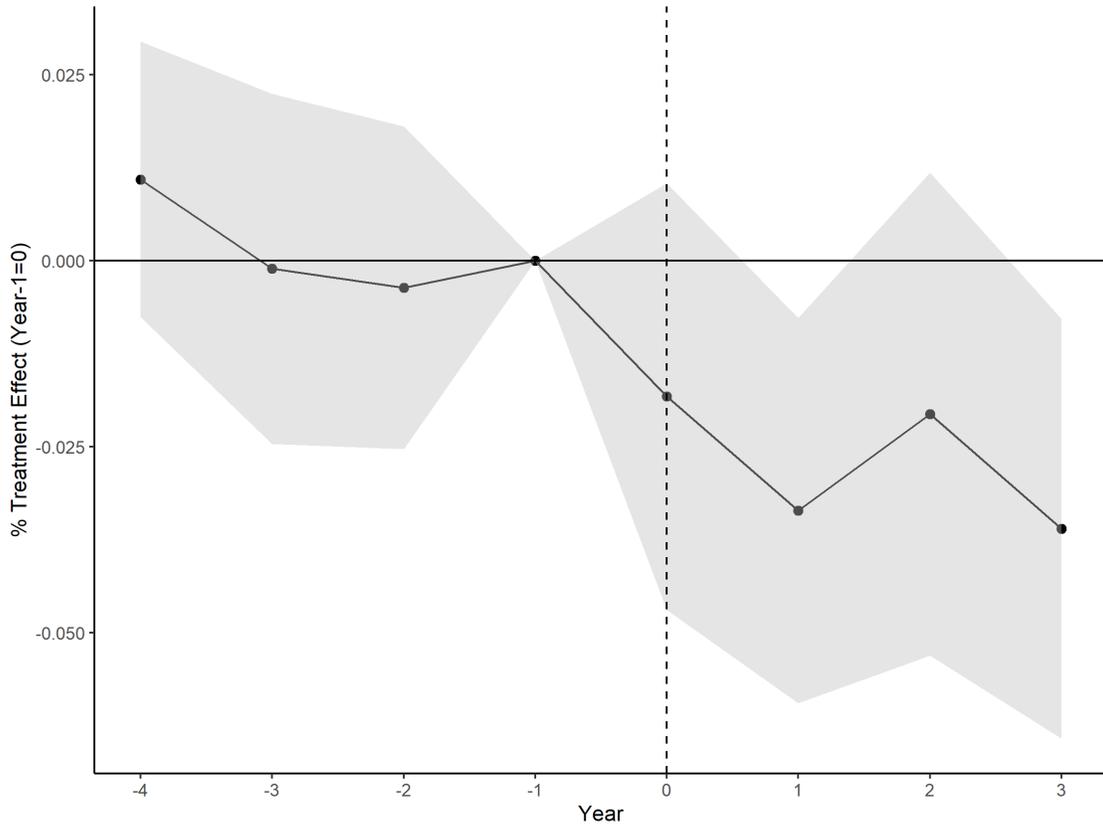
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for β_n from Equation (3). Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside the 3km radius) at a given period (also reported in Table A10). The omitted time category is the year before a city joined the automated monitoring program. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, and $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure 4: Effects of Automation on lnPM$_{2.5}$ at Different Distances from Monitors
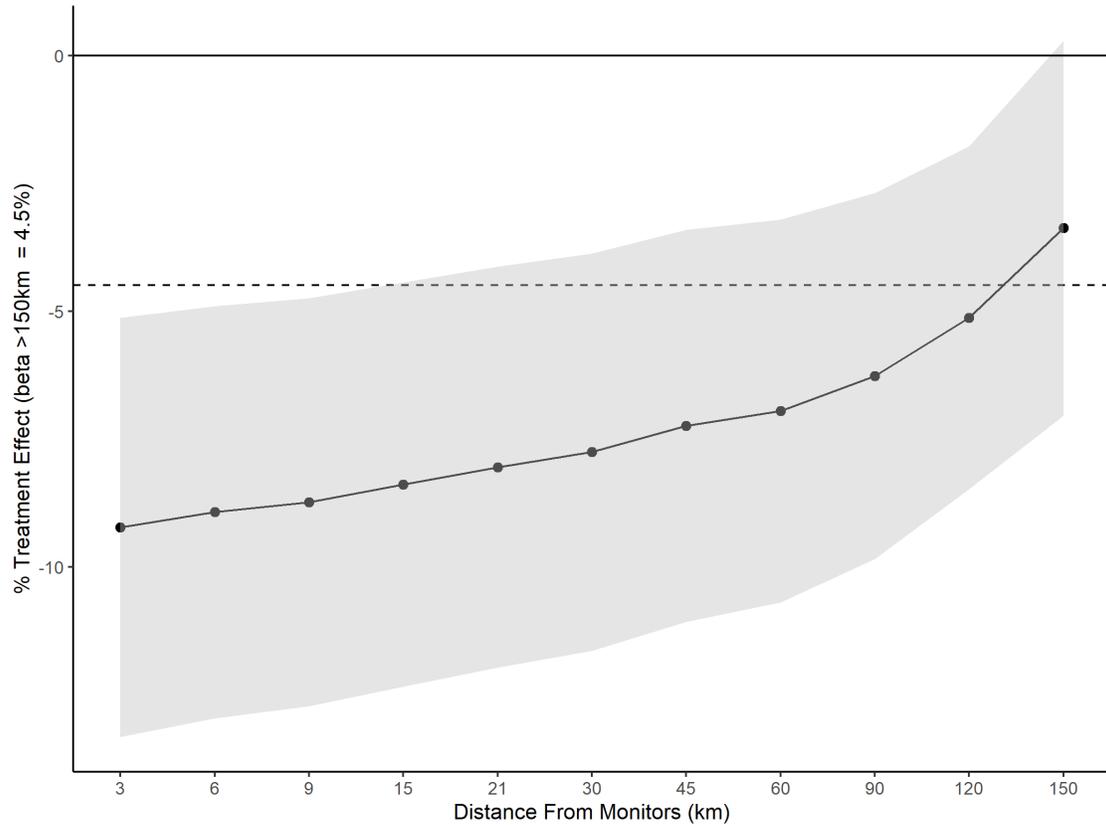
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the effects of monitor automation on the satellite-based lnPM$_{2.5}$ at different distance bins from the monitor. Each point estimate represents the pollution change in each distance bin relative to the baseline group at the outer range (distance to monitor >150 km), which on average experiences a 4.5% pollution increase. The absolute effect becomes positive above the dotted line. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, PM$_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure 5: The Location of Thermal Anomaly Hotspots and Plants

Enterprises subject to Intensive
Monitoring and Control of the State
in 2016

Thermal Anomalies Point in 2016

(a) Industrial Plants

Power Plants in 2014
Thermal Anomalies in 2014

Power Plants in 2017
Thermal Anomalies in 2017
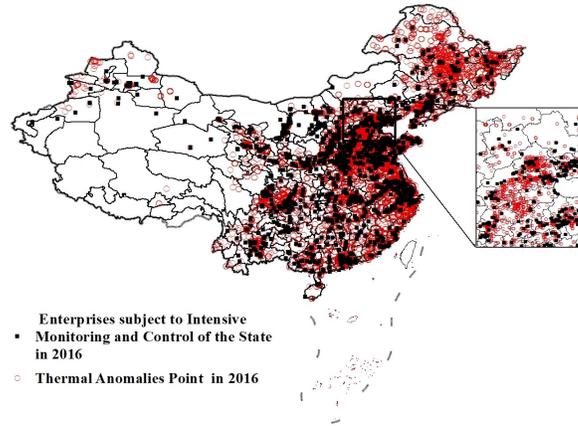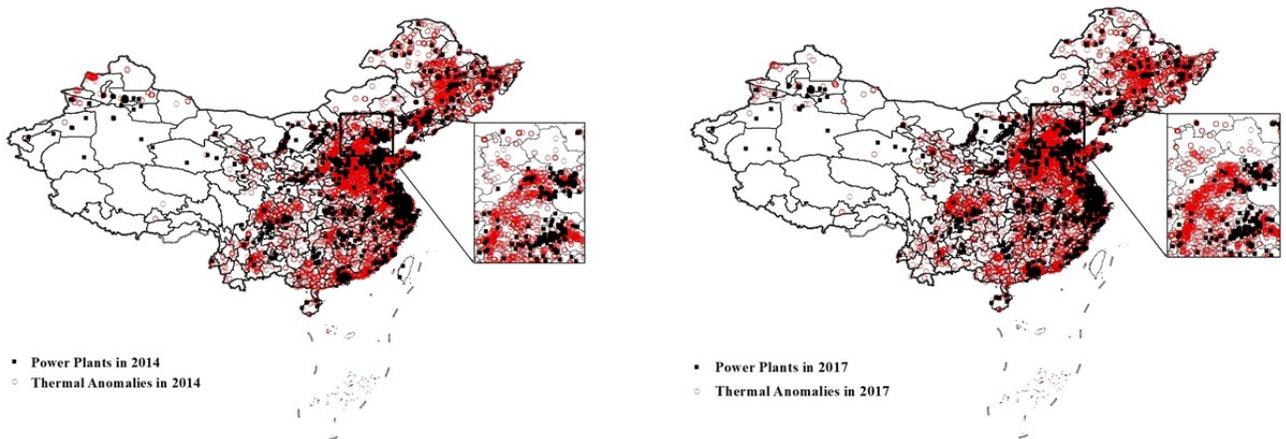
(b) Power Plants

Notes: These figures compare the locations of thermal anomaly hotspots and ma-
jor polluting plants. Panel (a) compares the location of 1,829 heavily polluting
industrial plants to the presence of 20,134 satellite-based thermal anomaly 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 anomaly points in 2014 and 2017, respectively. The power
plant locations 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 inset represents the Beijing-Tianjin-Hebei
Urban Agglomerates.

Figure 6: Heterogeneity Analysis by Pre-Automation Status: Event Study Results
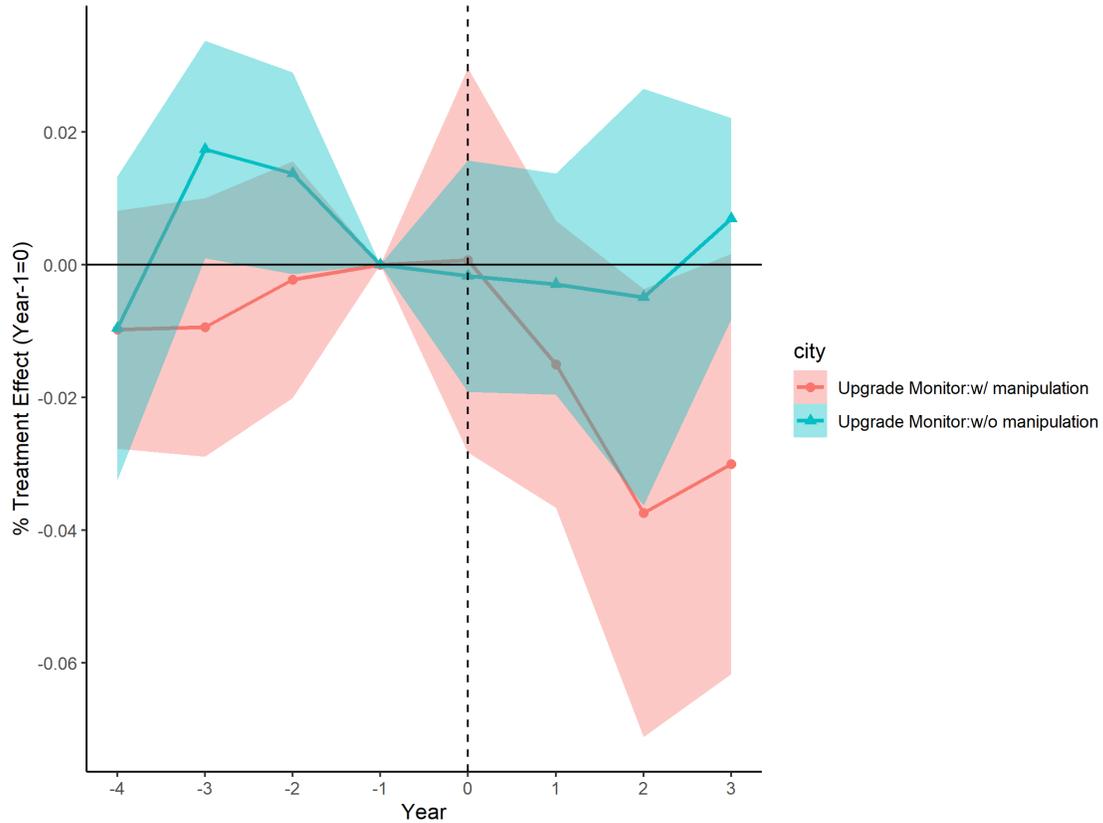
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the interaction terms between year-specific automation treatment effects and the indicators for three subsamples of cities. Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells 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. 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. (2022). "Upgraded Monitor: w/ manipulation" refers to cities that experienced 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. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure 7: Heterogeneity Analysis by Local Government Incentives
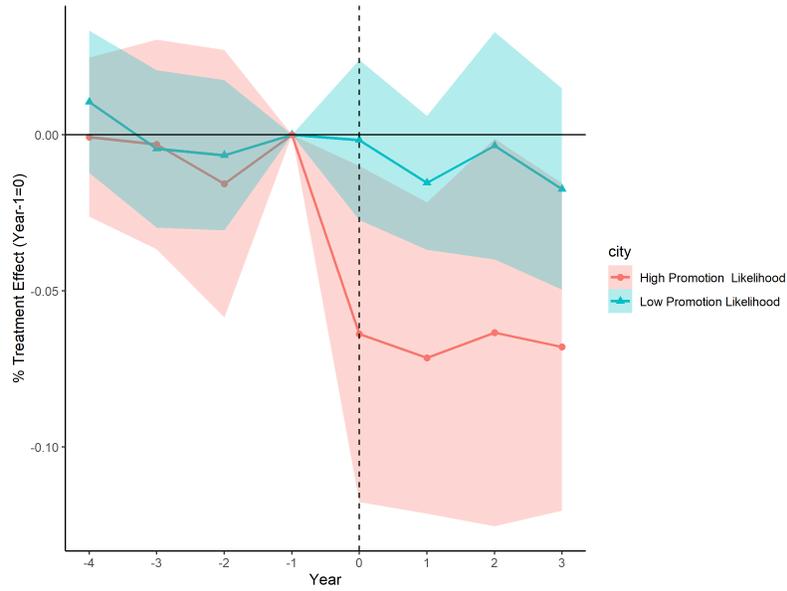
(a) Promotion Likelihood
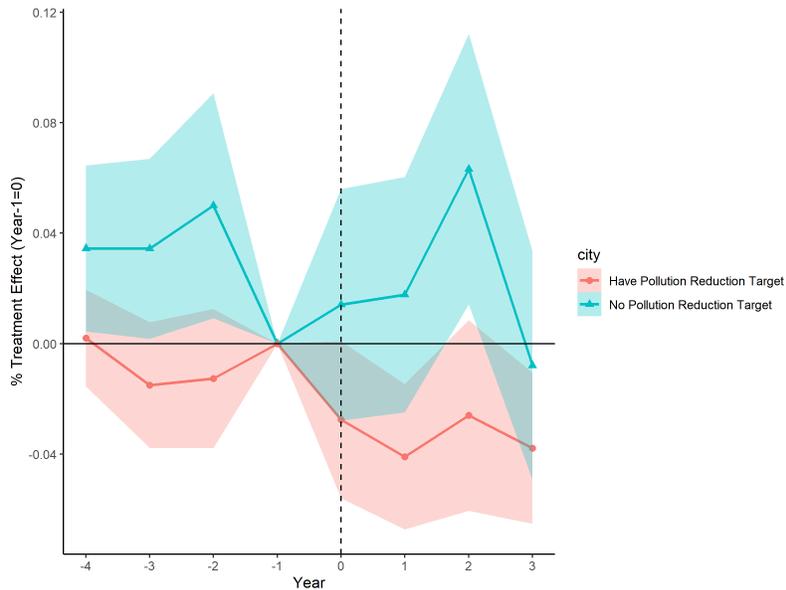
(b) Pollution Reduction Target

Notes: This figure shows the estimated coefficients and 95% confidence intervals for the year-specific automation treatment effects in two subsamples of cities based on their leaders' likelihood of promotion and the pollution reduction targets assigned. Each estimate represents the difference in $PM_{2.5}$ between monitored and unmonitored areas for a given period and subgroup. We estimate the likelihood of promotion for city leaders following Wang, Zhang and Zhou (2020), where leaders (party secretaries) above/below the provincial median are considered of high/low promotion likelihood. In the bottom figure, cities are split into two groups based on pollution reduction targets from 2013 to 2017. "Have Pollution Reduction Target" refers to the city with a target specified in government documents, while "No Pollution Reduction Target" refers to cities that are only required to maintain their baseline pollution levels. The omitted time category is one year before a city joined the automatic monitoring program. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city). Standard errors are clustered at the city level.

Figure 8: Heterogeneity Analysis by Local Public Pressure: Event Study Results
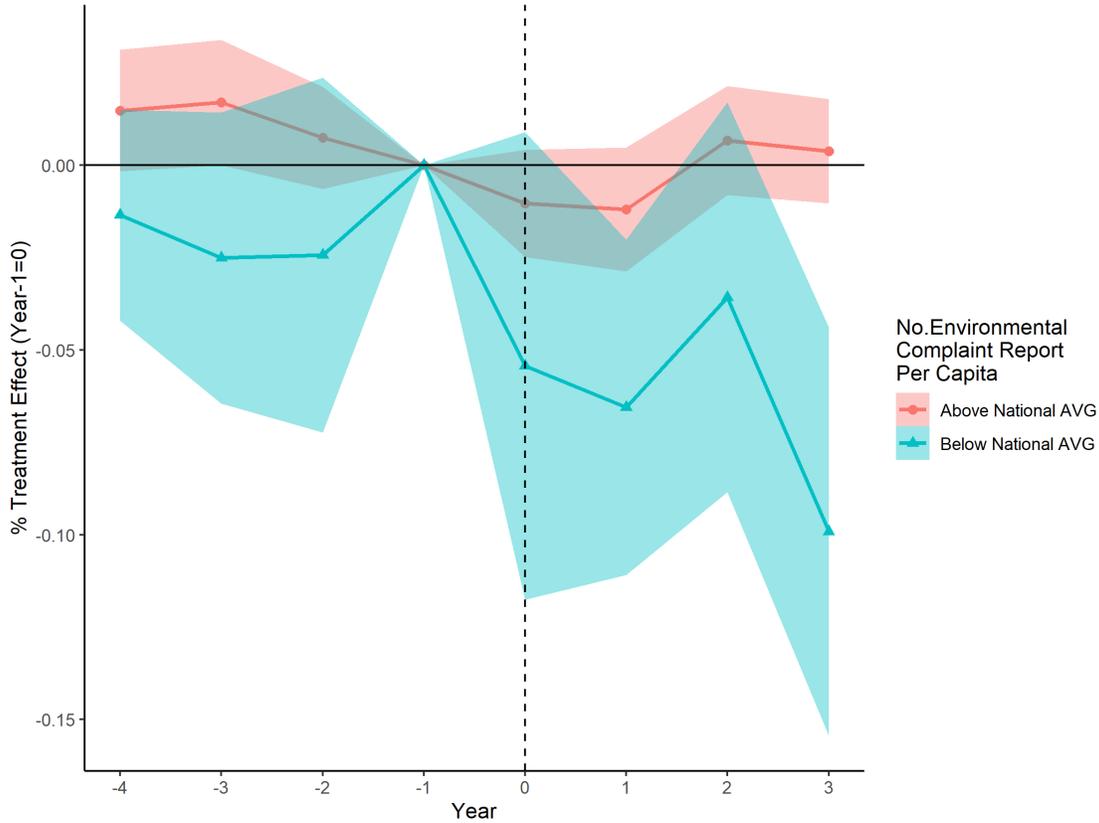
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the interaction terms between year-specific automation treatment effects and indicators for two subsamples of cities classified by local residents' attention to environmental issues. Each estimate represents 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 with above-median per capita complaints about environmental issues in 2017. The omitted time category is one year before a city joined the automatic monitoring program. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Table 1: Descriptive Statistics

Variables	N	Mean	Sd	Min	Max
	(1)	(2)	(3)	(4)	(5)
Panel A: Cells and Satellite-Based Pollutants					
$PM_{2.5}$ (ug/m^3)	10,413,717	35.219	27.234	0.111	223.889
Distance to monitor (km)	10,413,717	160.423	149.877	0.031	873.008
Auto	10,413,717	0.438	0.496	0	1
(0-3km)	10,413,717	0.003	0.059	0	1
Panel B: Satellite-Based Economic Activity Measures					
Fire Radiative Power(FRP)	10,413,717	0.070	1.392	0	794.400
Days	10,413,717	0.018	0.984	0	554.000
1(TAP)	10,413,717	0.006	0.078	0	1.000
Panel C: City Socio-Economic Characteristics					
Population in 2015 (person)	10,413,647	907.720	4,585.402	0.000	
GDP$^{2008-2011}$ (1 billion¥)	330	75.408	111.389	1.443	1,097.079
Number of monitors in a city	330	10.552	6.845	0	46.000
Maximum distance between monitors and cells (km)	330	153.722	123.273	18.187	873.008
Environmental priority city indicator	330	0.336	0.473	0	1.000
$PM_{2.5}$ reduction targets	1,340	3.300%	8.700%	0.000	55.000%
PM_{10} reduction targets	1,340	3.400%	4.700%	0	25.000%
Pre-automation manipulation indicator	123	0.561	0.498	0	1.000
High Promotion likelihood	327	0.468	0.500	0	1.000
Number of environmental complaints reported	31	19,963.100	21,098.130	28	105,601

Notes: Observations are at the cell-year level. Each cell has a 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 equals one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one 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, which is related to the rate at which fuel is consumed and smoke emissions are released. Days measure the annual number of days with active thermal anomalies in each cell. 1(TAP) denotes 1(Thermal Anomalies Presence), which is a dummy variable that equals one if there are thermal-related economic activities in a given cell in a given year. In Panel C, the population in 2015 (*person*) denotes the cell-level population. GDP$^{2008-2011}$ (in 1 billion¥) denotes the average city-level GDP from 2008 to 2011. Number of monitors in a city denotes the details of the treatment received. Maximum distance between monitors and cells denotes a city's geographical size. Environmental priority city indicator is a dummy variable denoting whether or not a city is an environmental priority city. PM_{10} reduction targets and $PM_{2.5}$ reduction targets denote concurrent policies—-city-level $PM_{2.5}$ and PM_{10} pollution control targets from the Air Pollution Prevention and Control Action Plans from 2014 to 2017. Pre-automation manipulation indicator is a dummy variable denoting whether or not a city's environmental data had previously been manipulated (before automation), as calculated using the algorithm in Greenstone et al. (2022). High Promotion likelihood is a dummy variable that indicates the likelihood of promotion for city leaders is above the provincial median, as calculated using the algorithm in (Wang, Zhang and Zhou 2020). Number of environmental complaints reported denotes the number of complaints from the public on environmental issues(at the provincial level due to data availability).

Table 2: Localized Cleanup Response to Monitoring Program Automation

	Dependent variable: $\ln(PM_{2.5})$					
	(1)	(2)	(3)	(4)	(5)	(6)
Unmonitored Areas:	>3km	>3km	>3km	>3km	>3km	>60km
Auto	0.052**	-0.001	0.052**	-0.010	-0.0004	-0.004
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.032)
(0-3km)×Auto			-0.091***	-0.032***	-0.032***	-0.054***
			(0.022)	(0.010)	(0.011)	(0.021)
CellFE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Year FE × citypopulation$^{2008-2011}$		X		X	X	X
Year FE × $PM_{2.5}^{2008-2011}$		X		X	X	X
Year FE × Other City-level Controls		X			X	X
Concurrent Policy		X			X	X
Observations	10,413,717	10,413,717	10,413,717	10,413,717	10,413,717	7,582,459
R^2	0.966	0.975	0.966	0.974	0.975	0.976

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 equals one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if the cells are located within a 3km radius of a city's monitoring stations. Columns (1)–(6) use cells within 3km of the monitor as the monitored group, comparing them with different unmonitored groups: cells beyond 3km from the monitors in columns (1)–(5) and 60km from the monitors in column (6), respectively. $PM_{2.5}^{2008-2011}$ is average city-level $PM_{2.5}$ during the 2008–2011 period and citypopulation$^{2008-2011}$ is the average city population from 2008 to 2011. Other city-level controls are the average city-level GDP from 2008 to 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. 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	>3km
Auto	0.077***	0.018	-0.001	0.020
	(0.026)	(0.024)	(0.023)	(0.017)
(0-3km)×Auto	-0.086***	-0.031*	-0.022**	-0.031***
	(0.031)	(0.015)	(0.011)	(0.021)
Auto×Dirtier	-0.048***	-0.035**		
	(0.018)	(0.017)		
(0-3km)×Auto×Dirtier	0.014	0.013		
	(0.022)	(0.019)		
CellFE	X	X	X	X
Year FE	X	X	X	X
Dirtier × Time dummy			X	
Year FE × citypopulation$^{2008-2011}$		X	X	
Year FE × $PM_{2.5}^{2008-2011}$		X	X	
Year FE × Other City-level Controls		X	X	
Concurrent Policy		X	X	
Entropy Balancing				X
Observations	10,413,717	10,413,717	10,413,717	9,362,369
R^2	0.966	0.975	0.976	0.954

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 one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if cells are located within 3km of a city's monitor. Dirtier is a dummy variable that equals one if the $\ln PM_{2.5}$ of a cell is higher than the city's average $\ln PM_{2.5}$ in 2008 (the 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 re-weighted with entropy balance weights. All regressions control for cell-fixed effects, year-fixed effects, and time dummy interactions. $PM_{2.5}^{2008-2011}$ is average city-level $PM_{2.5}$ during the 2008–2011 period and citypopulation$^{2008-2011}$ is average city population from 2008 to 2011. Other city-level controls are the average city-level GDP between 2008 and 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table 4: Mechanism of Localized 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***	-0.0348	-0.0526	0.000768	0.000523
	(0.0111)	(0.0376)	(0.0414)	(0.0142)	(0.0131)
Marginal Effect	*0.0904****	*-0.00969*	*-0.0335*		
	(0.00264)	*(0.0105)*	*(0.0264)*		
(0-3km)×Auto	-0.410***	-0.307***	-0.253***	-0.0917***	-0.0170
	(0.0435)	(0.0397)	(0.0380)	(0.0212)	(0.0157)
Marginal Effect	*-0.101****	*-0.0856****	*-0.161****		
	(0.0107)	*(0.0111)*	*(0.0242)*		
Cell FE	X	X	X	X	X
Year FE	X	X	X	X	X
Year FE × citypopulation$^{2008-2011}$		X	X	X	X
Year FE × PM$_{2.5}^{2008-2011}$		X	X	X	X
Year FE × Other City-level Controls		X	X	X	X
Concurrent Policy		X	X	X	X
Model	Logit	Poisson	Poisson	OLS	OLS
Sample	All	All	All	1(TAP)	1(TAP)
Observations	227,750	233,760	233,750	52,099	52,099
R-squared				0.721	0.552

Notes: This table reports the effects of the monitor automation program on 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 using 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 for 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 for the average intensity of thermal anomalies, denoted by ln(FRP+1). FRP is defined as the rate of radiant heat output, which is related to the rate at which fuel is 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 grid cell-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. Auto is the treatment indicator that takes the value of one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if cells are located within 3km of a city's monitoring stations. PM$_{2.5}^{2008-2011}$ is average city-level PM$_{2.5}$ over the 2008–2011 period and citypopulation$^{2008-2011}$ is average city population from 2008 to 2011. Other city-level controls are the average city-level GDP between 2008 and 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table 5: A Re-evaluation of Policy Goals Set in the "Air Ten" Action Plan

Region	$PM_{2.5}$ in 2013	$PM_{2.5}$ in 2017	Change of $PM_{2.5}$
	(1)	(2)	(3)
Panel A: Based on Ground-monitor Readings			
Whole Country	47.012	35.451	-24.59%
Beijing-Tianjin-Hebei	73.802	48.377	-34.45%
Panel B: Based on City-wide Satellite Data			
Whole Country	43.485	42.358	-2.59%
Beijing-Tianjin-Hebei	68.706	58.649	-14.63%
Panel C: Based on Monitored Area Satellite Data			
Whole Country	51.768	49.163	-5.03%
Beijing-Tianjin-Hebei	82.101	64.188	-21.82%
Panel D: Based on Satellite Data in (3-30)km Cell			
Whole Country	46.212	44.331	-4.07%
Beijing-Tianjin-Hebei	73.071	60.552	-17.13%
Panel E: Based on Satellite Data in (30-60)km Cell			
Whole Country	43.828	42.445	-3.16%
Beijing-Tianjin-Hebei	67.059	58.453	-12.83%
Panel F: Based on Satellite Data in (60-90)km Cell			
Whole Country	41.386	40.437	-2.29%
Beijing-Tianjin-Hebei	55.036	52.012	-5.49%
Panel G: Based on Satellite Data in (>90)km Cell			
Whole Country	37.495	37.387	-0.288%
Beijing-Tianjin-Hebei	55.324	54.556	-1.39%

Notes: This table presents pollution reduction estimates from 2013 to 2017 for the whole country and the Beijing-Tianjin-Hebei region specifically. Columns (1) and (2) report residualized pollution outcomes (unit: $\mu g/m^3$), which are obtained from an OLS regression of city-level $PM_{2.5}$ concentration on city fixed effects in 2013 and 2017, respectively. Column (3) denotes the change in air pollution during the "Air Ten" Action Plan period. Panel A is based on satellite data of grid cells within 3km of monitors. Panel B uses citywide satellite data. Panel C shows reductions in satellite-based pollution directly at the monitored areas, while Panels D to G report pollution reductions in rings at different distances from the monitored areas.

Appendices

A1 Institutional Details and Evidence on Strategic Cleaning

A1.1 Environmental Regulations in China

In this section, we review three particular regulatory policies that change the evaluation rules for local officials and thus directly affect their incentives for pollution control. The major policy document that set the stage for China's "war on pollution" is the "Action Plan for Air Pollution Prevention and Control" (Air Ten hereafter), announced in September 2013. To better implement the Air Ten at the local level, the Ministry of Ecology and Environment (MEE), Chinese equivalence of EPA, signed a "Target Responsibility Agreement (*mubiao zerenshu*) for Atmospheric Pollution Prevention and Control" (Target hereafter)—essentially performance contracts–with 31 provinces after Air Ten was issued. Another key document, titled "Notice of the General Office of the State Council on Performance Assessment Measures for Air Pollution Prevention and Control Action Plan" (Document NO. GUOBANFA[2014]21, Assessment hereafter), issued on April 30th, 2014, provides more details on environmental performance assessment metrics with respect to the Air Ten action plan. Table A1 of the Appendix contains further detail on the three documents.

The key points of these documents are summarized below:

(1) **Air Ten**—Air Ten set the national guidelines on air quality improvement targets and laid out ten tasks. These tasks include industrial upgrading, clean production, management of coal and oil sources, regulation of coal-power plants, vehicle pollution control, and so on.

(2) **Target**—Each provincial government signed its Target Responsibility Agreement with the MEE. In the agreements, provincial leaders pledge to attain certain air pollution reduction targets for the 2013-2017 period using the 2012 pollution level as the base. There are two major components to the target: air quality targets and the progress

with the ten tasks. Notably, the air quality targets vary across provinces: for some, the focus was on reducing $PM_{2.5}$ concentration levels, while for others, the goal was set to reduce PM_{10}. To ensure accountability and implementation, the provincial targets are further decomposed and allocated to city governments, again through the target responsibility system.

(3) **Assessment**—The Assessment lays out metrics of both the air quality targets and tasks progress. The final score will be a minimum of two (both standardized to have a full mark of 100).

  (a) Air quality improvement is measured as the annual average concentration reduction rate of $PM_{2.5}$ (PM_{10}).The annual concentration of $PM_{2.5}$ (PM_{10}) of a city is measured as the annual arithmetic mean concentrations across its central monitors.

  (b) Key tasks on the prevention and control of air pollution: a scoring system with deduction points from violations to performance standards of the ten tasks, which could also vary by provinces (An example of the deduction point is: "gas stations in a region will be randomly checked for qualified fuel supply. Non-compliant fuel sales will result in a penalty of 1 point.").

(4) **Assessment**—Regions that fail to pass the annual assessment will face the following penalties:

  (a) Local leaders would be summoned for questioning by upper-level officials from the province or other departments.

  (b) Financial penalties will be imposed, such as a reduction in the central grants to the local governments.

  (c) The procedures for approving new projects that have environmental impacts will be suspended.

(5) **Assessment**—Falsification of monitoring data during the assessment results in a disqualification result, followed by a serious investigation by the Supervision Organs.

A1.2 Local Government Documents and News articles Related to Strategic Cleaning

In this section, we perform additional analysis on 121 local government documents that mandate strategic cleaning in their pollution control action plans. In the upper panel of Figure A1, we create a heat map to visualize the spatial distribution of those official documents. In the lower panel, we plot the histogram of the related distance range mentioned.

We further present collaborative evidence from news articles in support of the channels of strategic cleaning. First, some local governments have targeted air pollution near the ground monitors for strategic intervention using water or water vapor. Since the monitor locations are well known by local officials, they sometimes sprayed water in adjacent areas or targeted fog cannons at the monitors (a high-risk yet effective approach) or toward other subjects near the monitors (lower risk, but less effective). As a case in point, in January 2018, it was reported that the Environmental Protection Agency's office building in Shizhuishan, a city of Ningxia Province, where a central monitor is located, was turned into an "ice sculpture" after being targeted by fog cannons.[28] The next set of strategies involves various traffic restriction policies targeted at the monitored areas. An official report by Tianjin's environmental inspection team documented the strategic use of temporary traffic control plans by the local agency.[29] The media also reported incidents in which the gas stations near the monitors in Pingdingshan City were temporarily shut down, again a carefully calibrated approach taken by the local government to improve air quality in the immediate area around

---

[28]Source: CCTV.

[29]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: The Paper (a leading Chinese digital newspaper).

the monitors.[30] A longer-term strategy was to shut down major sources of pollution to unmonitored suburban areas.[31]

A1.3 Examples of Government Documents on Strategic Cleaning

1.Spraying Water

标题：白银市人民政府办公室关于印发白银市大气污染防治 2017 年度实施方案的通知
文件号：市政办发〔2017〕39 号
发文日期：2017 年 3 月 27 日
相关内容： 3.全面落实扬尘管控措施
（12）加大道路扬尘防治力度，全面落实道路洒水喷雾降尘作业，对重点路段，**特别是两个监测点周边一公里区域内**，要增加洒水频次，保持路面湿润。在沙尘天气开始前和结束后，要全方位进行洒水喷雾降尘。白银区在东南西北城市出入口开展环境卫生治理工作，选取合理位置设置洗车点，禁止施工车辆带泥上路，同时对出入口易起尘路段，采取洒水降尘、建立绿化带、全面清理路旁垃圾、定期保洁等措施有效解决扬尘污染问题。（白银区政府负责，实施时限：全年）

Title: Notification on the Issuance of the Implementation Plan for the Prevention and Control of Air Pollution in Baiyin City in 2017 by the Office of the People's Government of Baiyin City

Document Number: Municipal Office [2017] No. 39

Issue Time: 2017.03.27

Related Content: 3. Comprehensive Implementation of Dust Control Measures (12) Enhance the efforts in dust control on roads, **fully implement road sprinkling and spraying operations to reduce dust, especially within a one-kilometer radius around two monitoring stations,** and increase the frequency of sprinkling to keep the road surface moist. Before and after sand and dust weather events, carry out comprehensive sprinkling

---

[30]See news and media coverages of the existing manipulation strategies by Xinhua, Chinanews, and The Economic Daily for more details.

[31]See "Linfen Data Falsification Case: One Year Later, Part of Shanxi's Environmental Information Still Undisclosed" (dated May 10, 2019) for an example. Source: Chinatimes.

and spraying to reduce dust. Baiyin District will conduct environmental sanitation management work at the entrances and exits of the city in the southeast, northwest, and other directions. Reasonable locations will be selected for car washing points, and construction vehicles carrying mud will be prohibited from the roads. Additionally, measures such as sprinkling to reduce dust, establishing green belts, comprehensive roadside garbage cleaning, and regular cleaning will be implemented to solve the problem of dust pollution effectively. (Responsibility: Baiyin District Government, Implementation timeframe: Full year)

2.Ban the Coal-fired boiler/ polluted firms and Spray Water

标题：徐州市人民政府关于印发徐州市 2014 年大气污染防治工作任务分解方案的通知
文件号：徐政发〔2014〕47 号
发文日期：2014 年 7 月 17 日
相关内容：38. 对市区 7 个空气质量监测站点周边 1 公里范围内全面取缔燃煤锅炉、露天烧烤，餐饮企业符合环保要求，推行立体式绿化和喷淋措施，加强交通疏导，推行湿法保洁，减少二次扬尘和机动车排气污染。

Title: Notification on the Issuance of the Task Decomposition Plan for the Prevention and Control of Air Pollution in Xuzhou City in 2014 by the People's Government of Xuzhou City

Document Number: Xu Zhengfa [2014] No. 47

Issue Time: 2014.07.17

Related Content: 38. **Completely ban coal-fired boilers and open-air barbecues within a one-kilometer radius around seven air quality monitoring stations in the urban area.** Encourage catering enterprises that meet environmental protection requirements to implement three-dimensional greening and sprinkler measures, strengthen traffic diversion, promote wet cleaning, and reduce secondary dust and vehicle exhaust pollution.

3.Vehicle Restriction

Title: Notification on the Issuance of the Special Action Plan for the Control of Cooking

Issue Time: 2017.07.11

Related Content: Oil Fume Pollution in the Central Urban Area of Ji'an City and the Special Action Plan for the Control of Motor Vehicle Exhaust Pollution in Ji'an City by the

标题： 吉安市人民政府办公室关于印发《吉安市中心城区油烟污染专项整治行动方案》《吉安市机动车排气污染专项整治行动方案》的通知

发文日期： 2017 年 7 月 11 日

相关内容：（五）科学组织货车绕行、禁行。制定**国控监测点周边 3 公里范围**内主要道路货车禁行、限行管控方案，规定禁行、限行区域、时段和车型，并设置标志。对过境市中心城区的重型货运车辆实施远端绕行，禁止穿越主城区。对参与城市建设的水泥罐车、渣土运输车、专项作业车等工程车，严格限制运行时间和路线；对关乎民生的保障货车，公安交管部门对车辆进行严格审批，按照"避开高峰，远离中心，夜间进入"的原则，核准通行时间路线。

Office of the People's Government of Ji'an City

(5) Organize the diversion and prohibition of trucks in a scientific manner. **Develop control plans for the prohibition and restriction of trucks on major roads within a 3-kilometer radius of the central stations.** The plans will specify prohibited and restricted areas, time periods, and vehicle types, and appropriate signage will be installed. Heavy-duty freight vehicles passing through the central urban area will be required to take detours at remote locations, and crossing the main urban area will be strictly prohibited. Cement tankers, construction waste transporters, and specialized vehicles involved in urban construction will be subject to strict limitations on operating time and routes. For freight vehicles related to public livelihood, the Public Security Traffic Management Department will conduct a rigorous vehicle approval process, following the principles of "avoiding peak hours, staying away from the city center, and entering during nighttime" to approve the designated times and routes for passage.

4.Ban Open-Air Barbecues

标题： 贵阳市人民政府关于印发贵阳市 2015 年大气污染防治工作方案的通知

文件号： 筑府发[2015]26 号

发文日期： 2015 年 6 月 23 日

相关内容：（3）加强夜市烧烤摊点规范管理。着重整治露天烧烤，严格限制和规范建城区露天烧烤，重点依法取缔**空气自动监测站点周边 1 千米范围内**无组织排放烧烤点。

Title: Notification on the Issuance of the Work Plan for the Prevention and Control of Air Pollution in Guiyang City in 2015 by the People's Government of Guiyang City

Document Number: Zhu Fu Fa [2015] No. 26

Issue Time: 2015.06.23

Related Content: (3) Strengthen the standardized management of night market barbecue stalls. Place a special emphasis on regulating open-air barbecues and strictly restrict and regulate open-air barbecues in the urban area. **Particularly, take decisive actions to prohibit unregulated emission of pollutants from barbecue sites within a one-kilometer radius of automatic air monitoring stations.**

A2   Additional Data Details and Robustness Checks

A2.1   Satellite-based PM$_{2.5}$ Data

Our main dependent variable is the annual AOD-based PM$_{2.5}$ data compiled by Van Donkelaar et al. (2016b). We note, however, that the monthly level data were also made available in a more recent data release by Van Donkelaar et al. (2021).

As detailed in the reference, annual and monthly ground-level fine particulate matter (PM$_{2.5}$) for 1998–2021 were estimated by combining Aerosol Optical Depth (AOD) retrievals from the NASA MODIS, MISR, and SeaWIFS instruments with the GEOS-Chem chemical transport model, and subsequently calibrating to global ground-based observations using a Geographically Weighted Regression (GWR). Meanwhile, raw satellite AOD data are available at the daily level, but their temporal resolution largely depends on satellite coverage. For example, the NASA MODIS instrument collects AOD data every 1 to 2 days from two satellites, Terra and Aqua, which only record AOD on cloud-free days and are sensitive to light surfaces and other weather conditions, leading to missing values at the daily or even weekly level.

Here we assess the pros and cons associated with more- and less-aggregated data. In particular, we examine the correlation between ground-based and AOD-based PM$_{2.5}$ data at the annual and monthly levels, respectively. As shown in Figure A2, there appears to be a stronger correlation at the annual level. This is possibly due to the fact that satellite-based

PM$_{2.5}$ measures are subject to idiosyncratic measurement errors from time to time, such as cloud covers, light interference, and other temporal meteorological variations, but can be smoothed out over longer time periods.

Relatedly, we also assess the quality of satellite-based PM data by checking the correlation between AOD-based and ground-based pollution measures before and after automation. However, note that monitoring stations in China only recorded PM$_{10}$, not PM$_{2.5}$ prior to automation, so we could only investigate the correlation between AOD-based PM$_{2.5}$ and ground-based PM$_{10}$ throughout the sample period. As shown in Figure A3, the correlation becomes strong after automation. We interpret it as evidence of data quality change: PM$_{10}$ data before automation were subject to tampering and were not reliable, as documented by (Greenstone et al. 2022). Automation has improved the accuracy and reliability of ground station data.

A2.2 Robustness Checks to Address Satellite-base PM$_{2.5}$ Measurement Errors

Using Raw AOD Data As discussed above, the AOD-based PM$_{2.5}$ data were derived from raw satellite images, and the calibration procedure also required information from ground-based monitoring stations. Specifically, the Geographical Weighted Regression method assigns larger weights to areas closer to ground monitors and smaller weights to areas that are farther away. 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, we conduct our own analysis with raw AOD measurements as an alternative outcome indicator. The pertinent results are reported in Table A2 and Figure A4. Reassuringly, they are largely consistent with the baseline estimates.[32]

---

[32]The estimated coefficients are smaller with AOD as the outcome, largely because the raw satellite images are sensitive to meteorological conditions (e.g., cloud coverage). In the case of extreme conditions such as haze and fog events, which tend to be associated with heavily-polluted time periods and regions, AOD data may be missing or become unreliable.

Using Satellite-based Pollution Measures at Finer Temporal Variation   We further explore the potential aggregation bias, also known as the "ecological fallacy", associated with annual data (Banzhaf, Ma and Timmins 2019). To this end, we re-run analyses on monthly AOD-based $PM_{2.5}$ data and other monthly moments (the maximum and minimum value of daily observations in a month) of raw AOD data.

In order to assess the relevance of aggregation bias, consider the daily regression form of Equation (1) below:

$$\ln(PM_{2.5icdt}) = \alpha Auto_{ct} + \beta Near_i \times Auto_{ct} + \gamma X_{ct} + Cell_{id} + Year_t + \varepsilon_{ict} \qquad (1)$$

Where i and c denote the cell and the city, as previously defined; d denotes the day, and t denotes the year. Aggregating the daily regression to the yearly level means that we can no longer control for the cell-by-day-of-year-fixed effects (Cell $\times$ day of the year), which essentially capture location-specific within-year daily or seasonal patterns of pollution–for example, cells near a factory would be polluting during working days, especially in the summer, but not during weekends. Because of this distinction, estimation biases would arise in the aggregated vis-à-vis disaggregated level when these within-year daily patterns differ (non-causally) before and after automation. We conduct our own analysis to partially evaluate the extent of this omission. In particular, we use a variant of Equation (1) that is based on monthly data. We then estimate β with and without the inclusion of cell-by-calender-month fixed effects ($Cell_{id}$). The results are reported in Table A3 (also in Figure A5). As shown, they are consistent with our baseline estimates obtained using the annual data, as reported in Table 2. The estimated β with our preferred specification (Column 5) is -0.025, compared to -0.032 in the baseline specification.

Thus, without considering such spatial and temporal meteorological variations, the annual average effect estimated from the daily AOD observations (which partially average out) is likely to be small.

Another issue about analyses based on the annual data is that temporal aggregation might discard information: it averages out the rich variation in the heterogeneity of β across clean and dirty days. The same average treatment effect at the aggregated level could represent very different compositions of individual treatment effects across different days. And the different compositions might entail different welfare implications. Just to take one example (conceptually), if local governments attempted to reduce pollution during clean days but increase pollution during dirty days, the resulting health costs could be much greater than those if pollution was reduced during dirty days and increased during clean days, even though the annual average change in pollution levels remained the same under these two scenarios. To shed light on the potential distributional effects, we collect daily raw AOD data, which is the highest resolution possible, and collapse them to the monthly level. This allows us to obtain monthly summary statistics, including the mean, maximum, and minimum values, which can offer a deeper understanding of the within-month distributional response to automation. The results are reported in Table A4. Our findings indicate that monitoring station automation reduced the AOD levels across all three measures. However, the largest log-point change was observed for max AOD. This result indicates that automation likely leads to a reduction in pollution across the entire distribution, but with greater strategic cleaning efforts observed around the monitors during extreme pollution days.

Bias Correction with Multiple Imputation Lastly, we address measurement errors due to remote sensing by following the lead of Proctor, Carleton and Sum (2023) in utilizing multiple imputation methods to establish the relationship between AOD-imputed $PM_{2.5}$ values and ground-based monitoring data. Specifically, we employ bootstrap sampling to randomly select 70% of the ground-based monitoring data, and then generate the remaining 30% of the sample through multiple imputations for 100 times. We then utilize that sample of 70% original data and 30% imputed observations to perform regression analysis and simulate the relationship between satellite $PM_{2.5}$ values and their corresponding ground-based

readings.

Following that, we predict $PM_{2.5}$ values for all grids in our main dataset using the satellite data and the regression model derived in the previous step. Subsequently, we employ the corrected $PM_{2.5}$ values to repeat our baseline regressions. To account for both sample and imputation uncertainty, we repeat the process of random sampling and prediction 100 times. The results are robust, as reported in Appendix Table A5.

A2.3 Heterogeneous Treatment with Staggered DiD

This section discusses potential biases arising from the traditional two-way fixed effects (TWFE) estimator in a staggered difference-in-differences design. These biases can arise when early treated units are used as control groups for later treated ones, particularly when there are heterogeneous dynamic treatment effects (Goodman-Bacon 2021; De Chaisemartin and d'Haultfoeuille 2020; Sun and Abraham 2021; Callaway and Sant'Anna 2021).

To address this issue, we first employ a Goodman-Bacon decomposition of the DiD estimation that regresses the pollution gap between monitored and unmonitored areas on the staggered treatment of monitor automation in a city-year panel. Figure A6 plots the average treatment effect against the weight of each of the six 2×2 comparison groups in the present study. It appears that the average effects of the three early versus later treated groups (i.e., Wave 1 vs. 2, Wave 2 vs. 3, and Wave 1 vs. 3) concentrate around -0.003 and are very close to the TWFE estimator of -0.0039. The sum of the weights of all the earlier versus later groups adds up to more than 50%. Of the three later versus earlier treated groups, which tend to introduce biases to the TWFE estimate, the share of Wave 3 versus Wave 1 and that of Wave 3 vs. Wave 2 add up to 26.7% and 10.3%, respectively. Therefore, we exclude monitors treated in Wave 3 from the treated units and repeat the analysis, again using the triple-difference specification and $lnPM_{2.5}$ concentrations as the outcome variable.[33] The estimation results are presented in Table A6. Specifically, in our preferred specification,

---

[33]To maximize identification power from the earlier versus later treated comparison, we switch the treatment status of all monitors treated in Wave 3 to untreated in years 2014

Column (3) shows that the triple difference estimator, i.e., the coefficient of (0-3km)×Auto, increases from -0.032 of the baseline to -0.039 and remains statistically significant. In a similar vein, the results on thermal anomalies are reported in Table A7.

In addition, we employ an alternative estimation approach proposed by Callaway and Sant'Anna (2021), which estimates the group-time average treatment effects (ATTgt) separately for all treatment cohorts (each group/cohort corresponds to units treated in the same period) relative to never-treated or not-yet-treated control units, and aggregates all of them into simpler parameters. Since the method only applies to a DiD setting, we modify our triple-difference model into a DiD specification, using the pollution gap as the outcome variable. As presented in Figure A7, the result is in line with that of our main triple difference analysis: the pollution gap evolves in parallel between the treated and control monitors before the treatment, and drops significantly in the post-treatment period, implying that the observed improvement in air quality around the monitors relative to the entire city is not driven by biases in the TWFE estimators.

A2.4 Validation of Thermal Anomalies Data

Before proceeding to the empirical specification, we conduct a set of external validation exercises. We start by assessing the geographical correlation between thermal anomalies and polluting activities. To do so, we obtain two lists of major polluting plants: the first is drawn from the MEE's Key Centrally Monitored Polluting Enterprises database, which consists of 1,829 heavily polluting industrial firms. The other composes of 10,491 power plants, sourced from the China Emissions Accounts for Power Plants (CEAP) database. Figure 5 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,

and 2015, based on the assumption that treatment effects would not be realized in the first year (most were treated in December 2014). We also drop all the post-2015 observations.

although those industrial firms and power plants are more spatially dispersed. On that basis, we argue that thermal anomalies provide sufficiently comprehensive coverage of major polluting sources. At the extensive margins, Table A8 shows that, for each 10km-by-10km cell, the presence of any thermal anomaly increases the probability of the presence of a polluting plant by 99%. At the intensive margins, Column (2) of Table A9 shows that for the sample of plant sites, 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 the thermal anomaly data.

We also test if short-run variations in thermal anomalies respond to temporary drastic government measures on pollution. As a case in point, a series of strict emission control policies were adopted in Beijing and the surrounding regions to ensure blue skies during the 2014 Asia-Pacific Economic Cooperation (APEC) summit. Figure A8 presents the time series of two measures of thermal anomalies one month prior to and one month following 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 synchronized pattern again highlights the validity of the thermal anomaly measure in measuring temporal variation in local pollution.

Ring Analysis with Thermal Anormalies Figure A9 further reports the impact of monitor automation on thermal anomalies across different distance bins. The estimated magnitudes of the responses of thermal-based industrial activities to automation decrease as the distance from the monitors increases, which is consistent with the uneven pollution reduction pattern documented in Section 4.3. However, the effect appears to be more localized, with the reduction in industrial activities approaching zero at around 20 km away from the monitors, compared to around 100 km for $PM_{2.5}$ reductions. One possible explanation is that the spatial impact of localized shutdowns of industrial sources on pollution could

extend tens of kilometers away as pollutants disperse.

A3   Dynamic Monitor Representativeness

Using fine-scale pollution data and spatial information from the national environmental air monitoring network, we examine the spatial representativeness of these monitors, defined as the difference between the monitor-based and satellite-based city average $PM_{2.5}$. 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 A10 (a) shades cities according to representativeness errors (i.e., how well the monitors capture average air pollution). The blue shading denotes that the monitor under-represents a city's average pollution. At the base year of our study period (i.e., 2008), the monitoring system was indeed fairly representative for most Chinese cities, and monitor locations are unlikely to change once they are placed.[34] However, the spatial representativeness of air quality monitoring is not static but an involving process that can be profoundly shaped by local interventions that target monitored areas. Recall the estimate for the localized pollution reductions in monitored areas: grid cells within a 3km radius of monitors experience a 3.2% greater reduction in $PM_{2.5}$ concentrations than those farther away. Using this central estimate and the last year of the sample period (i.e., 2017), we calculate the projected pollution levels for the five-year period from 2018 to 2022, as shown in Figure A10 (b).[35] The forecast suggests that some previously over-representative monitors

[34]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 counties that did not have them, the existing monitor locations have not changed since then.

[35]We do not extend the extrapolation beyond 2022 because of the large uncertainty and the possibility of new regulations.

seem to move closer to a city's average air quality. However, the most striking result is that monitors in approximately 52 cities are predicted to under-represent overall air pollution by the end of 2022, having been greatly affected by dynamic local strategic conduct.

Figure A1: Government Documents Mentioning "Strategic Cleaning"

(a) Spatial Distribution
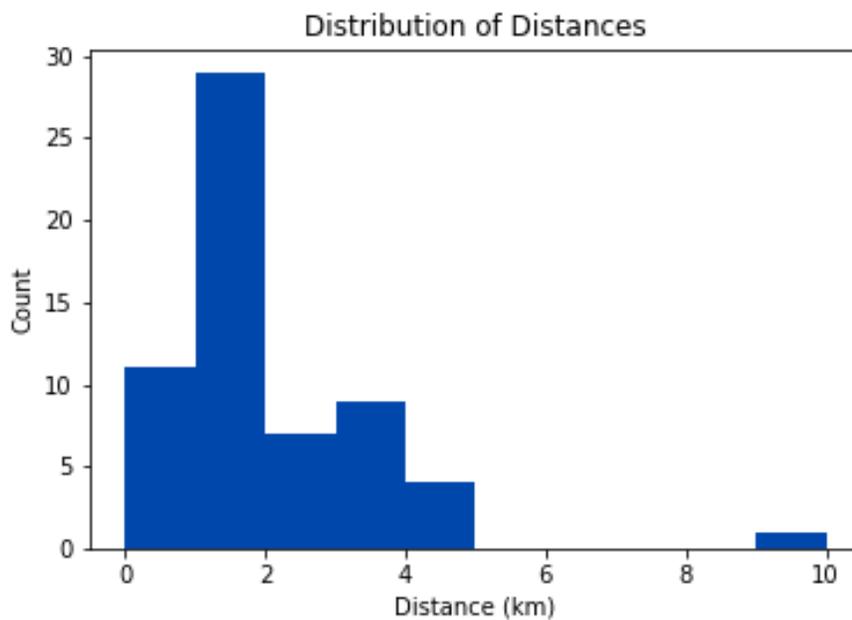
(b) Strategic Cleaning Distance Range

Notes: Panel A shows the spatial distribution of 121 documents that directly mentioned strategic cleaning around monitors, while Panel B shows the histogram of cleaning range (distance from the monitored) stated in 42 of them.

Figure A2: Correlation between AOD-based and Ground-based PM$_{2.5}$ Measurements
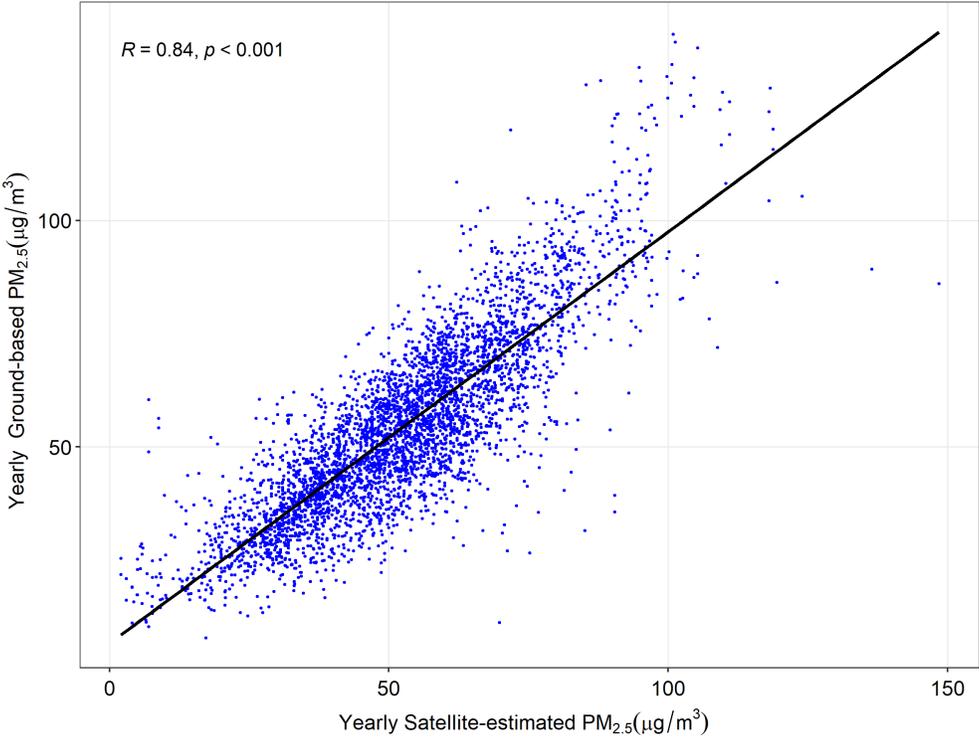
(a) Yearly Data
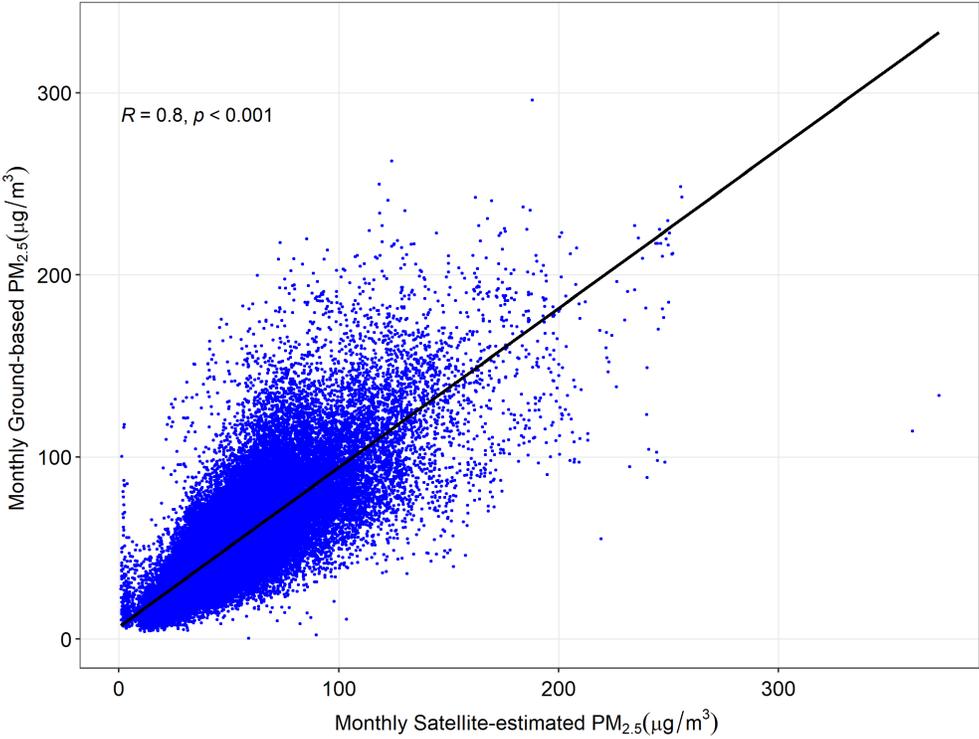
(b) Monthly Data

Notes: This figure depicts the correlation between AOD-based PM$_{2.5}$ and ground reading data. Panel A displays the correlation at the yearly level, while Panel B shows the correlation at the monthly level. Ground reading data for PM$_{2.5}$ only became available after automation.

Figure A3: The Correlation between AOD-based $PM_{2.5}$ and Ground-based PM_{10} Before and After Automation
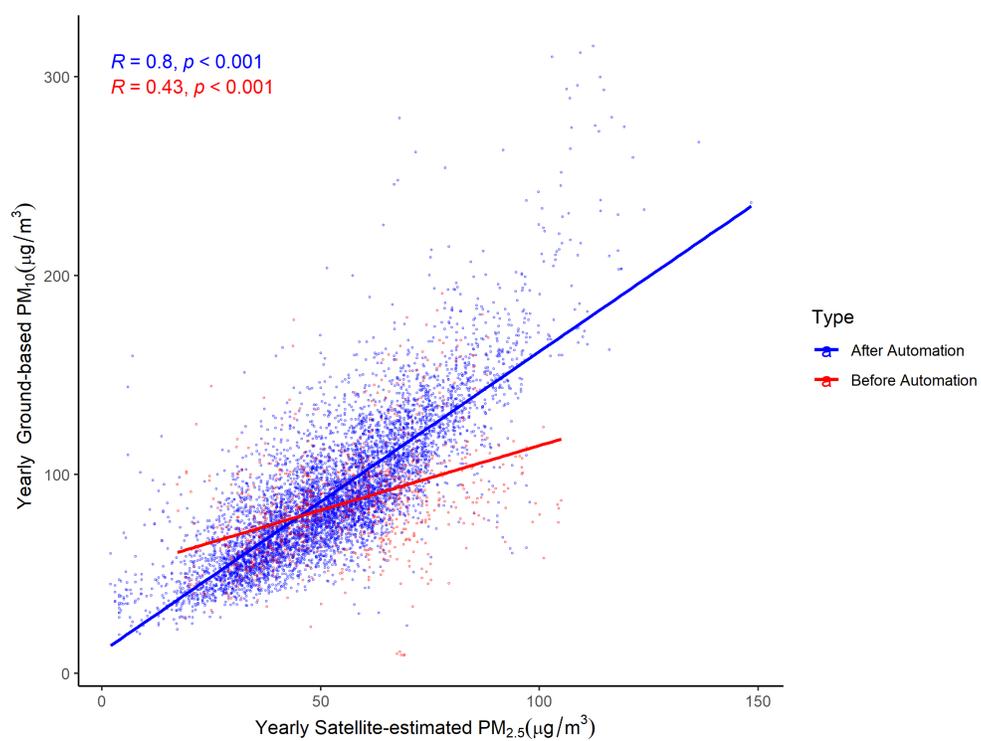
Notes: This figure shows the correlation between yearly AOD-based $PM_{2.5}$ and Ground-based PM_{10} reading data. The red and blue lines represent the fitted linear relationship before and after automation, respectively.

Figure A4: Robustness Check: Event Study of Monitor Automation, using AOD Raw Data
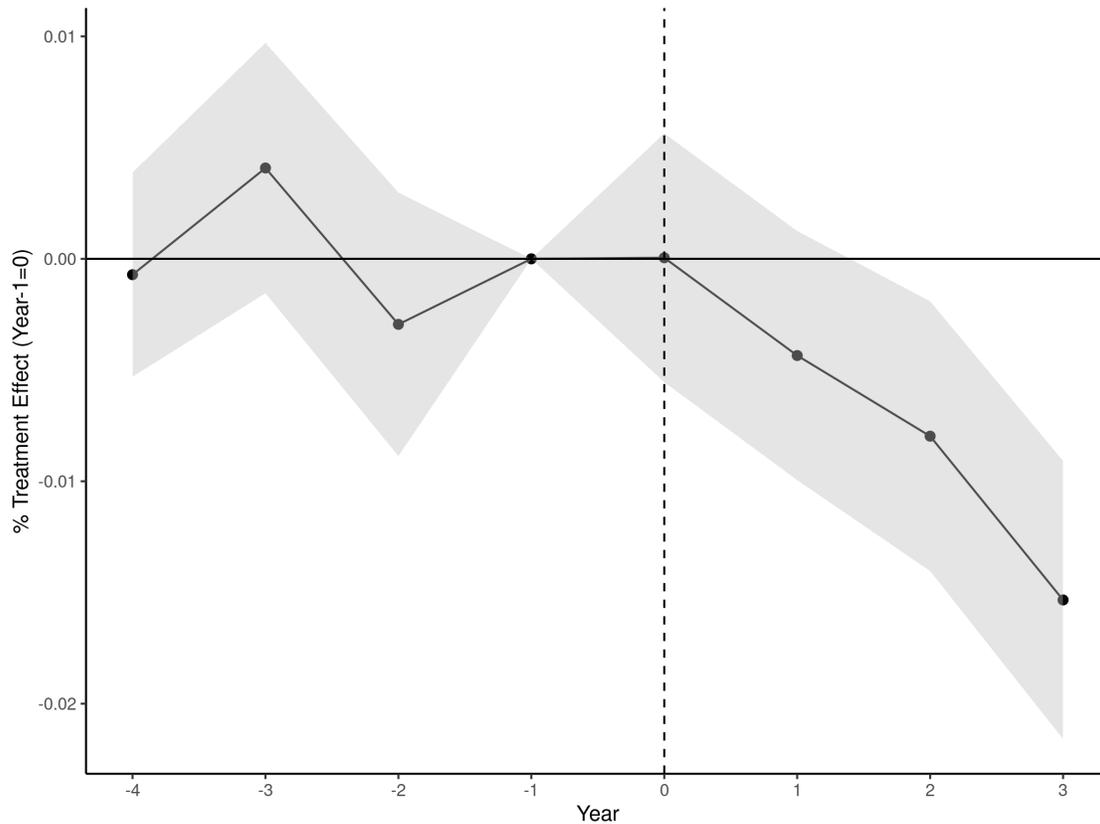
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for β_n from Equation (3), replacing $PM_{2.5}$ with raw AOD data. Detailed estimates are reported in Table A10. The omitted time category is one year before a city joined the automatic monitoring program. Each estimate represents the difference in ln(AOD) between monitored areas (cells within a 3km radius of a monitor) and unmonitored areas (cells outside a 3km radius around a monitor) at a given period. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure A5: Event Study: The Effect of Automation on Air Pollution within 3km (Monthly)
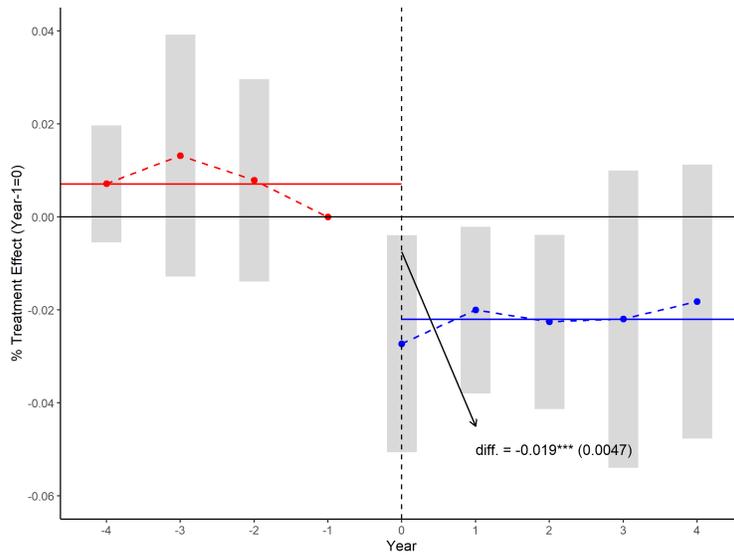
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for β_n from Equation (3) by using monthly $PM_{2.5}$ Data. Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside the 3km radius) at a given period (also reported in Table A3). The omitted time category is the year before a city joined the automated monitoring program. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, and $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure A6: Robustness Check: Effects of Automation on Pollution Gap of Monitors, Bacon Decomposition for Difference-in-differences
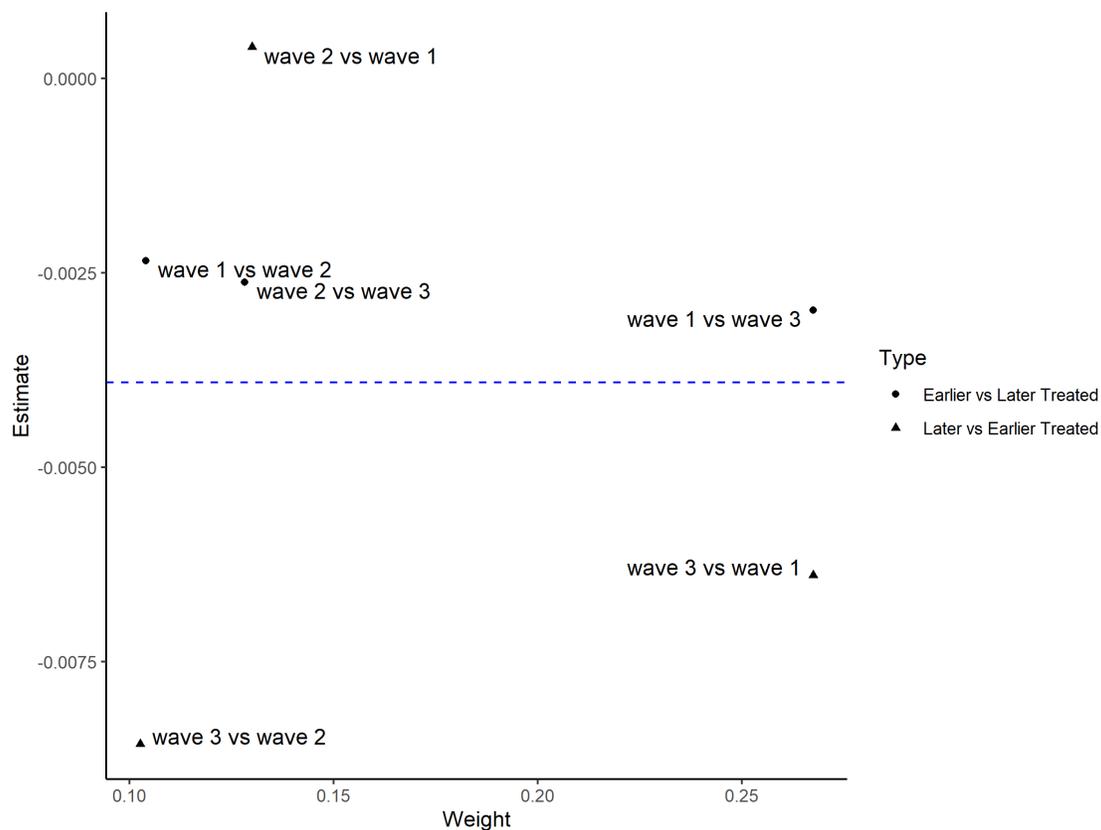
Notes: This figure shows each 2x2 DD estimate from the Bacon decomposition Goodman-Bacon (2021) against their weight for the automation impact analysis. The outcome variable is the pollution gap, defined as the difference between the average pollution within a 3km radius of a monitor and the city's average pollution level. The horizontal dashed line is the difference-in-difference estimate with the pollution gap as the dependent variable (-0.0039 at the 1% significance level). In the Bacon decomposition, the estimate of the "Later vs. Earlier Treated" groups equals -0.0051 and the weights are 0.50.

Figure A7: Robustness Check: Event Study of Monitor Automation on Pollution Gap, Group-Time Average Treatment Effect Estimation
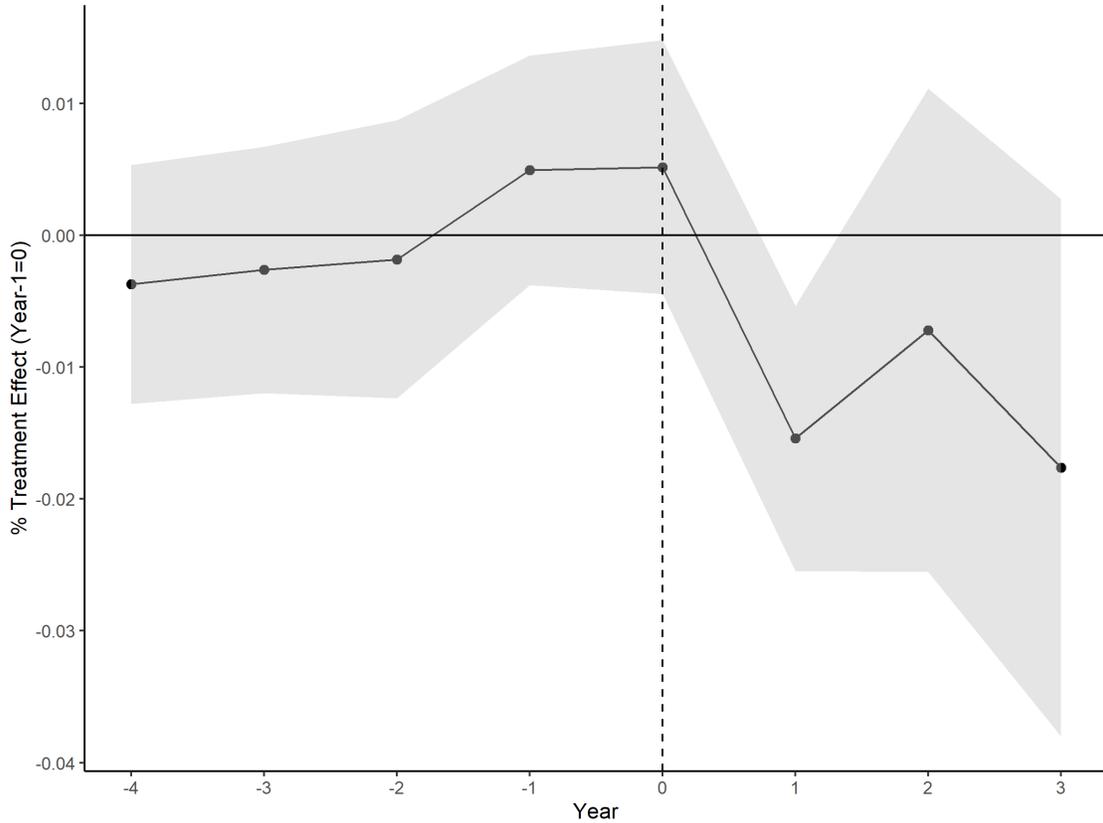
Notes: This figure shows the aggregate event study result following the approach of Callaway and Sant'Anna (2021). The sample includes the period from 2008 to 2015 and sets Wave 3 as the never treated group. The outcome variable is the pollution gap, defined as the difference between the pollution level within 3km of a monitor and the city's average pollution level. All regressions control for cell fixed effects, year fixed effects, and interactions between year dummies and the average city population, and average city-level $PM_{2.5}$ over the 2008–2011 period. Standard errors are clustered at the city level.

Figure A8: Validation of Thermal Anomalies Using the APEC Event
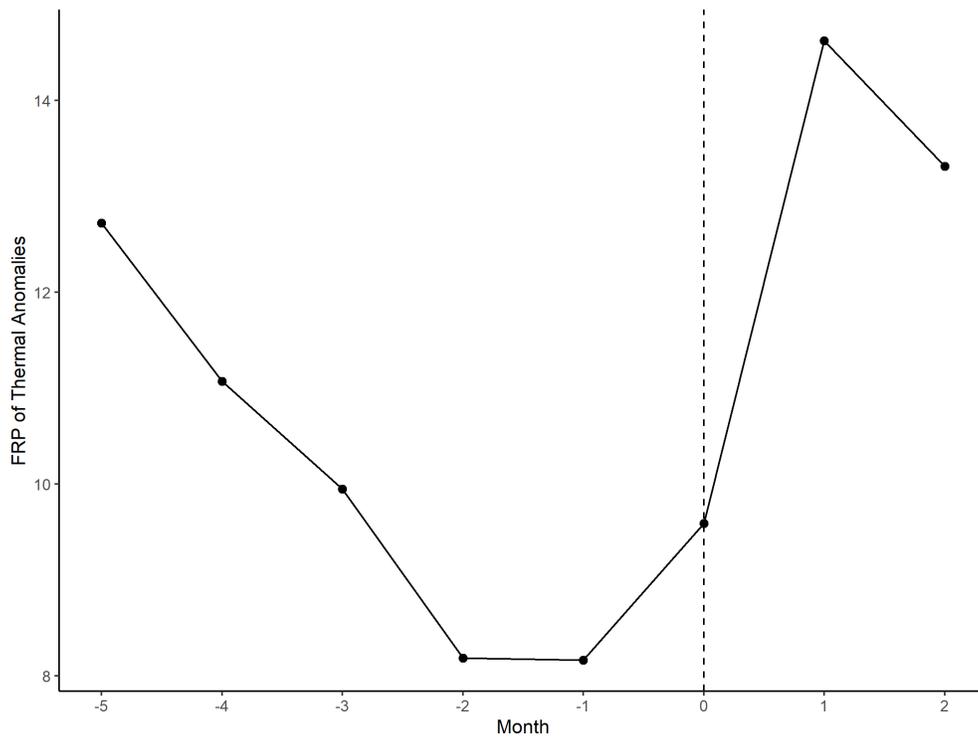
Notes: This figure shows the time series of the thermal anomalies measured shortly before and after the APEC event. Month 0 denotes the month APEC was held (November 2014). FRP is defined as the rate of radiant heat output, which is related to the rate at which fuel is consumed, and smoke emissions are released.

Figure A9: Effects of Automation on Thermal Anomalies at Different Distances from Monitors
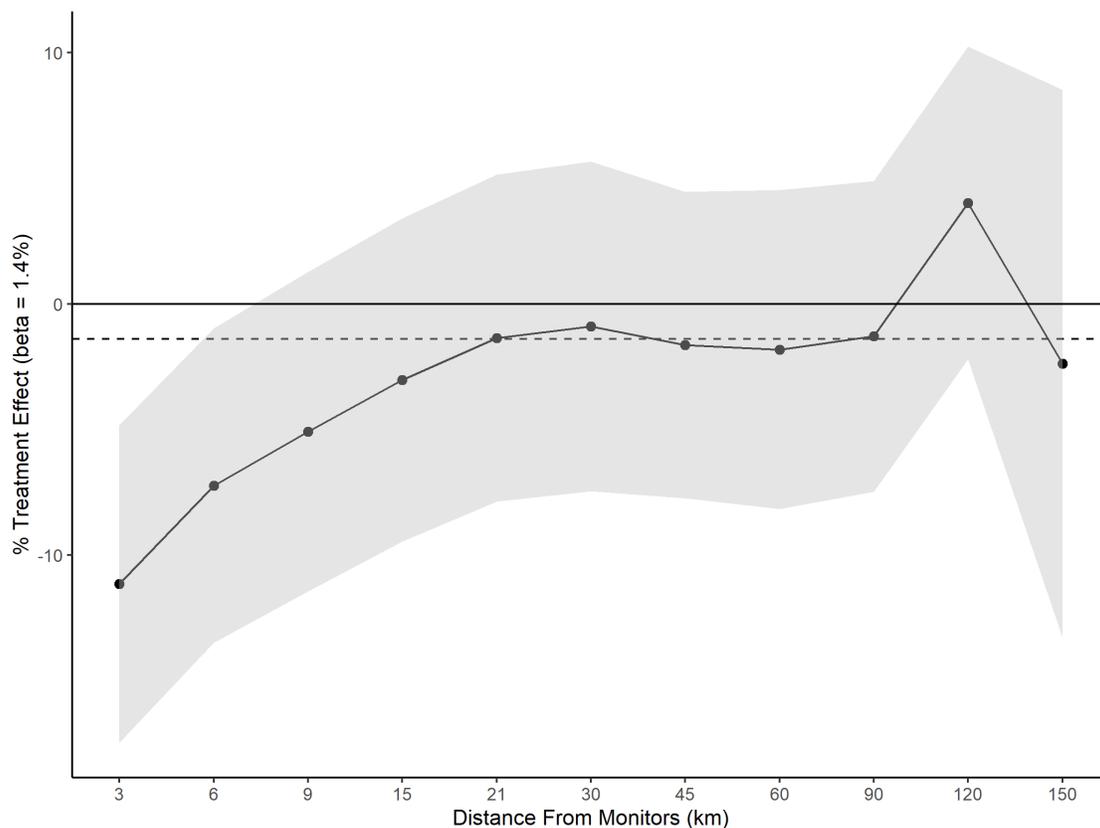
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the monitor automation effects on the number of days with active thermal anomalies across different distance bins from the monitor. Each point estimate represents the pollution change in each distance bin relative to the baseline group at the outer range (distance to monitor >150 km), which on average experiences a 0.5% pollution increase. The absolute effect becomes positive above the dotted line. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure A10: Monitor Representation Errors: All Cells vs. Monitored Cells
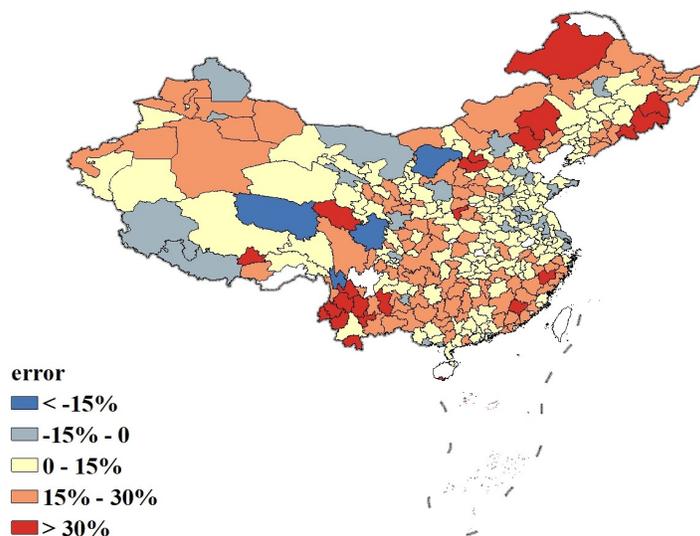
(a) Representation Errors in 2008
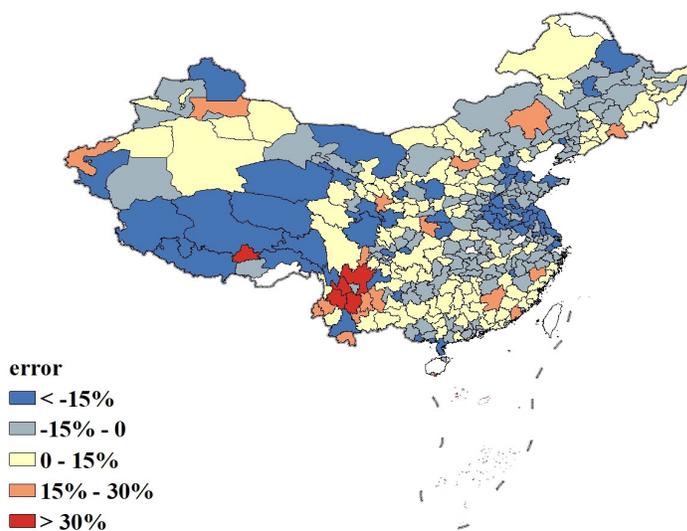
(b) Representation Errors in 2022

Notes: This figure presents the monitor representation errors. Panel A shows monitor representation errors at the base year of the study period (2008). Panel B 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 3.2% greater reduction in air pollution relative to unmonitored areas), and are projected beyond the last year of the sample period (2017). The representation error is defined as the percentage difference between the population-weighted, satellite-based average pollution in monitored cells and the average pollution across all cells within the city boundary. Negative measures indicate under-representation by monitors.

Figure A11: Change in Environmental Authorities' Responsibilities
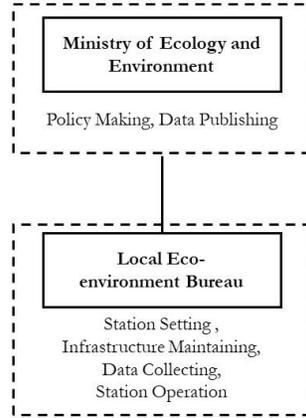
(a) Before
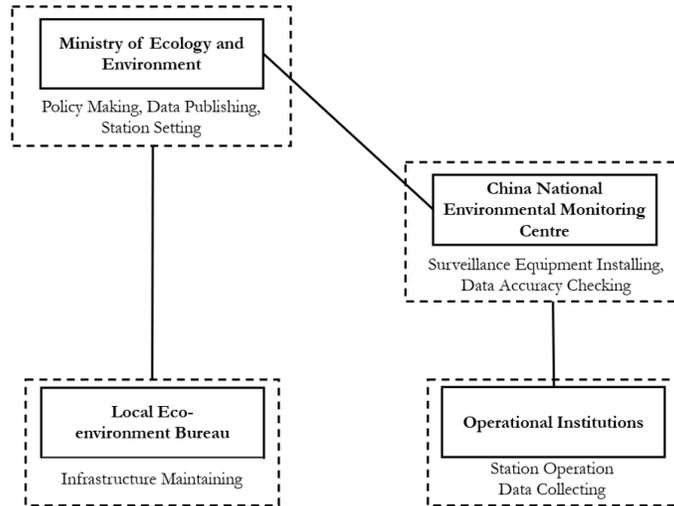
(b) After

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 the various responsibilities, Infrastructure Maintenance refers to ensuring the supply of electricity and communications, and Data Accuracy Checking denotes checking the anomaly data.

Figure A12: Robust Analysis of Event Study on PM$_{2.5}$
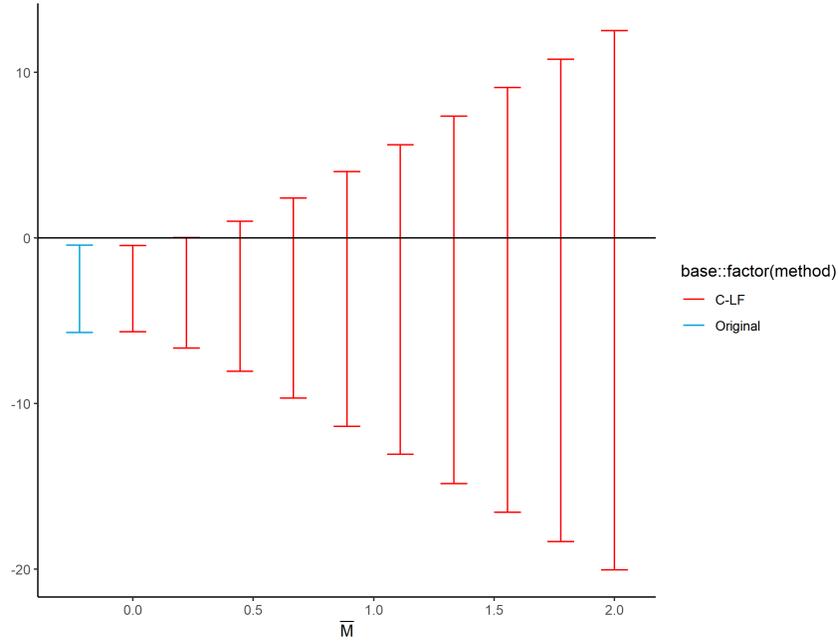
Notes: This figure shows the sensitivity analysis of estimated effects on PM$_{2.5}$ to potential violations of the parallel trends assumption following the methods proposed by Rambachan and Roth (2019). The blue bar represents the 95% confidence interval of the DiD estimate for the last period ($\tau = 3$) from the estimation of Equation (3). The red bars represent corresponding 95% confidence intervals when allowing for per-period violations of parallel trends up to M, which is the largest allowable change in the slope of an underlying linear trend between two consecutive periods.

Figure A13: Cities' Maximum Distance between Cells and Monitors
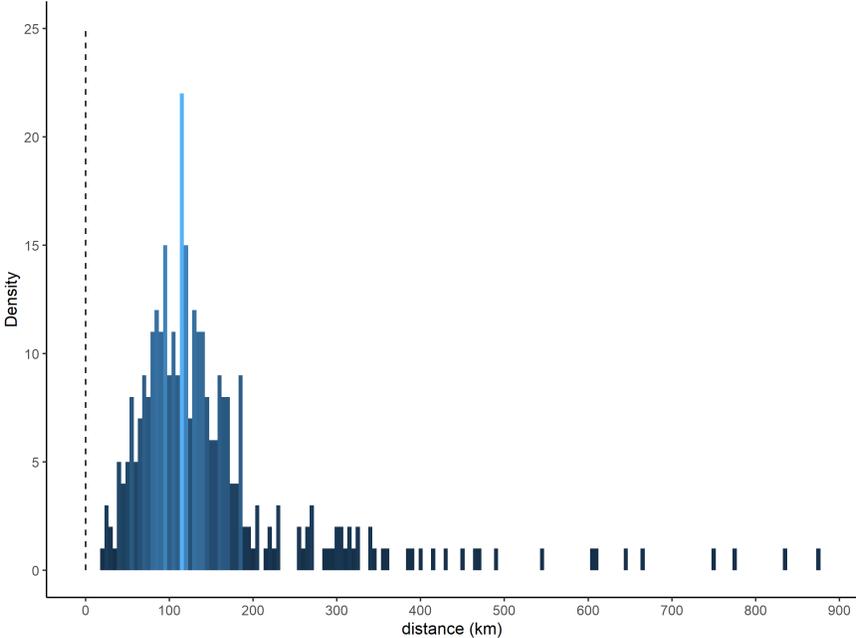
Notes: This figure shows the distribution of the maximum distance between cells and monitors (a proxy for city's geographical size) across cities. The maximum distance ranges from 18 km to 873 km, and the average maximum distance is 152 km.

Figure A14: Robustness Check: Effects of Automation on ln(AOD) across Distances from Monitors
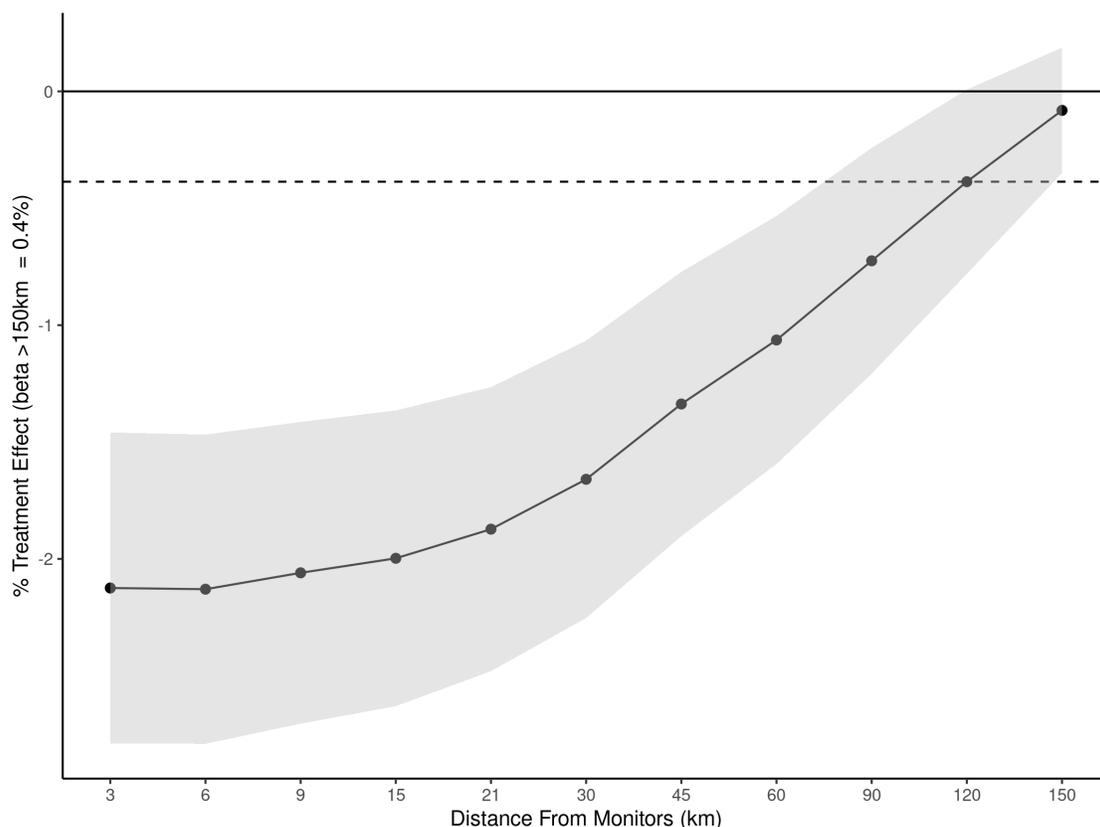
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the effects of monitor automation on the satellite-based ln(AOD) at different distance bins from the monitor. Each point estimate represents the pollution change in each distance bin relative to the baseline group at the outer range (distance to monitor >150 km), which on average experiences a 0.2% pollution increase. The absolute effect becomes positive above the dotted line. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure A15: Placebo Tests: Randomizing Treatment Timing and Locations
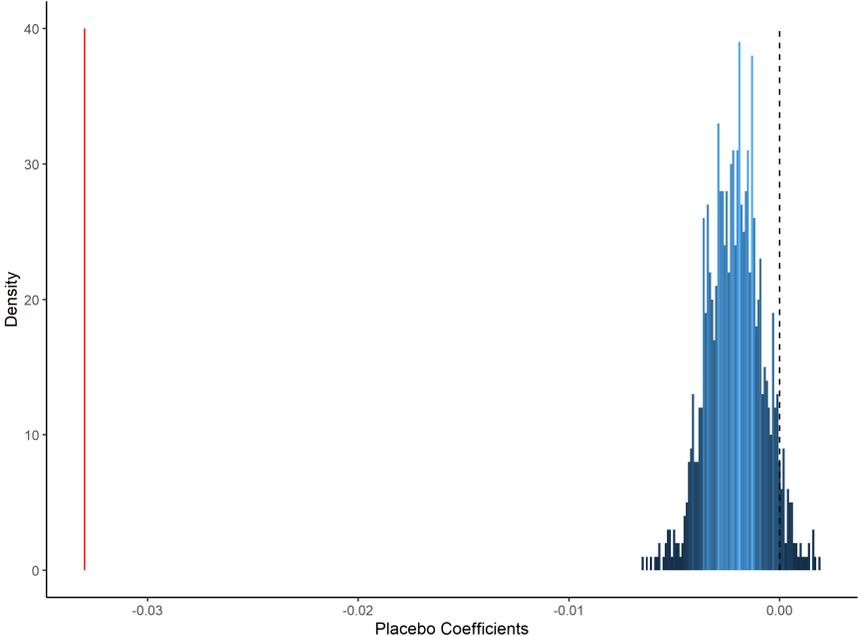
(a) Random Automation Years
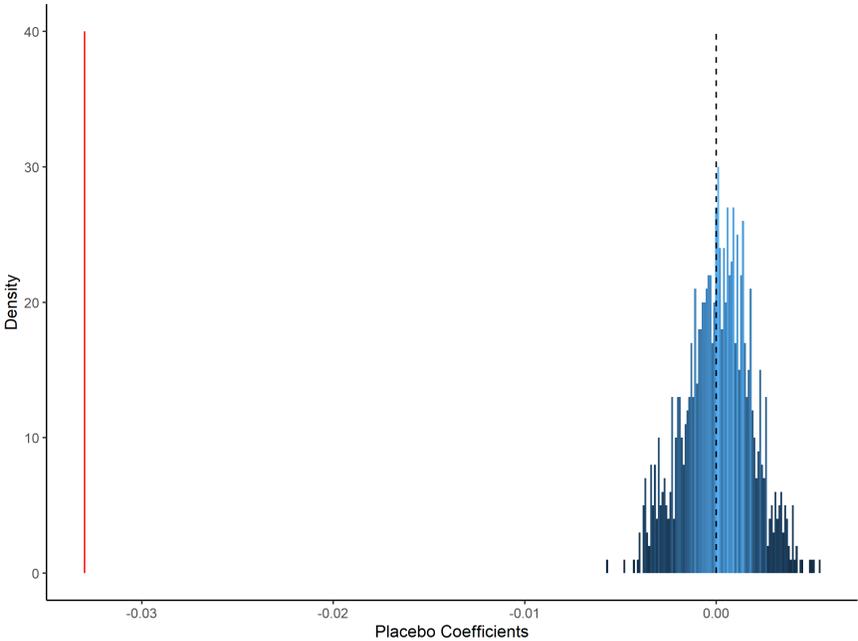
(b) Random Monitor Locations

Notes: This figure presents the results of two placebo tests (See Equation (1)). 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 A16: Event Study: The Effect of Automation on Air Pollution within 3km Across Different Types of Monitors
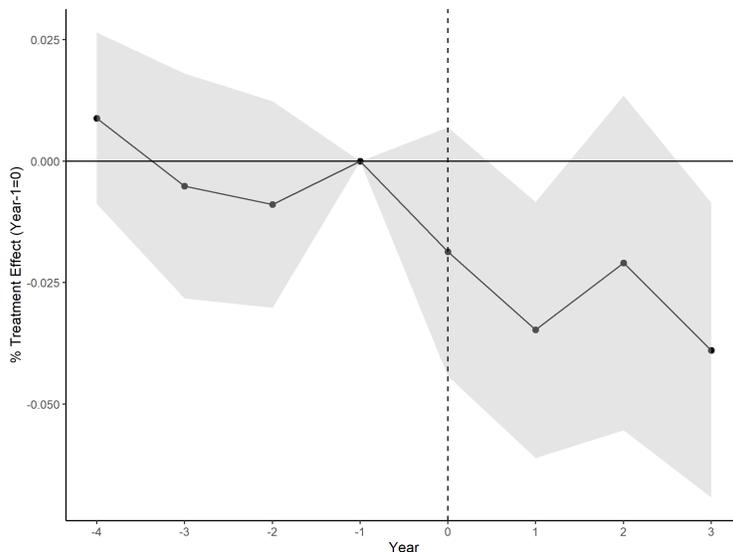
(a) Regional Assessing Monitors
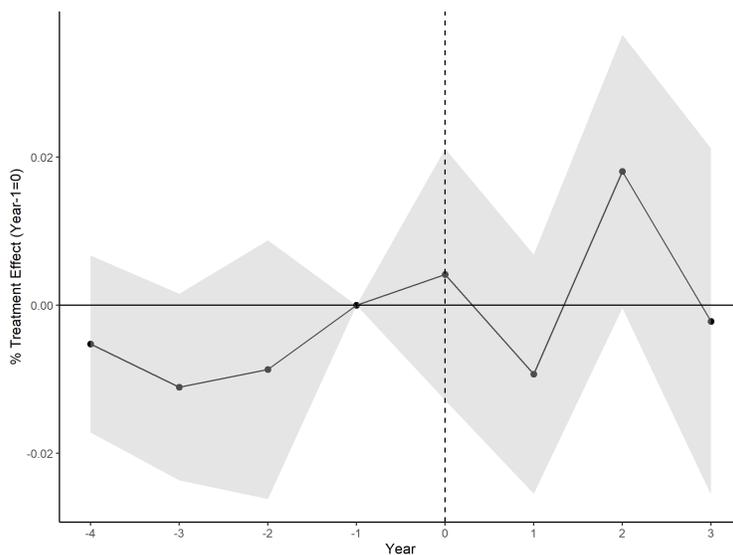
(b) Background Monitors

Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for β_n from Equation (3) by different types of monitors (Minitry of Ecology and Environment 2013). Panel A uses regional assessing monitors that are used to measure a city's pollution level. Panel b uses background stations that are placed far away from pollution sources and urban areas to serve as a reference. Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside the 3km radius) at a given period (also reported in Table A10). The omitted time category is the year before a city joined the automated monitoring program. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, and $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure A17: Share of the Population Exposed to Unhealthy Pollution Levels at Different Distances from Monitors
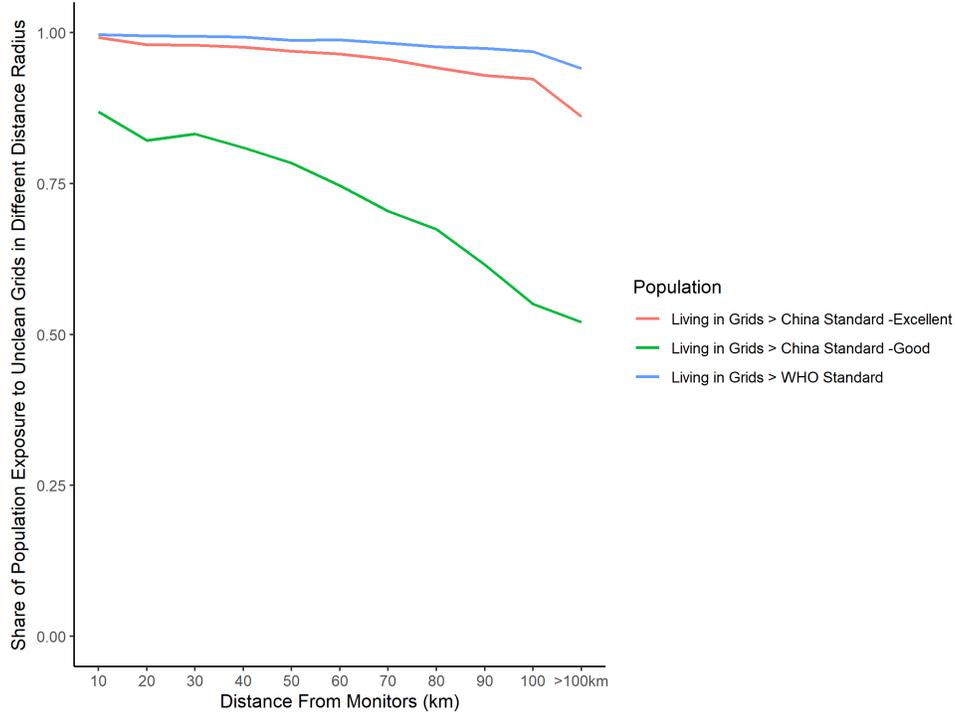
Notes: This figure displays the proportion of the population exposed to unhealthy levels of pollution at varying distances from the monitoring stations. The exposure to unhealthy levels of pollution is defined as residing in grid cells where the concentration of PM2.5 exceeds the established air quality standards–The World Health Organization (WHO) recommends a standard of 10 ug/m^3, while China specifies good air quality as 35 ug/m^3 and excellent air quality as 15 ug/m^3. To infer the health threshold of AOD-based PM$_{2.5}$ from the 10/15/35 ug/m3 standards with the ground monitoring data, we follow a two-step process. Firstly, we establish the relationship between AOD-imputed PM$_{2.5}$ and ground-based PM$_{2.5}$ using a regression model. Secondly, we pin down the AOD-based PM$_{2.5}$ levels when ground-based PM$_{2.5}$ takes on the value of 10/15/35 ug/m3.

Figure A18: Uneven Distribution of Air Quality Monitors Across Counties
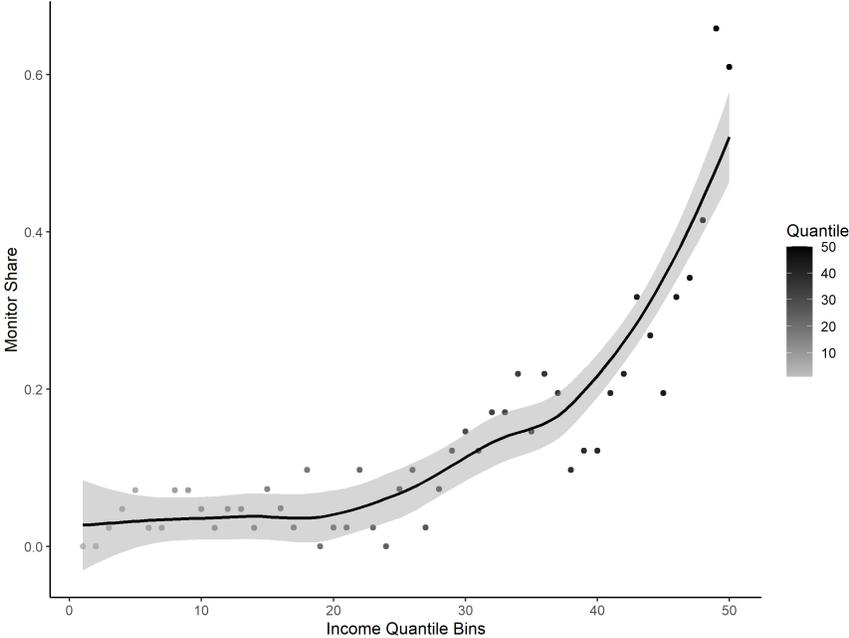
(a) By Income
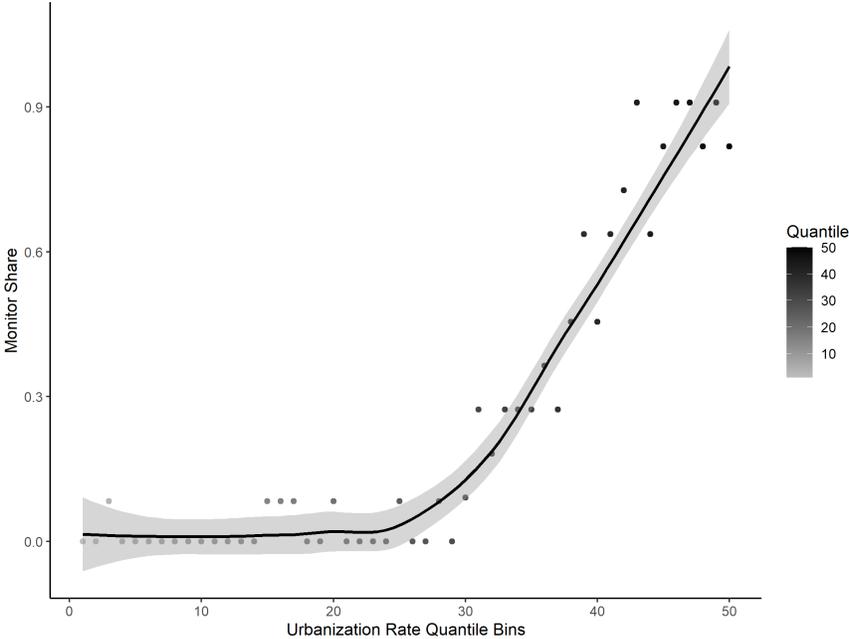
(b) By Urbanization Rate

Notes: This figure presents the monitor share across quantiles of counties. Panel A divides counties into 50 groups according to their GDP per capita during the pre-policy period (before 2012). Panel B categorizes counties into 50 groups based on their urbanization rate. The monitor share is defined as the percentage of counties with air quality monitors within their corresponding groups.

Figure A19: Event Study: The Effect of Monitor Automation on Air Pollution within 3km of a Monitor (2008–2015)
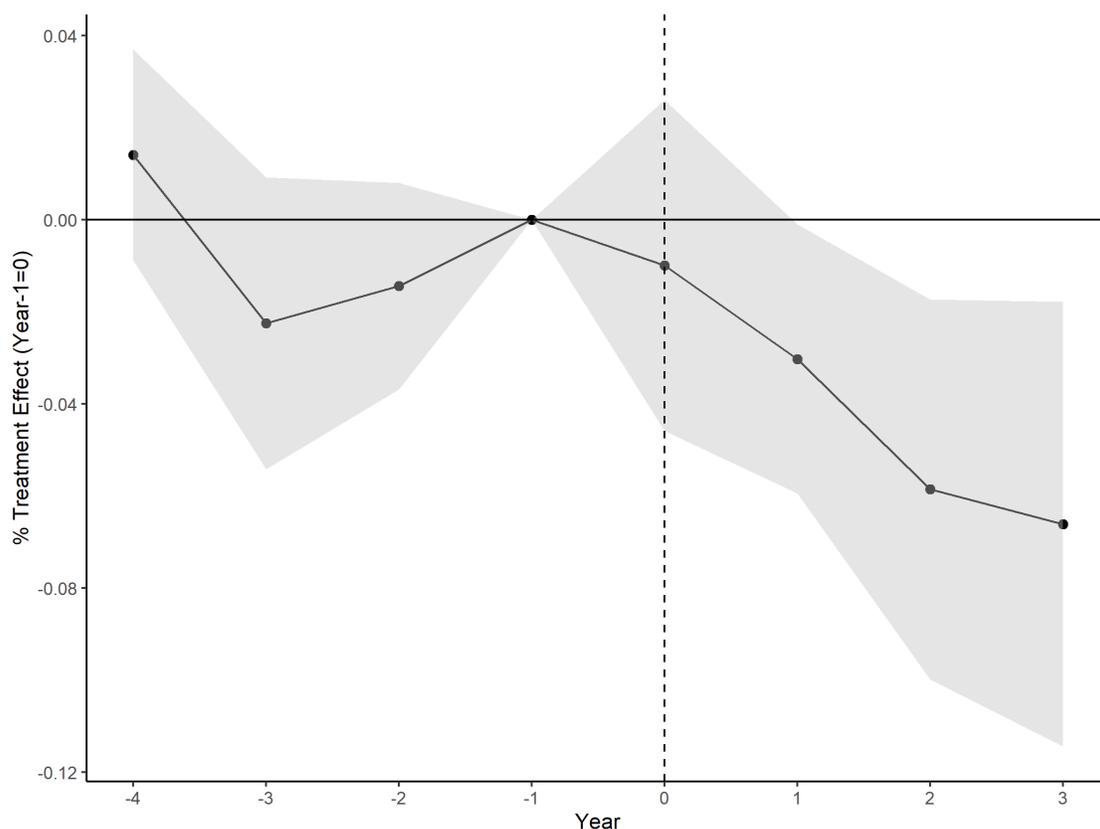
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for β_n from Equation (3). This figure depicts the periods from 2008 to 2015 and sets the sample in Wave 3 as the never treated group. Each estimate represents the difference in $PM_{2.5}$ between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period (also reported in Table A10). 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 flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, $PM_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Figure A20: Effects of Automation on lnPM$_{2.5}$ at Different Distances from Monitors
(Monthly)
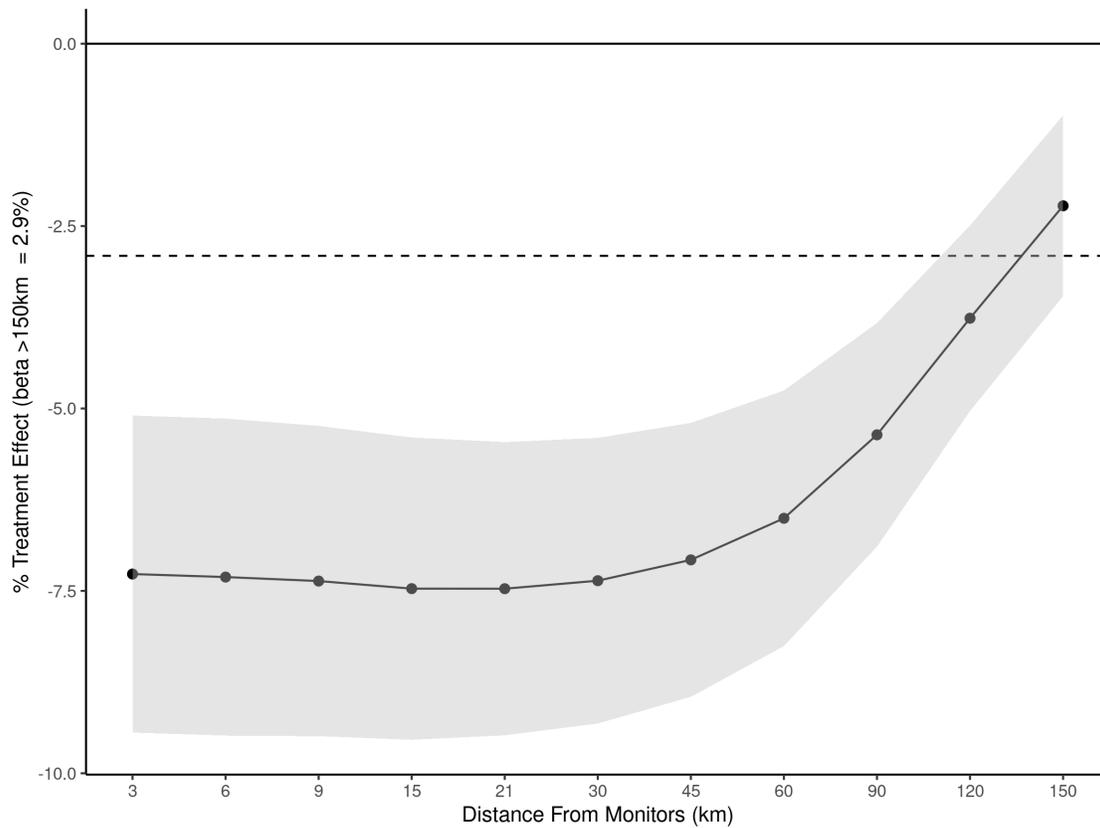
Notes: This figure plots the estimated coefficients and their 95% level confidence intervals for the effects of monitor automation on the satellite-based lnPM$_{2.5}$ at different distance bins from the monitor by using Monthly AOD-based PM$_{2.5}$ data. Each point estimate represents the pollution change in each distance bin relative to the baseline group at the outer range (distance to monitor >150 km), which on average experiences a 3.2% pollution increase. The absolute effect becomes positive above the dotted line. The regression includes cell fixed effects and year fixed effects, along with flexible interactions between year dummies and an array of pre-treatment city characteristics (such as average GDP, population, PM$_{2.5}$ at the city level from 2008 to 2011, the maximum distance between cells and monitors within a city and a dummy indicator for an environmental priority city), and city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level.

Table A1: Key government policy documents about Air Pollution Prevention during 2013–2017

Policy Name	Short Name	Issue Time
Action Plan for Air Pollution Prevention and Control	Air Ten	2013.09
Target Responsibility Agreement (mubiao zerenshu) for Atmospheric Pollution Prevention and Control	Target	2013.10
Notice of the General Office of the State Council on Performance Assessment Measures for Air Pollution Prevention and Control Action Plan	Assessment	2014.04

Notes: This table shows the key government policy documents on air pollution prevention and performance assessment during 2013–2017 and their issuing time. All are issued by the Ministry of Ecology and Environment (MEE) in China.

Table A2: Robustness Check: Localized Cleanup Response to Monitoring Program Automation, using AOD Raw Data

	Dependent variable: ln(AOD)			
	(1)	(2)	(3)	(4)
Unmonitored Areas:	>3km	>3km	>3km	>60km
Auto	-0.0044	0.0024	-0.0044	-0.0053*
	(0.0027)	(0.0027)	(0.0027)	(0.0031)
(0-3km)×Auto		-0.0198***	-0.0090***	-0.0180***
		(0.0028)	(0.0016)	(0.0032)
CellFE	X	X	X	X
Year FE	X	X	X	X
Year FE × citypopulation$^{2008-2011}$	X		X	X
Year FE × PM$_{2.5}^{2008-2011}$	X		X	X
Year FE × Other City-level Controls	X		X	X
Concurrent Policy	X		X	X
Observations	10,865,784	10,888,044	10,865,784	7,399,416
R^2	0.945	0.930	0.945	0.934

Notes: This table reports the effects of the monitor automation program on the satellite-based ln(AOD). Auto is the treatment indicator that switches on after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if the cells are located within a 3km radius of a city's monitoring stations. Columns (3) and (4) use cells within 3km of a monitor as the monitored group and compare them with different unmonitored groups: cells outside 3km and 60km of the monitors. PM$_{2.5}^{2008-2011}$ is the average city-level PM$_{2.5}$ and citypopulation$^{2008-2011}$ is the average city population over the 2008–2011 period. Other city-level controls include the average city-level GDP between 2008 and 2011, the number of monitors for each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A3: Localized Cleanup Response to Monitor Automation (Monthly)

	Dependent variable: $\ln(PM_{2.5})$					
	(1)	(2)	(3)	(4)	(5)	(6)
Unmonitored Areas:	>3km	>3km	>3km	>3km	>3km	>60km
Auto	0.0164	-0.0203	0.0168	-0.0154	-0.0202	-0.0168
	(0.0116)	(0.0140)	(0.0116)	(0.0115)	(0.0140)	(0.0233)
(0-3km)×Auto			-0.0622***	-0.0302***	-0.0247***	-0.0374**
			(0.0093)	(0.0073)	(0.0079)	(0.0157)
CellFE	X	X	X	X	X	X
Year Month FE	X	X	X	X	X	X
CellFE X Month FE		X		X	X	X
Year FE × citypopulation$^{2008-2011}$		X		X	X	X
Year FE × $PM_{2.5}^{2008-2011}$		X		X	X	X
Year FE × Other City-level Controls		X			X	X
Concurrent Policy		X			X	X
Observations	124,594,942	124,594,942	124,594,942	124,594,942	124,594,942	90,429,709
R^2	0.961	0.962	0.961	0.963	0.956	0.955

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 monthly satellite-based $PM_{2.5}$. Auto is the treatment indicator that equals one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if the cells are located within a 3km radius of a city's monitoring stations. Columns (1)–(6) use cells within 3km of the monitor as the monitored group, comparing them with different unmonitored groups: cells beyond 3km from the monitors in columns (1)–(5) and 60km from the monitors in column (6), respectively. $PM_{2.5}^{2008-2011}$ is average city-level $PM_{2.5}$ during the 2008–2011 period and citypopulation$^{2008-2011}$ is the average city population from 2008 to 2011. Other city-level controls are the average city-level GDP from 2008 to 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM_{10} and $PM_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A4: Localized Cleanup Response to Monitoring Program Automation: AOD (Monthly)

Dependent variable:	$\ln(\text{AOD}^{MonthlyMax})$		$\ln(\text{AOD}^{MonthlyMin})$		$\ln(\text{AOD}^{MonthlyMean})$	
	(1)	(2)	(3)	(4)	(5)	(6)
Auto	0.0014	-0.0101	-0.00009.33	-0.0028	0.0006	-0.0054**
	(0.0064)	(0.0069)	(0.0023)	(0.0022)	(0.0023)	(0.0025)
(0-3km)×Auto	-0.0372***	-0.0141*	-0.0147**	-0.0093	-0.0278***	-0.0139**
	(0.0096)	(0.0079)	(0.0062)	(0.0059)	(0.0072)	(0.0064)
CellFE	X	X	X	X	X	X
Year Month FE	X	X	X	X	X	X
CellFE X Month FE		X		X	X	X
Year FE × citypopulation$^{2008-2011}$		X		X		X
Year FE × AOD$^{2008-2011}$		X		X		X
Year FE × Other City-level Controls		X		X		X
Concurrent Policy		X		X		X
Observations	114,062,258	113,802,919	114,062,258	113,802,919	114,062,258	113,802,919
R^2	0.684	0.692	0.721	0.726	0.769	0.777

Notes: This table reports the effects of the monitor automation program on the satellite-based lnAOD. We construct different monthly AOD statistics based on the daily AOD data within each month. $\ln(\text{AOD}^{MonthlyMax})$ and $\ln(\text{AOD}^{MonthlyMin})$ is the natural logarithm of the cell-level maximum and minimum AOD value of each month, and $\ln(\text{AOD}^{MonthlyMean})$ is the natural logarithm of the cell-level average AOD value of each month. Auto is the treatment indicator that equals one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if the cells are located within a 3km radius of a city's monitoring stations. AOD$^{2008-2011}$ is average city-level AOD$_{MonthlyMax}$ during the 2008–2011 period and citypopulation$^{2008-2011}$ is the average city population from 2008 to 2011. Other city-level controls are the average city-level GDP from 2008 to 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A5: Robustness Check: Correction Via Multiple Imputations

Dependent Variable:	$\ln(PM_{2.5})$		
	(1)	(2)	(3)
Variables			
Auto	0.0250	-0.00989	0.00208
	(0.0147)	(0.0153)	(0.0164)
(0-3km)×Auto	-0.0586***	-0.0247***	-0.0234***
	(0.0112)	(0.00653)	(0.00701)
CellFE	X	X	X
Year FE	X	X	X
Year FE × citypopulation$^{2008-2011}$		X	X
Year FE × ln(PM2.5)$^{2008-2011}$		X	X
Year FE × Other City-level Controls			X
Concurrent Policy			X
Observations	10,413,717	10,413,717	10,413,717

Notes: This table reports the effects of the monitor automation program on the natural logarithm of the cell-level yearly satellite-based lnPM$_{2.5}$ with correction. Following the lead of Proctor, Carleton and Sum (2023), we employ bootstrap sampling to randomly select 70% of the ground-based monitoring data and then generate the remaining 30% of the sample through multiple imputations. We then utilize that sample of the 70% original data, and 30% imputed observations to perform regression analysis and simulate the relationship between satellite PM$_{2.5}$ values and their corresponding ground-based readings. Following that, we predict PM$_{2.5}$ values for all grids in our main dataset using the satellite data and the regression model derived in the previous step. This process is repeated 100 times. The parameters presented represent the means calculated from this distribution of bootstrap samples. Auto is the treatment indicator that equals one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if the cells are located within a 3km radius of a city's monitoring stations. PM$_{2.5}^{2008-2011}$ is average city-level PM$_{2.5}$ during the 2008–2011 period and citypopulation$^{2008-2011}$ is the average city population from 2008 to 2011. Other city-level controls are the average city-level GDP from 2008 to 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A6: Robustness Check: Localized Cleanup Response to Monitor Automation (2008–2015)

	Dependent variable: $\ln(\text{PM}_{2.5})$			
	(1)	(2)	(3)	(4)
Unmonitored Areas:	>3km	>3km	>3km	>60km
Auto	-0.029	-0.002	-0.029	-0.057*
	(0.021)	(0.024)	(0.021)	(0.029)
(0-3km)×Auto		-0.032***	-0.039***	-0.077***
		(0.012)	(0.012)	(0.027)
CellFE	X	X	X	X
Year FE	X	X	X	X
Year FE × citypopulation$^{2008-2011}$	X		X	X
Year FE × PM$_{2.5}^{2008-2011}$	X		X	X
Year FE × Other City-level Controls	X		X	X
Concurrent Policy	X		X	X
Observations	8,330,086	8,330,086	8,330,086	6,067,588
R^2	0.980	0.977	0.980	0.980

Notes: This table reports the effects of the monitor automation program on the satellite-based lnPM$_{2.5}$. The sample covers the 2008–2015 period, setting monitors treated in Wave 3 as the never treated ones. lnPM$_{2.5}$ 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 has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if the cells are located within a 3km radius of a city's monitoring stations. Columns (3) and (4) use cells within 3km of the monitor as the monitored group and compare it with different unmonitored groups: cells outside 3km and 60km of the monitors, respectively. PM$_{2.5}^{2008-2011}$ is the average city-level PM$_{2.5}$ and citypopulation$^{2008-2011}$ is the average city population over the 2008–2011 period. Other city-level controls are the average city-level GDP between 2008 and 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A7: Robustness Check: Mechanism of Localized Cleaning (2008-2015)–Thermal Anomalies

VARIABLES	(1) 1(TAP)	(2) ln(Days+1)	(3) ln(FRP+1)	(4) ln(Days+1)	(5) ln(FRP+1)
Auto	0.336***	-0.0469	-0.0325	0.00160**	0.00329**
	(0.0150)	(0.0414)	(0.0445)	(0.000651)	(0.00150)
Marginal Effect	*0.0839****	*-0.0144*	*-0.0230*		
	(0.00402)	*(0.0127)*	*(0.0315)*		
(0-3km)×Auto	-0.377***	-0.263***	-0.233***	-0.00676	-0.00528
	(0.0589)	(0.0448)	(0.0455)	(0.00558)	(0.0125)
Marginal Effect	*-0.0660****	*-0.0808****	*-0.165****		
	(0.0144)	*(0.0138)*	*(0.0322)*		
Cell FE	X	X	X	X	X
Year FE	X	X	X	X	X
Year FE $\times$ citypopulation$^{2008-2011}$		X	X	X	X
Year FE $\times$ PM$_{2.5}^{2008-2011}$		X	X	X	X
Year FE $\times$ Other City-level Controls		X	X	X	X
Concurrent Policys		X	X	X	X
Model	Logit	Poisson	Poisson	OLS	OLS
Sample	All	All	All	1(TAP)	1(TAP)
Observations	127,288	165,040	165,040	39,125	39,125
R-squared				0.743	0.583

Notes: This table reports the effects of the monitor automation program on thermal anomalies. The analysis uses the sample from 2008 to 2015 and sets monitors automated in Wave 3 as the never treated ones. 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 and Poisson regression models, the marginal effects are also reported. Column (1) reports the results for 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 for 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 for the average intensity of thermal anomalies, denoted by ln(FRP+1). FRP is defined as the rate of radiant heat output, which is related to the rate at which fuel is 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 grid cell-year observations when 1(TAP) 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. lnPM$_{2.5}$ is the natural logarithm of the cell-level yearly satellite-based PM$_{2.5}$. Auto is the treatment indicator that takes the value of one after a city has joined the automatic monitoring program. (0-3km) is a dummy variable that equals one if cells are located within 3km of a city's monitoring stations. PM$_{2.5}^{2008-2011}$ is the average city-level PM$_{2.5}$ and citypopulation$^{2008-2011}$ is the average city population over the 2008–2011 period. Other city-level controls are the average city-level GDP from 2008 to 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A8: 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	24.35***	44.26***	22.91***	37.70***
	(0.1384)	(0.000)	(0.1408)	(0.000)
Marginal Effect	*0.997****	*0.996****	*0.995****	*0.992****
	(0.000)	*(0.000)*	*(0.001)*	*(0.000)*
City FE		X		X
Observations	95,168	68,161	380,672	240,188

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 if there are any polluting plants within a 10km-by-10km cell. The polluting plants come from 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 is 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 that equals 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 A9: Validating Thermal Anomalies Data: Intensive Margins

	Dependent variable: $\ln(PM_{2.5})$	
Sample	Polluting firms	Power plants
	(1)	(2)
$\ln(FRP+1)$	0.129***	0.136***
	(0.009)	(0.004)
Observations	1,806	10,491
R-squared	0.108	0.102

Notes: This table shows the relationship between the intensity of the thermal anomalies observed and the satellite-derived pollution levels of firms or power plants at the intensive margins. The samples are restricted to only those grid cells with polluting firms or power plants. The polluting plants are defined using 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 within a 10km radius of polluting firms, which is based on 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.

Table A10: Event Study: The Effect of Monitor Automation on Air Pollution within 3km of a Monitor

	Dependent variable:	
	$\ln(\text{PM}_{2.5})$	$\ln(\text{AOD})$
	(1)	(2)
(0-3km)×before4	0.011	-0.0007
	(0.009)	(0.0023)
(0-3km)×before3	-0.001	0.0037
	(0.012)	(0.0028)
(0-3km)×before2	-0.004	-0.0030
	(0.011)	(0.0031)
(0-3km)×after0	-0.018	-0.0006
	(0.015)	(0.0029)
(0-3km)×after1	-0.033**	-0.0048*
	(0.013)	(0.0028)
(0-3km)×after2	-0.020	-0.0081***
	(0.017)	(0.0031)
(0-3km)×after3	-0.036**	-0.0163***
	(0.014)	(0.0031)
CellFE	X	X
YearFE	X	X
Year FE × citypopulation$^{2008-2011}$	X	X
Year FE × PM$_{2.5}^{2008-2011}$	X	X
Year FE × Other City-level Controls	X	X
Concurrent Policy	X	X
Observations	10,413,717	10,407,855
R-squared	0.975	0.964

Notes: The table reports the event study results of monitor automation on air pollution with different dependent variables. Column (1) shows the effect of monitor automation on $\ln(\text{PM}_{2.5})$ of monitored areas (within 3km of a monitor; Figure 3), and column (2) shows the effect of monitor automation on the annual $\ln(\text{AOD})$ in monitored areas (within 3km of a monitor; Figure A4). All regressions control for cell-fixed effects, year-fixed effects, and time dummy interactions. PM$_{2.5}^{2008-2011}$ is average city-level PM$_{2.5}$ and citypopulation$^{2008-2011}$ is the average city population over the 2008–2011 period. Other city-level controls are the average city-level GDP from 2008 to 2011, the number of monitors in each city, the maximum distance between cells and monitors within a city, and a dummy variable that indicates whether or not a city is an environmental priority city. The Concurrent Policy refers to the city-level concurrent PM$_{10}$ and PM$_{2.5}$ reduction targets. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A11: Placebo Effects of Pseudo Automation Treatment on Cities without Monitors

	(1) Environment Bureau	(2) Government Offices	(3) All	(4) Environment Bureau	(5) Government Offices	(6) All
	Dependent variable: $\ln(PM_{2.5})$					
Monitor Area:						
Automation Wave:		2			3	
(0-3km)×Auto	-0.070	-0.022	-0.059	-0.063	-0.016	-0.054
	(0.052)	(0.021)	(0.033)	(0.053)	(0.014)	(0.034)
CellFE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Year FE × citypopulation$^{2008-2011}$	X	X	X	X	X	X
Year FE × PM$_{2.5}^{2008-2011}$	X	X	X	X	X	X
Observations	163,746	163,746	163,746	163,746	163,746	163,746
R^2	0.980	0.980	0.980	0.980	0.980	0.980

Notes: This table presents the placebo effects of monitor automation on cities that had never received the monitoring automation treatment. lnPM$_{2.5}$ is the natural logarithm of the cell-level yearly satellite-based PM$_{2.5}$. To make them a comparable control group to our treatment group, we identify "placebo" monitor spots in these nine cities. By checking the existing monitor siting rules, we assigned the counterfactual monitor/s to the location of 1) the municipal Environmental Protection Bureau, 2) the municipal government building, or 3) both. Further, we assigned their fake automation timing to be either in Wave 2 or 3, denoted by Auto. (0-3km) is a dummy variable that equals one if the cells are located within a 3km radius of a city's "pseudo" monitoring stations. In Column (1) and Column (4), monitors are sited in the environment bureau buildings, while in Column (2) and Column (5), they are sited in government office buildings. In Columns (3) and (6), monitors are assigned to both environment bureau and government office buildings. PM$_{2.5}^{2008-2011}$ is average city-level PM$_{2.5}$ during the 2008–2011 period and citypopulation$^{2008-2011}$ is the average city population from 2008 to 2011. Standard errors are clustered at the city level. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table A12: Summary of Government Documents Mentioning Strategic Cleaning

Type	Policy Measures	Number of Government Documents
Coal and other energy pollution control	Clean energy replacement	2
	Boiler renovation	16
Transportation pollution control	Yellow-label vehicles (high-emission vehicles)	2
	Travel restrictions	2
Dust pollution control	Spraying Water	16
	Windproof and dust suppression nets	6
	Wet cleaning	2
	Dust suppression/suction vehicles	3
Agricultural and other pollution control	Banning open burning	4
	Banning outdoor cookings	22
Industrial pollution control	Shutdown	3
	key monitoring enterprises	2

Notes: This table reports measures for strategic cleaning which mention in the government documents and the number of documents for each kind of strategic cleaning measures. There are 121 government documents that mention strategic cleaning in total.

Table A13: Mechanism: Effect of Automation on Local Relative Humidity (Monthly)

Dependent Variable:		ln(Humidity)			
	(1)	(2)	(3)	(4)	(5)
Unmonitored Areas:	All	All	All	Winter	Summer
Variables					
Auto	-0.0053	-0.0054	0.0025	-0.0057	0.0064
	(0.0065)	(0.0065)	(0.0077)	(0.0100)	(0.0084)
(0-3km)×Auto		0.0288***	0.0210**	0.0335***	0.0094
		(0.0082)	(0.0083)	(0.0084)	(0.0077)
Cell	X	X	X	X	X
Yearmonth	X	X	X	X	X
Climate Controls			X	X	X
Fit statistics					
R^2	0.92948	0.92948	0.95959	0.96254	0.95638
Observations	11,222,880	11,222,880	11,222,880	4,676,200	6,546,680

Notes: This table presents the effects of the monitor automation program on satellite-based relative humidity (monthly) with meteorological data from He et al. (2020). The variable "Auto" is a treatment indicator that equals one after a city has joined the automatic monitoring program, while the dummy variable "(0-3km)" equals one if the cells are located within a 3km radius of a city's monitoring stations. The climate controls include temperature, precipitation, and wind. Columns (1) to (3) reports the results for the whole sample period. Column (4) reports the results for the winter period from October of one year to February of the next year, while Column (5) reports the estimation results for the summer period from March to September of a year. Standard errors are clustered at the city level, and significance levels are indicated by asterisks: *p<0.1; **p<0.05; ***p<0.01.

Table A14: Mortality and Morbidity Impacts of Uneven Pollution Control

Distance from Monitors	Population ($\times 10^3$ person)	Excess Death Reduction ($\times 10^3$ person)	Total Monetary Value of Excess Death Reduction (billion $)	Total Healthcare Spending Savings from PM_2.5 Reduction (billion $)	Per capita Monetary Value of Excess Death Reduction ($)	Per capita Healthcare Spending Savings from PM_2.5 Reduction ($)
	(1)	(2)	(3)	(4)	(5)	(6)
(0-3km)	118970.12	13.06	15.02	0.25	126.22	2.10
(3-6km)	92787.05	9.22	10.6	0.18	114.24	1.90
(6-9km)	47720.46	4.29	4.93	0.08	103.41	1.72
(9-15km)	72816.85	5.94	6.83	0.11	93.85	1.56
(15-21km)	63986.63	4.81	5.53	0.09	86.41	1.44
(21-30km)	87327.55	6.05	6.95	0.12	79.63	1.33
(30-45km)	127270.47	7.29	8.39	0.14	65.88	1.10
(45-60km)	107054.99	5.29	6.08	0.1	56.84	0.95
(60-90km)	128741.45	4.15	4.77	0.08	37.09	0.62
(90-120km)	54978.79	0.51	0.59	0.01	10.67	0.18
(120-150km)	21358.18	-0.34	-0.4	-0.01	-18.52	-0.31
(>150km)	22651.89	-1.54	-1.77	-0.03	-78.36	-1.30
Total	945664.44	58.72	67.53	1.12	677.37	11.28

Notes: This table presents various distance bins from the monitors, the monetized health benefits of automation— the value of $PM_{2.5}$-attributable death reduction and healthcare spending saved annually. Column (1) shows the corresponding population of each distance bin. Columns (2)–(6) report the benefits, as shown in the heading. These outcomes are computed using the pollution reduction from automation, which is denoted by $\Delta PM_{2.5}^{welfare} = \alpha \times Auto \times PM_{2.5}^{pre} + \sum_{n=(0-3)}^{(>150)} \beta_n Auto_{ct} \times Bin_n \times PM2.5_{2.5}^{pre}$. $PM_{2.5}^{pre}$ is the average cell-level yearly satellite-based $PM_{2.5}$ from 2008 to 2011 (the pre-treatment period). *Auto* is the treatment indicator that equals one after a city has joined the automatic monitoring program. The annual reduction of excess deaths in column (2) is equal to the $PM_{2.5}$-attributable monthly mortality rate 3.25% from He, Liu and Zhou (2020) (i.e., a 10 $\mu g/m^3$ increase in $PM_{2.5}$ increases monthly mortality by 3.25%) × pollution reduction calculated above $(\Delta PM_{2.5})/10 \times$ population$_{2015} \times$ 12 months. Based on Fan, He and Zhou (2020), the average value of a typical Chinese person's statistical life was around 1.15 million USD in 2015. Column (3) then infers the monetary value of lives saved from $PM_{2.5}$ reduction. According to Barwick et al. (2018), a medium-run reduction of 10 $\mu g/m^3$ in daily $PM_{2.5}$ would lead to \$22.4 annual savings in healthcare spending per household. Given that the average household size is 3 people (source: National Bureau of Statistics, UNICEF China, UNFPA China, 'Population Status of Children in China in 2015: Facts and Figures', 2017) and using population data in 2015, the healthcare spending savings are thus calculated as population$_{2015}/3 \times \Delta PM_{2.5}/10 \times$ \$22.4 in Column (4). Columns (5) and (6) report the per capita health benefits from reduced mortality and morbidity.