DISCUSSION PAPER SERIES

DP15622

INFORMED ENFORCEMENT: LESSONS FROM POLLUTION MONITORING IN CHINA

Sebastian Axbard and Zichen Deng

DEVELOPMENT ECONOMICS

ISSN 0265-8003

INFORMED ENFORCEMENT: LESSONS FROM POLLUTION MONITORING IN CHINA

Sebastian Axbard and Zichen Deng

Discussion Paper DP15622
Published 31 December 2020
Submitted 30 December 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

INFORMED ENFORCEMENT: LESSONS FROM POLLUTION MONITORING IN CHINA

Abstract

Government regulations are often imperfectly enforced by public officials. In this study, we investigate if real-time monitoring of policy outcomes can improve enforcement of existing regulations by exploring the introduction of air pollution monitors in China. Exploiting assignment criteria established by the central government and new geo-referenced data on local enforcement activities, we show that monitoring: 1) increases enforcement against local firms, 2) improves the targeting of enforcement, and 3) reduces aggregate pollution. These effects are driven by officials facing performance incentives and are stronger when there is limited scope for data manipulation, suggesting that real-time monitoring improves top-down accountability.

JEL Classification: O13, Q53, Q58

Keywords: accountability, Regulatory Enforcement, pollution, China

Sebastian Axbard - s.axbard@qmul.ac.uk
Queen Mary, University of London and CEPR

Zichen Deng - zichen.deng@nhh.no
NHH Norwegian School of Economics

Informed Enforcement
Lessons from Pollution Monitoring in China*

Sebastian Axbard† Zichen Deng‡

This Version: Wednesday 30^{th} December, 2020

Abstract

Government regulations are often imperfectly enforced by public officials. In this study, we investigate if real-time monitoring of policy outcomes can improve enforcement of existing regulations by exploring the introduction of air pollution monitors in China. Exploiting assignment criteria established by the central government and new geo-referenced data on local enforcement activities, we show that monitoring: 1) increases enforcement against local firms, 2) improves the targeting of enforcement, and 3) reduces aggregate pollution. These effects are driven by officials facing performance incentives and are stronger when there is limited scope for data manipulation, suggesting that real-time monitoring improves top-down accountability.

Keywords: Accountability, Regulatory Enforcement, Pollution, China
JEL: O13, Q53, Q58

*We are grateful to seminar participants at University of Zurich, University of York, BI Norwegian Business School, Helsinki, CEIBS, NHH, University of Amsterdam, Renmin University, VU Amsterdam, Queen Mary University of London, Warwick, TI Jamboree, University of East Anglia, University of Gothenburg, EBRD Conference on Corruption and Anti-Corruption Policies (Kyiv) and NEUDC 2019 (Northwestern University) for many useful comments and suggestions. We also thank Pengzhan Qian, Dan Xie, and Jordan Ashmore for excellent research assistance.

†Queen Mary, University of London & CEPR, s.axbard@qmul.ac.uk
‡NHH Norwegian School of Economics, zichen.deng@nhh.no

1 Introduction

Across the globe, there is a substantial discrepancy between central government regulations and actual enforcement of those regulations at the local level. This gap exists across a wide range of policy areas and is particularly severe in low- and middle-income countries (CIPE, 2012; World Bank, 2017). A common practice to address the principal–agent problem inherent in the delegation of authority to lower levels of government is to provide high-powered incentives to implementing officials to ensure that their interests are better aligned with those of the policymaker.[1] However, such incentive schemes require reliable information on the actions of agents or local policy outcomes. In many settings, such information is either not widely available, of poor quality, or could easily be manipulated by local officials who have an interest in misreporting due to the incentives they face (Jacob and Levitt, 2003; Figlio and Winicki, 2005; Figlio and Getzler, 2006; Banerjee, Duflo, and Glennerster, 2008; Sandefur and Glassman, 2015; Fisman and Wang, 2017; Greenstone et al., 2019; Acemoglu et al., 2020).

This paper explores how a technology that enables the central government to directly monitor local policy outcomes in real time can overcome the gap in enforcement. More specifically, we study the rollout of air pollution monitors in China – a setting where local officials face strong incentives to reduce pollution under centrally set targets – and investigate how that affects local governments' enforcement of air pollution regulations as well as local pollution levels. Our focus on environmental policy is motivated by recent reporting from the United Nation (2019) arguing that a lack of enforcement of environmental regulations is one of the greatest obstacles that needs to be overcome in order to combat climate change and pollution. Despite international efforts in recent years to improve air quality, more than 90% of the world's population in 2016 (WHO, 2016) still lived in areas where air pollution exceeded World Health Organization guidelines with far-reaching consequences for both health and productivity (Neidell and Currie, 2005; Greenstone and Hanna, 2014; Ebenstein et al., 2017; Jia, 2017; Barwick et al., 2018). A large part of this population lives in emerging economies, including China, where pollution levels have exceeded the highest levels ever recorded in rich countries.

We begin by investigating how a central government-led program that introduced 552 pollution monitors in 2015 has shaped the enforcement activities of prefecture-level gov-

[1] The theoretical literature has focused on how incentives could be designed to ensure the motivation of agents while decreasing any distortionary impact on effort (Holmström, 1979; Holmström and Milgrom, 1991; Baker, Gibbons, and Murphy, 1994).

ernments in China.[2] To conduct this analysis, we collect more than 55,000 environmental enforcement records from local governments. We then classify these records and identify the firm involved, the type of regulation violated, and the punishment imposed. Using this information, we estimate a flexible difference-in-differences model, which compares firms located close to a monitor with firms located further away from the monitor but within the same city. The results show an increase in the probability of enforcement by 60% for firms located within 10 km of a monitor, consistent with anecdotal evidence suggesting that cities stepped up enforcement activities close to the monitors after their introduction (see discussion in Section 4.1 and Figure D9 in the Appendix). The main threat to identification – potential endogenous placement of monitors – is mitigated in this setting because the placement of monitors followed strict guidelines issued by the central government. We support this claim by documenting that the placement of monitors is unrelated to prior enforcement activity and that there are no differential pre-trends for firms located at different distances from the monitor. In addition, we show that air pollution monitoring does not affect enforcement of other environmental regulations related to water or solid waste pollution.

To shed further light on how government actions are affected, we investigate how the type of enforcement carried out changes in the presence of monitoring. We document that local governments target and impose stricter punishment against high-polluting firms. City governments also become more responsive to local pollution levels once monitors have been introduced. To show this, we exploit exogenous shocks to local pollution induced by fluctuations in rainfall. We show that enforcement is higher when rainfall is low (and pollution is high) in the presence of monitoring, but that no such relationship exist when there is no monitoring. This suggests that monitors can ensure a more efficient response by local enforcement agencies – mitigating concerns that our results are driven by a uniform increase in enforcement around all monitors.

Building on the above evidence that local enforcement efforts against firms increases in the vicinity of monitors, we move on to study the pollution monitoring program's citywide effects. The focus on the city level allows us to capture the aggregate impact on pollution (including any within city spillover or displacement effect).[3] By exploiting plausibly exogenous variation in the number of monitors installed in different cities, we can assess the impact of more extensive monitoring (covering a larger share of the local firm distribution)

[2]Our sample focuses on the 177 prefecture-level cities that received a monitor for the first time in 2015. The prefecture-level city is an administrative division ranking below a province and above a county. Figure D8 shows an image of the type of monitors that we study.

[3]As depicted in Panel B of Figure D1 cities are large geographical units. Due to the administrative structure in China and the large distance between the urban centres of different cities, we are not concerned about across city spillovers.

on total enforcement and pollution outcomes. To capture overall pollution changes at the city-level, we follow previous literature and use satellite data on the aerosol optical depth (AOD).[4] The AOD data enable us to measure pollution across the whole city both before and after the introduction of monitors and provide us with a reliable data source that cannot be manipulated by local officials. To address potential endogenous installation of monitors, we exploit the strict rules established by the central government, which assign monitors to cities based on their population and geographical size. Using this information, we employ three different empirical strategies: a standard difference-in-differences specification for cities with a different number of monitors, an instrumental variable approach that instruments the number of monitors by the assigned number, and a regression discontinuity specification that exploits assignment cutoffs. All three empirical strategies produce consistent estimates and show that one additional monitor increases enforcement activities by about 20% and reduces pollution by about 3%. This is a sizable effect given that the median number of monitors assigned to a city in our sample is 3.

Our preferred interpretation of the above results is that monitors improve the central government's ability to hold local officials accountable for their actions. In this setting, local mayors face promotion incentives (as discussed above, this is a common approach to address the principal–agent problem) and are specifically evaluated on their ability to achieve predefined pollution reduction targets set by the central government. To empirically assess the validity of this interpretation, we follow Xi, Yao, and Zhang (2018) and exploit discontinuities in promotion incentives caused by the age of local mayors at the time of the National People's Congress. Estimating our baseline empirical model for mayors facing different promotion probabilities, we find evidence suggesting that monitoring is only effective when mayors face performance incentives. Hence, this finding is in line with pollution monitoring strengthening top-down accountability and through that making existing performance incentives more effective.

An alternative mechanism explaining our results is that monitors improve bottom-up accountability.[5] This is possible in this setting because the real time air pollution data is made publicly available on the website of the Ministry of Environmental Protection. To evaluate this mechanism, we investigate whether additional monitors strengthen local awareness of pollution by studying data on city-level online searches for pollution-related keywords. However, we find limited evidence suggesting that monitoring increases citizens' awareness of pollution and therefore conclude that this is unlikely to be an important mechanism in

[4]We provide validation of the satellite data using ground station measures in subsequent periods.

[5]Previous empirical studies (Chen, Pan, and Xu, 2016; Meng, Pan, and Yang, 2017) have found that authoritarian regimes are also responsive to citizen pressure. It is plausible that citizens better informed about pollution will pressure the local government to act in our setting.

this setting.

Finally, as discussed above, there are two main reasons why information about policy outcomes may be lacking or of poor quality in low- and middle-income countries: capacity constraints and misreporting. The policy we study is potentially reducing both of these factors at the same time. To shed some light on the relative importance of the two factors, we take advantage of an additional policy shift – the reassignment of control of the monitors from the local government to external third parties. This reassignment decouples the information provision responsibility from the enforcement of regulation responsibility and was conducted after it was discovered that several local governments tried to manipulate the data from the monitors. By exploiting information from the monitors as well as our satellite-based measure of pollution, we show that the monitor recordings are more strongly correlated with the satellite data when they are under the control of a third party – consistent with a reduction in manipulation. Following this logic, we further document that when monitors are under the control of the independent third party, the effect of an additional monitor on enforcement and pollution is substantially larger. This provides suggestive evidence that not only the capacity to collect information is important for top-down accountability, but also the way in which this information is provided.

This paper contributes to three strands of literature. First, it relates to a growing empirical literature studying policies aimed at reducing pollution in developing countries. Prior work has documented that regulatory changes can bring about pollution reduction (Greenstone and Hanna, 2014; Tanaka, 2015; Ebenstein et al., 2017) and that the incentives faced by both local leaders (Kahn, Li, and Zhao, 2015) and auditors matter for policy outcomes (Duflo et al., 2013). However, the literature also emphasizes that enforcement of environmental regulations is a major challenge (see, e.g., discussion in Greenstone and Hanna, 2014) and that we know little about how to improve it in developing countries (Shimshack, 2014). For example, simply increasing the rate of environmental inspections does not seem to have any substantial impact on compliance and environmental outcomes due to the importance of regulatory discretion (Duflo et al., 2018). Our findings suggest that improved monitoring of local pollution – a policy that strengthens top-down accountability without reducing regulatory discretion – could be an effective way of addressing the enforcement gap and reducing pollution. Hence, our work suggests that automatic pollution monitoring could be an effective policy instrument to address high levels of pollution in developing countries. We also relate to two concurrent studies that investigate other dimensions of the same pollution monitoring program (Greenstone et al., 2019; Barwick et al., 2020). Barwick et al. (2020) investigate the impact of sharing air pollution information with the public and show how that leads to avoidance behavior, while Greenstone et al. (2019) study how the updating of monitors in major cities (as opposed to the rollout of new monitors in smaller cities that we study) improved air pollution data quality and reduced the scope for manipulating the data. An additional related concurrent paper is He, Wang, and Zhang (2020), which studies how water pollution monitoring affect firm performance and document that firms immediately upstream of a water monitor have lower productivity than those immediately downstream. Our work complements these studies by showing how air pollution monitoring affect the enforcement behavior of local governments and aggregate pollution levels.

Second, we contribute to an extensive literature showing that monitoring and the provision of information can improve accountability and government performance (Besley and Burgess, 2002; Olken, 2007; Snyder and Strömberg, 2010; Reinikka and Svensson, 2005, 2011; Kosack and Fung, 2014; Avis, Ferraz, and Finan, 2018). While the broader literature has considered the impact of media as well as of audits, we are most closely aligned with recent work showing how information technology affects government performance and efficiency (Duflo, Hanna, and Ryan, 2012; Muralidharan, Niehaus, and Sukhtankar, 2016; Dhaliwal and Hanna, 2017; Banerjee et al., 2020). Proponents of such technological innovations have argued that they could increase efficiency, reduce the scope for manipulation and be implemented at a relatively low cost. Our study differs from previous work by focusing on monitoring of the final policy outcome (pollution), rather than intermediate inputs in policy production – such as public official attendance (Duflo, Hanna, and Ryan, 2012; Dhaliwal and Hanna, 2017) or transfer of funds (Muralidharan, Niehaus, and Sukhtankar, 2016; Banerjee et al., 2020). While the monitoring of final policy outcomes might not always be feasible, it could mitigate concerns about multitasking (Holmström and Milgrom, 1991) associated with intermediate monitoring. We show that policy outcome monitoring can indeed be effective in the context of pollution. In addition, we expand prior work by studying how enforcement of regulations as opposed to public service provision is affected by monitoring.

Third, we relate to a literature investigating the potentially distorting effect of high-powered incentives on data reporting (Banerjee, Duflo, and Glennerster, 2008; Fisman and Wang, 2017; Acemoglu et al., 2020), including manipulating pollution data (Andrews, 2008; Chen et al., 2013; Ghanem and Zhang, 2014; Oliva, 2015). We contribute to this literature by studying how control over the information infrastructure (shifting from local governments to external firms) is correlated with the quality of information as well as government actions and actual policy outcomes. While we are cautious when interpreting the results from this analysis due to the strong assumptions required for causal inference, it has the benefit that we can observe both potentially manipulated data from monitors as well as satellite data independent of government influence (and therefore also policy impact).

The paper is structured as follows. Section 2 describes the context as well as the rollout of the pollution monitoring program we investigate. After that, the data used in this study is described (Section 3). The first analysis, which explores firm-level evidence on enforcement, is presented in Section 4.1. The causal effect of pollution monitoring on enforcement and actual pollution at the city level is reported in Section 4.2. These two sections present both the respective empirical strategies and results. The analysis of the mechanisms is discussed in Section 5. Finally, Section 6 offers concluding remarks.

2 Institutional Context

This section provides background information and describes the context in which the national monitoring program studied in this paper was introduced. In the first subsection 2.1, we describe the environmental policies in place in China during this period and discuss the local leaders' role in achieving them. After that, the program rolled out to monitor these policies' implementation is described in subsection 2.2.

2.1 Environmental Policies in China

While the Chinese government's priority during the past decades has largely been to stimulate economic growth, attention has lately shifted towards environmental policies (Zheng and Kahn, 2017).[6] Starting in 2013, the National Air Quality Action Plan was set up to improve air quality by the end of 2017. As a part of China's successful "war on pollution" (Greenstone and Schwarz, 2018), the plan laid out the general goal for the whole country and set differentiated goals for each region. In January 2014, the Ministry of Environmental Protection (MEP) entered into "contracts" with all 31 provinces and set up a three-year air quality plan to decrease the concentration of particulate matter (PM) in the whole country. In each "contract", an air quality target for 2017 was set – resulting in different percentage reduction targets of $PM_{2.5}/PM_{10}$ for each province relative to the 2012 level.[7]

These centrally set targets are implemented by local government officials, who are incentivized to fulfill them through performance-based promotions. Promotions are the key instrument used in China to ensure that local officials carry out policies in line with the goals set by the central government (see Zheng and Kahn, 2013, 2017, for further discussion of this topic). For a long time, the central government focused on economic performance

[6]The concentration of air pollutants in China is among the world's highest and is a problem with serious health consequences. Average $PM_{2.5}$ (particulate matter with a diameter of 2.5 μm or less) concentrations in 2013 were 91 $\mu g/m^3$, which is nine times the amount the World Health Organization considers safe. Estimates by Greenstone and Schwarz (2018) suggest that if these levels of pollution are sustained, it will result in a 6.5 year decline in life expectancy for the average resident.

[7]For the list of targets by province, see Table C2 in Appendix C.

and emphasized economic growth as the key evaluation criteria for local officials' promotion (Chen, Li, and Lu, 2018). However, from the 12^{th} Five-Year Plan onward, the central government have used the fulfilment of environmental performance targets as a requirement for the promotion of local mayors (Zheng and Kahn, 2013).

2.2 National Monitoring System

To address issues raised about limited coverage and quality of existing pollution data, the central government introduced a new monitoring system as a part of its 2013 National Air Quality Action Plan. This new system expanded coverage to all of China – introducing monitors in prefecture-level cities that previously had no systematic air pollution monitoring in place. In addition, cities with existing monitors received new updated monitors that could capture the wider range of pollutants included in the revised air pollution standards (notably, $PM_{2.5}$, widely regarded as the key measure of ambient air pollution, was included for the first time). One of the key features of the new system is that all monitoring stations report six pollutants (SO_2, NO_2, CO, PM_{10}, $PM_{2.5}$, and O_3) to the central government in real time (Greenstone et al., 2019). Hourly pollution data is then automatically published online by the central government.

The new monitors were installed in three separate phases. The first phase was conducted in 2013 and focused on 74 major cities that represented the country's key population and economic centers.[8] The second phase was implemented in 2014 and focused on an additional 87 cities, that were covered either because they were Environmental Improvement Priority Cities or because they were part of a larger policy package aimed at improving environmental outcomes in the Shandong province.[9] The primary aim of the first two phases was to automate old manual monitors.[10] The main expansion phase, which is the one we focus on in this paper, was carried out in the following year when all 177 remaining prefecture-level cities (53% of all prefecture-level divisions in China) installed monitors. After this final expansion, all prefecture-level cities had at least one air quality monitor. These monitors all started transmitting information to the central government from January 1, 2015.

The MEP provided detailed instructions for how many monitors should be installed and where they should be located. All the monitors were installed in the so called "built-up area" – the main urban center of the prefecture-level city. The number of monitors installed in each city was determined by the city's population size and the geographical size of the

[8] The Beijing–Tianjin–Hebei Metropolitan Region, the Yangtze River Delta, the Pearl River Delta, directly administered municipalities, and provincial capitals.

[9] See this link for a description of the policy package.

[10] 113 out of the 161 cities in the first two waves had manual pollution monitoring in place before the new monitors were introduced.

built-up area. The detailed assignment criterion, which we use for identification, is presented in Table C1. Each monitor's precise location was chosen by a simulation method that took surrounding buildings, traffic, and the direction of seasonal winds into account to make sure that the monitors captured a fair representation of local pollution. The location of these monitors is depicted in the map in Panel A of Figure D1.

The funding for the monitors was provided by the province-level environmental bureaus. Once all equipment had been put in place, the city-level environmental bureau were made responsible for the maintenance and operation of all monitors within the city. The local governments, who have incentives to report low levels of pollution because of the performance targets they face, could potentially do this by manipulating the recordings from the monitors. Such manipulation was facilitated by the direct control of the monitors that the local governments were given. Indeed, many media sources have reported that such manipulation did occur.[11] Figure D10 shows an example from a newspaper article documenting such manipulation, where the pollution monitor is being sprayed with water to reduce the recordings.

Realizing that the data provided by local environmental protection bureaus might not be reliable, the MEP decided to contract the operation of the monitor stations to private companies through a procurement process. According to official documents from the MEP, all of the monitors were operated by private companies from November 1, 2016. Monitors were procured through twelve contracts. Each contract was designed to involve monitors in different provinces spread out over the country, to make it difficult for firms to select a given area. Six companies were selected, and each of them won two contracts. Importantly, after the monitors' operation was taken over by the firms, all the operation costs are paid by the MEP instead of the local government.

In addition to the regular monitors in the built-up area of each city, half of the cities were also assigned one background monitor. There are two main differences between the background monitors and the regular monitors: background monitors are installed outside of the built-up area of the city and are usually placed in a local scenic area; more importantly, the readings from the background monitors are not used in the performance evaluation of local officials. Due to the different nature of the background monitors, we are not including them in the main analysis.[12] In section 5.3, we show suggestive evidence that background monitors were subject to less manipulation.

[11]See `https://p.dw.com/p/32jqR` and `http://www.xinhuanet.com/politics/2018-08/09/c_1123244676.htm`, for two examples.

[12]Including them in the analysis does not alter any of our results. This is due to the fact that there is a limited number of firms located close to the background monitors. We also check robustness of our main results to controlling for whether a city has a background monitor.

2.3 Conceptual Framework

As discussed in the previous section, the central government regulates (e.g., sets pollution standards), while the local government is responsible for enforcing these regulations (e.g., by issuing fines to firms' violating existing regulations). Our interest is in understanding to what extent the introduction of monitors helps the central government hold the local government accountable for their actions and how that affects enforcement behavior and pollution at the local level.

To capture this effect we focus on cities that face the same regulations and receive monitoring for the first time.[13] Figure 1 illustrates how the introduction of monitors changes access to information on pollution both within and between cities. Within cities, monitors provide information on pollution for firms located close to the monitor, but not for those located further away.[14] Between cities, information on a larger share of overall pollution will be available for those cities that were assigned a greater number of monitors (indicated by the thicker dashed arrow from city B in Figure 1).[15]

Hence, the monitoring program that we study changes the capacity of the central government to collect information about pollution. This capacity changes both at the extensive margin (covering some firms but not others) and at the intensive margin (covering a larger vs. smaller share of firms in a city). In addition to the change in monitoring capacity in 2015, the reassignment of monitors from the local government to external third parties in 2016 changes the information provision process and decouples the responsibility of providing information with the responsibility to enforce regulations. The intention of the central government is that this shift should improve data quality and reduce the scope for manipulation. Because third parties are paid directly by the MEP, their incentives are arguably more aligned with those of the central rather than the local government.[16] In our analysis

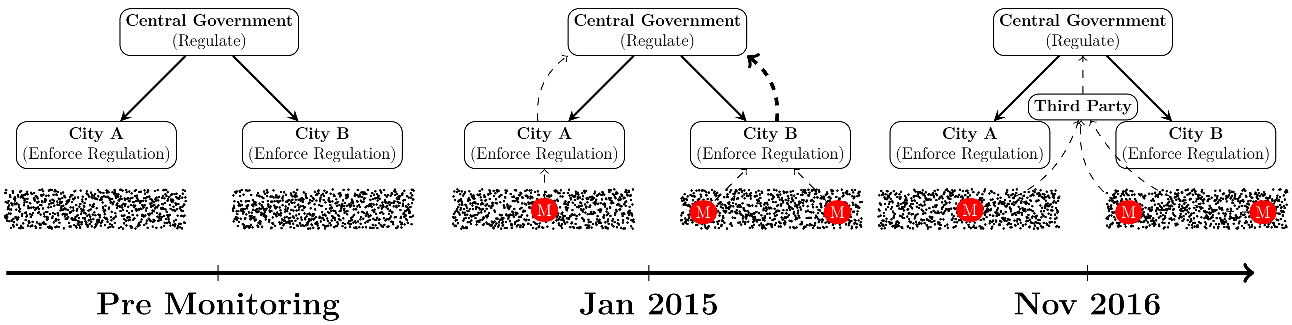
[13] As discussed above, pollution reduction targets differ across some regions in China. We implement a number of strategies to ensure that we do not capture differences in pollution targets, including controlling for target by time fixed effects in all specifications and ensuring that these targets are balanced in our regression discontinuity design.

[14] There is no exact cutoff for how far away from the monitor pollution could be picked up. For example, anecdotal evidence discussed in Appendix D suggests that environmental officials are concerned with pollution from firms within 5 km of a monitor. Schlenker and Walker (2015) show that health effects can be picked up 20 km from a polluting source, suggesting that monitor might be able to pick up differences at such a distance. We take a flexible approach in our analysis and let the data inform us about this cutoff.

[15] This is because a larger share of the potential polluters in a city will be covered. Note that this does not need to be mechanically true because the central government guidelines reported in Table C1 stipulate that larger cities are assigned a larger number of monitors. We document that additional monitors imply greater coverage in Figure 4, which shows the relationship between the assigned number of monitors and the average distance from a firm to its closest monitor. As shown in the figure, the average distance drops monotonically with the number of assigned monitors.

[16] This follows a similar logic to Duflo et al. (2013), who document that the incentives that third-party auditors face affect their reporting.

Figure 1. Monitors, Coverage and Flow of Information



Notes: This figure describes how the flow of information changes with the introduction of monitors. While responsibilities are unchanged – the central government regulates and the local government enforces these regulations – the quality of information changes differently between cities. Starting in January 2015, a different number of monitors transfer pollution recordings via the cities to the central government. Following the retraction of the monitors in November 2016, the recordings from the monitors are transferred to the central government via external third parties.

we will mainly focus on the overall effect of the monitors. However, in Section 5.3 we will shed some light on the potential importance of who is responsible for information provision.

3 Data

In this article, we combine several data sources that provide comprehensive information on the enforcement of environmental regulation and air pollution performance in cities that introduced air pollution monitors in 2015. Section 3.1 describes the new data on local air pollution enforcement that we collect and digitize. After that, Section 3.2 describes the two sets of data that we use to measure air pollution: a satellite-based measure of the AOD and data from the monitoring stations. Finally, Section 3.3 discusses the summary statistics for our three main samples. Additional details on data processing and on supplementary data sets used are provided in Appendix A.

3.1 Enforcement Records and Firm Data

To fully understand the impact of new air quality monitors on enforcement activities and the consequences of those activities, we face some data-related empirical challenges: first, the need to measure the quantity (and the quality) of governments' enforcement activities, and second, the need to link enforcement activities to the location of air quality monitors. We address these challenges by constructing a new data set on local enforcement of air pollution regulation in China using records collected from local environmental bureaus by the Institute of Public & Environmental Affairs (IPE). To the best of our knowledge, this is the first attempt to fully track enforcement activities carried out by local environmental bureaus in China. To identify where these enforcement activities occur, we geo-reference all major manufacturing firms in China using the Annual Survey of Industrial Firms (ASIF) and link these to the IPE records.[17]

Enforcement Records We collected all 55,000 enforcement records carried out from 2010 to 2017 in the 177 prefecture-level cities in our sample. Figure A1 in Appendix A provides an example of what these records look like and the type of information they contain. Each record includes details about the violating firm, a description of the violation, a reference to the regulation that has been violated, and the local environmental bureau's enforcement action. Using a classification algorithm described in detail in Appendix A.1, we categorize enforcement records in two dimensions. First, we identify what type of violation has been logged and whether this relates to air pollution, water pollution, waste pollution, or procedural violations. In total, we classify 22,000 records as being related to violations of air pollution regulations. Second, we identify what type of action has been taken by the local environmental bureau. For 95% of the enforcement records related to air pollution, the actions belong to one or several of the following four categories: suspending production (53%), ordering replacement/upgrading of the equipment (55%), levying fines (48%) or issuing a warning (17%).

Firm Data and Geo-referencing To be able to track where and against which firms that local environmental bureaus choose to enforce regulations, we use data from the 2013 ASIF. This survey is conducted by the National Bureau of Statistics (NBS). It includes all state-owned industrial enterprises (SOEs) and all private industrial enterprises with annual sales exceeding 5 million Chinese yuan. This corresponds to about 90% of all manufacturing firms in China and thus covers all major industrial polluters.[18] Previous versions of the ASIF data have been used in a number of papers (see, e.g., Song, Storesletten, and Zilibotti,

[17]There are two main reasons why we think these records accurately reflect the actions of local governments and are subject to limited misreporting. First, these records are only used for local administrative purposes and are not tied to central government performance evaluations. IPE collect records directly from local government agencies, since they are not held by the central government. Hence, local governments do not face incentives to misreport enforcement actions. Second, any misreporting is made difficult by the nature of the records since they capture public information on actual punishments imposed on local firms.

[18]According to the economic census 2004, firms in the ASIF represent 89.5% of the total revenue of all manufacturing firms in China.

2011; Brandt, Van Biesebroeck, and Zhang, 2012; Huang et al., 2017). We focus on the 2013 version of the survey, which is the latest available, to gain an understanding of the underlying distribution of manufacturing firms at the time of the introduction of monitors. Before linking the data to the enforcement records, we use detailed firm address information to identify the exact geographical location of all firms in the data. The process used for this geo-referencing is outlined in Appendix A.1. Panel C in Figure D1 shows the location of all the ASIF firms in our sample. Finally, we link our collection of enforcement records to the underlying distribution of manufacturing firms in the ASIF. Out of our 55,000 records, 52% of them refer to enforcement actions against firms in the ASIF data. Panel D in Figure D1 shows the geographical distribution of enforcement activities against these manufacturing firms.

3.2 Air Pollution Data

Monitor Data: $PM_{2.5}$, PM_{10} & AQI Air pollution data for the 552 monitoring stations in the 177 prefecture-level cities in our sample is published online by the MEP from the introduction of the monitors in January 2015.[19] The MEP website reports hourly data of SO_2, NO_2, CO, PM_{10}, $PM_{2.5}$, and O_3. An air quality index (AQI) based on these six pollutants is also constructed and reported.[20] The AQI ranges from 0 to 500. It is further divided into six ranges: $0-50$, $51-100$, $101-150$, $151-200$, $201-300$ and $301-500$. In public reports, these are categorized as excellent, good, lightly polluted, moderately polluted, heavily polluted, and severely polluted, respectively. We scrape pollution data from the MEP website and focus on the two main indicators used as targets in the National Air Quality Action Plan (PM_{10} and $PM_{2.5}$) as well as the AQI. To facilitate comparison with our other pollution measure described below, we aggregate the monitor data at the monthly level.

Satellite Data: Aerosol Optical Depth (AOD) Before the expansion of the monitoring system, none of the cities in our sample had any pollution monitoring. To obtain an objective measure of pollution both before and after monitor construction, we use data on AOD captured by the NASA MODIS satellites. AOD measures the degree to which aerosol particles prevent the transmission of light by absorption or scattering and can therefore be used as a measure of local pollution. Formally, Aerosol Optical Depth is defined as the neg-

[19] http://106.37.208.233:20035/

[20] The AQI is calculated using the following equation: $AQI = \max\{IAQI_1, IAQI_2, ...IAQI_6\}$, where each Individual Air Quality Index (IAQI) is given by $IAQI_i = \frac{I_h - I_l}{C_h - C_l}(C - C_l) + I_l$. The formula to compute IAQI is the same one used in the United States, but with differences in parameters (C_h, C_l, I_h, and I_l). C is the pollutant concentration measured by the air quality monitor. C_h and C_l are the concentration breakpoints, and I_h and I_l the index breakpoints. More details about these parameters can be found here https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/jcffbz/201203/t20120302_224166.shtml.

ative of the natural logarithm of the fraction of radiation (e.g., light) that is not scattered or absorbed. Hence, estimates of AOD in this paper can be interpreted as percentage changes. Monthly information on AOD is available at 0.1 by 0.1 degrees since 2000. In this project, we combine measures from the MODIS Aqua and Terra satellites to calculate the mean of AOD in a given month and city. To deal with potential within-city spillovers in pollution, we calculate this measure based on the whole prefecture-level city polygon, as depicted in Panel B of Figure D1. This figure shows the distribution of average AOD in 2010, the first year of our analysis, across all cities in our sample. As indicated in the figure there is substantial cross-sectional variation in pollution in our sample. The mean of AOD in the data is 0.34, and the standard deviation is 0.23.

AOD has been shown to be highly correlated with ground-based measures of pollution (see, e.g., Wang and Christopher, 2003; Gupta et al., 2006).[21] While AOD data has been used in various studies to measure air pollution (see, e.g., Chen et al., 2013; Jia, 2017), only a few studies have internally verified the correlation between AOD and local ground-based measures. To ensure the validity of AOD data in our setting, we take advantage of the ground-based measures of pollution that are available after the expansion to study the correlation between the AOD data and the two most common measures of air pollution ($PM_{2.5}$ and PM_{10}) as well as the joint Air Quality Index (AQI).[22] In Table C3, we report results from regressions controlling for monitor fixed effects, time fixed effects as well as precipitation, temperature, and mayor's age. Column (1) shows the estimate for $PM_{2.5}$, which is 0.33. This is largely comparable with the correlations reported by Gupta et al. (2006). Estimates for PM_{10} and AQI are smaller but of a broadly similar magnitude. Taken together, this suggests that AOD is a suitable measure for local air pollution and that it most strongly reflects changes in $PM_{2.5}$.

3.3 Main Sample and Summary Statistics

To supplement our analysis, we collect additional data on: monthly weather conditions, résumés of all mayors during our sample period and city level aggregates of citizens' online searches for a set of keywords related to pollution. Appendix A.3 describes this additional data in detail and the procedure used for collecting it. Using the data on pollution and enforcement described above together with these additional sources, we construct three main

[21] Wang and Christopher (2003) find that the correlation coefficient between the monthly means of AOD and $PM_{2.5}$ is around 0.7 using data in Alabama in 2002. Using much more comprehensive data, Gupta et al. (2006) find that the correlation ranges from 0.14 to 0.6 for a number of cities across the world.

[22] For this analysis we match monitors with AOD data from the intersecting pixel (0.1 by 0.1 degrees). If data is missing for this pixel, we interpolate and calculate the average AOD measure for all surrounding pixels. All results are robust to using data at the city level instead.

samples for our analysis – all covering the 177 prefecture-level cities that installed monitors in 2015. Summary statistics for these three samples is presented in Table 1.

Table 1. Summary Statistics

	Mean	Std. dev.	Obs.	Periods	Freq.
Panel A: Firm-Level Data					
Any Air Pollution Enforcement	0.013	0.11	288848	2010-2017	Yearly
Suspension	0.0073	0.085	288848	2010-2017	Yearly
Upgrading	0.0078	0.088	288848	2010-2017	Yearly
Fine	0.0067	0.081	288848	2010-2017	Yearly
Warning	0.0016	0.040	288848	2010-2017	Yearly
Any Water Pollu. Enforce.	0.0091	0.095	288848	2010-2017	Yearly
Any Solid Waste Pollu. Enforce.	0.0031	0.056	288848	2010-2017	Yearly
Any Procedure Pollu. Enforce.	0.0075	0.086	288848	2010-2017	Yearly
Monitor within 10 km	0.40	0.49	36106	2013	Cross Sec.
Distance to Monitor (km)	19.2	15.4	36106	2013	Cross Sec.
Owner: SOEs	0.094	0.29	36106	2013	Cross Sec.
Owner: Private	0.82	0.38	36106	2013	Cross Sec.
Owner: Foreign	0.041	0.20	36106	2013	Cross Sec.
Owner: Other	0.040	0.20	36106	2013	Cross Sec.
Year Started	2003.2	7.92	36106	2013	Cross Sec.
Employment	434.8	1076.4	36106	2013	Cross Sec.
Revenue	278716	1656811	36106	2013	Cross Sec.
Panel B: City-Level Data					
Number of Monitors	2.75	1.08	16992	2010-2017	Monthly
Urban Population (10,000)	33.8	21.0	16992	2010-2017	Monthly
Size of Built-up Area (km^2)	46.8	27.2	16992	2010-2017	Monthly
Age of the Mayor	51.7	6.32	16992	2010-2017	Monthly
Precipitation (mm)	3.38	4.27	16992	2010-2017	Monthly
Mean Temperature	10.5	11.4	16992	2010-2017	Monthly
Aerosol Optical Depth	0.34	0.23	16319	2010-2017	Monthly
# Firms with Any Air Pollu. Enforce.	10.38	24.94	1416	2010-2017	Yearly
# ASIF Firms with Any Air Pollu. Enforce.	4.28	7.65	1416	2010-2017	Yearly
Search Index: air pollution	1.89	4.38	14610	2011-2017	Monthly
Search Index: haze/smog	18.10	30.62	14610	2011-2017	Monthly
Search Index: $PM_{2.5}$	0.20	1.63	14610	2011-2017	Monthly
Search Index: air mask	5.53	8.51	14610	2011-2017	Monthly
Search Index: air purifier	22.40	25.97	14610	2011-2017	Monthly
Panel C: Monitor-Level Data					
Particulate Matter 2.5 ($PM_{2.5}$)	44.8	26.9	19185	2015-2017	Monthly
Particulate Matter 10 (PM_{10})	79.0	50.9	19185	2015-2017	Monthly
Air Quality Index (AQI)	71.0	33.1	19185	2015-2017	Monthly

Notes: The table presents summary statistics for the samples used in our analyses. The data cover the 177 cities that installed monitors in 2015. Panel A reports the summary statistics for the data of firm-level enforcement. We rely on the Annual Survey of Industrial Firms (ASIF) 2013 and restrict the sample to include only firms set up before 2010 and located within 50 km of an air quality monitor. Panel B reports the summary statistics for city-level analysis. Panel C reports the summary statistics of three monthly pollution indicators. The monthly data is averaged from real-time readings of 552 monitors in 177 cities.

Panel A reports information for the firm-level data. We rely on the 2013 ASIF and restrict the sample to firms that started operating before 2010 (the first year of our analysis) and that are located within 50 km of an air quality monitor.[23] This leaves us with a total sample of 36,106 firms. The majority of these firms are private (82%) and cover a wide range of different industries (Table C4 reports the industry composition for our sample).[24] On average, the firms in our sample are located 19 km from a monitor. However, as depicted in Figure D6 the spatial distribution of firms is skewed and 40% of firms are located within 10km from a monitor. For a given firm in our sample, the probability of receiving an air pollution related enforcement action is 1.3%. Such an enforcement action most commonly requests the firm to upgrade their equipment, but suspension of operation and issuing fines are also common. Violations relating to water pollution regulations or conducting a procedural violation are less common, but of a comparable magnitude (0.9% and 0.75%, respectively). Most (more than 75%) of the enforcement actions were taken after the introduction of air quality monitors.

Panel B reports the summary statistics for the city-level data. For this sample we consider pollution as well as enforcement at the aggregate city level.[25] The cities we study are small by Chinese standards and have an average population of around 340,000. The average size of our sample (measured by both the urban population and the size of build-up area) are one third of the country average. However, the air pollution level in our sample (measured by AOD) is only slightly lower (10%) than the country average.[26] On average the cities in our sample have 2.75 monitors installed and about 10 firms face an environmental enforcement actions related to air pollution per year (on average 1.77 per year before 2015 and 22.73 per year afterwards).

Panel C reports the summary statistics for the monitor-level data for the three pollution measures we use. This data is aggregated at the monitor-month level and covers data from all 552 monitors installed in the 177 cities that we study. The sample period for this data starts in January 2015, when all the monitors have been installed.

4 Impact of the Monitors

This section describes the impact of monitoring on local government enforcement activities and pollution. First, we conduct a firm-level analysis in Section 4.1. We start by describ-

[23] Note that while we have yearly information on enforcement actions, our information on firm characteristics is from the 2013 ASIF and therefore cross-sectional.

[24] Note that this table reports 2-digit industry codes, while we use 4-digit industry codes when estimating industry fixed effects in our analysis.

[25] Hence, this sample is not restricted to firms within 50 km from a monitor and covers the whole city polygon as depicted in Figure D1.

[26] Appendix A.2 discusses additional details regarding the representativeness of our sample and compares it to other cities in China.

ing the general patterns before presenting the main empirical strategy and results showing how monitors affect enforcement activities. Thereafter, we investigate how the information provided by the monitors shape the enforcement response. In Section 4.2, we move from studying the local effects of monitoring to the aggregate city effects – exploiting differences in the number of monitors induced by the new monitoring program.

4.1 Firm-Level Evidence

Before conducting a formal analysis, we start by investigating the spatial distribution of enforcement activities and how these change with the introduction of monitors. Figure 2a shows a binned scatter plot of the probability that a firm has any enforcement record related to air pollution in a year on the distance to the closest monitor. Black dots indicate the mean probability during the period before air quality monitors were introduced, red diamonds show the mean probability in the post-period. The two lines represent the linear fit of the data before and after the introduction of monitors. The graph shows that the average yearly probability of a firm receiving any air pollution-related enforcement action is around 0.0067 before 2015 and that this probability does not seem to depend on the distance to the (planned) monitor (i.e., there is no gradient in enforcement activity in the pre-period). This provides some first evidence suggesting that monitors are not endogenously placed in localities with differential enforcement activities.[27] However, during the post-period we see a substantial increase in the enforcement activity – in particular enforcement against firms close to the monitor. Figure 2b estimates the gradient in enforcement activities nonparametrically and shows that enforcement activities increase by about 1 percentage point within 0–5 km from the monitor and by about 0.6 percentage points 10–15 km from the monitors, while there is no statistically significant impact on enforcement beyond this point.[28] This change in spatial pattern is also noticeable by visual inspection of the raw data.[29]

[27] While this is reassuring, the identification assumption for our main analysis is on the trends as opposed to the levels of enforcement activity. We provide a test of this assumption in the following section.

[28] Formally, we estimate the following equation:

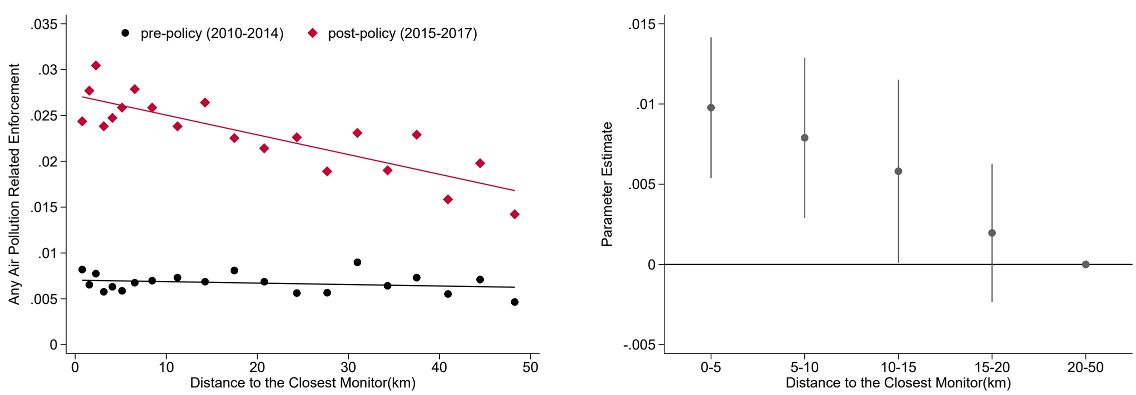
$$y_{ijpt} = \delta_i + \theta_{jt} + \eta_{pt} + \sum_{d=0-5km}^{15-20km} \beta_d m_{it}^d + \epsilon_{ijpt}$$

where m_{it}^d is an indicator for there being a monitor within distance d from firm i in year t; and all other variables are the same as in Equation 1. Hence, we are here estimating the average change in enforcement in the post-period relative to the pre-period (instead of estimating effects by year).

[29] Figure D2 in the Appendix shows a map of cities in central China depicting the location of air quality monitors, the underlying distribution of manufacturing firms as well as the geographical location of enforcement activities related to air pollution before (in blue) and after (in red) the introduction of monitors in 2015.

The above results are consistent with extensive media reporting that local environmental bureaus step up environmental inspections close to the monitors. We document some of this evidence in Figure D9 in Appendix D, which shows a list of news articles generated from a search on the Chinese search engine Baidu using the keywords “monitors”, “surrounding area”, and “check”. The list includes a large number of articles discussing how local governments organize their environmental inspections around the monitors. Some examples include cities that draw special zones around their air quality monitors and send teams of inspectors to those zones, to ensure that firms comply with national environmental regulations. Other sources mention that city governments hire volunteers from the public to inspect air pollution from venues (such as restaurants) within a certain distance from the monitors. Finally, several sources suggest that mayors take a special interest in these inspections by, for example, directly appointing officials to this task or by visiting surrounding areas. This further underlines the weight that mayors put on the recordings from the monitors because of the performance incentives that they face.

Figure 2. Air-Pollution-Related Enforcement and Distance to a Monitor



(a) Binned Scatter Plot

(b) Nonparametric Estimates

Notes: Figure 2a shows a binned scatter plots of the relationship between enforcement activity and distance to the closest monitor. Black dots indicate the mean probability of air pollution-related enforcement before introducing the air quality monitors, while red diamonds show the mean probability after the introduction of monitors. Figure 2b shows the relative increase in enforcement for each distance bin after 2015. Error spikes represent 95 percent confidence intervals.

Firm Level: Event Study

To investigate the relationship between monitors and enforcement formally, we estimate a flexible nonparametric event study specification. If we denote a generic firm by i, with $i \in j, p$, where j denotes a 4-digit industry, p denotes a province and t a generic year, our model can be written as:

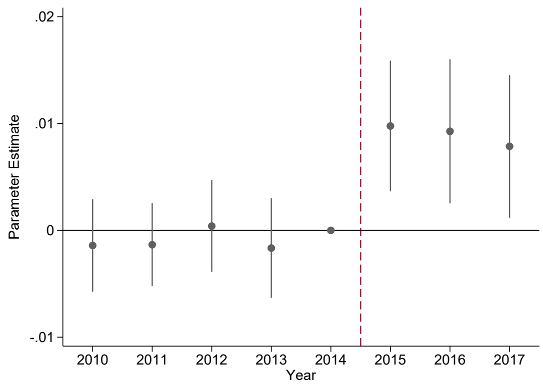
$$y_{ijpt} = \delta_i + \theta_{jt} + \eta_{pt} + \sum_{\substack{d=0-5km \\ d\neq 20-50km}}^{15-20km} \sum_{\substack{k=2010 \\ k\neq 2014}}^{2017} \beta_{dk} m_i^{dk} + \lambda X_{it} + \epsilon_{ijpt} \tag{1}$$

where y_{ijpt} is an indicator for enforcement, δ_i is a firm fixed effect, θ_{jt} and η_{pt} represent, respectively, industry-by-year and province-by-year fixed effects, m_i^{dk} is an indicator for any monitor being within d km from a firm in year k, X_{it} denotes weather controls and ϵ_{ijpt} is the error term. Because we condition on firm as well as on industry-by-year and province-by-year fixed effects, parameter estimates capture the average (across industries and provinces) effect of monitoring on the differential change in enforcement across firms in the same industry or province. This specification addresses two important concerns. First, we ensure that we estimate the impact of monitoring within the same regulatory environment (pollution reduction targets vary across provinces as discussed in Section 2). Second, we allow for different enforcement trends depending on local industrial composition at baseline. We use the year before the introduction of the monitors and firms 20–50 km from the monitor as reference categories and estimate β_{dk} for $d \in \{$0–5 km, 5–10 km, 10–15 km, 15–20 km$\}$. Equation 1 allows us to estimate the temporal and spatial relationship between monitors and enforcement activity. Hence, it is informative about the key identification assumption for our analysis (parallel trends in enforcement for firms located at different distances from the monitors) as well as the spatial reach of monitors. We cluster standard errors at the city level to account for correlation of errors across firms and time within cities.[30]

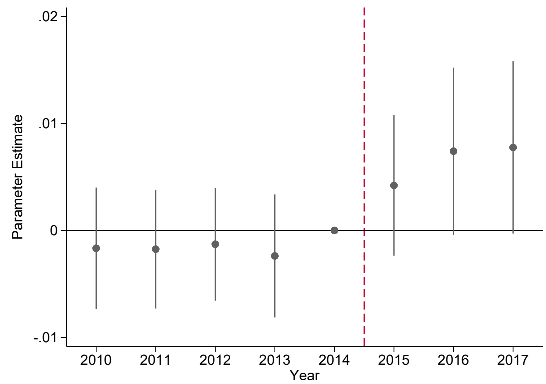
Figure 3 reports the results from estimating Equation 1. We present the estimates in four separate event study graphs each showing how enforcement activity changes around the introduction of monitors for firms within 0–5 km, 5–10 km, 10–15 km and 15–20 km from the monitors relative to firms 20–50km from the monitors (the reference category). In all four graphs, there is no evidence of any differential trends leading up to the intervention – lending credibility to the main identification assumption of parallel trends. After the introduction of the monitors we see a substantial increase in enforcement activity close to the monitors. This step-up in enforcement is particularly pronounced within 0–5 km from the monitors, but is noticeable also for firms 5–10 km and 10–15 km from the monitor. For firms 15–20 km from the monitor there is no differential change in enforcement activity during our sample period. These results mirror the gradient observed in Figure 2.

[30] As a robustness check, we also report standard errors based on the spatial HAC variance estimator proposed by Conley (1999), which allows for correlation between areas that are geographically close but belong to different administrative units (See Panel A of Table 2). These standard errors are smaller, but overall similar, to our baseline standard errors. We focus on the city-level clustered standard errors since these are more conservative.

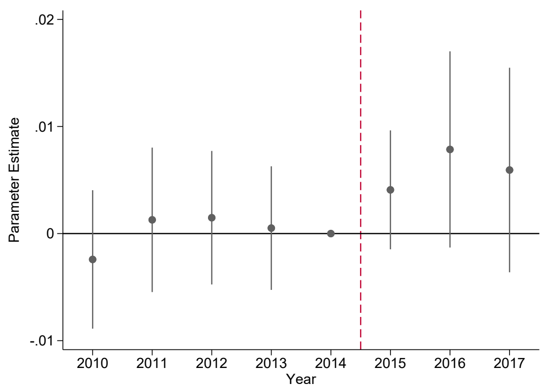
Figure 3. Nonparametric Event Study



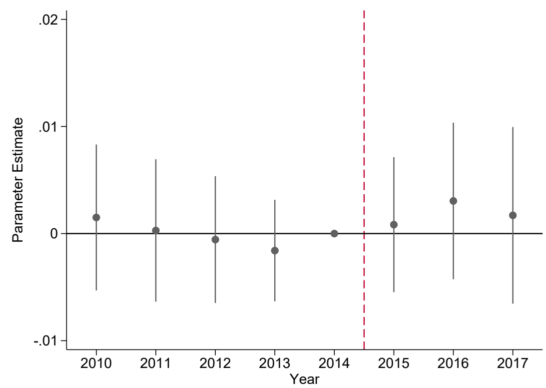
(a) 0–5 km



(b) 5–10 km



(c) 10–15 km



(d) 15–20 km

Notes: The figure shows the estimates of the nonparametric event study using Equation 1. The sub-figures report event studies for firms within each distance bin. The reference group is firms located 20–50 km from the closest monitor. Error spikes represent 95 percent confidence intervals, calculated using robust standard errors clustered at the city level.

Firm Level: Main Results

Guided by the results in the previous section, we use a simplified difference-in-differences specification to provide an aggregate estimate of the magnitude of the effect. This specification compares firms within and beyond 10 km from a monitor.[31] The results from estimating this specification are shown in Table 2. The first column of Panel A reports estimates on

[31] Formally, we estimate:

$$y_{ijpt} = \delta_i + \theta_{jt} + \eta_{pt} + \beta m_{it}^{10km} + \lambda X_{it} + \epsilon_{ijpt},$$

, where m_{it}^{10km} is an indicator for a firm having a monitor within 10 km and all other variables are the same as in Equation 1. To the extent that firms beyond 10 km from the monitors are also affected, this specification provides a lower bound for the true causal effect.

whether any air pollution-related enforcement took place (i.e., the same outcome as in figures 2 and 3). Results suggest that the probability of a firm within 10 km from a monitor receiving an enforcement action in a year is 0.78 percentage points higher compared to firms further away from the monitor. This suggests that a monitor increases the probability of an air pollution-related enforcement activity occurring by 60% compared to the average yearly probability of enforcement (1.3). The remaining columns of Panel A in Table 2 shed light on what type of action that was taken by the local government, by estimating the same model for the four most common enforcement classifications we identify in the data ("suspension" – suspending production of the whole factory or part of the factory; "upgrading" – ordering replacement/upgrading of the equipment, levying a "fine" or issuing a "warning"). We find similar estimates for the first three categories and no effect for the last type ("warning"). These results suggest that the local environmental bureau is responding to the monitors by implementing costly punishments on local firms.

One potential concern with the above results is whether monitors are placed in strategically important locations within the city where the local government has a greater interest in enforcing environmental regulations after 2015 (e.g., because of greater health benefits to the local population or because of lower costs of enforcement for the environmental bureau). The patterns observed in Figure 2a suggest that at least in terms of pre-policy enforcement that is not the case. To further investigate this concern we conduct two additional tests. First, we use the same baseline specification to look at environmental enforcement that is not related to air pollution. The results are reported in Table C5 in the Appendix. For enforcement related to water pollution, solid waste pollution, and procedure violation, estimates are small and statistically insignificant. Second, we conduct placebo tests in which we estimate the specification used to produce Figure 2b, but instead of the minimum distance to the monitor we use the distance to the local environmental bureau or the distance to the city's firm centroid.[32] Figure D3 of Appendix D shows that there are no detectable changes in the gradients of enforcement activity pre and post 2015. To further validate our main results, we include the distances to both the environmental bureau and the firm centroid in our main specification and interact the distance bins with the time fixed effects. Due to the high correlation between monitor location and these measures our results are slightly less precise, but largely unaffected when including this full set of controls.[33] Taken together, these results suggest that the step-up in enforcement behavior that we observe is indeed

[32]We do this to investigate whether enforcement is guided by the spatial distribution of firms. To calculate centroids for each city, we use the geographical distribution of all ASIF firms. The firm centroid is a single point representing the barycenter of all firms.

[33]The main estimate in Column (1) of Panel A in Table 2 changes from 0.0078 to 0.0051 and is significant at the 5%-level.

Table 2. Monitors and Enforcement Activities

	(1)	(2)	(3)	(4)	(5)
Panel A: Any Enforcement Action					
Outcome	Air	Suspension	Upgrading	Fine	Warning
$\text{Mon}_{<10km}$ × Post	0.0078***	0.0043***	0.0041***	0.0043***	-0.00012
	(0.0016)	(0.0012)	(0.0011)	(0.0011)	(0.00030)
	[0.0012]	[0.0010]	[0.0010]	[0.0010]	[0.00026]
Mean of dependent variable	0.013	0.0073	0.0078	0.0067	0.0016
Observations	288696	288696	288696	288696	288696
Panel B: Intensity and Leniency					
Outcome	Air	Low Intensity	High Intensity	Lenient	Strict
$\text{Mon}_{<10km}$ × Post	0.0068***	0.0071***	-0.00025	0.0027***	0.0015*
	(0.0016)	(0.0014)	(0.00050)	(0.00060)	(0.00088)
$\text{Mon}_{<10km}$ × Post × High polluter	0.030*	0.0054	0.025**	-0.017**	0.040***
	(0.016)	(0.012)	(0.010)	(0.0067)	(0.011)
Mean of dependent variable	0.013	0.011	0.0023	0.0026	0.0045
Observations	288696	288696	288696	288696	288696
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of the impact of air pollution monitoring on the probability of being subject to different air pollution-related enforcement actions by the local government. Each coefficient is from a separate difference-in-differences regression. All regressions control for fixed effects specific to firm, industry-by-year interactions, and province-by-year interactions. Robust standard errors clustered on the city in parentheses. In Panel A, standard errors based on the spatial HAC technique suggested by Conley (1999) are reported in brackets, using a bartlett kernel and bandwidth of 100 kilometers. Panel B reports heterogeneity for firms identified as high polluters according to ESR during the pre-period. The outcome "low intensity" ("high intensity") corresponds to a dummy variable indicating that a firm received only one (at least two) enforcement actions in a year. The outcome "lenient" is a dummy variable that equals one if only one punishment (among "suspension", "upgrading", and "fine") is issued against a firm in a year. In contrast, the dummy variable "strict" is defined as one if all three types of punishments are issued against a firm in a year. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

driven by the monitors.

Firm Level: Targeting and Enforcement Efficiency

In the previous sub-section, we showed that local governments respond to monitoring by increasing the probability of enforcement. In this sub-section we will explore whether monitoring also affect other aspects of enforcement activities. We do this in two ways. First, we study which firms that are targeted by the local governments and if the intensity and strictness of enforcement that they face changes with the introduction of monitors. Second, we investigate how monitors shape the responsiveness of enforcement actions to local pollution levels.

To better understand which firms that local governments target and whether this targeting changes with the introduction of monitors, we study actions against a set of high polluting firms. We rely on the Environmental Survey and Reporting Database (ESR) to identify these firms. The ESR is put together by the central government and includes firms that are considered to be major polluters (in total responsible for 65% of local emissions).[34] In Panel B of Table 2, we estimate the differential enforcement response against these firms. The estimates in Column (1) suggest that there is a larger increase in the overall probability of enforcement against these firms (significant at the 10% level). The following four columns report what type of enforcement these firms receive. We start by differentiating between low and high enforcement intensity, where we define low as receiving one enforcement action in a year and high as receiving more than one action. The results in columns (2)-(3) show that low-intensity enforcement is not significantly different between low and high polluting firms, but that all high-intensity enforcement focuses on key polluters in the presence of monitoring. Next, we consider the strictness of enforcement action. To capture this, we construct two additional dummy variables that classify enforcement records as either lenient or strict. Since there is no clear ranking of the three main punishment types discussed above ("suspension", "upgrading", and "fine") and enforcement records often include multiple punishments, we consider the two extreme cases where either one (lenient) or all three (strict) punishments are issued against a firm in a year. The last two columns in Table 2 reports the results and show that high polluting firms are less likely to receive lenient treatment and more likely to receive strict enforcement action. Taken together, these results suggest that local governments respond to monitoring by shifting both the intensity and the strictness of enforcement towards high polluters.

We conduct an additional analysis to investigate whether monitors make local governments' enforcement efforts more responsive to local pollution levels. The main empirical challenge inherent in studying this is the endogeneity of local pollution. To overcome this challenge, we take advantage of local rainfall shocks, which are important determinants of local pollution levels since heavy rainfall makes pollution less severe.[35] We start by documenting this empirically. For each monitor–month pair, we construct an indicator ($Rain_{>\tilde{x}}$) for whether rainfall is above the median in that pair or not. Hence, this variable captures year-on-year fluctuations in rainfall (i.e., comparing a rainy January with a dry January). We focus on yearly variation because this is the highest temporal frequency that we have

[34] The ESR database has been used in several recent paper (see, e.g. He, Wang, and Zhang, 2020). We use the ESR firms identified between 2010 and 2014, the period before introducing air quality monitors. In total, this corresponds to 1,445 of the firms in our baseline firm sample.

[35] An alternative strategy used in previous literature is to exploit inversion episodes (Arceo, Hanna, and Oliva, 2016; Jans, Johansson, and Nilsson, 2018; Deschenes et al., 2020). While inversion episodes are strong predictors of pollution levels with high frequency data in our setting (hourly or daily), they are not predictive of annual pollution levels, which we need to use due to the yearly frequency of the enforcement data.

for the enforcement records.[36] Columns (1) - (4) in Table 3 show estimates of the relationship between pollution recordings and rainfall shocks from a regression controlling for monitor and time fixed effects. Results show that average pollution recordings are consistently about 5% lower in years with above median rainfall. These effects are substantially stronger at higher levels of pollution, which are arguably more important for local policy response, where, e.g., the share of days that are considered heavily polluted (AQI>200) are reduced by 17% (.0038/.022).

We then explore how enforcement activities respond to monitors in the presence of rainfall shocks by estimating an augmented version of our baseline model that interacts the effect of monitors with rainfall shocks. We prefer taking this reduced form approach instead of instrumenting local pollution levels. The reason for this is that polluting firms might endogenously respond to rainfall shocks, e.g., by polluting more during heavy rainfall leading the above results to be lower bounds. Reduced form interactions are reported in Column (5) of Table 3. There are two main takeaways from these results. First, they document that enforcement activities do not respond to rainfall shocks in the absence of monitors. Second, it shows that the effect of monitors on enforcement is about half the size (0.011–0.0051=0.0059) when rainfall levels are above the median – i.e., when pollution levels are lower – as opposed to when rainfall levels are below the median (0.011) – when pollution levels are higher. These results suggest that the information captured by the monitors is important for the enforcement actions taken by the local governments and that the monitors make the local government more responsive to changes in local pollution.

The results above could be driven by two channels: 1) that local governments respond more strongly to higher pollution levels close to the monitor in the post-period or 2) that they respond more strongly to higher pollution levels in general after the introduction of the monitors. Both these channels could contribute to increased responsiveness as long as the recordings from the monitors are informative about the overall pollution level in the city. To shed further light on which of these two mechanisms is driving the results, we estimate the model with the full set of interactions.[37] Column (6) in Table 3 reports the results from estimating this model and shows suggestive evidence (significant at the 10%-level) that both the channels discussed above seem to be at play. There is a weaker response in general during

[36] An alternative way of doing this is to instead aggregate the data to the yearly level. This produces similar but somewhat less precisely estimated effects.

[37] Formally we estimate:

$$y_{ijpt} = \delta_i+\theta_{jt}+\eta_{pt}+\beta_0 r_{it}^{>\tilde{x}}+\beta_1 m_i^{10km}\times Post_t+\beta_2 m_i^{10km}\times Post_t\times r_{it}^{>\tilde{x}}+\beta_3 m_i^{10km}\times r_{it}^{>\tilde{x}}+\beta_4 Post_t\times r_{it}^{>\tilde{x}}+\lambda X_{it}+\epsilon_{ijpt}$$

where m_i^{10km} is an indicator for a firm being located within 10 km from a monitor, $Post_t$ indicates the post-period (i.e. from 2015 onward) and $r_{it}^{>\tilde{x}}$ is an indicator for rainfall being above the median in a city. All other variables are the same as in Equation 1.

the post-period when rainfall is high (low pollution), but this effect is even stronger closer to the monitor. The results also provide additional support for the validity of this exercise by showing that there is no differential response to rainfall shocks in the pre-period in areas close to the monitors compared to areas further away.

Table 3. Rainfall, Pollution Recordings and Enforcement Response

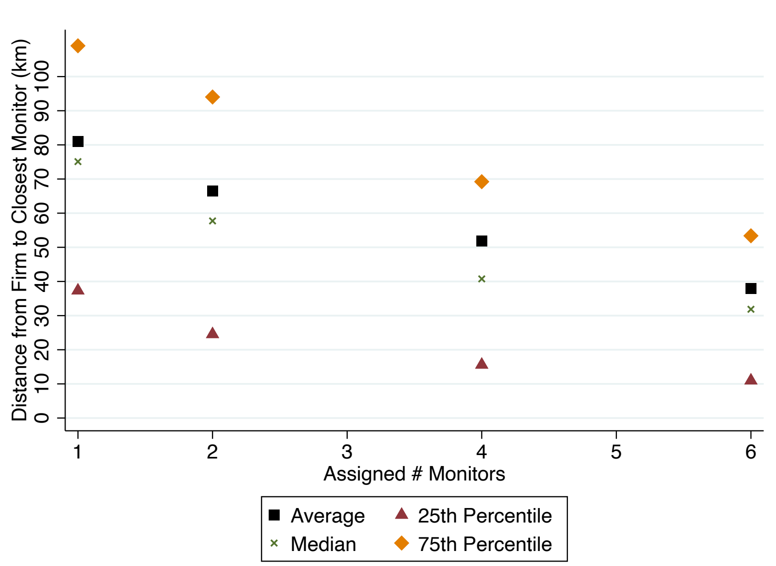
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:		$\log(\bar{x})$		Share of Days	Indicator (0/1)	
	$PM_{2.5}$	PM_{10}	AQI	AQI>200	Any Enforcement	
$\text{Rain}_{>\tilde{x}}$	-0.049***	-0.053***	-0.042***	-0.0038**	-0.00044	0.00074
	(0.0065)	(0.0060)	(0.0048)	(0.0017)	(0.0011)	(0.00077)
$\text{Mon}_{<10km} \times \text{Post}$					0.011***	0.010***
					(0.0021)	(0.0024)
$\text{Mon}_{<10km} \times \text{Post} \times \text{Rain}_{>\tilde{x}}$					-0.0051***	-0.0043*
					(0.0018)	(0.0026)
$\text{Mon}_{<10km} \times \text{Rain}_{>\tilde{x}}$						0.0011
						(0.00098)
$\text{Post} \times \text{Rain}_{>\tilde{x}}$						-0.0041*
						(0.0024)
Monitor FE	Yes	Yes	Yes	Yes	No	No
Time FE	Yes	Yes	Yes	Yes	No	No
Firm FE	No	No	No	No	Yes	Yes
Industry-Year FE	No	No	No	No	Yes	Yes
Province-Year FE	No	No	No	No	Yes	Yes
Mean Outcome	3.64	4.23	4.17	0.022	0.012	0.012
Observations	19185	19185	19185	19185	288720	288720
R-squared	0.77	0.77	0.78	0.42	0.24	0.24

Notes: This table reports the effect of precipitation shocks on pollution (columns 1–4) and the differential effect of monitors on enforcement activities during precipitation shocks (columns 5–6). $Rain_{>\tilde{x}}$ is an indicator variable identifying time periods when precipitation is above the median rainfall in a city during the main sample period. To ensure comparability between the monthly temporal resolution in the first analysis (columns 1–4) and the yearly temporal resolution in the second, the first analysis calculates precipitation shocks within months across years (i.e. comparing a rainy January with a dry January). This analysis investigates the relationship between precipitation and four monitor-based measures of air pollution: $PM_{2.5}$, PM_{10}, the combined AQI, and the share of days when the air quality index is above the critical value (200). Each column is from a separate regression. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

4.2 City-Level Evidence

In this section we move from studying the extensive margin impact of monitors at the firm level to investigate the aggregate intensive margin effects at the city level. This approach allows us to infer the overall impact of a more extensive monitoring program on both enforcement and pollution. To conduct this analysis, we exploit the criteria set up by the central government when implementing the monitoring program (listed in Table C1 in Appendix C) and compare outcomes in cities that installed different numbers of monitors. The argument behind this approach is that a greater number of monitors will cover a larger share of the firms in a city and therefore hold officials accountable for a greater share of the overall potential polluters. Figure 4 illustrates this point for our data by showing how the number of monitors assigned to a city is related to the minimum distance between firms and monitors. The figure shows that the distance to the closest monitor is substantially smaller in cities with a larger number of monitors. Considering firms that are closest to the monitor (the 25th percentile) this distance drops by 75% from 40 km to 10 km for firms that are assigned 6 rather than 1 monitor.

Figure 4. Number of Monitors and Firm Coverage



Notes: This figure shows the relationship between the distance to the closest monitor and the number of monitors assigned to a city. Each symbol represents a bin showing the average distance across all cities assigned that number of monitors (these numbers corresponds to those reported in Table C1 in Appendix C). Four different distances are reported: average distance for all firms in a city, the median distance to all firms in the city, the distance to the 25th percentile (i.e. the closest firms) and the distance to the 75th percentile (i.e. the firms furthest away).
Source: Based on authors' own calculations using the assigned number of monitors reported in Table C1 and the geographical location of manufacturing firms described in section 3.1.

City Level: Event Study

To study the effects of monitoring at the city level, we first estimate a standard event study specification. If we denote a generic city by c, with $c \in r$, where r denotes a pollution reduction target group in Table C2, and t is a generic time period, our model can be written as:

$$y_{crt} = \delta_c + \gamma_{rt} + \sum_{\substack{k=2010 \\ k \neq 2014}}^{2017} \beta_k m_c^k + \lambda X_{ct} + \epsilon_{crt}, \tag{2}$$

where y_{crt} is either an aggregate measure of a city's monthly AOD or the total yearly number of firms that receive any enforcement related to air pollution regulations, m_c^k is either the actual number of monitors in the city, or the predicted number of monitors according to Table C1, in a given year k, δ_c are city fixed effects and γ_{rt} are pollution target group by time fixed effect (month–year for the pollution specification and year for the enforcement specification). The variable X_{ct} represents time-varying characteristics of each city including: monthly precipitation, monthly average temperature and the age fixed effects of the mayor in office.[38] The error term is denoted by ϵ_{crt}, which we cluster at the city level to account for potential serial correlation of the errors over time. Because we condition on city as well as on pollution target-by-year fixed effects, parameter estimates capture the average effect of monitoring on the differential change in pollution/enforcement across cities with the same pollution reduction target.

To causally identify the impact of an additional monitor, we rely on common trends across cities assigned different numbers of monitors. To assess the validity of this assumption, we start by investigating the AOD trend for each group. In Figure 5a, we plot demeaned city-level AOD trends in four groups, which are determined according to the minimum number of monitors assigned by the central government. Two important patterns can be noted. First, there is a relatively flat AOD trend in cities assigned one monitor, suggesting that there was no major change in pollution in these cities and that it is therefore a suitable control group.[39] Second, and more importantly, raw AOD data in all four groups share a common trend before 2015, after which AOD diverges – with a more substantial reduction for cities assigned a larger number of monitors.

To formally test this, we estimate Equation 2 – setting the average pollution in the year before monitors were installed as the baseline. Estimates from this specification are shown in Figure 5b. We first estimate a standard event study specification using the actual number

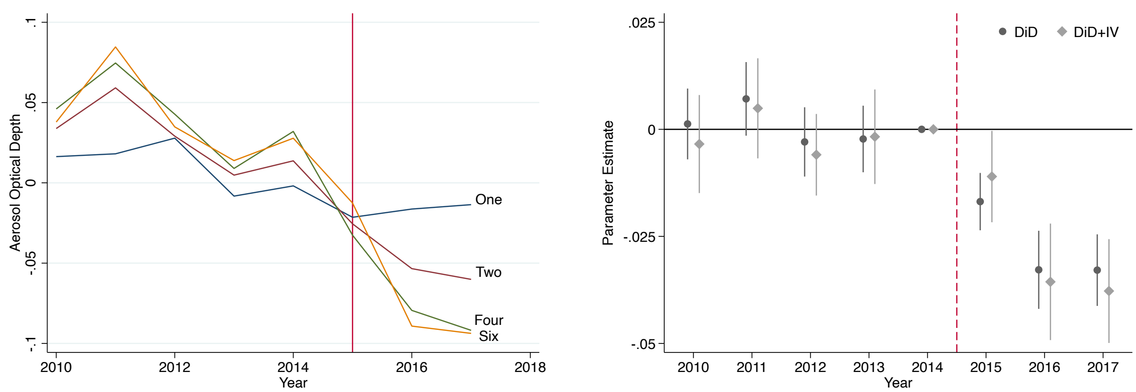
[38] The inclusion of weather controls is motivated by the fact that ambient pollution has been shown to be affected by local weather conditions in previous work (Schlenker and Walker, 2015; Barwick et al., 2020).

[39] As discussed in Section 2.3, all cities in China were assigned at least one monitor.

of monitors installed in a city as our independent variable of interest. These estimates are reported by the black dots on the graph. Results corroborate the findings above that there are no differential trends in AOD leading up to the intervention. We also see a substantial drop in pollution in the post-period for cities that installed additional monitors.[40] These effects are even stronger in the second and third year.

One potential concern with the above specification is that it might lead to biased results if cities were able to influence the number of monitors installed. The estimates would be biased if, for example, cities that expected lower pollution in the future installed a larger number of monitors. To address this concern, we use the minimal number of monitors set by the MEP as an instrument for the actual number of monitors (m_c).[41] The instrumental variable estimates are marked by diamonds in Figure 5b. The coefficients follow the OLS estimates closely, but are slightly less precisely estimated. Again, there is no evidence of differential trends leading up to the intervention, supporting the common trend assumption between cities of different sizes.

Figure 5. Event Study



(a) AOD Trends **(b)** Event Study

Notes: Panel (a) presents demeaned city-level AOD trends in four groups. Groups are determined according to the minimum number of monitors assigned according to the regulation. The red line marks the introduction of air quality monitors. Panel (b) presents the estimates from equation (2) using two different specifications (DiD, DiD+IV). Black dots represent the coefficients from DiD, whereas diamonds represent DiD+IV estimates. Error spikes represent 95 percent confidence intervals. AOD is formally defined as the negative of the natural logarithm of the fraction of light that is not scattered or absorbed. Hence, these estimates can be interpreted as percentage changes in pollution.

[40] Note that monitors are operational from January 1, so all periods in the year of adoption are treated. We report estimates by year rather than by month to facilitate comparison with the enforcement results.

[41] We report the first-stage estimates in Table 4.

City Level: Main Results

Following the same structure as in the firm-level analysis, we use a simplified difference-in-differences specification to provide aggregate estimates of the magnitude of the effect.[42] Table 4 summarizes the main results from estimating different versions of this specification. Panel A shows the effect of monitoring on air pollution measured by aerosol optical depth. The first two columns use the difference-in-differences strategy by comparing the change in pollution before and after the policy between cities that installed a different number of monitors. The last two columns show the instrumental variable estimates. We only control for city fixed effects and time fixed effects in the first and third column. We add pollution target-by-year fixed effects and control for time-varying weather conditions and fixed effects for the age of the mayor in office in the second and fourth column. The estimates consistently show that one additional monitor leads to a 2.7–3 % decrease in air pollution as measured by aerosol optical depth.[43]

Panel B of Table 4 reports the results from using the same specification as above to estimate the impact on our aggregate measure of enforcement activities: the logarithm of the total number of firms that are subject to any air pollution-related enforcement in the city in a year (or log(# firms) for short). We show results using both the simplified difference-in-differences specification and the instrumental variable models with different controls. All estimates are of a similar magnitude and show that one additional monitor leads to an increase in air pollution-related enforcement of 13–19%. Event study results are reported in Figure D4 in Appendix D and provide evidence suggesting that the parallel trends assumption is valid for this analysis as well. Following the logic in the firm analysis in Section 4.1, we further show that the increase in enforcement is driven by enforcement against firms located within 10 km from a monitor. Table C6 in Appendix C documents these results. As a robustness check, we further investigate the impact of monitoring on the total number of firms that face any enforcement in a city in a year (this includes firms that are not covered by the ASIF survey). Panel B in Table C6 reports the results and show that the overall effects are slightly smaller than the results for our baseline sample. This is driven by a weaker impact on non-ASIF firms, which are much smaller entities (such as local restaurants) that

[42] Formally we estimate:

$$y_{crt} = \delta_c + \gamma_{rt} + \beta m_{ct} + \lambda X_{ct} + \epsilon_{crt},$$

where m_{ct} is the number of monitors installed in city c in year t and all other variables are the same as in Equation 2.

[43] Aerosol optical depth is formally defined as the negative of natural logarithm of the fraction of radiation (e.g., light) that is not scattered or absorbed. Hence, these estimates can be interpreted as percentage changes in pollution.

arguably contribute less to aggregate pollution.[44]

Table 4. Impact of Monitoring

	DiD		DiD+IV	
	(1)	(2)	(3)	(4)
Panel A: Outcome - Aerosol Optical Depth				
# Mon × Post	-0.027***	-0.027***	-0.030***	-0.028***
	(0.0037)	(0.0039)	(0.0049)	(0.0055)
Mean of dependent variable	0.34	0.34	0.34	0.34
Observations	16319	16312	16319	16312
First stage: Dependent variable is # Mon × Post				
Min # Mon × Post			0.60***	0.56***
			(0.047)	(0.048)
F-stat of excl. instrument			161.5	134.7
Panel B: Outcome - log(# firms receiving any air pollution enforcement)				
# Mon × Post	0.13***	0.14***	0.19***	0.19***
	(0.045)	(0.041)	(0.071)	(0.065)
Mean of dependent variable	1.79	1.79	1.79	1.79
Observations	1416	1416	1416	1416
First stage: Dependent variable is # Mon × Post				
Min # Mon × Post			0.57***	0.57***
			(0.060)	(0.060)
F-stat of excl. instrument			90.0	91.9
City FE	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	No
Target-Time FE	No	Yes	No	Yes
Additional controls	No	Yes	No	Yes

Notes: This table reports estimates of one additional monitor's effects on both air pollution and city-level enforcement using our simplified baseline specification. Given the different temporal aggregations of the data, time controls are at the monthly level in Panel A, and at the yearly level in Panel B. Additional controls include weather controls: precipitation and average temperature at the respective time level, and mayor controls: mayor's age. Robust standard errors clustered on the city in parentheses. *, **, *** indicates significance at the 10%, 5% and 1% level respectively. AOD is formally defined as the negative of the natural logarithm of the fraction of light that is not scattered or absorbed. Hence, these estimates can be interpreted as percentage changes in pollution.

[44] We are able to include the non-ASIF firms in this analysis since we can match them to cities, even if we don't know the exact geographic location within the city.

City Level: Specifications and Sample Definitions

In this section, we explore some additional specification checks to make sure that our estimates from the previous section can be interpreted as the causal effects of the monitoring program. Table C7 presents these additional results.

Panels A and B of the Table C7 report the results for AOD and log(# firms) respectively. In columns (1) and (3), we drop data from the provinces Xinjiang and Tibet because the areas covered by cities in these two provinces are much larger than for the rest of the country. The estimates for AOD using the restricted sample is slightly smaller (2.3–2.5%), but of a comparable magnitude to our baseline estimates (2.7–3%) and still highly statistically significant. In columns (2) and (4), we add controls for the following baseline city characteristics interacted with year dummies: city GDP in 2010, whether a city is assigned a background monitor, and the size of the build-up area. The estimated effects are slightly larger (2.8–3.5%) when these controls are included, but again of a similar magnitude to what we find in the baseline model. As shown in Panel B, estimates for log(# firms) are also very close to the baseline estimates. Taken together this evidence suggests that both the difference-in-differences and the instrumental variable model that we estimate are robust to changes in the sample and to the inclusion of additional controls.

City Level: Regression Discontinuity Evidence

While all results from the above implemented empirical strategies and specification checks suggest that we capture the causal effect of the number of monitors on pollution and enforcement, a potential remaining concern is that we are comparing cities of different sizes. This could be an issue if the incentives to reduce pollution changes differently across cities of varying sizes after 2015. We do not have any reason to suspect that this is the case. However, to formally address this potential concern, we conduct an additional analysis in which we explore the variation caused by discontinuities in the number of monitoring stations determined by cutoffs set up in the central government's assignment criteria.

The strategy is essentially a fuzzy regression discontinuity design, where the identification relies on the assumption that all other city characteristics change smoothly at the cutoffs. Table C1 in Appendix C shows the criteria determining the minimum number of monitors for each city in our sample. Compared to the standard regression discontinuity design using one running variable and cutoff, we have two running variables and multiple cutoffs. However, in practice, we document that population size does not predict the realized number of monitors.

Hence, we only use the size of the build-up area as the running variable.[45] To improve the statistical power of our regression discontinuity analysis, which suffers from the low number of cities in our sample, we pool all observations, and make inferences as in a standard regression discontinuity design with a single cutoff. As documented in Table C1, there are 3 cutoffs in total. Among the 177 cities, only 8 cities have a population larger than 1 million or a geographical size of the built-up area that is larger than 100 km^2. We, therefore, focus on the first two cutoffs.

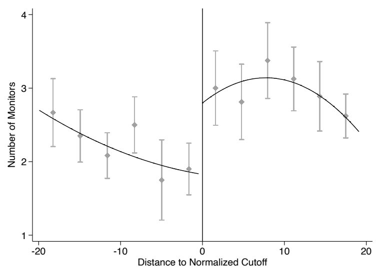
We start with a visual inspection of the data following the approach suggested by Calonico, Cattaneo, and Titiunik (2014). First, we investigate whether the actual number of monitors installed differs for cities on opposing sides of the assignment cutoffs. Figure 6a illustrates the results by showing a binned scatter plot of the number of monitors in each city on the geographical size of the city's build-up area, with negative values for cities below the closest cutoff and positive values for cities above the closest cutoff. Cutoff fixed effects are absorbed before plotting the data and the graph also reports a fitted second degree polynomial. The number of monitors exhibits a sharp jump when moving from the left to the right of the threshold. The first-stage estimates show that cities just above the threshold have installed approximately 1.5 additional monitors. Figure 6b and 6c use the same approach as above and show the reduced form estimates on AOD and log(# firms) in the post period. We see clear jumps in both AOD and log(# firms) when moving from the left to the right of the threshold.

Table 5 quantifies the graphical findings in Figure 6 using the bias corrected local linear regressions approach suggested by Calonico, Cattaneo, and Titiunik (2014), controlling for cutoff fixed effects and average AOD in 2010 (our baseline year). We report estimates for the optimal bandwidth suggested by the same authors in the table and show robustness to alternative bandwidths in Figure D5. The first three columns report the RD regression results using different kernel weighting methods. The last column reports the estimate from a difference-in-discontinuities regression proposed by Grembi, Nannicini, and Troiano (2016), which also exploits the longitudinal nature of the data. This approach entails an estimation of the standard nonparametric RD model with every term being interacted with dummy variables indicating the post period.[46] In all specifications, we cluster standard errors on the city. The evidence shows that cities just above the threshold have substantially lower satellites-measured AOD than cities just below the threshold, with a slightly smaller size of the build-up area. The size of these coefficients are of a similar magnitude to those found

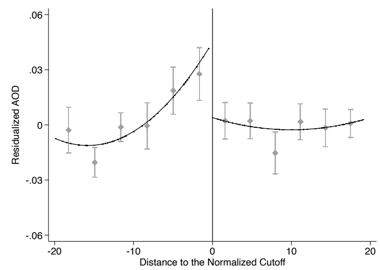
[45] An alternative strategy that uses the shortest distance to any of the two running variables, produces similar results.

[46] More details about the difference-in-discontinuities strategy can be found in Appendix B.1.

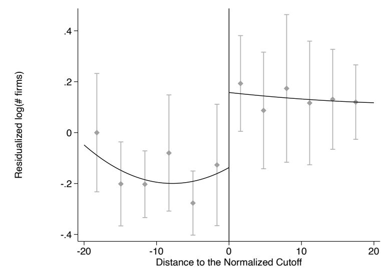
Figure 6. Regression Discontinuity Plots



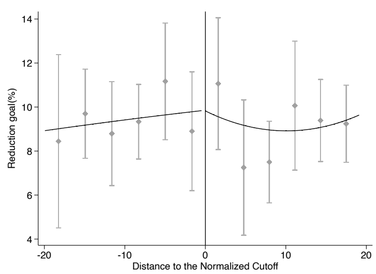
(a) # Monitors (First Stage)



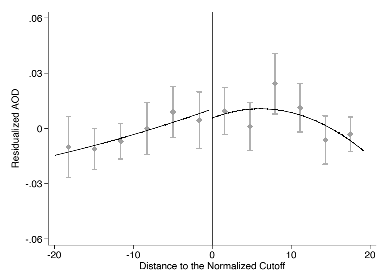
(b) AOD (2015–2017)



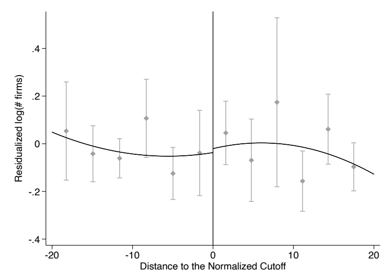
(c) log(# firms) (2015–2017)



(d) Reduction Targets



(e) AOD (2012–2014)



(f) log(# firms) (2012–2014)

Notes: These graphs use the geographical size of the city's build-up area as the running variable and pool data for the first two cutoffs. The cutoff fixed effect is absorbed before plotting the regression discontinuities in Panel 6a and 6d. Baseline (2010) AOD and log(# firms) are also controlled for in Panel 6b, 6c, 6e, and 6f. Robust standard errors are clustered at the city level. Error spikes represent 95 percent confidence intervals.

when estimating the difference-in-differences and instrumental variables strategies discussed above. Cities above the threshold also have significantly higher enforcement levels in the post period. Although less precise, all regression discontinuity estimates confirm the findings in the previous analysis.

The above estimation results rest on the standard assumption that there is no manipulation of the running variable and that other characteristics of cities are smooth at the thresholds. If mayors were able to manipulate the size of the built-up area and sort below the threshold to avoid an additional monitor, our estimates would still suffer from selection bias. Figure D6 in Appendix D is reassuring about the absence of manipulation, as there is no jump in the distribution at any threshold. To test whether municipalities could have manipulated the running variable, we take advantage of the McCrary (2008) observation that in the absence of manipulation, the density of the running variable should be continuous around the threshold. To formally test whether the density of the running variable is continuous at the threshold, we use the local polynomial density estimator and test statistic as described in Cattaneo, Jansson, and Ma (2018). Figure D6e plots the estimated empirical density. The graphical representation clearly suggests that the running variable is contin-

Table 5. Regression Discontinuity Estimates

	RD			Diff-in-Disc
	(1)	(2)	(3)	(4)
Panel A: Outcome - Aerosol Optical Depth				
# Monitors	-0.024**	-0.027*	-0.028**	-0.025*
	(0.012)	(0.014)	(0.014)	(0.014)
Observations	3960	2376	2448	9216
Bandwidth	12.8	8.57	8.91	12.8
First stage	1.50***	1.61***	1.79***	1.50***
	(0.26)	(0.22)	(0.25)	(0.26)
Panel B: Outcome - log(# firms receiving any air pollu. enforce.)				
# Monitors	0.27**	0.27**	0.26**	0.27**
	(0.13)	(0.13)	(0.13)	(0.13)
Observations	258	258	285	688
Bandwidth	10.3	10.7	11.4	10.3
First stage	1.67***	1.68***	1.65***	1.67***
	(0.26)	(0.25)	(0.25)	(0.26)
Kernel	Uniform	Epanechnikov	Triangle	Uniform

Notes: This table reports results from the regression discontinuity design. The first three columns report estimates using different kernel weighting methods. The discontinuities at the normalized cutoff are estimated using local linear regressions and MSE-optimal bandwidth proposed by Calonico, Cattaneo, and Titiunik (2014) for respective kernel weighting method. RD regressions control for cutoff fixed effects and baseline (2010) AOD/log(# firms). The last column reports the Diff-in-Disc regression proposed by Grembi, Nannicini, and Troiano (2016). Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

uous at the threshold. The p-value for the null hypothesis that the density of the running variable is continuous at the threshold is 0.791.

To test the second assumption, we study the main threat to this identification strategy, i.e., that cities with a different number of monitors face different pollution reduction targets. We look at targets for cities close to the thresholds using the same cross-sectional specification we used above to estimate the first-stage impact on the number of monitors. Figure 6d reports the results from this exercise and shows that pollution reduction targets are smooth around the thresholds. This suggests that differential pollution reduction targets do not drive our results. As additional checks, we present RD plots (Figures 6e and 6f) of AOD and log(# firms) for the pre-policy periods (2012-2014). Contrary to the post-policy periods (2015-2017), we see no jumps at the threshold of the normalized running variable. If there is any jump in other characteristics of cities at the thresholds, the violation of the second assumption would likely be reflected in these two figures.

The regression discontinuity estimates corroborate the results from our baseline specification discussed above, indicating that the introduction of the air quality monitors affected local enforcement and pollution. While the regression discontinuity approach has the key advantage of requiring weaker assumptions for causal inference, the power of this analysis is lower and it rests on a limited sample close to the threshold. We, therefore, focus on the panel specifications in the following section, where we explore potential mechanisms.

5 Mechanisms

In this section, we investigate the potential channels through which the information captured by the monitors strengthens enforcement and reduces pollution. In sections 5.1 and 5.2, we explore whether monitors improve top-down and/or bottom-up accountability. Thereafter, we explore how a change in the information provision process that separates the responsibility to provide information from the responsibility to enforce regulations affects our results.

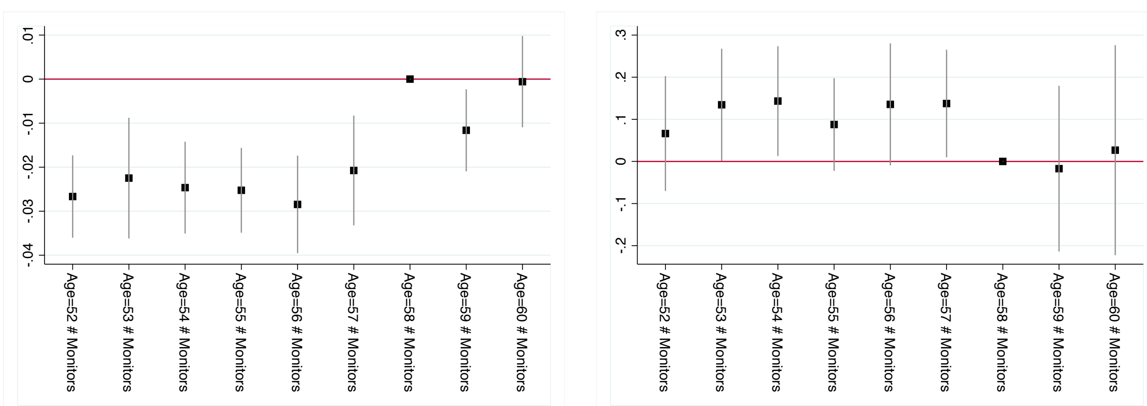
5.1 Top-Down Accountability: Performance Incentives

As discussed in Section 2.1, pollution reduction is one of the criteria that local leaders are evaluated on and their performance determines their probability of promotion. Hence, a natural interpretation of our main findings is that monitors improve the central government's ability to evaluate how well local officials perform. In this section, we investigate this proposed mechanism more directly by exploiting heterogeneity in the promotion incentives faced by local officials. To get a plausibly exogenous measure of local promotion incentives, we use two unique features of the Chinese political system. First, we use the timing of the National People's Congress (NPC), which is held every five years and determines when political promotions are made in China. As documented in Xi, Yao, and Zhang (2018), the average probability of promotion for a city official in the last year of a political cycle (when the NPC is held) is nearly three times that of the first year in a cycle. We then combine this information with two official requirements for mayors of prefecture-level cities: that they retire at age 60 and serve for at least three years in a post. This means that city officials above the age of 57 at the time of the NPC face a discontinuously lower probability of being promoted and, therefore, weaker performance incentives (as documented in Xi, Yao, and Zhang, 2018).

To conduct this analysis, we collect data on all mayors in office during our sample period and calculate their age at the 13^{th} National People's Congress (NPC), which was held in March 2018. If the information provided by the monitors strengthens the ability of the central government to hold local officials accountable, we would expect smaller effects of monitoring for cities with mayors that will be above 57 years of age at the time of the congress. Mayors who are not facing promotion incentives are arguably less likely to work to achieve stricter enforcement of regulations.

To test our hypothesis about promotion incentives formally, we employ the simplified difference-in-differences version of Equation 2 and add an interaction term between the number of monitors in a city and the age of the mayor at the time of the congress. As mayors' work experience might confound our analysis, we use a similar idea to the RD design and plot the differential effects (i.e., the interaction terms) of an additional monitor on both pollution and enforcement by the age of the mayor at the time of the congress in Figure 7. We normalize the effect to 0 for cities with a mayor who would be 58 years old. A distinctive feature of both graphs is that the average effects are not distinguishable from 0 if the mayor is older than 58. At age 57, we see a substantial jump of the estimates in both graphs. The fact that estimates jump at 57 and are then consistent, suggests that our results are indeed driven by performance incentives and not by work experience or other age-related characteristics (for which we would not expect a jump at age 57). We conclude from this analysis that a pre-existing incentive scheme similar to those that are typically proposed to address the principal–agent problem is key in order for monitoring to have an impact on enforcement and air pollution.

Figure 7. Main Results by Performance Incentives



(a) Age vs Decreases in Pollution **(b)** Age vs Increases in Enforcement

Notes: Figure 7 displays the effects of an additional monitor on both enforcement and pollution by mayors' age in 2018. The effect is normalized to 0 for cities with a mayor who would be 58 years old in 2018. Figure 7a is the effect on the increase of enforcement. Figure 7b is the effect on the pollution reduction.

We report the regression results of a simplified version of the results presented in Figure 7 in Table C8 in Appendix C, where we instead interact the number of monitors with whether a mayor is above or below the age cutoff at the time of the NPC. Panel A displays the results for air pollution, and Panel B displays the results on enforcement. In the first column, we use the full sample from our main analysis in Section 4.2 and we then subsequently restrict the sample to mayors closer to the performance age cutoff (again following the regression discontinuity logic). The coefficients for the number of monitors are very similar to those obtained in our main analysis. The interaction terms have the opposite sign and are substantial – suggesting that effects are reduced by between 50–100% for mayors that are not facing performance incentives.

5.2 Bottom-Up Accountability: Citizen Information Acquisition

An alternative explanation for our main results, is that monitors strengthen bottom-up accountability. We believe this is possible in this setting because monitors could inform the population about local pollution levels, noting that awareness of the health impact of $PM_{2.5}$ has increased since 2010 (Barwick et al., 2020). A growing literature (Chen, Pan, and Xu, 2016; Meng, Pan, and Yang, 2017) has further found that authoritarian regimes can be responsive to societal actors.[47] Anecdotal evidence further suggests that local governments could respond to threats of collective action, which are often seen in environment-related issues.[48] Such engagement could potentially further drive up efforts to protect the environment after the air pollution information is released. To formally test this hypothesis, we estimate the simplified version of Equation 2 to identify the causal effects of monitoring on online search behavior. Table 6 shows the estimates for five pollution-related keywords. Columns (1) and (2) show that searches on air pollution increase by around 3% in cities that installed one additional monitoring station.[49] However, no other estimates are significantly different from 0. We believe these effects are likely too small to explain the decrease in air pollution that we document in the previous section – especially considering that real action or engagement would probably be lower than the increase in information acquisition indicated by online searches.

[47] Chen, Pan, and Xu (2016), for example, document that approximately one third of local county governments in China respond to citizen demands expressed online.

[48] See https://www.bbc.com/zhongwen/simp/china/2015/06/150624_shanghai_chemicalplant for example.

[49] Barwick et al. (2020) find larger effects on citizen awareness and behavior when studying the introduction of monitors (extensive margin) in provincial capitals in 2013. The difference in results is explained by the fact that our analysis compares citizens in cities with different numbers of monitors – as opposed to with and without monitors – and focus on a sample of smaller cities.

Table 6. Impact of Monitoring on Online Searches

	(1)	(2)	(3)	(4)	(5)
Outcome:	log(key word)				
Key words:	air pollution	haze/smog	$PM_{2.5}$	air mask	air purifier
# Mon × Post	0.027***	0.042*	-0.0012	0.025	0.0020
	(0.0069)	(0.025)	(0.0015)	(0.016)	(0.030)
Mean of dependent variable	-0.60	-0.18	-0.68	-0.46	-0.021
Observations	14603	14603	14603	14603	14603
City FE	Yes	Yes	Yes	Yes	Yes
Target-Time FE	Yes	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of one additional monitor's effects on online search behavior. Each column is from a separate regression estimating the impact on a specific keyword. Additional controls include weather controls: precipitation and average temperature at respective time level; and mayor controls: mayor's age. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

5.3 Changing Information Provision

Although providing incentives for performance is a common approach to deal with the principal–agent problem, it has long been recognized that high-powered incentives can also distort the type of effort exerted or even encourage various harmful activities focused on improving indicators of performance (Figlio and Winicki, 2005; Banerjee, Duflo, and Glennerster, 2008; Fisman and Wang, 2017; Acemoglu et al., 2020). Manipulating data on which performance is evaluated is one strategy that has been documented in a series of studies (Jacob and Levitt, 2003; Figlio and Getzler, 2006; Banerjee, Duflo, and Glennerster, 2008; Sandefur and Glassman, 2015; Greenstone et al., 2019). In this section, we study whether the structure of the information provision system could mitigate such concerns. In particular, our interest lies in understanding whether a separation of the agent responsible for providing information from the agent responsible for enforcing regulations affects the quality of information and whether such quality improvements can, in turn, strengthen accountability and government performance (i.e. change behavior of the enforcement agent).

Several media sources have reported on manipulation of the pollution data from the monitors by local government officials.[50] Such manipulation took many different forms – ranging from directly adjusting the numbers to spraying the monitors with water, as shown in Figure D10 in Appendix D. Following this reporting, the central government decided to

[50] See `https://p.dw.com/p/32jqR` and `http://www.xinhuanet.com/politics/2018-08/09/c_1123244676.htm` for example.

reassign the control of monitors to external parties, as documented above. In this section, we take advantage of this reassignment policy to see whether increasing the cost of manipulation for the local government is an effective way to improve monitoring, reduce manipulation, and through that, to enforce environmental policy.

As discussed in Section 2, all monitors in our sample were reassigned to third parties at the same time in 2016.[51] Hence, we are not able to exploit any policy variation to estimate the causal effect of the information provider. Instead, we focus on a descriptive analysis and discuss potential implications. First, we study how the AOD elasticity of $PM_{2.5}$ changes when the way information is provided changes (I_t).[52] More specifically, we estimate:

$$log(PM_{2.5})_{mt} = \delta_m + \gamma_t + \beta_1 AOD_{mt} + \beta_2 AOD_{mt} \times I_t + \epsilon_{mt}, \tag{3}$$

where $log(PM_{2.5})_{mt}$ is the logarithm of monthly average concentrations of $PM_{2.5}$ reported from monitor m at time t, δ_m and γ_t represent fixed effects for monitors and time. The variable AOD_{mt} captures the average monthly AOD for pixels covering monitor m.[53] I_t is a dummy variable indicating whether the data is reported after the reassignment. Therefore, the main coefficient of interest is β_2. If information is more accurate when monitors are controlled by the third party, we would expect that AOD and $PM_{2.5}$ measures are more aligned after the reassignment and therefore that $\beta_2 > 0$. Note that this coefficient captures how the alignment between AOD and $PM_{2.5}$ changes over time, while still allowing for pollution levels to change over time.

The results from estimating Equation 3 are reported in Table 7. As a point of reference, we start by estimating the elasticity for all monitors without any interaction term (this replicates the results in the first column of Table C3). We then restrict our analysis to the main monitors used in our study and find a positive estimate for the interaction term. This shows that the elasticity is 0.12 larger after the third party takes over the monitoring stations (corresponding to a 40% increase compared to the pre-period when local governments control the monitors). This evidence is consistent with less manipulation and higher quality information during the period when the information provision responsibility is separated from the enforcement responsibility.

One alternative explanation for the above results is that the AOD data is better able to

[51] Note that there were in total only 6 companies responsible for taking over the operation of the monitors.

[52] We focus on $PM_{2.5}$ because this is the pollutant most strongly correlated with AOD (see Table C3) and because it is the primary indicator that officials are facing incentives to reduce (see Table C2). Martinez (2018) studies the manipulation of GDP data by autocratic leaders using a similar specification.

[53] To deal with the fact that data is sometimes missing for the pixel just above the monitor, due to cloud coverage, the value for missing pixels is interpolated from surrounding cells. All results are robust to using data at the city level instead.

capture changes in pollution after the reassignment (this could, e.g., be due to changes in the composition of pollution over time or changes to the satellite instruments). To make sure that the changes we observe are due to improved monitor data rather than satellite data, we conduct a placebo analysis using the background monitors discussed above.[54] The readings from these monitors are not used by the central government to evaluate the performance of the local government. Hence, there are weaker incentives for officials to manipulate this information. Columns (4) and (5) report the results. We notice that the overall elasticity between air pollution measures reported from monitors and satellites is larger for this sample. When looking at the reassignment, we find that the elasticity change is about half in magnitude and not statistically distinguishable from zero. Taken together this evidence is consistent with less manipulation of the background monitors from the start and no change after the reassignment. This supports our conclusion above that the change in elasticity that we observe for the main sample is driven by changes in the data reported from the monitors. However, we are careful about not to draw too strong conclusions from these patterns since the estimates for the background monitors are imprecisely estimated and not statistically different from those for the main monitors.

Table 7. Monitor Reassignment, Data Quality and Policy Impact

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample:	Full	Main		Background			
Outcome:	$\log(PM_{2.5})$					log(# firms)	AOD
AOD	0.33***	0.33***	0.29***	0.42***	0.40***		
	(0.032)	(0.033)	(0.040)	(0.054)	(0.068)		
AOD × Reassigned			0.12**		0.062		
			(0.053)		(0.13)		
# Mon × Post						0.11***	-0.023***
						(0.041)	(0.0040)
# Mon × Post × Reassigned						0.039**	-0.010***
						(0.019)	(0.0026)
Monitor FE	Yes	Yes	Yes	Yes	Yes	No	No
Time FE	Yes	Yes	Yes	Yes	Yes	No	No
City FE	No	No	No	No	No	Yes	Yes
Target-Time FE	No	No	No	No	No	Yes	Yes
Observations	19185	16987	16987	2198	2198	1416	16319
R-squared	0.78	0.78	0.78	0.77	0.77	0.67	0.63

Notes: This table reports the AOD elasticity of $PM_{2.5}$. Each column is from a separate regression. Columns (1)–(5) control for average temperature, rainfall, mayor's age, and fixed effects specific to monitor and time (month by year). Columns (6) and (7) control for average temperature, rainfall, mayor's age, fixed effects specific to city, and target group by time (6: month by year; 7: year) fixed effect. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

[54] They are usually installed in a local scenic area that is outside the build-up area.

The next exercise we carry out is to check whether local governments exert more effort to decrease pollution after monitors have been reassigned (i.e., when manipulation is more difficult and they thus have less opportunity to decrease recordings through this method). The results are reported in columns (6) and (7) of the Table 7 and show that effects are indeed stronger after monitors have been reassigned. Column (6) shows a 3.9% greater increase in enforcement and column (7) a 1% larger reduction in pollution per monitor installed after the retraction. These pieces of evidence are consistent with local governments switching from data manipulation towards exerting more effort to enforce environmental regulations. Again, we emphasize that these results must be interpreted with caution because we are only exploiting temporal variation and thus need to assume that there are no other simultaneous changes causing these results. An alternative interpretation is that these results capture a lagged impact of the introduction of the monitors. However, we see no apparent reasons for why that would affect the relationship between satellite and ground-based measures of pollution discussed above.

6 Concluding Remarks

This study uses the introduction of a nationwide program in China to investigate the impact of pollution monitoring on local government enforcement of environmental regulations. Exploiting geo-referenced firm data matched with enforcement records, we find that enforcement is stepped up against firms located within 10 km of a monitor. We also document that monitoring increases enforcement efficiency by altering which firms are targeted by local governments and by strengthening the responsiveness of enforcement to local pollution levels.

To study the aggregate response to the policy, we conduct a city-level analysis and compare enforcement and pollution levels in cities assigned different numbers of monitors. This analysis shows that one additional monitor leads to about a 20% increase in the number of firms that face regulatory enforcement and a subsequent 3% reduction in city-level pollution. Given that the policy assigned a median of 3 monitors per city, this corresponds to a substantial reduction in overall pollution. Our estimates suggest a 0.41–0.64 $\mu g/m^3$ reduction in average $PM_{2.5}$ per additional monitor.[55] Previous literature suggests that such a decrease in pollution could have significant implications for both health and economic out-

[55]We arrive at the estimate of 0.41 (0.64) $\mu g/m^3$, the lower (upper) bound, as follows. We multiply 2.8% (3.0%) from Table 4 by 0.33 (0.41 – the elasticity with truthful reporting) from Table 7 to obtain percentage changes in $PM_{2.5}$ per additional monitor. We then multiply by 44.8 (52), the average $PM_{2.5}$ in our sample (average $PM_{2.5}$ in 2015, the first year for which we have $PM_{2.5}$ data), to estimate the implied unit change in $PM_{2.5}$.

comes. For example, Ebenstein et al. (2017) find that a 10 $\mu g/m^3$ increase in PM_{10} reduces life expectancy by 0.64 years in China. The medical costs of heavy air pollution are also substantial – Barwick, Li, Lin, and Zou (2020), for example, document that a permanent decrease of 10 $\mu g/m^3$ in China leads to annual savings of more than 10 billion dollars in health spending. Another cost of heavy air pollution in developing countries is the loss of productivity – Chang et al. (2016); He, Liu, and Salvo (2019) find that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ leads to about 0.5 to 6% drop in productivity and labor cost saving. Combining the evidence from the above literature, the cost savings generated by the reduction in air pollution caused by the monitoring program would likely exceed the related costs in the short run.[56]

An examination of possible mechanisms suggests that the monitoring program is effective because it enables the central government to hold local government accountable for their actions. We support this claim by showing that monitoring is effective in localities where local officials face performance incentives, but is largely ineffective in cities where such incentives are not in place. On the contrary, we find little evidence that monitors strengthen bottom-up accountability. Finally, we document suggestive evidence showing that monitors deliver more reliable information when local governments are not involved in information reporting and are solely responsible for enforcement. When such an information provision structure is in place, the effect of an additional monitor both enforcement of regulations and the level of pollution is about 40% larger.

We believe our findings not only show that pollution monitoring could be an effective policy tool to combat ambient air pollution, but it also offers some general lessons on how to approach the problem of lacking enforcement of government regulations caused by the principal–agent problem. Our findings suggest that reliable real-time monitoring of policy outcomes at the local level could contribute to closing the enforcement gap as long as local officials face performance incentives. However, the existence of such performance incentives could at the same time distort the behavior of local officials towards data manipulation. Therefore, the information provision system would need to be carefully designed to ensure accurate top-down accountability – e.g., by ensuring that information provision and enforcement responsibilities are sufficiently separated.

[56] There are two main costs to consider when installing a monitor: the cost of the equipment and operation costs. According to the government procurement website, the cost of equipment per monitor ranges from $ 200,000 to $ 400,000, while yearly operation costs are about $ 20,000.

References

Acemoglu, Daron, Leopoldo Fergusson, James A Robinson, Dario Romero, and Juan F Vargas. 2020. "The Perils of High-Powered Incentives: Evidence from Colombia's False Positives." *American Economic Journal: Economic Policy* 12 (3):1–43.

Andrews, Steven Q. 2008. "Inconsistencies in Air Quality Metrics: 'Blue Sky' Days and PM10 Concentrations in Beijing." *Environmental Research Letters* 3 (3):034009.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *The Economic Journal* 126 (591):257–280.

Avis, Eric, Claudio Ferraz, and Frederico Finan. 2018. "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *Journal of Political Economy* 126 (5):1912–1964.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *The Quarterly Journal of Economics* 109 (4):1125–1156.

Banerjee, Abhijit, Esther Duflo, Clément Imbert, Santhosh Mathew, and Rohini Pande. 2020. "E-governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India." *American Economic Journal: Applied Economics* 12 (4):39–72.

Banerjee, Abhijit V, Esther Duflo, and Rachel Glennerster. 2008. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association* 6 (2-3):487–500.

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

Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur. 2018. "The Morbidity Cost of Air Pollution: Evidence From Consumer Spending in China." *Working paper* .

Beraja, Martin, David Y. Yang, and Noam Yuchtman. 2020. "Data-intensive Innovation and the State: Evidence from AI Firms in China." *Working paper* .

Besley, Timothy and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India*." *The Quarterly Journal of Economics* 117 (4):1415–1451.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. “Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing.” *Journal of Development Economics* 97 (2):339 – 351.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs.” *Econometrica* 82 (6):2295–2326.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2018. “Manipulation Testing Based on Density Discontinuity.” *The Stata Journal* 18 (1):234–261.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. “The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design.” *The Quarterly Journal of Economics* 125 (1):215–261.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. “Particulate Pollution and the Productivity of Pear Packers.” *American Economic Journal: Economic Policy* 8 (3):141–69.

Chen, Jidong, Jennifer Pan, and Yiqing Xu. 2016. “Sources of Authoritarian Responsiveness: A Field Experiment in China.” *American Journal of Political Science* 60 (2):383–400.

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, Yvonne Jie, Pei Li, and Yi Lu. 2018. “Career Concerns and Multitasking Local Bureaucrats: Evidence of a Target-based Performance Evaluation System in China.” *Journal of Development Economics* 133:84–101.

CIPE. 2012. “Improving Public Governance: Closing the Implementation Gap Between Law and Practice.” Tech. rep., Center for International Private Enterprise.

Conley, T.G. 1999. “GMM Estimation with Cross Sectional Dependence.” *Journal of Econometrics* 92 (1):1 – 45.

Deschenes, Olivier, Huixia Wang, Si Wang, and Peng Zhang. 2020. “The Effect of Air Pollution on Body Weight and Obesity: Evidence from China.” *Journal of Development Economics* 145:102461.

Dhaliwal, Iqbal and Rema Hanna. 2017. "The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics* 124:1 – 21.

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.

———. 2018. "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India." *Econometrica* 86 (6):2123–2160.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4):1241–78.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 114 (39):10384–10389.

Figlio, David N and Lawrence S Getzler. 2006. "Accountability, Ability and Disability: Gaming the System." *Advances in Applied Microeconomics* 14:35–49.

Figlio, David N and Joshua Winicki. 2005. "Food for Thought: The Effects of School Accountability Plans on School Nutrition." *Journal of public Economics* 89 (2-3):381–394.

Fisman, Raymond and Yongxiang Wang. 2017. "The Distortionary Effects of Incentives in Government: Evidence from China's "Death Ceiling" Program." *American Economic Journal: Applied Economics* 9 (2):202–18.

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.

Greenstone, Michael, He Guojun, Jia Ruixue, and Liu Tong. 2019. "Can Technology Solve the Principal-Agent Problem? Evidence from Pollution Monitoring in China." *Working paper* .

Greenstone, Michael and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *American Economic Review* 104 (10):3038–72.

Greenstone, Michael and Patrick Schwarz. 2018. "Is China Winning its War on Pollution?" Tech. rep., Energy Policy Institute at the University of Chicago (EPIC).

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. “Do Fiscal Rules Matter?” *American Economic Journal: Applied Economics* 8 (3):1–30.

Gupta, Pawan, Sundar A Christopher, Jun Wang, Robert Gehrig, YC Lee, and Naresh Kumar. 2006. “Satellite Remote Sensing of Particulate Matter and Air Quality Sssessment Over Global Cities.” *Atmospheric Environment* 40 (30):5880–5892.

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

He, Jiaxiu, Haoming Liu, and Alberto Salvo. 2019. “Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China.” *American Economic Journal: Applied Economics* 11 (1):173–201.

Holmström, Bengt. 1979. “Moral Hazard and Observability.” *The Bell Journal of Economics* 10 (1):74–91.

Holmström, Bengt and Paul Milgrom. 1991. “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design.” *Journal of Law, Economics, & Organization* 7:24–52.

Huang, Zhangkai, Lixing Li, Guangrong Ma, and Lixin Colin Xu. 2017. “Hayek, Local Information, and Commanding Heights: Decentralizing State-Owned Enterprises in China.” *American Economic Review* 107 (8):2455–78.

Jacob, Brian A and Steven D Levitt. 2003. “Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating.” *The Quarterly Journal of Economics* 118 (3):843–877.

Jans, Jenny, Per Johansson, and J. Peter Nilsson. 2018. “Economic status, air quality, and child health: Evidence from inversion episodes.” *Journal of Health Economics* 61:220 – 232.

Jia, Ruixue. 2017. “Pollution for Promotion.” *Working paper* .

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

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):223–42.

Kosack, Stephen and Archon Fung. 2014. "Does Transparency Improve Governance?" *Annual Review of Political Science* 17 (1):65–87.

Lemieux, Thomas and Kevin Milligan. 2008. "Incentive Effects of Social Assistance: A Regression Discontinuity Approach." *Journal of Econometrics* 142 (2):807 – 828.

Martinez, Luis R. 2018. "How Much Should We Trust the Dictator's GDP Estimates?" *Working paper* .

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2):698 – 714. The regression discontinuity design: Theory and applications.

Meng, Tianguang, Jennifer Pan, and Ping Yang. 2017. "Conditional Receptivity to Citizen Participation: Evidence from a Survey Experiment in China." *Comparative Political Studies* 50 (4):399–433.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2016. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review* 106 (10):2895–2929.

Neidell, Matthew and Janet Currie. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?*." *The Quarterly Journal of Economics* 120 (3):1003–1030.

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

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2):200–249.

Pettersson-Lidbom, Per. 2012. "Does the Size of the Legislature Affect the Size of Government? Evidence from Two Natural Experiments." *Journal of Public Economics* 96 (3):269 – 278.

Qin, Yu and Hongjia Zhu. 2018. "Run Away? Air Pollution and Emigration Interests in China." *Journal of Population Economics* 31 (1):235–266.

Reinikka, Ritva and Jakob Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3 (2-3):259–267.

———. 2011. "The Power of Information in Public Services: Evidence from Education in Uganda." *Journal of Public Economics* 95 (7-8):956–966.

Sandefur, Justin and Amanda Glassman. 2015. "The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics." *The Journal of Development Studies* 51 (2):116–132.

Schlenker, Wolfram and W Reed Walker. 2015. "Airports, Air Pollution, and Contemporaneous health." *The Review of Economic Studies* 83 (2):768–809.

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

Snyder, James M. and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118 (2):355–408.

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

Tanaka, Shinsuke. 2015. "Environmental Regulations on Air Pollution in China and Their Impact on Infant Mortality." *Journal of Health Economics* 42:90 – 103.

United Nation. 2019. "Environmental Rule of Law: First Global Report." Tech. rep., United Nations.

Wang, Jun and Sundar A Christopher. 2003. "Intercomparison Between Satellite-Derived Aerosol Optical Thickness and PM2.5 mass: Implications for air quality studies." *Geophysical Research Letters* 30 (21).

WHO. 2016. "Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease." Tech. rep., World Health Organization.

World Bank. 2017. "World Development Report 2017: Governance and the Law." Tech. rep., World Bank.

Xi, Tianyang, Yang Yao, and Muyang Zhang. 2018. "Capability and Opportunism: Evidence from City Officials in China." *Journal of Comparative Economics* forthcoming.

Zheng, Siqi and Matthew E Kahn. 2013. "Understanding China's Urban Pollution Dynamics." *Journal of Economic Literature* 51 (3):731–72.

———. 2017. "A New Era of Pollution Progress in Urban China?" *Journal of Economic Perspectives* 31 (1):71–92.

Online Appendix

Appendix A Data Details

A.1 Enforcement Data Processing

The analysis in this paper relies on new geo-coded data on the enforcement actions carried out by local officials. This data is constructed in two steps. First, information from all enforcement records in a city is extracted and categorised. Second, these records are matched to the annual survey of industrial firms, which we have geo-referenced. The following two sections describe the procedure in detail.

Encoding of Records

We rely on enforcement records collected by The Institute of Public & Environmental Affairs. Figure A1 provides an example of what these records look like and the type of information they contain. In the record, we can identify which regulation the firm has violated and the local government's response to that violation. For each record, we extract whether the violation refers to air pollution, water pollution, solid-waste pollution, or procedural issues[57]; and the punishment imposed by the local government. Our algorithm follows this step-wise procedure:

1. We first check whether the record contains multiple firms:
   - if the record only contains one firm, we extract the whole record;
   - if the record contains multiple firms, we extract only the relevant block.
2. Once the relevant information has been extracted, our categorization by type first distinguishes between enforcement related to air pollution and three other type of violations: water, solid waste, and procedure. The categorization is done by identifying the keywords listed below:[58]
   - keywords for air pollution: NO, PM, SO2, 气, 烟, 尘, 脱硝, 脱硫, 炉;
   - keywords for water pollution: COD, 污水, 水污染, 沉淀, 沟, 渠;
   - keyword for solid waste pollution: 固体;
   - keywords for procedural violation: 未批先建, 批建不符, 未验先投
3. For records related to air pollution, we separately identify the following punishment types: suspension, equipment replacement/upgrading, fine, and warning. The categorization is done by identifying the keywords listed below:

[57] The violation of a procedure usually refers to installation or production before receiving the required license.

[58] Note that one record could contain several different violations.

- keywords for suspension: 停;
- keywords for upgrading: 改, 维修;
- keywords for fine: 罚款, 经济处罚, 万元;
- keywords for warning: 监测情况, 超标

For the vast majority of records, we use a python algorithm to extract the above information. However, about 1500 records are stored as pictures. For these we have manually extracted the information.

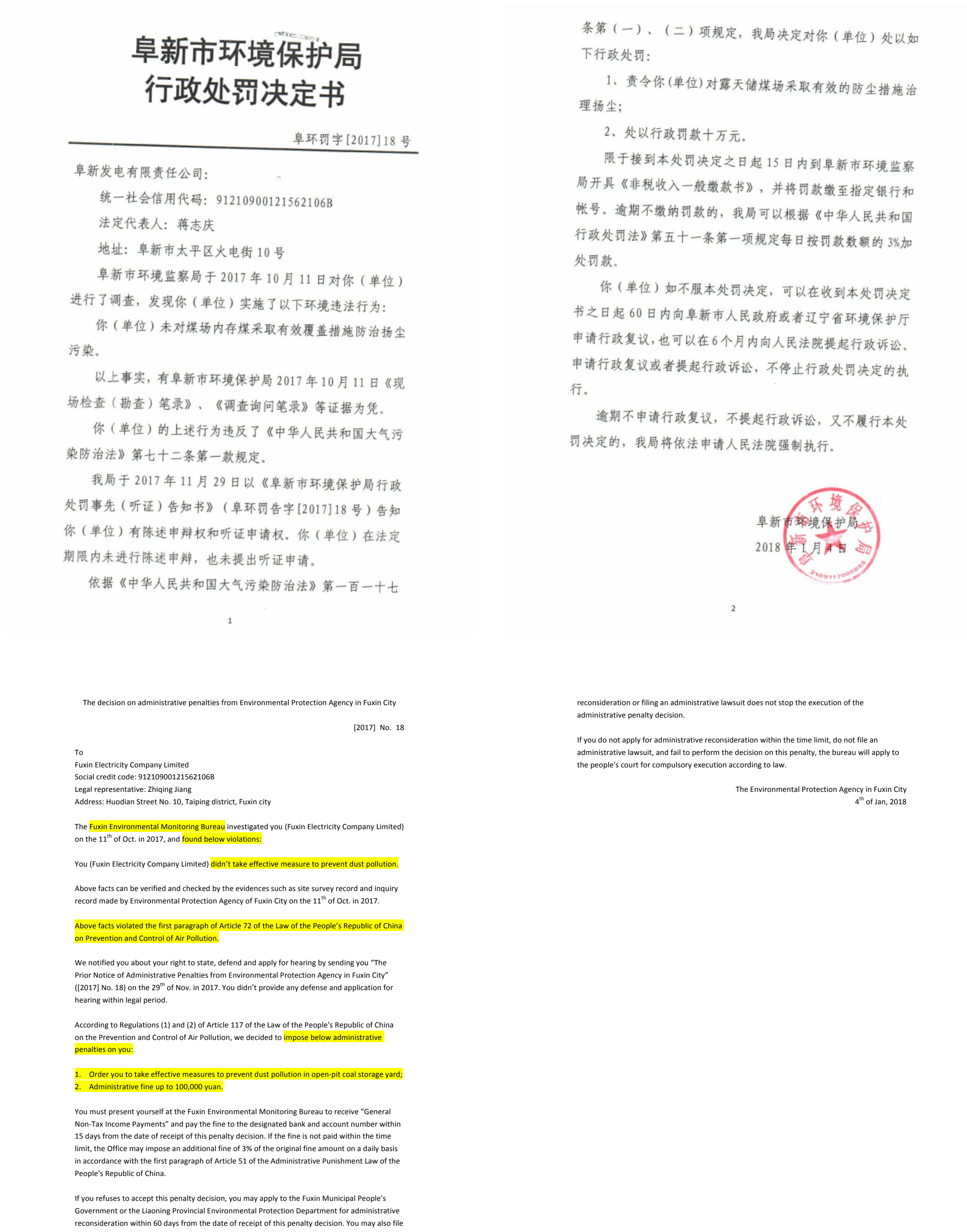
Geo-coding Firm Location

We collect information on all active manufacturing firms using the Annual Survey of Industrial Firms in 2013, the most recent wave. The ASIF data includes private industrial enterprises with annual sales exceeding 5 million RMB and all the state-owned industrial enterprises (SOEs). The data is collected and maintained by the National Bureau of Statistics and contains a rich set of information obtained from these firms' accounting books, such as inputs, outputs, sales, taxes, and profits. Essential for our analysis, the data also includes information about the address of the firm. However, this address information is not always detailed enough to identify an exact geographic location. If this is the case, we rely on two additional sources to complement the ASIF data. First, we follow the recent literature (Beraja, Yang, and Yuchtman, 2020) and use the Tianyancha firm registration database to identify the precise coordinates. If the precise coordinates are not available in the Tianyancha database, we use the Google Maps API to identify the coordinates by using the firm's full name. We then cross-reference the information generated by Google Maps to ensure that it corresponds to the general location provided in the Tianyancha database. For around 4,000 firms, we are unable to pinpoint the exact geographic location using the above approach. For these firms, we manually collect the address information from other internet sources. In the end, we have the precise geographic information for 98.7% of firms.

A.2 Representativeness of Main Analysis Sample

Our sample contains the 177 cities that installed new monitors in 2015. This focus is motivated by three main reasons. First, we do not want to mix cities that had monitors in the past with those that got a monitor for the first time. The key reason for this is that the new information gained from an updated monitor is different since it captures recordings on a much wider set of pollutants. Second, cities with old monitors are dramatically different from cities with new monitors. In Table A1 in the appendix we compare the descriptive statistics of cities with new monitors to those that had air quality monitors before the reform. We

Figure A1. An Enforcement Issued by Fuxin Government



阜新市环境保护局
行政处罚决定书

阜环罚字[2017]18 号

阜新发电有限责任公司：

统一社会信用代码：91210900121562106B

法定代表人：蒋志庆

地址：阜新市太平区火电街 10 号

阜新市环境监察局于 2017 年 10 月 11 日对你（单位）进行了调查，发现你（单位）实施了以下环境违法行为：

你（单位）未对煤场内存煤采取有效覆盖措施防治扬尘污染。

以上事实，有阜新市环境保护局 2017 年 10 月 11 日《现场检查（勘查）笔录》、《调查询问笔录》等证据为凭。

你（单位）的上述行为违反了《中华人民共和国大气污染防治法》第七十二条第一款规定。

我局于 2017 年 11 月 29 日以《阜新市环境保护局行政处罚事先（听证）告知书》（阜环罚告字[2017]18 号）告知你（单位）有陈述申辩权和听证申请权。你（单位）在法定期限内未进行陈述申辩，也未提出听证申请。

依据《中华人民共和国大气污染防治法》第一百一十七条第（一）、（二）项规定，我局决定对你（单位）处以如下行政处罚：

1、责令你(单位)对露天储煤场采取有效的防尘措施治理扬尘；

2、处以行政罚款十万元。

限于接到本处罚决定之日起 15 日内到阜新市环境监察局开具《非税收入一般缴款书》，并将罚款缴至指定银行和帐号。逾期不缴纳罚款的，我局可以根据《中华人民共和国行政处罚法》第五十一条第一项规定每日按罚款数额的 3%加处罚款。

你（单位）如不服本处罚决定，可以在收到本处罚决定书之日起 60 日内向阜新市人民政府或者辽宁省环境保护厅申请行政复议，也可以在 6 个月内向人民法院提起行政诉讼。申请行政复议或者提起行政诉讼，不停止行政处罚决定的执行。

逾期不申请行政复议，不提起行政诉讼，又不履行本处罚决定的，我局将依法申请人民法院强制执行。

阜新市环境保护局
2018 年 1 月 4 日

The decision on administrative penalties from Environmental Protection Agency in Fuxin City

[2017] No. 18

To
Fuxin Electricity Company Limited
Social credit code: 91210900121562106B
Legal representative: Zhiqing Jiang
Address: Huodian Street No. 10, Taiping district, Fuxin city

The Fuxin Environmental Monitoring Bureau investigated you (Fuxin Electricity Company Limited) on the 11th of Oct. in 2017, and found below violations:

You (Fuxin Electricity Company Limited) didn't take effective measure to prevent dust pollution.

Above facts can be verified and checked by the evidences such as site survey record and inquiry record made by Environmental Protection Agency of Fuxin City on the 11th of Oct. in 2017.

Above facts violated the first paragraph of Article 72 of the Law of the People's Republic of China on Prevention and Control of Air Pollution.

We notified you about your right to state, defend and apply for hearing by sending you "The Prior Notice of Administrative Penalties from Environmental Protection Agency in Fuxin City" ([2017] No. 18) on the 29th of Nov. in 2017. You didn't provide any defense and application for hearing within legal period.

According to Regulations (1) and (2) of Article 117 of the Law of the People's Republic of China on the Prevention and Control of Air Pollution, we decided to impose below administrative penalties on you:

1. Order you to take effective measures to prevent dust pollution in open-pit coal storage yard;
2. Administrative fine up to 100,000 yuan.

You must present yourself at the Fuxin Environmental Monitoring Bureau to receive "General Non-Tax Income Payments" and pay the fine to the designated bank and account number within 15 days from the date of receipt of this penalty decision. If the fine is not paid within the time limit, the Office may impose an additional fine of 3% of the original fine amount on a daily basis in accordance with the first paragraph of Article 51 of the Administrative Punishment Law of the People's Republic of China.

If you refuses to accept this penalty decision, you may apply to the Fuxin Municipal People's Government or the Liaoning Provincial Environmental Protection Department for administrative reconsideration within 60 days from the date of receipt of this penalty decision. You may also file an administrative lawsuit with the People's Court within 6 months. Applying for administrative reconsideration or filing an administrative lawsuit does not stop the execution of the administrative penalty decision.

If you do not apply for administrative reconsideration within the time limit, do not file an administrative lawsuit, and fail to perform the decision on this penalty, the bureau will apply to the people's court for compulsory execution according to law.

The Environmental Protection Agency in Fuxin City
4th of Jan, 2018

see that the urban population and the size of the built-up area in cities with old monitors are 5-6 times larger. The size of the economy is also substantially different, as captured by the lights at night data. Finally, we exclude cities that received monitors at an earlier stage due to targeting related to environmental policies as discussed in Section 2.2. This leaves us with our final sample in column (3) of Table A1.

Table A1. Summary Statistics

	Old Cities	Targeted Cites	Our Sample
Aerosol Optical Depth (AOD)	0.45	0.48	0.34
	(0.24)	(0.25)	(0.23)
Number of Monitors	6.26	3.69	3.12
	(2.75)	(0.79)	(1.26)
Urban Population (10,000)	195.1	59.9	33.8
	(303.7)	(34.6)	(21.0)
Size of the Build-up Area (km^2)	257.7	108.4	46.8
	(345.1)	(140.2)	(27.2)
New City	0	1	1
Number of observations	10848	4608	16319

Notes: Author's tabulations.

A.3 Additional Data

Local Leader Characteristics Information on local officials is collected from the database compiled by Jiang (2018). The database contains extensive demographic and career information for over 4,000 key city, provincial and national leaders in China from the late 1990s until 2015. For each leader, the database provides standardized information about the time, place, organization, and rank of every job assignment listed in their curriculum vitae. The data is collected from government websites, yearbooks, and other trustworthy Internet sources. We use the database to calculate the age of city mayors in our sample, which can be used to infer the promotion incentives faced by the mayor, as discussed above. Since our analysis stretches beyond 2015, we expand the database and collect information about the characteristics of mayors up until 2018.

Baidu Search Index To study the impact of new air pollution information, we collect data about local awareness of air pollution information from the Baidu Search Index. Similar to Google Trends (GT), Baidu Search Index provides a measurement of the search volume of a keyword in a given period from both computers and mobile devices. The Index is constructed by summing the weighted frequencies of all search queries for a specific keyword by city and by day. However, the exact algorithm of the Baidu Index is confidential and unknown to the public. Previous studies (Qin and Zhu, 2018; Barwick et al., 2020) argue that the correlation between the Index and actual online search volume is linear. To match the frequency of our analysis on the air pollution data, we collect the monthly search volume from the Baidu Search Index of each city for the following keywords (in Chinese): air pollution, haze/smog, PM2.5, air mask, and air purifier.[59]

Weather Variables To control for local weather conditions, which are important determinants of the concentration of air pollution in prior work, we collect temperature and precipitation data from the China Meteorological Administration. The data combines observations from 496 weather stations across China and uses the ANUSPLIN meteorological interpolation model to generate a monthly dataset with a geographic resolution of 0.5×0.5. We match this data to our prefecture-level cities to get a local measure of weather conditions.

[59] The Chinese translation of these five keywords are 空气污染, 雾霾, PM2.5, 口罩, 空气净化器.

Appendix B Additional Analysis

B.1 Difference-in-Discontinuities

We also exploit the longitudinal nature of our data using a "difference-in-discontinuities" (or Diff-in-Disc) design (Grembi, Nannicini, and Troiano, 2016).[60] This design essentially combines a difference-in-differences (comparing the air pollution in cities with a different number of monitors, before and after 2015) with a regression discontinuity design (comparing the air pollution of cities just above or below certain cutoffs). To estimate the Diff-in-Disc model, we follow the common practice of using local linear regression. More specifically, we estimate the following equation:

$$y_{it} = \delta_0 + \delta_1 D_i + S_i(\gamma_0 + \gamma_1 D_i) + T_t[\alpha_0 + \alpha_1 D_i + S_i(\beta_0 + \beta_1 D_i)] + \xi_{it}, \tag{4}$$

where S_i is a dummy for cities above cutoffs, T_t an indicator for the period after 2015, and D_i the normalized running variable. Standard errors are clustered at the city level. Treatment is captured by $T_t \times S_i$ and the coefficient of interest is therefore β_0. This is the Diff-in-Disc estimate and identifies the reduced-form effect of being just above the cutoff. We normalize the estimates to the treatment effect of one additional monitor by dividing β_0 by the first-stage RD estimates. We check the robustness of our results using different bandwidths. Results of the Diff-in-Disc regressions are shown in Column (4) of the Table 5.

[60]Several studies in the literature have exploited the longitudinal nature of the data in an RD framework, such as the fixed-effect RD estimator in Pettersson-Lidbom (2012), the first-difference RD estimator in Lemieux and Milligan (2008), or the dynamic RD design in Cellini, Ferreira, and Rothstein (2010).

Appendix C Additional Tables

Table C1. Monitor Assignment Criteria

Population (10,000)	Size of Build-Up Area (sq. km)	Min # Monitors	# Cities
< 25	< 20	1	26
$25 - 50$	$20 - 50$	2	82
$50 - 100$	$50 - 100$	4	61
$100 - 200$	$100 - 200$	6	8

Sources: Technical regulation (2013) for selection of ambient air quality monitoring stations (Ministry of Environmental Protection, see www.mee.gov.cn/ywgz/fgbz/bz/bzwb/jcffbz/201309/t20130925_260810.htm)

Table C2. Targets by Province

Targeted Pollutants	Target	Provinces
$PM_{2.5}$	-25%	Beijing, Tianjin and Hebei
$PM_{2.5}$	-20%	Shagxi, Shandong, Shanghai, Jiangsu, Zhejiang
$PM_{2.5}$	-15%	Guangdong, Chongqing
$PM_{2.5}$	-10%	Inner mongolia
PM_{10}	-15%	Henan, Shannxi, Qinghai, Xinjiang
PM_{10}	-12%	Gnasu, Hubei
PM_{10}	-10%	Sichuan, Liangning, Jilin, Hunan, Anhui, Ningxia
PM_{10}	-5%	Guangxi, Fujian, Jiangxi, Guizhou, Heilongjiang
PM_{10}	Keep improving	Hainan, Tibet, Yunnan

Notes: This table reports the pollution reduction targets stipulated by the central government for each province. The reduction targets correspond to the percentage reduction that should be achieved by the end of 2017 compared to 2012.

Source: The Ministry of Environmental Protection

Table C3. Validating Satellite Data

	(1)	(2)	(3)
Outcome:	log($PM_{2.5}$)	log(PM_{10})	log(AQI)
AOD	0.33***	0.28***	0.23***
	(0.021)	(0.020)	(0.015)
Monitor FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	19185	19185	19185
R-squared	0.78	0.79	0.80

Notes: This table reports the relationship between AOD and three monitor-based measures of air pollution: $PM_{2.5}$, PM_{10}, and the combined AQI. Each column is from a separate regression. All regressions control for average temperature, rainfall, mayor's age, and fixed effects specific to monitor and time (month by year). Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

Table C4. Industry Composition

Name of the Industry	Code (two digits)	Freq.	Pct.
Mining and Washing of Coal	6	1588	4.40
Extraction of Petroleum and Natural Gas	7	38	0.11
Mining and Processing of Ferrous Metal Ores	8	568	1.57
Mining and Processing of Non-Ferrous Metal Ores	9	244	0.68
Mining and Processing of Nonmetallic Mineral	10	560	1.55
Mining Support	11	23	0.06
Other Mining	12	4	0.01
Agricultural and Sideline Food Processing	13	3872	10.72
Fermentation	14	1241	3.44
Beverage Manufacturing	15	994	2.75
Tobacco Manufacturing	16	25	0.07
Textile Mills	17	1457	4.04
Wearing Apparel and Clothing Accessories Manufacturing	18	856	2.37
Leather, Fur and Related Products Manufacturing	19	654	1.81
Wood and Bamboo Products Manufacturing	20	994	2.75
Furniture Manufacturing	21	365	1.01
Products Manufacturing	22	768	2.13
Printing and Reproduction of Recorded Media	23	437	1.21
Education and Entertainment Articles Manufacturing	24	603	1.67
Petrochemicals Manufacturing	25	168	0.47
Chemical Products Manufacturing	26	2625	7.27
Medicine Manufacturing	27	999	2.77
Chemical Fibers Manufacturing	28	42	0.12
Rubber Products Manufacturing	29	1404	3.89
Plastic Products Manufacturing	30	3977	11.01
Non-Metallic Mineral Products Manufacturing	31	1449	4.01
Iron and Steel Smelting	32	450	1.25
Non-Ferrous Metal Smelting	33	1226	3.40
Fabricated Metal Products Manufacturing	34	1543	4.27
General Purpose Machinery Manufacturing	35	1537	4.26
Special Purpose Machinery Manufacturing	36	1268	3.51
Transport Equipment Manufacturing	37	238	0.66
Electrical machinery and equipment Manufacturing	38	1437	4.00
Electrical Equipment Manufacturing	39	553	1.53
Computers and Electronic Products Manufacturing	40	218	0.60
General Instruments and Other Equipment Manufacturing	41	134	0.37
Craft-works Manufacturing	42	118	0.33
Renewable Materials Recovery	43	26	0.07
Electricity and Heat Supply	44	1003	2.78
Gas Production and Supply	45	178	0.49
Water Production and Supply	46	222	0.61
Total		36106	100.00

Notes: Industrial classification for national economic activities (GB/T 4754—2002). The sample is from the 2013 Annual Survey of Industrial Firms and includes firms that were set up before 2010 and located within 50 km from an air quality monitor.

Table C5. Firm-Level Enforcement: Placebo

	(1)	(2)	(3)	(4)
Any Enforcement Action				
Outcome	Air	Water	Solid Waste	Procedure
$\text{Mon}_{<10km}$ × Post	0.0078***	0.0020	0.00050	0.0016
	(0.0016)	(0.0013)	(0.00085)	(0.0016)
Mean of dependent variable	0.013	0.0091	0.0031	0.0075
Observations	288696	288696	288696	288696
Firm FE	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes

Notes: This table reports estimates of air pollution monitoring on the probability of being subject to different types of environmental enforcement. Each column reports the estimate from the simplified version of Equation (1). All regressions control for fixed effects specific to firm, industry-by-year interactions, and province-by-year interactions. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

Table C6. City-Level Enforcement: Gradient and All Firms

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Outcome -log(# firms receiving any air pollution enforcement)						
Sample	ASIF	0-10km	10-20km	20-30km	30-40km	40-km
# Mon × Post	0.13***	0.18***	0.078***	0.033	0.046*	0.035
	(0.040)	(0.035)	(0.029)	(0.027)	(0.027)	(0.026)
Mean of dependent variable	1.11	0.50	0.22	0.19	0.21	0.16
Observations	1416	1416	1416	1416	1416	1416
Panel B: Outcome - log(# firms receiving any air pollution enforcement)						
Sample	All Firms		ASIF Firms		Non-ASIF Firms	
# Mon × Post	0.11***	0.16***	0.13***	0.18***	0.099*	0.11
	(0.045)	(0.072)	(0.040)	(0.065)	(0.054)	(0.090)
Mean of dependent variable	1.79	1.79	1.11	1.11	1.40	1.40
F-stat excl. instrument		91.9		91.9		91.9
Observations	1416	1416	1416	1416	1416	1416
Method	DiD	DiD+IV	DiD	DiD+IV	DiD	DiD+IV
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Target-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the effect of one additional monitor on city level enforcement. Each column reports the estimate from the simplified version of Equation (2). Additional controls include weather controls: precipitation and average temperature at respective time level; and mayor's age. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

Table C7. Robustness: Sample Restriction and Additional Controls

	DiD		DiD+IV	
	(1)	(2)	(3)	(4)
Panel A: Outcome - Aerosol Optical Depth				
# Mon × Post	-0.025***	-0.028***	-0.023***	-0.035*
	(0.0038)	(0.0062)	(0.0056)	(0.019)
Mean of dependent variable	0.35	0.35	0.35	0.35
Observations	14625	15931	14625	15931
First stage: Dependent variable is # Mon × Post				
Min # Mon × Post			0.55***	0.42***
			(0.051)	(0.12)
F-stat of excl. instrument			115.7	12.3
Panel B: Outcome - log(# firms receiving any air pollution enforcement)				
# Mon × Post	0.11**	0.14***	0.16**	0.17***
	(0.044)	(0.042)	(0.070)	(0.062)
Mean of dependent variable	1.18	1.18	1.18	1.18
Observations	1263	1415	1263	1415
First stage: Dependent variable is # Mon × Post				
Min # Mon × Post			0.55***	0.61***
			(0.061)	(0.053)
F-stat of excl. instrument			82.2	131.1
City FE	Yes	Yes	Yes	Yes
Target-Time FE	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes
Restricted sample	Yes	No	Yes	No
Baseline characteristics by year	No	Yes	No	Yes

Notes: This table reports the effect of one additional monitor on both air pollution and city level enforcement. Each column reports the estimate from the simplified version of Equation (2). Additional controls include weather controls: precipitation and average temperature at the respective time level, and mayor controls: mayor's age. In column (1) and (3), we drop data from the provinces Xinjiang and Tibet since the area covered by cities in these two provinces are much larger than for the rest of the country. In column (2) and (4), we add controls for the following baseline city characteristics interacted with year dummies: city GDP in 2010, whether a city is assigned a background monitor, and the size of build-up area. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

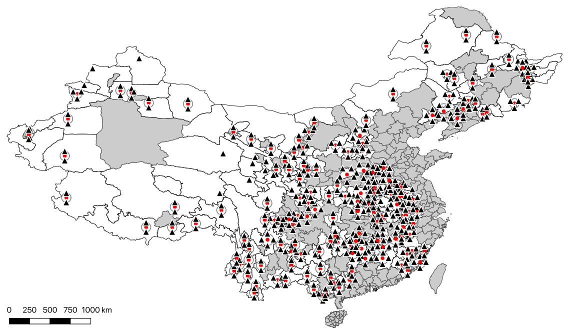
Table C8. Promotion Incentives

	(1)	(2)	(3)	(4)
Age bandwidth:	Full	±10 Years	±7 Years	±5 Years
Panel A: Outcome - Aerosol Optical Depth				
# Mon × Post	-0.026***	-0.026***	-0.027***	-0.022***
	(0.0035)	(0.0036)	(0.0038)	(0.0044)
# Mon × Post × Above 57	0.019***	0.018***	0.017***	0.021***
	(0.0046)	(0.0048)	(0.0048)	(0.0030)
Mean of dependent variable	0.35	0.33	0.30	0.32
Observations	16319	15208	14000	11084
Panel B: Outcome - log(# firms receiving any air pollution enforcement)				
# Mon × Post	0.14***	0.14***	0.14***	0.15***
	(0.046)	(0.046)	(0.049)	(0.052)
# Mon × Post × Above 57	-0.098**	-0.10**	-0.10**	-0.12**
	(0.048)	(0.049)	(0.049)	(0.050)
Mean of dependent variable	1.18	1.15	1.24	1.23
Observations	1416	1360	1248	1008
City FE	Yes	Yes	Yes	Yes
Target-Time FE	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes

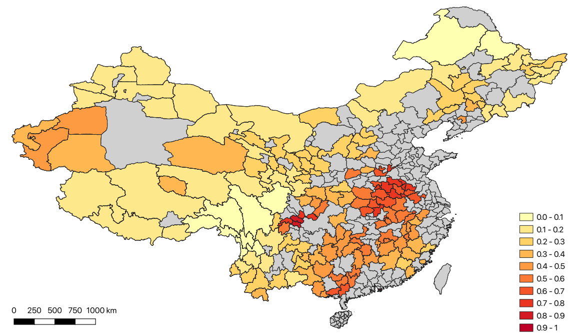
Notes: This table reports heterogeneous effects of monitoring on air pollution (Panel A) and environmental enforcement (Panel B). Each column reports the estimate from the simplified version of Equation (2). Additional controls include weather controls: precipitation and average temperature at the respective time level. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

Appendix D Additional Figures

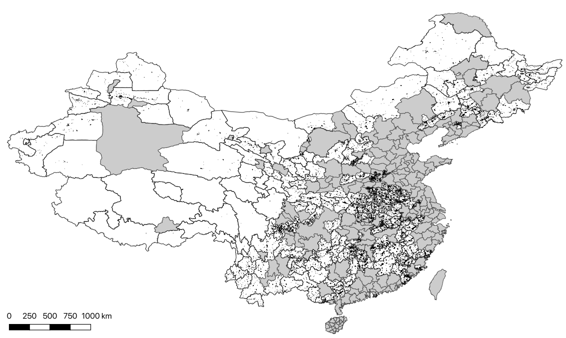
Figure D1. Geographical Distribution of Data



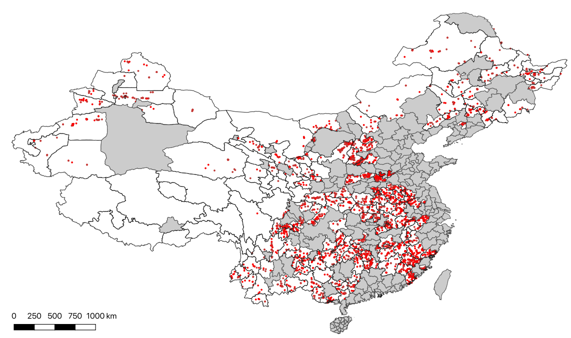
(a) Location of Pollution Monitors



(b) Distribution of AOD in 2010



(c) Location of Firms



(d) Location of Enforcement Activities

Notes: This figure shows the geographical distribution of the data used for analysis in this study. Panel A shows the location of pollution monitors (black triangles). To facilitate the reading of the map, overlapping monitors have been displaced, and the centroid of the overlapping monitors is displayed with a red circle. Panel B shows the average AOD for each prefecture-level city in 2010. Panel C shows the exact geographic location of manufacturing firms in the 2013 Annual Survey of Industrial Firms, and Panel D shows air-pollution related enforcement activities against these firms.

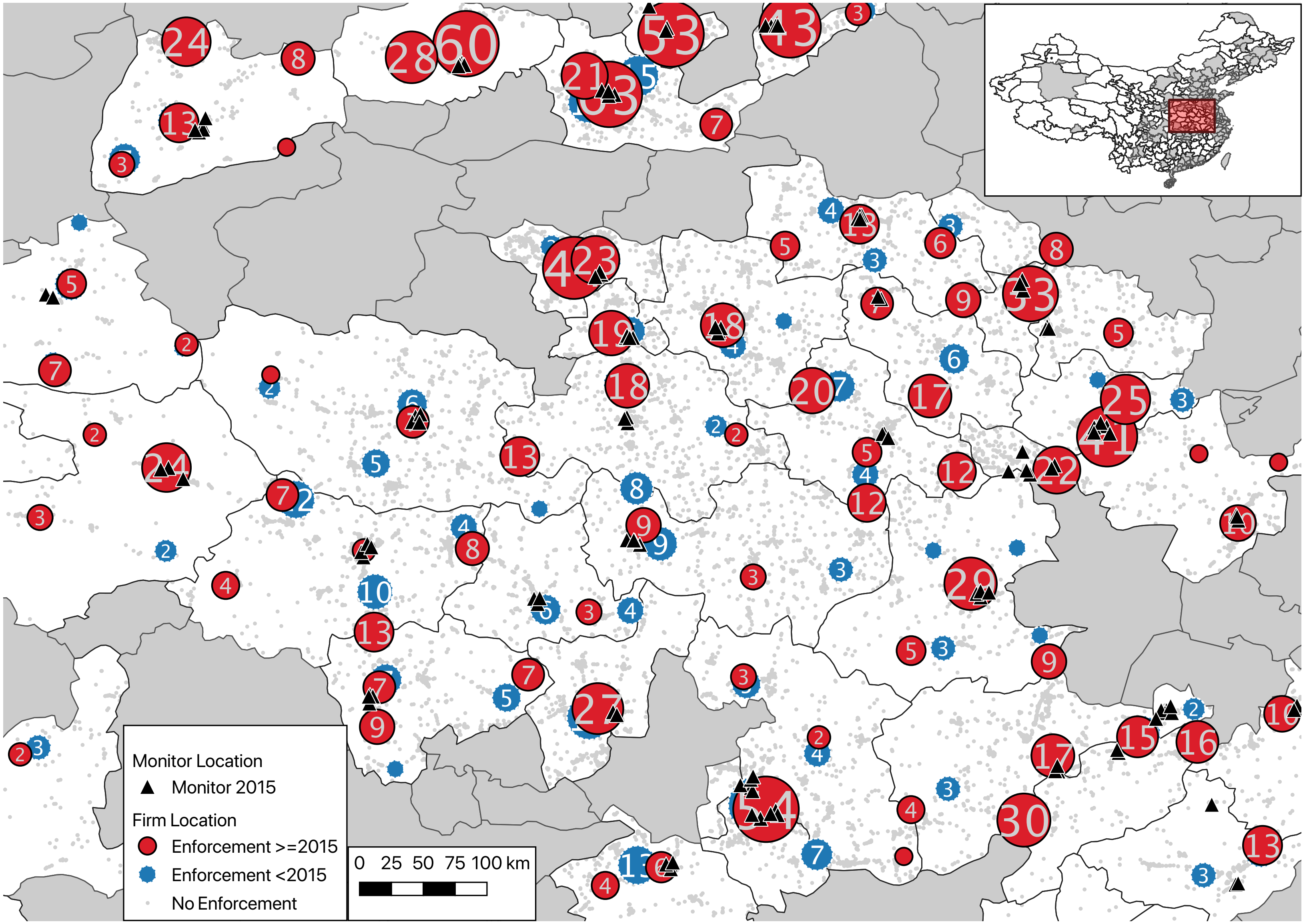
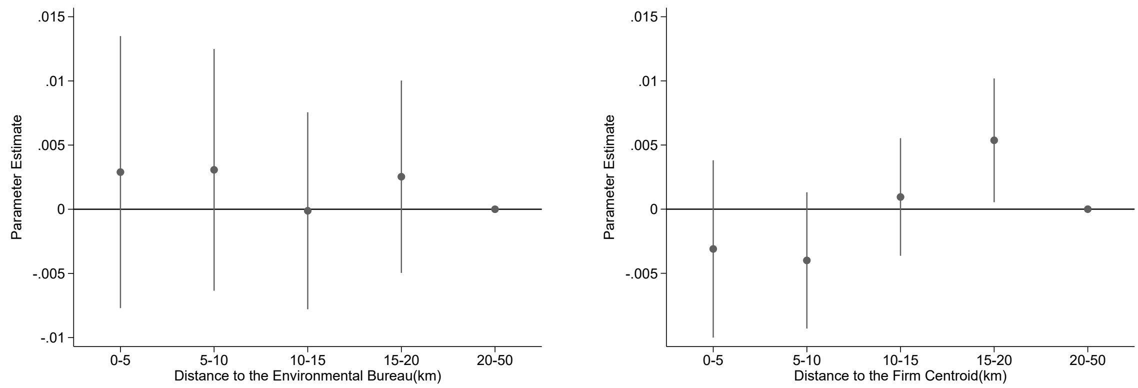


Figure D2. Monitors and Enforcement Activities in Central China

Notes: This figure shows the spatial relationship between monitors, manufacturing firms and enforcement activities for cities in central China. Enforcement activities are reported both for the period before monitors were installed (2010-2014) as well as for the period after installation (2015-2017). Red and blue bubbles mark clusters of enforcement. The size of the bubble corresponds to the number of firms that were issued at least one enforcement activity and the location of the bubble corresponding to the centroid of all enforcement activities occurring within 50 km. Two things emerge from this map. First, the number of enforcement activities clearly increases after 2015. Second, these enforcement activities tend to be located closer to where monitors are placed.

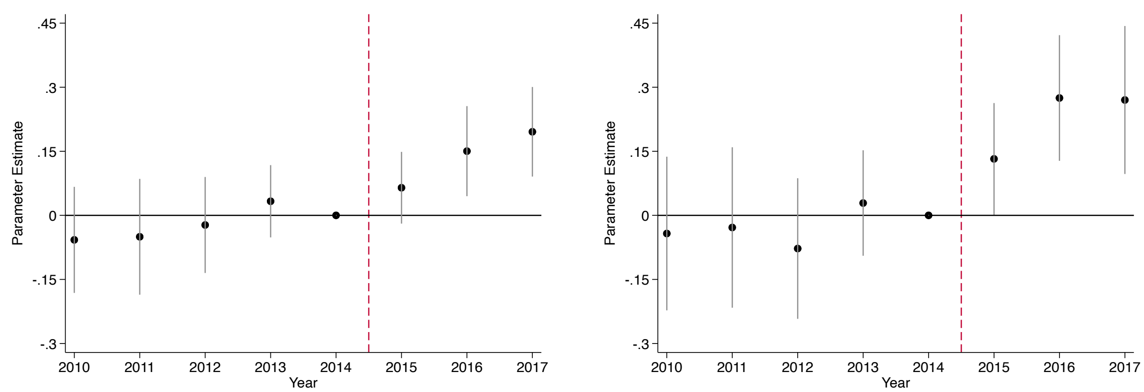
Figure D3. Placebo Gradient



(a) Environmental Bureau **(b)** Firm Centroid

Notes: The figures show the relative increase in enforcement after the introduction of monitors in 2015 for each distance bin to either the Environmental Bureau (Panel a) or the city's firm centroid (Panel b). Estimates are produced using the same specification as for Figure 2b in the paper. Error spikes represent 95 percent confidence intervals based on standard errors clustered on the city.

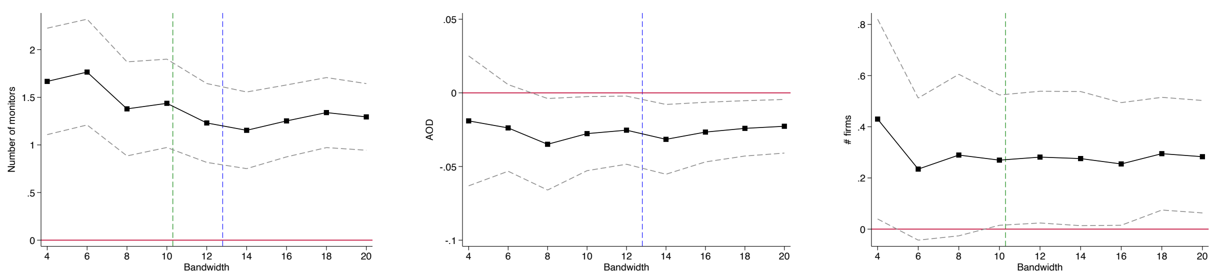
Figure D4. City-level Enforcement: Event Study



(a) DiD **(b)** DiD+IV

Notes: The figures present the estimates from Equation 2 of city-level enforcement (log(# firms)) using two different specifications (DiD, DiD+IV). Error spikes represent 95 percent confidence intervals based on standard errors clustered on the city.

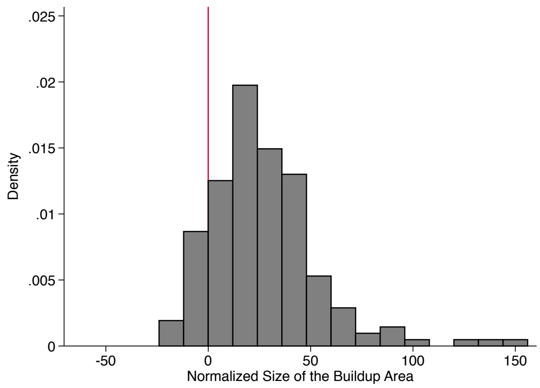
Figure D5. Alternative RD Bandwidths



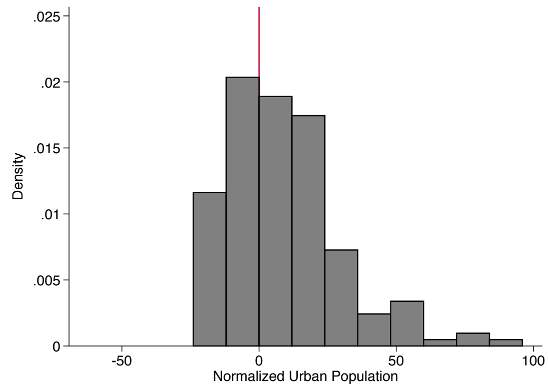
(a) RD in Number (First-stage) **(b)** RD in AOD **(c)** RD in log(# firms)

Notes: Vertical axis: RD coefficients. Horizontal axis: bandwidth used to estimate the reported RD coefficients. The blue dashed line marks the optimal bandwidth (12.8) for the pollution sample using the approach suggested by Calonico, Cattaneo, and Titiunik (2014). The green dashed line marks the optimal bandwidth (9.5) for the enforcement sample using the same approach.

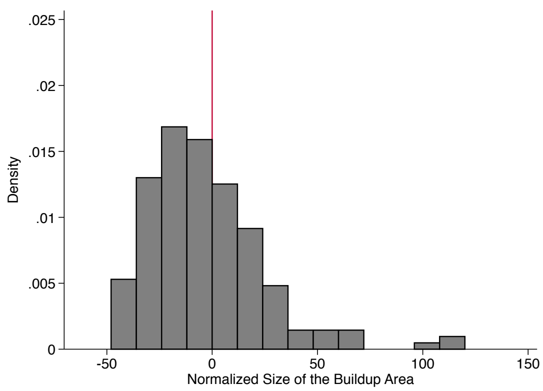
Figure D6. Histogram of Running Variables



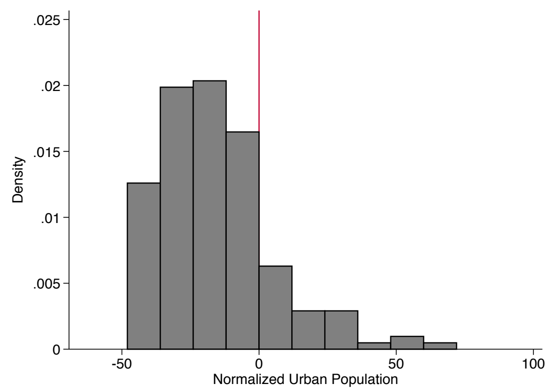
(a) Cutoff 1 – 20 (km^2)



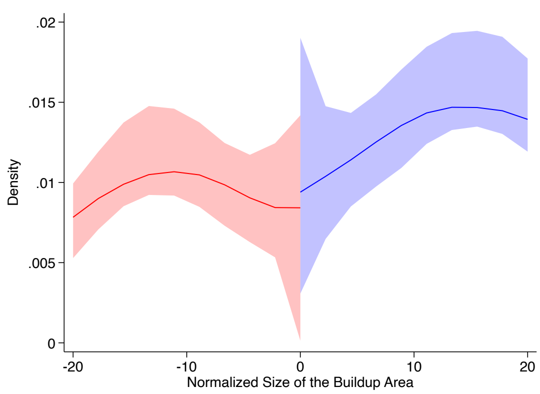
(b) Cutoff 1 – 25 (10,000)



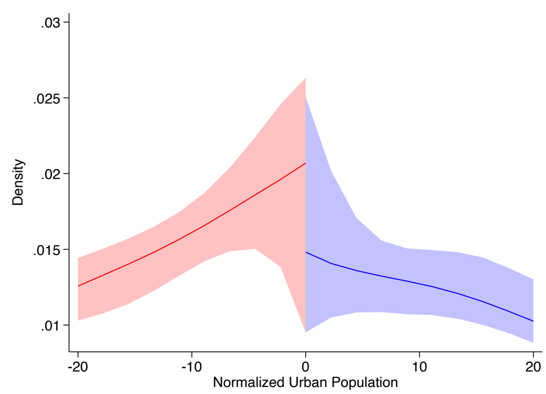
(c) Cutoff 2 – 50 (km^2)



(d) Cutoff 2 – 50 (10,000)



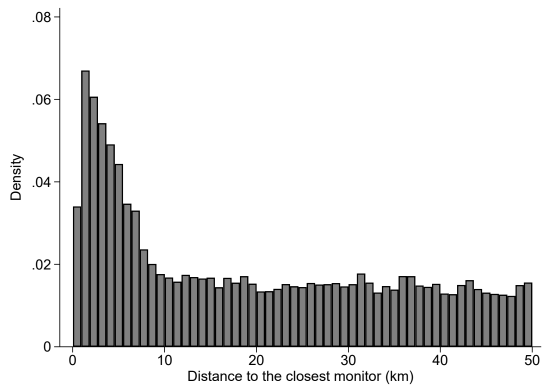
(e) Estimated Density



(f) Estimated Density

Notes: The figures provide histograms and estimated densities of the urban population and the size of the built-up area for our sample over the two cutoffs we use in the analysis. The size of the built-up area and the urban population have been normalized. The p-value for the null hypothesis that the density of the normalized size of the build-up area is continuous at the threshold is 0.791. The p-value for the null hypothesis that the density of the normalized urban population is continuous at the threshold is 0.312.

Figure D7. Distance to the Closest Monitor



Notes: This figure shows the distribution of the distance between ASIF firms and the closest monitor. The sample is restricted to firms that are located within 50 km from a monitor.

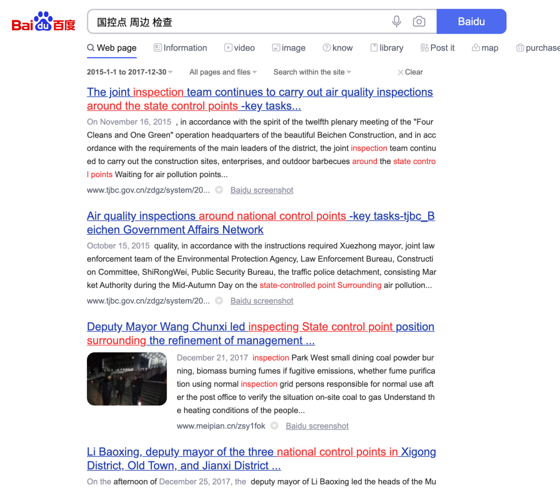
Figure D8. Air-Quality Monitor

Notes: This figure shows an example of the type of monitor that was installed as part of the program.

Source: https://new.qq.com/rain/a/20170124015370

Figure D9. Media Reporting on Enforcement Around Monitors

(a) Search Results in Chinese



(b) Translation

Notes: This figure includes a screenshot and the corresponding translation of a list of news articles generated from a search on the Chinese search engine Baidu using the keywords "monitors", "surrounding area", and "check". The list includes a large number of articles discussing how local governments organise their environmental inspections around the monitors. Some examples include cities that draw special zones around their air quality monitors and send teams of inspectors to those zones, whose task it is to ensure that firms comply with national environmental regulations. Other sources mention that city governments hire volunteers from the public to inspect venues (such as restaurants) within a certain distance from the monitors. Finally, several sources suggest that mayors take a special interest in these inspections by, e.g., directly appointing officials to this task or by visit surrounding areas.
Sources: http://www.baidu.com

Figure D10. Manipulation of Pollution Recordings

Notes: This figure shows a case of spraying water on a pollution monitor to reduce reported pollution.

Source: http://hsb.hsw.cn/2015-01/20/content_8562907.htm