Informed Enforcement: Lessons from Pollution Monitoring in China

BY Sebastian Axbard and Zichen Deng

DISCUSSION PAPER

NHH

Institutt for samfunnsøkonomi
Department of Economics

SAM 01/2021

ISSN: 0804-6824

May 2023

Informed Enforcement
Lessons from Pollution Monitoring in China*

Sebastian Axbard† Zichen Deng‡

This Version: Tuesday 25th April, 2023

Abstract

Government regulations are often imperfectly enforced by public officials. In this study, we exploit the introduction of air pollution monitors in China to investigate if real-time monitoring of policy outcomes affects the enforcement of existing regulations. Using 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 Bergen, University of Zurich, University of York, BI Norwegian Business School, University of Helsinki, CEIBS, NHH, University of Amsterdam, Renmin University, VU Amsterdam, Queen Mary University of London, Warwick, 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 Jordan Ashmore, Pengzhan Qian, Yu Xiao, and Dan Xie 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., 2022; 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 introduction 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 recorded in high income countries.

We begin by investigating how a central government-led program that introduced 552 pollution monitors in 2015 has shaped the enforcement activities of city governments in China. 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,

[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).

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 72% 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 4.2). 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 investigate the determinants of monitor location and document that the placement 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 related to other types of environmental regulations.

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 show that local governments impose stricter punishment against high-polluting firms and become more responsive to local pollution shocks once monitors have been introduced. To show the latter, we exploit exogenous variation in pollution induced by fluctuations in rainfall and wind direction. 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. Similarly, the enforcement response is stronger against firms whose emissions will impact the monitor recording – i.e. firms that are located upwind from a monitor. This suggests that monitors can ensure a more responsive enforcement of regulations – 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 of the policy (including any within city spillover or displacement effect).[2] By exploiting plausibly exogenous variation in the the number of monitors installed in different cities, and thus the share of polluting activity covered by the program, we can assess the impact on total enforcement and pollution. To capture overall pollution changes at the city-level, we follow previous literature and use satellite data on the aerosol optical depth (AOD). The AOD data enable us to measure pollution across the whole city both before and after the introduction of

[2]As depicted in Panel B of Figure D2 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.

monitors and provide us with a reliable data source that cannot be manipulated by local officials. Our analysis exploit the fact that the central government assigned monitors to cities based on their population and geographical size. Using this information, we employ four different empirical strategies: a difference-in-differences specification for cities that installed a different number of monitors, an instrumental variable approach that uses the assigned number of monitors, a regression discontinuity specification that exploits assignment cutoffs and a difference-in-discontinuity specification which combines the previous approaches. All four empirical strategies produce consistent estimates and show that one additional monitor, which increases coverage of high-pollution activity by about 30%, leads to a step-up in enforcement activities by between 15-28% and reduces pollution by 3.1-4.6%. We then investigate spillovers and document that while enforcement activities are concentrated in the high-pollution area close to the monitors, pollution is reduced across the city.

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 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 the most 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. Since air-pollution information is made available to the public online, an alternative mechanism explaining our results is that monitors improve bottom-up accountability. However, additional analysis suggests that this is unlikely to be the main driver behind our results.[3]

Finally, as discussed above, two reasons why information about policy outcomes may be lacking or of poor quality in low- and middle-income countries are: 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

[3] As discussed in Section 5, we do not suggest that the monitors did not improve access to information about local pollution to the public (which in turn may have an impact on government enforcement). However, the response in enforcement and pollution to more comprehensive monitoring that we document at the city level do not seem to be driven by increased dissemination of information to the public.

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., 2022; 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. (2022) study how the updating of monitors in large cities (as opposed to the introduction 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. Finally, a literature focusing on the US document strategic responses to air-quality monitoring (Auffhammer, Bento and Lowe, 2009; Grainger and Schreiber, 2019; Zou, 2021). Our work complements earlier studies by showing how air pollution monitoring affects the enforcement behavior of local governments and aggregate pollution levels within a fixed regulatory framework. We shed light on the consequences of monitoring for government performance – exploring how the responsiveness and targeting of enforcement is affected. In addition, we investigate the long term impact on an aggregate and objective measure of pollution (AOD).[4]

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 most of the 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

[4]He, Wang and Zhang (2020) uses a spatial regressions discontinuity design to investigate the impact on *self-reported* water pollution by firms (i.e. the analysis does not capture potential misreporting by firms and aggregate pollution effects including non-monitored firms). Greenstone et al. (2022) uses a temporal regression discontinuity design to study the short-term impact on air pollution from the upgrading of monitors. Auffhammer, Bento and Lowe (2009) document larger reductions in pollution close to monitors that violate air-quality standards. Grainger and Schreiber (2019) document the strategic placement of air-pollution monitors. Zou (2021) show that there is a temporal reduction in pollution as measured by satellite data during days when monitors under the Clean Air Act are turned on.

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 implementation of the pollution monitoring program we investigate. Then the data used in this study is described in Section 3. The first analysis, which explores firm-level evidence on enforcement, is presented in Section 4.2. The aggregate effect of monitoring on enforcement and pollution at the city level is discussed in Section 4.3. 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 Background & Policy Description

This section provides background information and describes the context in which the national monitoring program studied in this paper was introduced. In 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 infrastructure put in place to monitor these policies' implementation is described in subsection 2.2. Section 2.3 discusses how the information presented in this section guides our analysis.

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).[5] 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 ambition 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.[6]

[5]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.

[6]For the list of targets by province, see Table C3 in Appendix C.

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 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 has 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., 2022). 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.[7] As part of the National Air Quality Action Plan, these cities were simultaneously targeted by a number of additional policies aimed at reducing air pollution (MEP, 2013).[8] The second phase was implemented in 2014 for the 87 designated "environmental role model cities", which face stricter and more frequent evaluation of environmental performance (MEP, 2011; Brehm and Svensson, 2020). The primary aim of the first two phases was to automate old manual monitors, since 70% of these cities already had pollution monitoring in place. The main expansion phase, which is the one we focus on in this paper,

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

[8]This included, e.g., stricter pollution reduction targets by 2017, prohibiting the construction of new coal-fired power stations from 2015, different motor gasoline standards from 2013, reduced reliance on coal and increased construction of natural gas infrastructure, setting up regional environmental impact assessment and joint law enforcement, etc.

was carried out in the following year when all 177 remaining prefecture-level cities (53% of all prefecture-level divisions in China) received monitoring for the first time. 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 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, several media sources have reported that such manipulation did occur (e.g. by spraying the monitor with water to reduce the recordings).[9]

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. These firms were then paid directly by the MEP to operate the monitors.

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.[10]

2.3 Conceptual Framework and Sample Selection

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

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

[10]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.

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 the 177 cities that face the same regulations and receive monitoring for the first time.[11] Figure D1 illustrates how the introduction of monitors changes access to information on pollution both within and between cities. Within cities, monitor readings will be mostly affected by firms located close to and upwind from a monitor.[12] Between cities, information on a larger share of high-pollution activity will be available for those cities that installed a greater number of monitors (as shown in Figure 3).

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 polluting activity). 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. In our analysis we will mainly focus on the overall effect of the monitors. However, in Section 5.2 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 (MEP, 2014). 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

[11]We exclude the cities that received a monitor for the first time before 2015 from our analysis since they were specifically targeted and simultaneously affected by other central government policies aimed at reducing air pollution (as discussed in Section 2.2). This type of targeting is common practice in China (Wang and Yang, 2022). Wang and Yang (2022) emphasize that focusing on areas where policies are initially implemented risk leading to biased estimates of policy impact because of the way these areas are selected and the incentives faced by leaders exposed to early policy experimentation.

[12]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 a monitor would 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.

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) (IPE, 2017). 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) (National Bureau of Statistics, 2013) and link these to the IPE records.

3.1.1 Enforcement Records

We collected all 55,184 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 24,691 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 and in which quarter and year it was carried out. For 95% of the enforcement records related to air pollution, the actions belong to one or several of the following four categories: suspending production (52%), ordering replacement/upgrading of the equipment (54%), levying fines (48%) or issuing a warning (15%).

3.1.2 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.[13] Previous versions of the ASIF data have been used in a number of papers (see, e.g., Song, Storesletten and Zilibotti, 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, and restrict our sample to firms that started operating before 2010 (the first year of our analysis). This allows us to gain an understanding of the underlying distribution of manufacturing firms. 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 D2 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 55,184 records, 52% of them refer to enforcement actions against firms in the ASIF data. Panel D in Figure D2 shows the geographical distribution of enforcement activities against these manufacturing firms.

3.2 Air Pollution Data

3.2.1 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 (MEP, 2017).[14] 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.[15] 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

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

[14] The data can be accessed via this link: http://106.37.208.233:20035/

[15] 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.

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.

3.2.2 Satellite Data (AOD)

Before the expansion of the monitoring system, none of the cities in our sample had any consistent 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 (NEO, 2017). 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 negative 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. We combine measures from the MODIS Aqua and Terra satellites to calculate the mean of AOD in a given month. We aggregate AOD at different geographical levels – ranging from the pixel intersecting with the monitor (approximately 11km x 11km), to the city centre around the monitors (10 to 50km from a monitor) and surrounding areas (beyond 50km from a monitor). To deal with potential within-city spillovers in pollution, our baseline measure is based on the whole prefecture-level city polygon (ESRI, 2017) (as depicted in Panel B of Figure D2). 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.

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).[16] 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). In Table C4, we report results from regressions controlling for monitor fixed effects, time fixed effects as well as

[16] 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.

precipitation, temperature, and mayor's age. Column (1) shows the estimate for $PM_{2.5}$, which is 0.30. 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 (CMA, 2017), quarterly dominant wind direction (CMA, 2017), résumés of all mayors during our sample period (Jiang, 2017) and city level aggregates of citizens' online searches for a set of keywords related to pollution (Baidu, 2017). 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 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 C1.

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.[17] This leaves us with a total sample of 36,103 firms. The majority of these firms are private (81%) and cover a wide range of different industries (Table C5 reports the industry composition for our sample). On average, the firms in our sample are located 19 km from a monitor. However, as depicted in Figure D9 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 in a quarter is 0.5%. 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 of a comparable magnitude (0.29% and 0.52%, 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.[18] The cities we study are small by Chinese standards and have an average population of around 340,000. The average size

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

[18]Hence, this sample is not restricted to firms within 50 km from a monitor and covers the whole city polygon as depicted in Figure D2.

of our sample (measured by both the urban population and the size of built-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.[19] On average the cities in our sample have 2.8 monitors installed and about 4.2 firms face an environmental enforcement actions related to air pollution per quarter, of which 1.5 are against ASIF firms.

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. We start by discussing what determines where monitors are installed in Section 4.1, since this is key for our empirical analysis. Then, we present our firm-level empirical strategy and the corresponding results in Section 4.2. In Section 4.3, we move from studying the local effects of monitoring to the aggregate city effects – exploiting differences in the number of monitors installed.

4.1 Assignment and Location of Monitors

The Ministry of Environmental Protection provided detailed instructions for how many monitors that should be installed in each city 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 assigned to each city was determined by the city's population size and the geographical size of the built-up area. Cities are assigned to one of four groups (corresponding to the installation of 1, 2, 4 or 6 monitors), determined by the criteria that assigns it to the highest group. The detailed assignment criterion, which we use for identification in our city-level analysis, is presented in Table C2.

The location of these monitors is depicted in the map in Panel A of Figure D2. While the exact parameters used by the central government for deciding the precise geographic location of each monitor are unknown, official government documents report that the 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

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

of local pollution. To shed light on what determines the placement decision, we investigate how the location of monitors relates to the spatial distribution and trends in enforcement activities and pollution.

We start by investigating the spatial distribution of enforcement activities and how these change with the introduction of monitors. Figure 1a shows a binned scatter plot of the probability that a firm has any enforcement record related to air pollution in a quarter by the distance to the closest monitor. Black dots indicate the mean probability during the period before air quality monitors were introduced and red diamonds show the mean probability in the post-period. The graph shows that the average quarterly probability of a firm receiving any air pollution-related enforcement action is around 0.0021 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 levels of enforcement. Figure 1b shows another binned scatter plot that instead plots the trends in enforcement for our treatment and control firms (firms located within/beyond 10km from a monitor). The graph shows that both groups of firms face similar trends as well as levels of enforcement in the years before the monitors are installed.

Next, we investigate the pollution level and trends (measured by AOD) in areas where monitors are eventually installed and compare it with surrounding areas. Due to the nature of the AOD data we are not able to use as fine-grained spatial information as in the above analysis, but instead compare the AOD in the pixel where the monitor is located (which are approximately 11km x 11km) to the AOD in the city centre beyond the monitor (10-50km from a monitor) and the area surrounding the city centre (beyond 50km from a monitor). The first two groups roughly correspond to the treatment and control groups in Figure 1b. Figure 1c shows that the pollution level in the pixels where the monitors are located is above average in both the city centre and surrounding area (suggesting that the central government's algorithm is successfully targeting high-pollution areas). However, importantly these areas were on a similar trend to both other areas within the city center as well as the surrounding areas – suggesting that monitors are not placed in areas within the city based on local trends in pollution.[20]

4.2 Firm-Level Evidence

Figure 1a and 1b suggest that there is a significant increase in enforcement activity against firms located close to a monitor once the monitors have been installed in 2015. The above

[20] The pattern shown in Figure 1c is confirmed by a formal test of the pre-trends, which shows no significant differences between the three areas in the five years prior to monitor installment.

(a) Enforcement Gradient **(b)** Enforcement Trends **(c)** AOD Trends

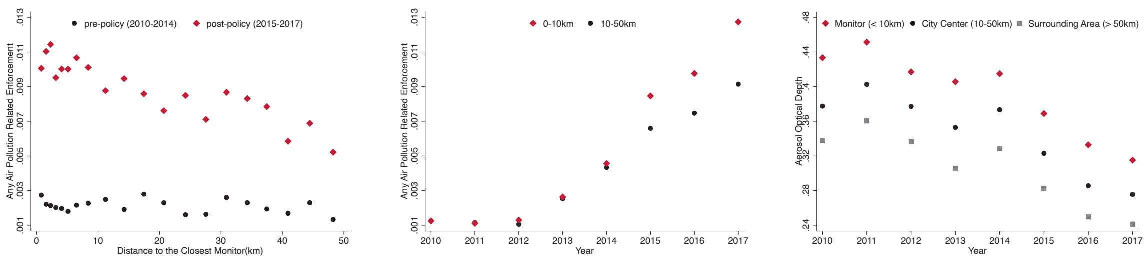


Figure 1. Monitor Location: Air-Pollution-Related Enforcement and Pollution

Notes: Figure 1a shows a binned scatter plots of the relationship between enforcement activity and distance to the closest monitor. Black dots indicate the average quarterly probability of air pollution-related enforcement before introducing the air quality monitors, while red diamonds show the probability after the introduction of monitors. Figure 1b shows a plot of the average share of firms with any air-pollution related enforcement for our treatment and control firms (within and beyond 10km from a monitor) over time. Figure 1c shows a plot of the average AOD over time in three areas: the pixel where the monitor is eventually placed, the city center (10-50km from a monitor) and the areas surrounding the center (>50km from a monitor) but within the boundaries of the prefecture-level city).

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 D10 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[21] 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 within a certain distance from the monitors. Finally, several sources[22] 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.

4.2.1 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 quarter, our

[21] www.163.com/dy/article/H4HC2IIH0534B975.html

[22] newpaper.dahe.cn/dhb/html/2017-12/28/content_212745.htm

model can be written as:

$$y_{ijpt} = \delta_i + \theta_{jt} + \eta_{pt} + \sum_{\substack{d=0-5km \\ d\neq 20-50km}}^{15-20km} \sum_{\substack{k=Q1/2010 \\ k\neq Q4/2014}}^{Q4/2017} \beta_{dk} m_i^{dk} + \epsilon_{ijpt} \tag{1}$$

where y_{ijpt} is an indicator for enforcement, δ_i is a firm fixed effect, θ_{jt} and η_{pt} represent, respectively, industry-by-time and province-by-time fixed effects, m_i^{dk} is an indicator for any monitor being within d km from a firm in quarter k and ϵ_{ijpt} is the error term. Because we condition on firm as well as on industry-by-time and province-by-time 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 quarter 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.[23]

Figure 2 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 from the monitor. For firms 15–20 km from the

[23] As a robustness check, we also report standard errors based on the spatial HAC variance estimator proposed by Conley (1999), following the implementation suggested by Hsiang (2010); Fetzer (2020), which allows for correlation between areas that are geographically close but belong to different administrative units (See Panel A of Table 1). 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.

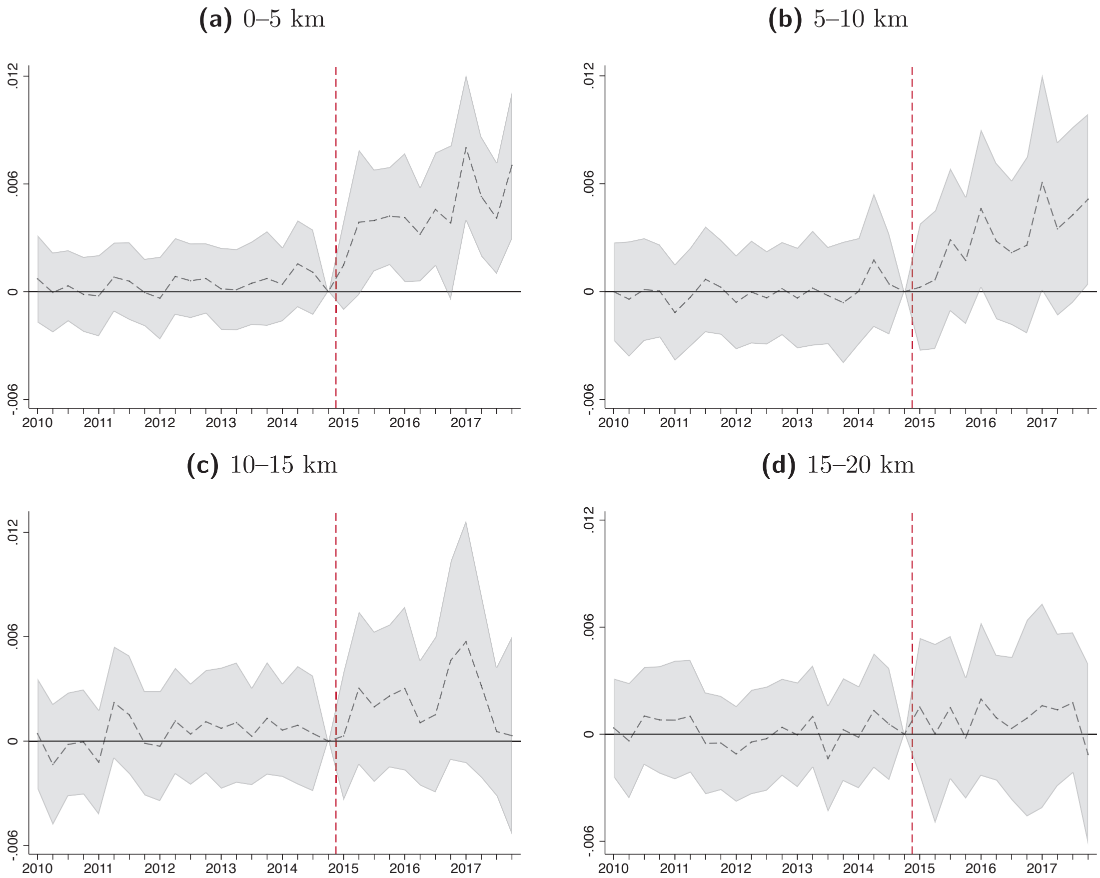


Figure 2. Firm Level: Nonparametric Event Study

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. The shaded area represents 95 percent confidence intervals, calculated using robust standard errors clustered at the city level.

monitor there is no differential change in enforcement activity during our sample period.[24]

4.2.2 Firm Level: Main Results

Guided by the results in the previous section, we use a simplified difference-in-differences specification to provide an estimate of the magnitude of the effect. This specification compares firms within and beyond 10 km from a monitor, but our results are robust to using 15km or 20km instead. Formally, we estimate:

$$y_{ijpt} = \delta_i + \theta_{jt} + \eta_{pt} + \beta m_{it}^{10km} + \epsilon_{ijpt}, \tag{2}$$

[24]In Figure D3 we estimate the aggregate change in the enforcement gradient using a simplified version of Equation 1, which reconfirms the main results reported in Figure 2.

where m_{it}^{10km} is an indicator for a firm having a monitor within 10 km. All other variables are the same as in Equation 1. The results from estimating this specification are shown in Table 1. The first column of Panel A reports estimates on whether any air pollution-related enforcement took place (i.e., the same outcome as in Figure 2). Results suggest that the probability of a firm within 10 km from a monitor receiving an enforcement action in a quarter is 0.33 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 72% compared to the average quarterly probability of enforcement (0.46). The remaining columns of Panel A in Table 1 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 for the firm; "upgrading" – ordering replacement/upgrading of 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.

4.2.3 Firm Level: Targeting

In the previous sub-section, we showed that enforcement increases following the introduction of monitors. 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 were 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 shocks. 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).[25] In Panel B of Table 1, we estimate the differential enforcement response against these firms. The estimates in Column (1) suggest that there is a larger increase in the number of enforcement actions against these firms (significant at the 1% 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 quarter and high as receiving

[25]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.

Table 1. Firm Level: Pollution Monitoring and Enforcement Activities

	(1)	(2)	(3)	(4)	(5)
Panel A: Any Enforcement Action Related to Air Pollution					
Outcome	Any Air	Suspension	Upgrading	Fine	Warning
$\text{Mon}_{<10km}$ × Post	0.0033***	0.0014***	0.0014***	0.0014***	-0.000058
	(0.00056)	(0.00045)	(0.00041)	(0.00043)	(0.00016)
Mean Outcome	0.0046	0.0024	0.0025	0.0022	0.00070
Observations	1155296	1155296	1155296	1155296	1155296
Conley SE	[0.00040]	[0.00031]	[0.00031]	[0.00030]	[0.00017]
Panel B: Intensity and Strictness of Enforcement Action Related to Air Pollution					
Outcome	# Air	Low Intensity	High Intensity	Lenient	Strict
$\text{Mon}_{<10km}$ × Post	0.0031***	0.0027***	0.00018	0.00028*	0.00070**
	(0.00064)	(0.00047)	(0.00013)	(0.00015)	(0.00034)
$\text{Mon}_{<10km}$ × Post × High Pollution	0.040***	0.0017	0.014***	-0.0037**	0.022***
	(0.011)	(0.0068)	(0.0028)	(0.0016)	(0.0058)
Mean Outcome	0.0052	0.0042	0.00040	0.00065	0.0016
Observations	1155296	1155296	1155296	1155296	1155296

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. All regressions control for fixed effects specific to firm, industry-by-time and province-by-time 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 quarter. 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 quarter. In contrast, the dummy variable "strict" is defined as one if all three types of punishments are issued against a firm in a quarter. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

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 1 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.

4.2.4 Firm Level: Responsiveness

Next, we investigate whether monitors make local governments' enforcement efforts more responsive to recorded pollution. The main empirical challenge inherent in studying this is the endogeneity of local pollution. To overcome this challenge, we exploit two different sources of plausibly exogenous variation: local rainfall shocks and wind direction.

We start by investigating rainfall shocks. First, we establishing that local rainfall shocks are important determinants of local pollution. For each city, we construct an indicator ($\text{Rain}_{>\tilde{x}}$) for whether the quarterly rainfall is above the median for that city or not. Table C6 shows estimates of the relationship between the average pollution recordings across the monitors in the city and this high rainfall indicator, controlling for city and time fixed effects. Results show that average pollution recordings are consistently about 7-9% lower in quarters 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 have any reading classified as heavily polluted (AQI>200) is reduced by 22% (.024/.11). We then explore how enforcement activities respond to pollution monitoring in the presence of rainfall shocks by estimating an augmented version of our baseline model (Equation 2) that introduces interactions with the quarterly rainfall shock.[26]

The results from this analysis are reported in Panel A of Table 2. Column (1) adds the rainfall shock and its interaction with our main treatment variable (m_{it}^{10km}). Two main take-aways emerge. First, enforcement activities do not respond to rainfall shocks in the absence of monitors – suggesting that there is no direct effect of rainfall on enforcement. Second, the effect of monitors on enforcement is substantially smaller (0.0043–0.0018=0.0025) when rainfall levels are above the median – i.e., when pollution levels are lower – as opposed to when rainfall levels are below the median (0.0043) – 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 responsive to changes in local pollution. To investigate whether this result reflect a general increase in the enforcement responsiveness in times of high pollution (as opposed to simply a stronger willingness to reduce the monitors' pollution recordings), we estimate a model with the full set of interactions. Column (2) reports the results from estimating this model and shows suggestive evidence (significant at the 10%-level) that, in the presence of monitoring, there is a general increase in enforcement during periods of high pollution also in areas further away from the monitor. However, as indicated by the triple interaction, the effect is even stronger for firms close to a monitor. These results also provide additional

[26] We prefer taking this reduced form approach instead of instrumenting local pollution levels since we only have pollution data from the monitors from 2015 on-wards.

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.

We then move on to explore how the direction of winds affect the enforcement response. Emissions from firms that are upwind from a monitor will be moved by the wind towards the monitor, while emissions from all other firms are moved away from the monitor. Since upwind firms arguably have a larger impact on the pollution recorded by the monitor, local government officials have a greater incentive to enforce regulations for these firms. To test this we follow previous work (Freeman et al., 2019) and identify a firm as "upwind" if it is within 45 degrees of the dominant quarterly wind vector which passes through the monitor. Figure D4 in the appendix illustrates how we classify upwind firms. Following the same approach as above, we interact whether a firm is upwind in a quarter with our main treatment variables. The results are reported in Panel B of Table 2. Column (3) shows that the coefficient for the upwind indicator is small and not significantly different from zero – suggesting that firms do not face differential enforcement by quarterly winds before the introduction of monitors. However, following the introduction of monitors, the enforcement response is stronger against upwind firms (as shown by the coefficient for $\text{Mon}_{<10km}$ $\times$ Post $\times$ Upwind). Upwind firms face an increase in enforcement when monitored that is more than twice as large ($0.0024 + 0.0035 = 0.0059$) as that faced by other firms (0.0024). Column (4) further documents that while this step up in enforcement is particularly pronounced for firms close to the monitor, it is also detectable for firms further than 10km from a monitor (as shown by the estimate for Post $\times$ Upwind). This suggests that more distant upwind firms also affect the monitor's recording of pollution and therefore face increased enforcement.

4.2.5 Firm Level Robustness: Time-Specific Shocks and Spillovers

Our interpretation of the above results is that monitors lead to an increase in enforcement of environmental regulations. In this section, we will discuss and investigate two threats to this interpretation of our results.

The first concern we will address is the risk that our results are affected by time-specific policies or shocks that co-inside with the introduction of monitors. This could be the case if firms close to a monitor respond differently to such a shock compared to firms further away. For this to bias our results, firms within the same industry/province would need to be differently affected depending on their distance to a monitor (since our baseline specification controls for both 4-digit industry-by-time and province-by-time fixed effects). While we are not aware of any such policies or shocks, we investigate this potential concern further by conducting three additional tests. First, we use Equation 2 to look at environmental

Table 2. Firm Level: Enforcement Response By Rainfall and Wind Direction

	(1)	(2)		(3)	(4)
Panel A: Rainfall			*Panel B: Wind Direction*		
	Any Air			Any Air	
$\text{Rain}_{>\tilde{x}}$	-0.00027	0.00031	Upwind	0.00033	0.000043
	(0.00033)	(0.00027)		(0.00023)	(0.00031)
$\text{Mon}_{<10km}$ × Post	0.0043***	0.0041***	$\text{Mon}_{<10km}$ × Post	0.0024***	0.0027***
	(0.00069)	(0.00074)		(0.00057)	(0.00067)
$\text{Mon}_{<10km} \times \text{Rain}_{>\tilde{x}}$		0.00014	$\text{Mon}_{<10km}$ × Upwind		-0.00035
		(0.00024)			(0.00045)
$\text{Mon}_{<10km} \times \text{Post} \times \text{Rain}_{>\tilde{x}}$	-0.0018***	-0.0015**	$\text{Mon}_{<10km}$ × Post × Upwind	0.0035***	0.0024**
	(0.00068)	(0.00076)		(0.00082)	(0.0011)
$\text{Post} \times \text{Rain}_{>\tilde{x}}$		-0.0018*	Post × Upwind		0.0015**
		(0.0011)			(0.00075)
Mean Outcome	0.0046	0.0046	Mean Outcome	0.0046	0.0046
Observations	1155296	1155296	Observations	1155296	1155296

Notes: This table reports results from augmented versions of our baseline model (Equation 2). Panel A (columns 1-2) adds interactions with an indicator for rainfall being above the median ($\text{Rain}_{>\tilde{x}}$) – i.e. the interaction captures the differential effect on enforcement when pollution is relatively low (see Table C6). Panel B (columns 3-4) adds interactions with an indicator for a firm being upwind from a monitor (Figure D4 illustrates this classification) – i.e. the interaction captures the differential effect for firms whose emissions are moved by the wind towards the monitor. All regressions control for fixed effects specific to firm, industry-by-time and province-by-time interactions. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

enforcement that is not related to air pollution. The results are reported in Panel A of Table C7 in the Appendix. For enforcement related to water pollution, solid waste pollution, and procedure violation, estimates are small and statistically insignificant. Second, Panel B of Table C7 gradually includes controls that interact baseline characteristics with time fixed effects to our main specification. First we address the fact that our sample period overlaps with the relocation of basic manufacturing away from coastal regions stipulated in the 12th 5-year plan. Column (2) addresses this issue by including the distance from the city to the coast interacted with time fixed effects. Then we address the fact that policies or shock exposure might be related to firm characteristics in column (3) and interact the number of employees and ownership status with time fixed effects. Finally, we control for city by time fixed effects in column (4). Estimates remain of a comparable magnitude and are highly statistically significant across all specifications. Third, we conduct placebo tests in which we estimate the specification used to produce Figure 2, but instead of the distance to the closest monitor we use the distance to the local environmental bureau or the distance to the city's firm centroid. We do this to investigate whether there is a differential response in enforcement after 2015 depending on plausible spatial determinants related to both the costs and returns from enforcement.[27] Figure D5 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.[28] Taken together, these results suggest that the step-up in enforcement behavior that we observe is indeed driven by the monitors.

The second concern we will investigate is whether the results we observe above is affected by spillovers. This could be the case if our control group (firms further away from the monitor) is also affected by the introduction of monitors, e.g. if firms observe and respond to enforcement actions against neighbours or if the local government reduce enforcement effort against firms far from a monitor. Depending on the nature of the spillovers these could lead us to either over or underestimate the impact of monitoring. Our first attempt to address this issue is to reexamine the patterns in the raw data presented in figures 1a and 1b above. The figures show that before 2015, enforcement levels and trends are the same in both treated and control firms. After 2015, there is a clear step-up of enforcement against firms within 10km from a monitor while firms further away continue on a similar upward-sloping enforcement trend. If these firms would have been affected by spillovers, we would have expected a change in their enforcement trend post 2015. While these patterns in the raw data are reassuring, they do not necessarily rule out spillovers since we do not know the counterfactual trends. Therefore, we move on to conduct an analysis at the city-level in the following section where we analyse aggregate outcomes that take any within city spillovers into account. In addition, we specifically investigate the impact of the intensity of monitoring on non-monitored areas following an approach similar in spirit to the work by Crépon et al. (2013).

4.3 City-Level Evidence

To study the impact of monitoring at the city-level, we exploit variation in the intensity of the monitoring program. This approach allows us to infer the overall impact of more extensive

[27] We argue that the costs of enforcement are likely to be lower close to the environmental bureau (where monitors are often placed) and that the returns from enforcement are likely greater in central areas (due to a greater concentration of population and economic activity). To calculate firm 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.

[28] The main estimate in Column (1) of Panel A in Table 1 changes from 0.0033 to 0.0017 and is significant at the 5%-level.

monitoring 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 C2 in Appendix C) and compare outcomes in cities that installed different numbers of monitors. The argument behind this approach is that a larger number of monitors will ensure that officials are held accountable for a greater share of the overall pollution in a city. Figure 3 illustrates this point for our data by showing how the number of monitors assigned to a city is related to the share of high-pollution activity that is covered by a monitor. We again rely on the ESR database to identify high-polluting firms. To adjust for firm size, our baseline measure of high-pollution activity covered calculates the share of a city's high-polluting firms' revenue that fall within 10km from a monitor.[29] The figure shows that the share of high-pollution activity covered within 10 km from a monitor increases monotonically with the number of monitors, from 20% for cities with 1 monitor to more than 80% for cities with 6 monitors. We report marginal effects in Table C8. Our baseline estimate suggests that an additional monitor increases coverage by 11 percentage point (a 30% increase compared to the mean). These results are robust to using alternative distances to the monitor, other indicators for firm size as well as exploiting variation in both the actual and assigned number of monitors (reported in Table C8).

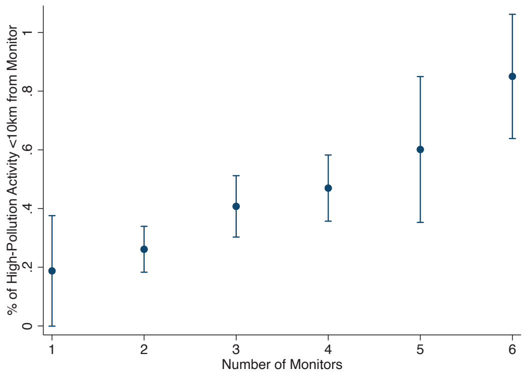


Figure 3. Number of Monitors and Coverage of High-Pollution Activity

Notes: This figure documents the relationship between the number of monitors in a city and the share of the city's high-pollution activity that is covered by monitors. This measure is constructed by calculating the share of the city's ESR firms' revenue that fall within 10km from a monitor. Marginal effects and alternative measures are reported in Table C8.

[29]The ESR database allows us to identify high/low-polluting firms (i.e., whether the firm is on the list or not). We use the revenue share to scale our measure by firm size and thereby get closer to the share of overall pollution that is covered by the monitors. Our results are robust to using alternative measures of firm size as reported in Table C8 in the appendix.

4.3.1 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 C3, and t is a generic time period (month for the pollution analysis and quarter for the enforcement analysis), our model can be written as:

$$y_{crt} = \delta_c + \gamma_{rt} + \sum_{\substack{k=Q1/2010 \\ k \neq Q4/2014}}^{Q4/2017} \beta_k m_c^k + \lambda X_{ct} + \epsilon_{crt}, \tag{3}$$

where y_{crt} is either an aggregate measure of a city's monthly AOD or the log of the total number of firms that receive any enforcement related to air pollution in a quarter, m_c^k is either the actual number of monitors in the city, or the predicted number of monitors according to Table C2, in a given quarter 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 quarter–year for the enforcement specification).[30] The variable X_{ct} represents time-varying city controls including: total precipitation, average temperature, the age of the mayor in office and the geographical and population size of the city at baseline interacted with the post variable. 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-time 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 with different numbers of monitors. To assess the validity of this assumption, we start by investigating the AOD trend for each group. In Figure 4a, 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. Second, 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 3 – setting the average pollution in the quarter

[30] We add +1 to the number of firms with any enforcement related to air pollution before taking the log to avoid generating missing observations. Our results are robust to using the inverse hyperbolic sine transformation instead of the log.

before monitors were installed as the baseline. Estimates from this specification are shown in Figure 4b. We first estimate a standard event study specification using the actual number of monitors installed in a city as our independent variable of interest. These estimates are reported by the black dashed line in the figure. Point estimates are imprecise in the early time period, but 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 – effects that are even stronger in the second and third year.[31] Results on enforcement follow the corresponding pattern: there is no evidence of differential pre-trends and a large increase in enforcement in cities with additional monitoring after 2015 (reported in Figure D6).

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). The instrumental variable estimates are marked by the dashed grey line in Figure 4b. The IV coefficients are less precise, but follow the OLS estimates closely in the pre-period and are even larger in the post-period. There is no evidence of differential trends leading up to the intervention, supporting the common trend assumption between cities assigned different numbers of monitors. Again these results are mirrored by estimates for enforcement (reported in Figure D6).

4.3.2 City Level: Regression Discontinuity Plots

While the analysis above suggests that we capture the causal effect of the number of monitors on pollution and enforcement, a potential remaining concern is that the above approach is not adequately capturing potential confounding effects of city size (i.e. our controls for population and the size of the built-up area are not sufficient). 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 discontinuities in the number of monitoring stations assigned by the central government. Cities are assigned to one of four groups (corresponding to the installation of 1, 2, 4 or 6 monitors) based on either the city's population or on the geographical size of the built-up area in 2014 (as shown in Table C2 in Appendix C), whichever assigns it to the highest group.[32]

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

[32] For example, a city with a population of 200,000 individuals and a built-up area of 21 square kilometers is assigned 2 monitors.

(a) AOD Trends (b) Event Study

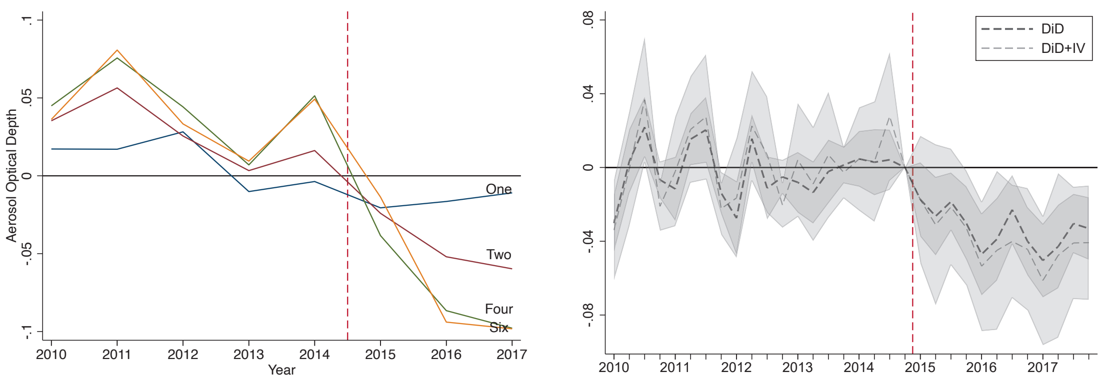


Figure 4. City Level: Impact of # Monitors on AOD

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 (3) using variation either the actual (DiD) or the assigned number of monitors (DiD+IV). Dashed black lines represent the coefficients from DiD, whereas grey dashed lines represent DiD+IV estimates. Shaded areas 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.

We exploit this variation to implement a fuzzy regression discontinuity design. Compared to the standard regression discontinuity design which uses one running variable and cutoff, this setting provides us with two potential running variables (population and geographical size) and three potential cutoffs per running variable (1-2, 2-4, 4-6). However, in practice, we document that there are no discontinuous changes in the number of monitors at the population thresholds (estimate 0.38 and p-value 0.23). We therefore only exploit variation in the size of the built-up area. To reduce the impact of outliers, we focus on the first two cutoffs since the final group only contains 8 cities (5 within the optimal bandwidth).[33] We pool all observations, and use the distance to the closest geographical threshold as the running variable. Following the suggestion by Cattaneo, Keele and Titiunik (2021), we control for the baseline value of the outcome as well as threshold fixed effects to improve precision.[34]

We start with a visual inspection of the data following the approach suggested by Calonico, Cattaneo and Titiunik (2014). First, we document the difference in the number of monitors installed for cities on opposing sides of the assignment cutoffs. Figure 5a illustrates the results by showing a binned scatter plot of the number of monitors in each

[33] All results are robust to include the third threshold in the analysis as well.

[34] The reasoning for introducing these controls follow the same logic as in experiments where baseline outcomes are commonly used to improve precision. We control for the AOD or the log of the number of firms facing any enforcement action in 2010, respectively. We show in Table C12 that estimates without controls are very similar, but less precise.

city on the geographical size of the city's built-up area, with negative values for cities below the cutoffs and positive values for cities above the respective cutoffs. Cutoff fixed effects, as well as the baseline value of the outcome variable for time-varying data, 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 1.3 additional monitors. Figure 5b and 5c use the same approach as above and show the reduced form estimates on average monthly AOD and the log of the quarterly number of firms that received any enforcement related to air pollution in the post period. We see clear jumps in both variables when moving from the left to the right of the threshold.

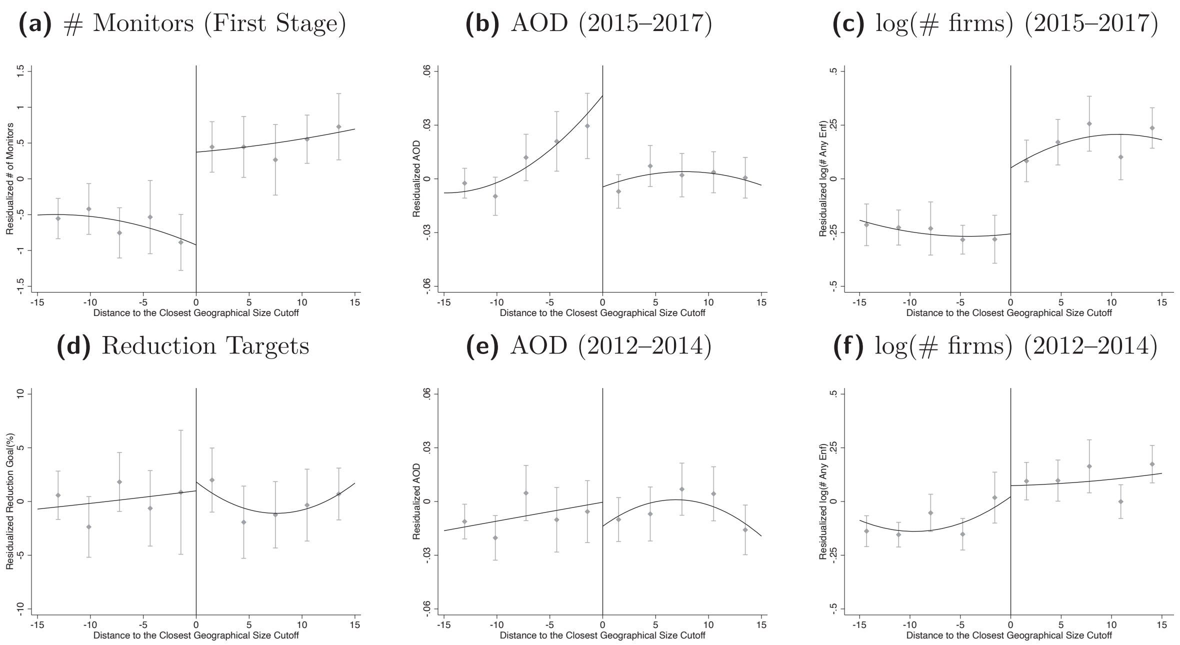


Figure 5. City Level: Regression Discontinuity Plots

Notes: This figure reports regression discontinuity plots using the procedure suggested by Calonico, Cattaneo and Titiunik (2014). The running variable is the distance to the closest geographical size threshold (as defined in Table C2). Figure (a) shows how the number of monitors differ for cities around the geographical size thresholds, while figure (d) shows the variation in pollution reduction target. The difference in AOD and log of the number of firms facing any enforcement after the monitors are installed are shown in figures (b) and (c), while the corresponding outcomes before the monitors are installed are shown in figures (e) and (f). To reduce sampling variability, we follow existing literature (Lee and Lemieux, 2010; Cattaneo, Keele and Titiunik, 2021) and residualize our dependent variables before plotting the RD graphs. We use cutoff fixed effects, and the baseline (2010-2011) value of the outcome variables (AOD and enforcement, respectively) for the time-varying data, to perform this residualization. Robust standard errors are clustered at the city level. Error spikes represent 95 percent confidence intervals.

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 could suffer from selection bias. Figure D8 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 D8c plots the estimated empirical density. The graphical representation clearly suggests that the running variable is continuous at the threshold. The p-value for the null hypothesis that the density of the running variable is continuous at the threshold is 0.642.

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 5d 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 5e and 5f) for our main outcomes during the pre-policy period (2012-2014). Contrary to the post-policy periods (2015-2017), we see no jumps at the threshold – suggesting that other relevant characteristics are smooth at the threshold.

4.3.3 City Level: Main Results and Robustness

We now turn to quantify the magnitude of the effect from the different strategies described above. First, we estimate aggregate DiD and IV effects using the following simplified version of Equation 3:

$$y_{crt} = \delta_c + \gamma_{rt} + \beta m_{ct} + \lambda X_{ct} + \epsilon_{crt}, \tag{4}$$

where m_{ct} is the number of monitors installed in city c at time t and all other variables are the same as in Equation 3. The first two columns of Table 3 report estimates of β using the DiD and IV strategy, respectively. Panel A shows the effect of monitoring on air pollution measured by aerosol optical depth and Panel B shows the results on the log of the number of firms that faced any air pollution related enforcement activity. The estimates show that one additional monitor leads to a 15-19% increase in enforcement and a 3.1–4.6 % decrease in AOD. To put these estimates in perspective we can calculate the enforcement response elasticity with respect to increased coverage of high-pollution activity – using our estimates from Figure 3. Since one additional monitor is associated with a 30% increase in coverage of high pollution activity our results imply an enforcement response elasticity between 0.5-0.7. However, we express caution when interpreting these estimates since our measure of pollution coverage is imprecise and does not necessarily capture all important pollution sources within a city.

Tables C9, C10 and C11 in the Appendix explore the robustness of these results. We start by interacting the time fixed effects with additional city characteristics in Table C9. First, we show that estimating a slightly more demanding specification where we interact baseline city population and the geographical size of the built-up area with time fixed effects instead of the post variable does not alter our main estimates. We then interact the time fixed effects with additional city characteristics: GDP in 2010 and whether a city is assigned a background monitor. Due to the large number of fixed effects, estimates on enforcement are somewhat smaller and less precise, but overall results are of a comparable to our main results. Table C10 report estimates when 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 both pollution and enforcement using the restricted sample are again of a comparable magnitude to our baseline estimates. As a final robustness check, we further investigate the impact of monitoring on the total number of firms that face any enforcement in a city in a quarter (this includes firms that are not covered by the ASIF survey). Panel A in Table C11 reports the results and show that the overall effects are close to the results for our baseline sample.[35]

Second, to quantify the regression discontinuity results above we use the bias corrected local linear regressions approach suggested by Calonico, Cattaneo and Titiunik (2014), using a uniform kernel and controlling for cutoff fixed effects and the 2010-2011 baseline value of our outcome variables.[36] We report the formal specification in Appendix B.1. Column (3) in Table 3 report our baseline estimate, which uses the optimal bandwidth suggested by the same authors. Results are comparable to our DiD and IV estimates discussed above and suggest a 26% increase in enforcement and a 3.9% reduction in AOD. We again investigate the robustness of these estimates in the appendix. We start with establishing in Table C10 that also these estimates are robust to excluding the large provinces Xinjiang and Tibet. Figure D7 shows that estimates are consistent across a range of alternative bandwidths and

[35]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.

[36]We explore the implications of these choices in Table C12 in the Appendix. This shows that the choice of kernel has limited impact on the results, while the inclusion of baseline controls is important for precision but less so for the size of the estimates.

Table C13 reports that estimates are similar across the two cutoffs.

Finally, we combine the two approaches above and estimate a difference-in-discontinuities regression following Grembi, Nannicini and Troiano (2016). This approach combines the standard nonparametric RD model with every term being interacted with dummy variables indicating the post period. We again rely on the optimal bandwidth approach suggested by Calonico, Cattaneo and Titiunik (2014). We report the formal specification in Appendix B.2. Estimates from this specification are reported in Column 4 of Table 3. Results are comparable to previous estimates, but less precisely estimated. Again we document the robustness of these results with respect to the sample studied, the firms considered and separate by cutoff in tables C10, C11, C13.

Table 3. City Level: Impact of Monitoring

	(1)	(2)	(3)	(4)
Empirical Strategy:	DiD	DiD+IV	RD	Diff-in-Disc
Panel A: Aerosol Optical Depth				
# Monitors	-0.031***	-0.046***	-0.039***	-0.029*
	(0.0069)	(0.013)	(0.015)	(0.021)
Observations	16335	16335	3209	8508
Panel B: log(# firms receiving any air pollution enforcement)				
# Monitors	0.15***	0.19**	0.26**	0.28*
	(0.046)	(0.098)	(0.10)	(0.19)
Observations	5664	5664	1116	2976
Panel C: # Monitors (First Stage)				
Estimate		0.72***	1.28***	1.28***
		(0.11)	(0.23)	(0.23)
Kernel			Uniform	Uniform
Bandwidth			11.3	11.3

Notes: This table reports the main results for the two main outcomes for each of the four different empirical strategies used in the city-level analysis. Panel A reports results for aerosol optical depth and Panel B for the log of the number of firms receiving any enforcement action related to air pollution. Columns (1) and (2) show the estimates from Equation 4, controlling for city fixed effects, time by pollution reduction target fixed effect, population and the geographical size of the built-up area at baseline interacted with the post variable, and time varying controls for total precipitation, average temperature and the age of the mayor. Column (1) exploits variation in the actual number of monitors installed, while column (2) instruments the actual number of monitors with the assigned number. The corresponding first stage estimate is reported in Panel C. Column (3) reports regression discontinuity estimates using local linear regression, a uniform kernel and the MSE-optimal bandwidth proposed by Calonico, Cattaneo and Titiunik (2014). We control for cutoff fixed effects and the average of the outcome in 2010-2011 – details in Appendix B.1. The first stage effect of being assigned to the group above the cutoff on the number of monitors installed is reported in Panel C. Column (4) reports estimates from the difference-in-discontinuity approach suggested by Grembi, Nannicini and Troiano (2016) – details in Appendix B.2. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

4.3.4 City Level: Spillovers

We now turn to separate the aggregate city-level effect of monitoring into direct and spillover effects. To perform this analysis we exploit the same four identification strategies as above, but calculate the outcomes separately for the monitoring station (within 10km), the city centre (10-50km from a monitor) and the surrounding areas (beyond 50km). This allows us to observe how enforcement behaviour and pollution changes in monitored as well as in non-monitored areas. Observing this is important since officials could potentially reallocate enforcement effort from non-monitored to monitored areas in response to the policy. If such behaviour took place we would expect to see a reduction in enforcement activity in non-monitored areas of cities that installed a greater number of monitors compared to cities that installed fewer monitors. Our approach is similar in spirit to the work by Crépon et al. (2013) – which identify spillovers by investigating the response of non-treated firms in regions with different treatment saturation.

Table C14 reports these results and show that while we see that a greater number of monitors led to a substantial increase in enforcement close to the monitors, estimates for the city centre and surrounding areas are still positive, though smaller and not significantly different from zero for most empirical strategies. This suggests that the increase in enforcement close to the monitors that we observe in the firm level analysis is not driven by a reallocation of enforcement effort from non-monitored areas. Estimates for pollution show reductions across the city. While DiD and IV estimates are comparable across areas, discontinuity estimates suggest smaller pollution reductions in surrounding areas (with point estimates ranging from from 49% to 78% of the size of the estimates for the area close to the monitor). These estimates suggest that there are positive spillovers (reduction in pollution) in areas not covered by a monitor. There are a number of potential explanations consistent with these results. One possibility is that there is a reduction in the spread of pollution from the high-pollution areas close to the monitors, where we document an increase in enforcement, to the non-monitored areas of the city. An alternative explanation is that the step up in enforcement in the city centre has deterrence effects affecting the pollution behaviour of non-covered firms as well. However, we are careful in not interpreting these estimates too strongly since satellite based measures of pollution are highly spatially correlated. Hence, these effects could partly reflect the inability of satellites to pick up differences in pollution at a high spatial resolution.

To sum up, the results in this section show that an increase in monitoring intensity (i.e. a greater coverage of high-polluters in a city) lead to enforcement against a larger number of firms (in particular those close to a monitor) and an overall reduction in AOD in the entire city. Point estimates are consistent across four different empirical strategies, but vary in how precise they are. The regression discontinuity approach has the key advantage of requiring weaker assumptions for causal inference, but the power of this analysis is lower and it rests on a limited sample close to the threshold. We therefore focus on the DiD specification, which produces the most conservative and precise estimates, in the following mechanism section.

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, we explore whether monitors improve top-down and/or bottom-up accountability. Thereafter, we document in Section 5.2 how a change in the information provision process that separates the responsibility to provide information from the responsibility to enforce regulations affects our results.

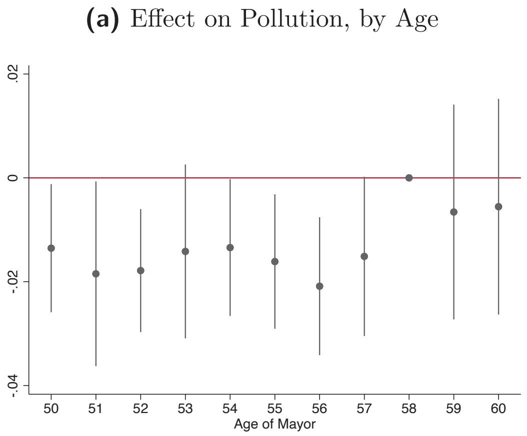
5.1 Top-Down Accountability: Performance Incentives

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 the main mayor in office in the year when the monitors were introduced 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 add an interaction term between the number of monitors in a city and the age of the mayor at the time of the congress to Equation 4. 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 6. We set the baseline to the mayor who is 58 years old at the time of the congress and estimate other coefficients relative to this. A distinctive feature of both graphs is that the effects are not distinguishable from the baseline 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 for lower ages, 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). To further ensure that this result is not driven by confounding factors we perform a corresponding analysis of city baseline characteristics in Figure D11 in the appendix. This balance test shows that the number of monitors, the size of the city, economics activity as proxied by lights at night as well as baseline pollution and enforcement activity is smooth around the 57 age threshold – suggesting that our results are driven by the difference in promotion incentives.

We report the regression results of a simplified version of the results presented in Figure 6 in Table C15 in Appendix C, where we instead interact the number of monitors with whether a mayor is 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 4.3 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 now correspond to the effect for the mayors with weaker performance incentives, while the interaction term shows the additional effect of monitors in the presence of promotion incentives. These terms are all statistically significant and suggest that the performance incentive increases the response by between 65-87%. A corresponding balance table is reported in Table C16. This shows that while there are some imbalances when we compare all cities with mayors above the threshold with all cities with mayors below the threshold, these differences are substantially smaller and no longer statistically significant when we move close to the threshold. 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 be effective.



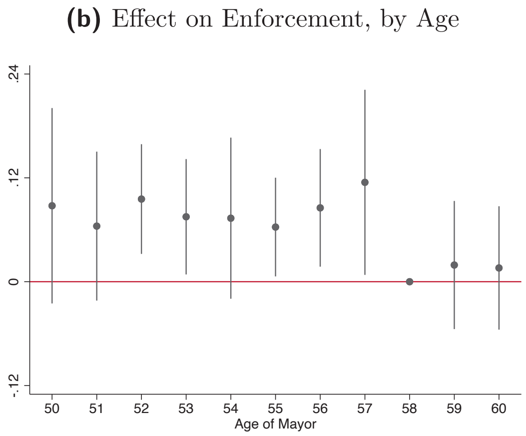


Figure 6. Main Results by Performance Incentives

Notes: This figure displays the effects of an additional monitor on both enforcement and pollution by mayors' age at the time of the NPC. Reported coefficients are relative to the effect for mayors who would be 58 years old at the time of the NPC (i.e. the baseline). Figure 6a reports estimates for pollution and Figure 6b reports estimates for enforcement. Error spikes represent 95 percent confidence intervals.

Since the information from the monitors is also provided to the public online, an alternative explanation for our main results is that monitors strengthen bottom-up accountability. To investigate whether this is a plausible driver of our results, we estimate the impact of the number of monitors on online search behaviour using Equation 4. We report results in Table C17, which show that while the point estimate on all 5 keywords are positive, they are small and none of them is significantly different from zero. Hence, we do not find any evidence suggesting that the number of monitors improve information acquisition by the local population. This test compares cities with more or less comprehensive monitoring and not cities with and without monitoring. The result does therefore not rule out that citizens with access to monitors may exert more pressure on local leaders than citizens without access to any monitoring. However, the impact of more comprehensive monitoring on enforcement and pollution that we document do not seem to be driven by increased dissemination of information to the public.

5.2 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., 2022). 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 enforcing agent).

Several media sources have reported on manipulation of the pollution data from the monitors by local government officials.[37] Following this reporting, the central government decided to 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. Hence, we are not able to exploit any cross-sectional 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).[38] More specifically, we estimate:

$$log(PM_{2.5})_{mt} = \delta_m + \gamma_t + \beta_1 AOD_{mt} + \beta_2 AOD_{mt} \times I_t + \epsilon_{mt}, \tag{5}$$

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.[39] 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 5 are reported in Table 4. As a point of reference,

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

[38]We focus on $PM_{2.5}$ because this is the pollutant most strongly correlated with AOD (see Table C4). Martinez (2018) studies the manipulation of GDP data by autocratic leaders using a similar specification.

[39]To deal with the fact that data is sometimes missing for the pixel just above the monitor, due to cloud coverage, we use the value from the closest neighbouring cell as long as this is within 20km from the monitor. All results are robust to using data at the city level instead.

we start by estimating the elasticity for all monitors without any interaction term (this replicates the results in the first column of Table C4). We then restrict our analysis to the incentivized monitors used in our study and find a positive estimate for the interaction term. This shows that the elasticity is 0.10 larger after the third party takes over the monitoring stations (corresponding to a 38% 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 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 described in Section 2.2. 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 cautious against drawing 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.

The next exercise we carry out is to check whether local governments exert more effort to decrease pollution after monitors have been reassigned, since manipulation is then a less viable option. The results are reported in columns (6) and (7) of the Table 4 and show that effects are indeed stronger after monitors have been reassigned. Column (6) shows a 1.4 percentage points greater reduction in pollution and column (7) a 6.5 percentage points larger increase in enforcement per monitor 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

Table 4. Monitor Reassignment, Data Quality and Policy Impact

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample:	All	Incentivized		Background			
Outcome:	log($PM_{2.5}$)					AOD	log(# firms)
AOD	0.30***	0.30***	0.27***	0.39***	0.37***		
	(0.031)	(0.033)	(0.038)	(0.050)	(0.057)		
AOD × Reassigned			0.10**		0.050		
			(0.047)		(0.11)		
# Monitors						-0.025***	0.12***
						(0.0070)	(0.045)
# Monitors × Reassigned						-0.014***	0.065**
						(0.0030)	(0.028)
Mean Outcome	3.68	3.71	3.71	3.42	3.42	0.34	0.58
Observations	17535	15496	15496	2039	2039	16322	5646

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 city fixed effects, time by pollution reduction target fixed effect, population and the geographical size of the built-up area at baseline interacted with the post variable, and time varying controls for total precipitation, average temperature and the age of the mayor. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

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 Discussion and Concluding Remarks

This study uses the introduction of a pollution monitoring program in China to investigate the impact 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. This resulted in an increase in enforcement in highly polluted areas in the city, which is where the monitors were placed. We further document that monitoring affect enforcement by altering which firms that are targeted by local governments and by strengthening the responsiveness to local pollution shocks.

To study the aggregate impact of the policy, we conduct a city-level analysis and compare enforcement and pollution levels in cities that installed different numbers of monitors and thus introduced differential coverage of local pollution activity. Our baseline analysis shows that one additional monitor (a 30% increase in coverage of high-pollution activity) leads to about a 15% increase in the number of firms that face regulatory enforcement and a subsequent 3% reduction in city-level pollution. The increase in enforcement is focused on the areas close to the monitor, while pollution is reduced across the city. 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 0.42–0.86 $\mu g/m^3$ reduction in average $PM_{2.5}$ per additional monitor.[40] Previous literature suggests that such a decrease in pollution could have significant health and economic benefits, which would likely exceed the cost of the program in the short run.[41]

An examination of possible mechanisms suggests that the monitoring program is effective because it enables the central government to hold local government officials accountable for their actions. We support this claim by showing that monitoring is substantially more effective in localities where local officials face performance incentives. 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 on both enforcement of regulations and the level of pollution is significantly 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.

[40] We arrive at the estimate of 0.42 (0.86) $\mu g/m^3$, the lower (upper) bound, as follows. We multiply 3.1% (4.6%) from Table 3 by 0.30 (0.36 – the elasticity with truthful reporting) from Table 4 to obtain percentage changes in $PM_{2.5}$ per 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 change in $PM_{2.5}$.

[41] 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 air pollution are also substantial – Panle Jia Barwick, Shanjun Li, Liguo Lin and Eric Zou (2020), e.g., 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. There are two main costs to consider for monitors: the cost of equipment and operation. According to the government procurement website, the cost of equipment/monitor is $200,000–$ 400,000, while yearly operation is $ 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.

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

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.

Baidu. 2017. "Search Index." https://index.baidu.com/v2/index.html#/.

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

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

Brehm, Stefan, and Jesper Svensson. 2020. "Environmental governance with Chinese characteristics: are environmental model cities a good example for other municipalities?" *Asia-Pacific Journal of Regional Science*, 4(1): 111–134.

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, Luke Keele, and Rocio Titiunik. 2021. "Covariate Adjustment in Regression Discontinuity Designs." *Handbook of Matching and Weighting in Causal Inference*, forthcoming.

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, 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." Center for International Private Enterprise.

CMA. 2017. "Meteorological Data." *China Meteorological Administration,* *https://data.cma.cn/index.html*.

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

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. 2013. "Do labor market policies have displacement effects? Evidence from a clustered randomized experiment." *The quarterly journal of economics*, 128(2): 531–580.

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.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 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.

ESRI. 2017. "China Prefecture City Boundaries." *Environmental Systems Research Institute,* *https://www.arcgis.com/home/item.html?id=4220d0d28777c46789bfe6baf344cee1f*.

Fetzer, Thiemo. 2020. "Can workfare programs moderate conflict? Evidence from India." *Journal of the European Economic Association*, 18(6): 3337–3375.

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.

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

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.

Freeman, Richard, Wenquan Liang, Ran Song, and Christopher Timmins. 2019. "Willingness to Pay for Clean Air in China." *Journal of Environmental Economics and Management*, 94: 188–216.

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.

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

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

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

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.

Hsiang, Solomon M. 2010. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." *Proceedings of the National Academy of sciences*, 107(35): 15367–15372.

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.

IPE. 2017. "Environmental Supervision Records." *The Institute of Public & Environmental Affairs*, https://www.ipe.org.cn/index.html.

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.

Jiang, Juanyan. 2017. "Chinese Political Elite Database." https://www.junyanjiang.com/data.html.

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

Jia, Ruixue. 2017. "Pollution for Promotion." *Working paper*.

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.

Lee, David S, and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281–355.

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.

MEP. 2011. "Model Cities." https://www.mee.gov.cn/gkml/hbb/bgt/201101/W020110125328042389677.pdf, Accessed: 2022-01-06.

MEP. 2013. "Air Ten." http://www.gov.cn/zwgk/2013-09/12/content_2486773.htm, Accessed: 2022-01-06.

MEP. 2014. "Monitor Program - wave 3." https://www.mee.gov.cn/gkml/hbb/bwj/201405/t20140509_273595.htm, Accessed: 2022-01-06.

MEP. 2017. "Mornitor Reading." http://106.37.208.233:20035, Accessed: 2022-01-06.

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

National Bureau of Statistics. 2013. "Annual Survey on Industrial Firms (ASIF)." *Ruisi Company [distributor]* http://www.resset.cn/ied, Accessed: 2017-09-01.

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.

NEO. 2017. "Aerosol Optical Thickness." *NASA Earth Observations*, https://neo.gsfc.nasa.gov/view.php?datasetId=MYDAL2_M_AER_OD.

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.

Reinikka, Ritva, and Jakob Svensson. 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." 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).

Wang, Shaoda, and David Y. Yang. 2022. "Policy Experimentation in China: The Political Economy of Policy Learning." *Working paper.*

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

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

Zheng, Siqi, and Matthew E Kahn. 2017. "A New Era of Pollution Progress in Urban China?" *Journal of Economic Perspectives*, 31(1): 71–92.

Zou, Eric Yongchen. 2021. "Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality." *American Economic Review*, 111(7): 2101–26.

Online Appendix for *Informed Enforcement: Lessons from Pollution Monitoring in China*

by Sebastian Axbard and Zichen Deng

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.

Data Collection and Validation

We rely on enforcement records collected by The Institute of Public & Environmental Affairs (IPE) from local environmental bureaus in China. 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.

Environmental bureaus are mandated by law to publicise all enforcement actions since 2008 (two years before our sample period starts).[42] IPE have compiled records from environmental bureaus at all levels of government using several different sources.[43] To validate the IPE data, we have conducted a manual validation using information that we have collected directly from local environmental bureaus. To perform this validation, we randomly select 1000 firms from our baseline sample (consisting of all firms in the Annual Survey of Industrial Firms in the cities that we study). We focus on enforcement records issued between 2015 and 2017, as bureaus are only required to keep records for 5 years. Our team manually went through all relevant websites of local environmental bureaus as well as their social media accounts. Using this approach, we were not able to identify a single enforcement action that was not already captured in the IPE data. We ended up classifying 957 (year 2015), 979 (year 2016), 992 (year 2017), firms in the same way as IPE (year 2015: 41 with any air

[42]Specified in the regulations for disclosure of environmental information, adopted at the first executive meeting of the State Environmental Protection Administration in 2007, available on this website.

[43]IPE collect information directly released by environmental bureaus in provinces, prefecture-level cities and counties. They also compile information communicated by government bodies in Chinese media and crawl official government Weibo accounts.

pollution enforcement, 916 without any enforcement; year 2016: 49 with any air pollution enforcement, 930 without any enforcement; year 2017: 88 with any air pollution enforcement, 904 without any enforcement). For the remaining 43 (year 2015), 21 (year 2016), 8 (year 2017) firms, the IPE identifies air pollution enforcement records that we are not able to identify manually. This could be due to the fact that the IPE cover a wider range of sources than we are able to check manually or because records had been removed from government websites at the time for our manually check in 2022. The fact that we are primarily missing records from years for which the archival requirement had passed at the time of our manual check suggests that the latter explanation may play an important role.

Encoding of Records

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[44]; 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:[45]
   - 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:
   - keywords for suspension: 停;

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

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

- 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 monitors for the first time in 2015. The majority of the remaining cities had some type of pollution monitoring before the reform and were simultaneously targeted by other policies as discussed in Section 2. In Table A1 in the appendix we compare the descriptive statistics of our sample with the average across all cities in China. We see that our cities are small by Chinese standards, with the urban population and the size of the built-up area being close to one third of the Chinese average. While the level of pollution in the cities that we focus on (as measured by AOD) is also lower in our main sample, it is closer to the average city AOD in China.

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

1

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

2

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

Table A1. Summary Statistics

	Our Sample	All Cities
AOD	0.333	0.394
	(0.177)	(0.191)
# Monitors	2.751	4.056
	(1.085)	(2.405)
Size of Built-up Area (km2)	44.82	125.0
	(27.64)	(229.0)
Urban Population (10,000)	33.92	91.49
	(22.03)	(191.9)
	177	338

Notes: Author's tabulations.

A.3 Additional Data

Local Leader Characteristics (Jiang, 2017) 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 cities, and 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 2017.

Baidu Search Index (Baidu, 2017) 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.[46]

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

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 (CMA, 2017) from the China Meteorological Administration. The data combines observations from 496 weather stations across China. We match this data to our prefecture-level cities to get a local measure of weather conditions.

Wind Direction To investigate whether firms upwind from a monitor face differential enforcement, we collected information about the dominant quarterly wind direction in each city. This data (CMA, 2017) is from the China Meteorological Administration and is based on readings from 496 weather stations across China. We calculate the angle between the locations of the firm and the quarterly prevailing direction of the wind vector passing through the closest monitor. As illustrated in Figure D4, a firm is defined as upwind of the closest monitor if the firms is within 45 degrees of the vector.

Appendix B Discontinuity Specifications

B.1 Regression Discontinuity

To explicitly consider the potential confounding effects of city size, we explore discontinuities in the number of monitoring stations assigned by the central government. We pool all observations post the introduction of monitors and rely on the local linear approach to estimate the following equation within the optimal bandwidth suggested by (Calonico, Cattaneo and Titiunik, 2014):

$$y_{cgt} = \gamma_g + \alpha r_c + a_c(\beta_0 + \beta_1 r_c) + \lambda X_c + \xi_{cgt} \tag{6}$$

where r_c is the value of the running variable for city c, which is the distance in sq km to the closest geographical size cutoff g listed in Table C2. The variable γ_g is a threshold fixed effect and a_c is an indicator variable for cities being above their closest cutoff. To improve precision, we follow Cattaneo, Keele and Titiunik (2021) and control for baseline characteristics indicated by X_c above. We include a control for average AOD in 2010-2011 in the pollution specification and for the 2010-2011 number of firms facing any enforcement related to air pollution for the enforcement specification. The coefficient of interest is therefore β_0, which captures the reduced form effect of being assigned to a group with a larger number of monitors. Standard errors are clustered at the city level.

To make the RD estimates comparable with the DiD/DiD+IV estimates, we normalize the estimates to the effect of one additional monitor by dividing β_0 by the first-stage RD estimates.[47] Our baseline estimates are reported in Column (3) of Table 3.

B.2 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).[48] This design essentially combines a difference-in-differences (comparing the outcomes in cities with a different number of monitors, before and after 2015) with a regression discontinuity design (comparing the outcomes 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

[47] This is essentially a fuzzy regression discontinuity design, and the estimates are implemented following Calonico, Cattaneo and Titiunik (2014).

[48] 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).

the following equation within the optimal bandwidth suggested by Calonico, Cattaneo and Titiunik (2014) and using data for all time periods:

$$y_{cgt} = \gamma_g + \mu_t + \alpha r_c + a_c(\beta_0 + \beta_1 r_c) + Post_t \times [\delta r_c + a_c(\theta_0 + \theta_1 r_c)] + \xi_{cgt}, \tag{7}$$

where $Post_t$ is an indicator for the period after 2015 and μ_t represent time fixed effects. All other variables are the same as in Equation 6. Standard errors are clustered at the city level. Treatment is captured by $Post_t \times a_c$ and the coefficient of interest is therefore θ_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 θ_0 by the first-stage RD estimates. Results of the Diff-in-Disc regressions are shown in Column (4) of the Table 3.

Appendix C Additional Tables

Table C1. Summary Statistics

	Mean	Std. dev.	Obs.	Periods	Freq.
Panel A: Firm-Level Data					
Any Air Pollution Enforcement	0.0046	0.068	1155296	2010-2017	Quarterly
Suspension	0.0024	0.049	1155296	2010-2017	Quarterly
Fine	0.0022	0.047	1155296	2010-2017	Quarterly
Upgrading	0.0025	0.050	1155296	2010-2017	Quarterly
Warning	0.00070	0.027	1155296	2010-2017	Quarterly
# Air Pollution Enforcement	0.0051	0.082	1155296	2010-2017	Quarterly
Any Water Pollu. Enforc.	0.0029	0.054	1155296	2010-2017	Quarterly
Any Solid Waste Pollu. Enforc.	0.00094	0.031	1155296	2010-2017	Quarterly
Any Procedure Pollu. Enforc.	0.0052	0.072	1155296	2010-2017	Quarterly
Upwind	0.25	0.44	1155296	2010-2017	Quarterly
Monitor within 10 km	0.40	0.49	36103	2013	Cross Sec.
Distance to Monitor (km)	19.2	15.4	36103	2013	Cross Sec.
Year Started	2003	7.92	36103	2013	Cross Sec.
Owner: SOEs	0.100	0.30	36103	2013	Cross Sec.
Owner: Private	0.81	0.39	36103	2013	Cross Sec.
Owner: Foreign	0.041	0.20	36103	2013	Cross Sec.
Owner: Other	0.048	0.21	36103	2013	Cross Sec.
Employment	434.8	1076.5	36103	2013	Cross Sec.
Revenue	278736.4	1656898.7	36103	2013	Cross Sec.
Panel B: City-Level Data					
# Monitors	2.76	1.09	16335	2010-2017	Monthly
Size of Built-up Area (km2)	44.8	27.3	16335	2010-2017	Monthly
Urban Population (10,000)	33.9	22.0	16335	2010-2017	Monthly
Age of the Mayor	50.7	3.46	16335	2010-2017	Monthly
Precipitation (mm)	77.0	93.2	16335	2010-2017	Monthly
Mean Temperature	13.8	10.3	16335	2010-2017	Monthly
Aerosol Optical Depth	0.34	0.23	16335	2010-2017	Monthly
# Firms Any Air Pollu. Enfor.	1.53	3.23	5664	2010-2017	Quarterly
# Firms Any Air Pollu. Enfor. (incl non-ASIF)	4.18	10.9	5664	2010-2017	Quarterly
Search Index: air pollution	2.01	4.24	14610	2011-2017	Monthly
Search Index: haze/smoke	18.2	28.4	14610	2011-2017	Monthly
Search Index: PM25	0.22	1.90	14610	2011-2017	Monthly
Search Index: air mask	5.97	9.36	14610	2011-2017	Monthly
Search Index: air purifier	23.4	26.5	14610	2011-2017	Monthly
Panel C: Monitor-Level Data					
Particulate Matter 2.5 ($PM_2.5$)	45.7	26.5	17535	2015-2017	Monthly
Particulate Matter 10 (PM_10))	81.1	51.4	17522	2015-2017	Monthly
Air Quality Index (AQI)	72.4	32.6	17541	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 firm-level data. 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 the city-level analysis. Panel C reports the summary statistics for the monitor-level data, which is monthly averages of the real-time readings from the monitors.

Table C2. Monitor Assignment Criteria

Group	Population (10,000)	Size of Built-Up Area (sq. km)	Min # Monitors	# Cities
1	< 25	< 20	1	26
2	$25 - 50$	$20 - 50$	2	86
3	$50 - 100$	$50 - 100$	4	57
4	$100 - 200$	$100 - 200$	6	8

Notes: Author's tabulations. **Source**: 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 C3. 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%	Gansu, 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 C4. Validating Satellite Data

	(1)	(2)	(3)
Outcome:	log($PM_{2.5}$)	log(PM_{10})	log(AQI)
AOD	0.30***	0.26***	0.20***
	(0.031)	(0.031)	(0.023)
Mean Outcome	3.68	4.26	4.20
Observations	17535	17522	17535

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 C5. 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	855	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	11.02
Non-Metallic Mineral Products Manufacturing	31	1449	4.01
Iron and Steel Smelting	32	450	1.25
Non-Ferrous Metal Smelting	33	1224	3.39
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	3.98
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		36103	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 C6. Rainfall Shocks and Monitor Recordings

	(1)	(2)	(3)	(4)
Outcome:				Share of Days
	log($PM_{2.5}$)	log(PM_{10})	log(AQI)	AQI>200
$Rain_{>\tilde{x}}$	-0.091***	-0.091***	-0.078***	-0.024***
	(0.018)	(0.015)	(0.012)	(0.0061)
Mean Outcome	3.63	4.24	4.16	0.11
Observations	2099	2099	2099	2099

Notes: This table reports the effect of precipitation shocks on monitor recordings of pollution. $Rain_{>\tilde{x}}$ is an indicator variable identifying time periods when precipitation is above the median rainfall in a city during the main sample period. We document the impact on four monitor-based measures of air pollution: $PM_{2.5}$, PM_{10}, the combined air quality index (AQI), and the share of days when the monitor reaches an air quality index that is above the critical value for heavily polluted (200). All regressions control for city fixed effects, time fixed effects, and average temperature. Robust standard errors clustered on the city are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

Table C7. Firm-level Robustness: Other Enforcement Actions and Additional Controls

	(1)	(2)	(3)	(4)
Panel A: Different Enforcement Actions				
Outcome	Air	Water	Solid Waste	Procedure
$\text{Mon}_{<10km} \times \text{Post}$	0.0033***	0.00055	0.00026	0.00082
	(0.00056)	(0.00041)	(0.00025)	(0.00066)
Mean Outcome	0.0046	0.0029	0.00094	0.0052
Observations	1155296	1155296	1155296	1155296
Firm FE	Yes	Yes	Yes	Yes
Industry-Time FE	Yes	Yes	Yes	Yes
Province-Time FE	Yes	Yes	Yes	Yes
Panel B: Additional Controls				
Outcome	Any Air Pollution Related Enforcement			
$\text{Mon}_{<10km} \times \text{Post}$	0.0033***	0.0034***	0.0032***	0.0031***
	(0.00056)	(0.00056)	(0.00057)	(0.00060)
Mean Outcome	0.0046	0.0046	0.0046	0.0046
Observations	1155296	1155296	1155296	1155296
Distance to coast-Time FE	No	Yes	Yes	No
Firm characteristics-Time FE	No	No	Yes	Yes
City-Time FE	No	No	No	Yes
Firm FE	Yes	Yes	Yes	Yes
Industry-Time FE	Yes	Yes	Yes	Yes
Province-Time FE	Yes	Yes	Yes	No

Notes: All regressions in both panels control for fixed effects specific to firm, industry-by-time interactions, and province-by-time interactions. Panel A reports results from estimating Equation 2 on the probability of being subject to different types of environmental enforcement. Panel B reports additional sensitivity analysis, by adding additional controls to Equation 2. Column (1) reports the baseline estimate from Table 1 as a point of reference. Column (2) adds distance to coast by time fixed effects to the estimation equation. Column (3) further includes interactions between the number of employees and firm ownership status (6 categories) with time fixed effects. Column (4) introduces city by time fixed effects (this drops distance to coast and province by time fixed effects since these are collinear). Robust standard errors clustered on the city are reported in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level.

Table C8. Number of Monitors and Coverage of High Pollution Activity

	(1)	(2)	(3)	(4)
Outcome	Share of high polluters			
Distance	Within 10 km		Within 5 km	
Measure	Revenue	Employment	Revenue	Employment
Panel A: DiD Estimates				
# Monitors	0.11***	0.084***	0.097***	0.073***
	(0.023)	(0.024)	(0.023)	(0.024)
Mean Outcome	0.37	0.36	0.26	0.26
Observations	160	160	160	160
Panel B: DiD + IV Estimates				
# Monitors	0.13***	0.11***	0.13***	0.11***
	(0.034)	(0.031)	(0.032)	(0.031)
Mean Outcome	0.37	0.36	0.26	0.26
Observations	160	160	160	160

Notes: This tables shows the results from a regression of different measures of the share of high pollution activity that occurs close to a monitor on the number of monitors in the city. This analysis is limited to the 160 cities for which we have at least one high polluter according to the ESR database. Panel A reports results on the actual number of monitors, while Panel B reports results on the assigned number of monitors. Columns (1)/(3) shows the relationship between the number of monitors and the share of a city's high polluter's revenue that is within 10/5km from a monitor. Columns (2)/(4) shows the relationship between the number of monitors and the share of a city's high polluter's employment that is within 10/5km from a monitor. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

Table C9. City-level Robustness: Additional Controls

	(1)	(2)	(3)	(4)	(5)	(6)
	DiD	DiD+IV	DiD	DiD+IV	DiD	DiD+IV
Panel A: Outcome - Aerosol Optical Depth						
# Monitors	-0.031***	-0.046***	-0.031***	-0.044***	-0.037***	-0.049***
	(0.0069)	(0.013)	(0.0070)	(0.013)	(0.0065)	(0.013)
Observations	16335	16335	16335	16335	16335	16335
Panel B: Outcome - log(# firms receiving any air pollution enforcement)						
# Monitors	0.15***	0.19**	0.15***	0.19*	0.11**	0.17
	(0.046)	(0.098)	(0.046)	(0.099)	(0.050)	(0.11)
Observations	5664	5664	5664	5664	5664	5664
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Target-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
City size × Post	Yes	Yes	No	No	No	No
City size-Time FE	No	No	Yes	Yes	Yes	Yes
City char.-Time FE	No	No	No	No	Yes	Yes
Weather	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates from adding additional controls to our baseline city-level specification. Columns (1) and (2) report our baseline estimate from Table 3. Columns (3) and (4) report estimates from a slightly more demanding specification where we interact baseline city population and the geographical size of the built-up area with time fixed effects instead of the post variable. Columns (5) and (6) add interactions between baseline GDP as well as an indicator for whether a city installed a background monitor with time fixed effects. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

Table C10. City-level Robustness: Sample Restrictions

	(1)	(2)	(3)	(4)
	DiD	DiD+IV	RD	Diff-in-Disc
Panel A: Outcome - Aerosol Optical Depth				
# Monitors	-0.030***	-0.041***	-0.032**	-0.026
	(0.0069)	(0.013)	(0.016)	(0.019)
Observations	14646	14646	2853	7566
Panel B: Outcome - log(# firms receiving any air pollution enforcement)				
# Monitors	0.14***	0.16*	0.26**	0.23
	(0.047)	(0.097)	(0.11)	(0.16)
Observations	5056	5056	984	2624
Kernel			Uniform	Uniform
Bandwidth			11.3	11.3

Notes: This table reports the results from estimating our four baseline specifications using a restricted sample that excludes data from the provinces Xinjiang and Tibet, which cover much larger geographical areas than other cities. All controls are the same as in Table 3. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

Table C11. City-level Robustness: Including Non-ASIF Firms

	(1)	(2)	(3)	(4)
	DiD	DiD+IV	RD	Diff-in-Disc
Outcome	log(# firms receiving any air pollution enforcement)			
Panel A: All firms (including Non-ASIF)				
# Monitors	0.13***	0.25**	0.29**	0.37**
	(0.049)	(0.11)	(0.14)	(0.14)
Observations	5664	5664	1116	2976
Panel B: Only Non-ASIF firms				
# Monitors	0.13***	0.27**	0.31**	0.40***
	(0.049)	(0.11)	(0.15)	(0.13)
Observations	5664	5664	1116	2976

Notes: This table reports the results from estimating our four baseline specifications for two alternative enforcement definitions: Panel A includes all firms in a city (i.e. also those that are not in the ASIF database) and Panel B focuses only on enforcement against firms that are not covered in the ASIF database. All controls are the same as in Table 3. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

Table C12. City-level Robustness: RD Kernels and Covariates

	(1)	(2)	(3)	(4)
Panel A: Outcome - Aerosol Optical Depth				
# Monitors	-0.039***	-0.038**	-0.036**	-0.028
	(0.015)	(0.015)	(0.015)	(0.050)
Observations	3209	3735	3807	4224
Bandwidth	11.3	12.3	12.5	13.8
First stage	1.28***	1.31***	1.28***	1.11***
	(0.23)	(0.22)	(0.22)	(0.32)
Panel B: Outcome - log(# firms receiving any air pollu. enforce.)				
# Monitors	0.26**	0.29***	0.28***	0.24
	(0.10)	(0.10)	(0.10)	(0.16)
Observations	1116	1392	1296	1116
Bandwidth	11.3	13.1	12.4	11.4
First stage	1.28***	1.28***	1.28***	1.16***
	(0.23)	(0.21)	(0.22)	(0.33)
Kernel	Uniform	Epanechnikov	Triangle	Uniform
Covariates	Yes	Yes	Yes	No

Notes: This table reports additional regression discontinuity results. Columns (1)-(3) report baseline estimates, controlling for cutoff fixed effects and baseline (2010-2011) AOD/log(# firms), using different kernel weighting methods. Column (4) reports results from our baseline specification, but without any controls. The discontinuities are estimated using local linear regressions and the MSE-optimal bandwidth proposed by Calonico, Cattaneo and Titiunik (2014) for respective kernel weighting method. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

Table C13. RD Estimates by Cutoff

	(1)	(2)	(3)	(4)
Outcome	AOD		log(# firms...)	
Method	RD	Diff-in-Disc	RD	Diff-in-Disc
Panel A: Cutoff 1				
# Monitors	-0.041**	-0.019	0.28	0.14
	(0.021)	(0.029)	(0.21)	(0.21)
Observations	1508	3992	528	1408
Bandwidth	11.3	11.3	11.3	11.3
First stage	0.87***	0.87***	0.87***	0.87***
	(0.27)	(0.27)	(0.27)	(0.27)
Panel B: Cutoff 2				
# Monitors	-0.034*	-0.038*	0.29**	0.19
	(0.018)	(0.021)	(0.12)	(0.13)
Observations	1701	4516	588	1568
Bandwidth	11.3	11.3	11.3	11.3
First stage	1.79***	1.79***	1.79***	1.79***
	(0.34)	(0.34)	(0.34)	(0.34)
Kernel	Uniform	Uniform	Uniform	Uniform
Bandwidth	11.3	11.3	11.3	11.3

Notes: This table reports regression discontinuity and difference in discontinuity results separately by threshold. Panel A reports estimates for geographical size cutoff 1 (20 sq. km) and Panel B reports estimates for geographical size cutoff 2 (50 sq. km). The discontinuities are estimated using local linear regressions and the MSE-optimal bandwidth proposed by Calonico, Cattaneo and Titiunik (2014). The RD specification controls for cutoff fixed effects and baseline (2010) AOD/log(# firms), while the Diff-in-Disc control for cutoff and time fixed effects. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

Table C14. City-level: Direct Effects vs. Spillover

	(1)	(2)	(3)	(4)
	DiD	DiD+IV	RD	Diff-in-Disc
Panel A: AOD, Monitor ($\leq$ 10km)				
# Monitors	-0.032***	-0.038**	-0.040**	-0.051**
	(0.0068)	(0.015)	(0.019)	(0.021)
Observations	14180	14180	2680	7115
Mean Outcome	0.39	0.39	0.33	0.33
Panel B: AOD, City Center (10-50km)				
# Monitors	-0.032***	-0.044***	-0.031**	-0.037*
	(0.0069)	(0.014)	(0.015)	(0.019)
Observations	14180	14180	2680	7115
Mean Outcome	0.35	0.35	0.31	0.31
Panel C: AOD, Surrounding Area (> 50km)				
# Monitors	-0.028***	-0.037***	-0.031**	-0.027
	(0.0068)	(0.013)	(0.014)	(0.017)
Observations	14180	14180	2680	7115
Mean Outcome	0.32	0.32	0.28	0.28
Panel D: Enforcement, Monitor ($\leq$ 10km)				
# Monitors	0.15***	0.22**	0.21***	0.30***
	(0.034)	(0.088)	(0.071)	(0.098)
Observations	5664	5664	1116	2976
Mean Outcome	0.26	0.26	0.23	0.23
Panel E: Enforcement, City Center (10-50km)				
# Monitors	0.068*	0.039	0.12	0.20
	(0.037)	(0.086)	(0.094)	(0.12)
Observations	5664	5664	1116	2976
Mean Outcome	0.26	0.26	0.22	0.22
Panel F: Enforcement, Surrounding Area (> 50km)				
# Monitors	0.0065	0.0021	0.12	0.020
	(0.033)	(0.053)	(0.089)	(0.083)
Observations	5664	5664	1116	2976
Mean Outcome	0.21	0.21	0.20	0.20
Kernel			Uniform	Uniform
Bandwidth			11.3	11.3

Notes: This table reports results for our main outcomes calculated separately for: the monitoring station (outcomes observed within 10km from a monitor, panels A and D), the city centre (outcomes observed 10-50km from a monitor, panels B and E) and the surrounding areas (outcomes observed beyond 50km from a monitor, panels C and F). Estimates from the four different empirical strategies used in the city-level analysis are reported. Panels A-C report results for aerosol optical depth and panels D-F for the log number of firms receiving any enforcement action related to air pollution. To ensure that estimates are comparable across the first three panels, we restrict the AOD analysis to cities for which we can consistently observe AOD across the three outcomes. The specifications used are the same as those reported in Table 3. Robust standard errors clustered on the city in parentheses. *, **, *** indicate significance at the 10%, 5% and 1% level respectively.

Table C15. Mechanism: Promotion Incentives

	(1)	(2)	(3)	(4)
Age bandwidth:	Full	±7 Years	±5 Years	±3 Years
Panel A: Outcome - Aerosol Optical Depth				
# Monitors	-0.020***	-0.022***	-0.021***	-0.026***
	(0.0069)	(0.0072)	(0.0075)	(0.0081)
# Monitors × Below 58	-0.015***	-0.015***	-0.014***	-0.017***
	(0.0046)	(0.0046)	(0.0048)	(0.0050)
Mean Outcome	0.34	0.33	0.32	0.32
Observations	16335	13805	12048	8835
Panel B: Outcome- log(# firms receiving any air pollution enforcement)				
# Monitors	0.089**	0.088**	0.086*	0.078
	(0.041)	(0.043)	(0.048)	(0.057)
# Monitors × Below 58	0.077***	0.067***	0.067***	0.060***
	(0.021)	(0.021)	(0.022)	(0.022)
Mean Outcome	0.58	0.55	0.55	0.55
Observations	5664	4800	4192	3072

Notes: This table reports heterogeneous effects of monitoring by promotion incentives on aerosol optical depth (Panel A) and the log number of firms receiving any enforcement action related to air pollution (Panel B). Each column reports the estimate from Equation (4) with an additional interaction for mayors being below 58 years at the time of the National Peoples' Congress. All specifications control for city fixed effects, time by pollution reduction target fixed effect, population and the geographical size of the built-up area at baseline interacted with the post variable, and time varying controls for total precipitation, average temperature and the age of the mayor. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

Table C16. Balance Table: Mayor's Age and City Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean			Difference			
	Full	58+	57-	Full	±7 Years	±5 Years	±3 Years
# Monitors	2.75	2.48	2.80	0.32	0.30	0.27	0.0033
	(1.08)	(1.33)	(1.03)	(0.22)	(0.22)	(0.23)	(0.25)
Size of buildup area	44.8	35.6	46.6	11.0**	10.2*	10.8*	6.31
	(27.6)	(20.1)	(28.6)	(5.57)	(5.55)	(5.82)	(5.86)
Urban population	33.9	28.9	34.9	6.06	5.86	6.44	2.53
	(22.0)	(18.8)	(22.5)	(4.46)	(4.47)	(4.65)	(4.56)
AOD before 2015	0.36	0.29	0.38	0.084**	0.076*	0.067*	0.064
	(0.20)	(0.17)	(0.20)	(0.040)	(0.040)	(0.040)	(0.045)
Night light before 2015	-1.17	-1.16	-1.17	-0.0045	-0.036	-0.022	-0.048
	(0.73)	(0.91)	(0.70)	(0.15)	(0.15)	(0.16)	(0.16)
log(# Firms) before 2015	0.35	0.33	0.36	0.026	0.0032	0.0097	-0.0100
	(0.27)	(0.28)	(0.26)	(0.054)	(0.052)	(0.053)	(0.058)
Observations	177	29	148	177	150	131	96
Joint Test (p-value)				0.19	0.15	0.29	0.37

Notes: This table reports the balance of baseline characteristics for cities with mayors of different age at the time of the National People's Congress. Column (1) reports averages for the full sample, while columns (2) and (3) split the sample into cities with mayors above and below the age cutoff. Columns (4)-(7) report differences between cities above and below the threshold for different bandwidths ranging from the full sample to cities with mayors 3 years above to 3 years below the threshold. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

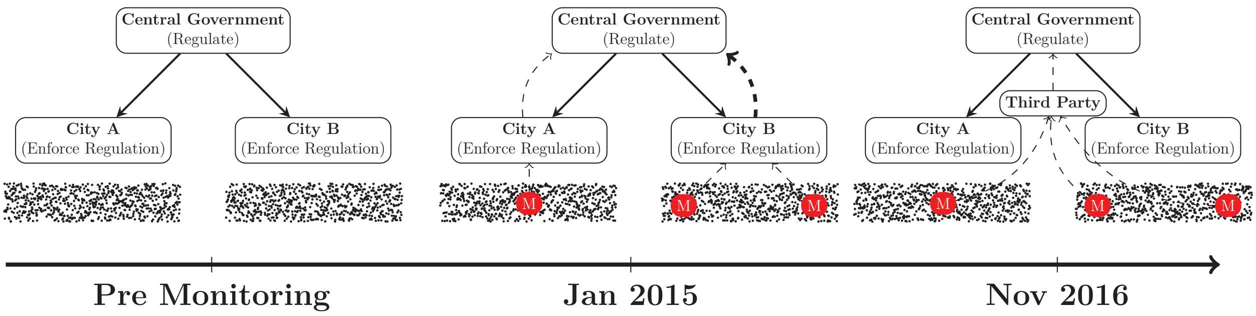
Table C17. Mechanism: Monitors and Online Searches

	(1)	(2)	(3)	(4)	(5)
Outcome:	log(key word)				
Key words:	air pollution	haze/smog	$PM_{2.5}$	air mask	air purifier
# Monitors	0.0097	0.011	0.0011	0.022	0.0064
	(0.0064)	(0.028)	(0.0016)	(0.018)	(0.029)
Mean Outcome	0.049	0.33	0.0052	0.13	0.43
Observations	14596	14596	14596	14596	14596

Notes: This table reports estimates from Equation (4) on city-level outcomes for online searches for pollution related keywords. All specifications control for city fixed effects, time by pollution reduction target fixed effect, population and the geographical size of the built-up area at baseline interacted with the post variable, and time varying controls for total precipitation, average temperature and the age of the mayor. Each column is from a separate regression estimating the impact on a specific keyword. Robust standard errors clustered on the city in parenthesis. *, **, *** indicates significance at the 10%, 5% and 1% level respectively.

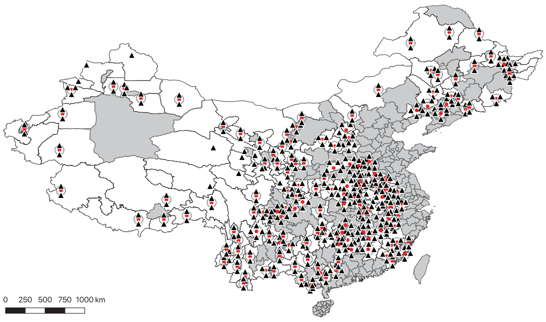
Appendix D Additional Figures

Figure D1. Monitors, Coverage and Flow of Information

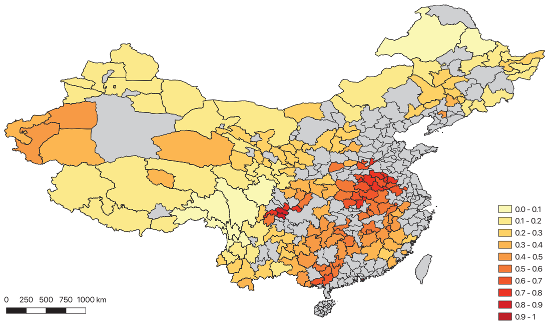


Notes: This figure describes how the flow of information changes with the introduction of monitors. While responsibilities are unchanged – the central government regulates and the local government enforces these regulations – the quality of information changes differently between cities. Starting in 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.

Figure D2. Geographical Distribution of Data

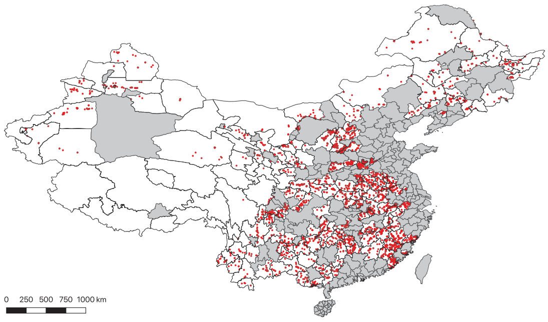


(a) Location of Pollution Monitors



(b) Distribution of AOD in 2010

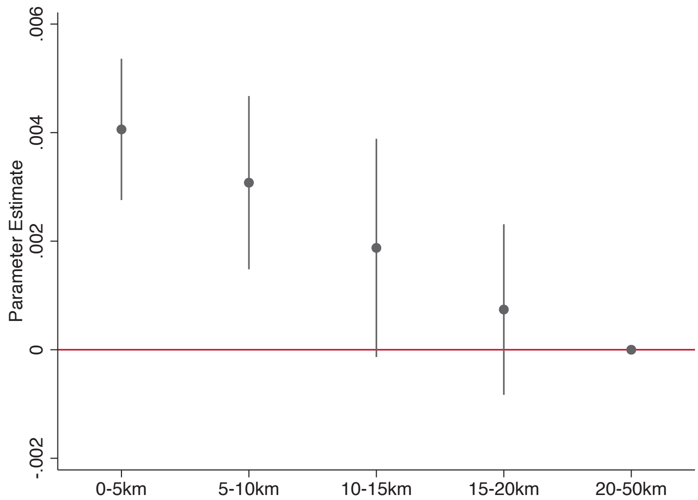
(c) Location of Firms



(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 D3. Firm-level: Enforcement Gradient

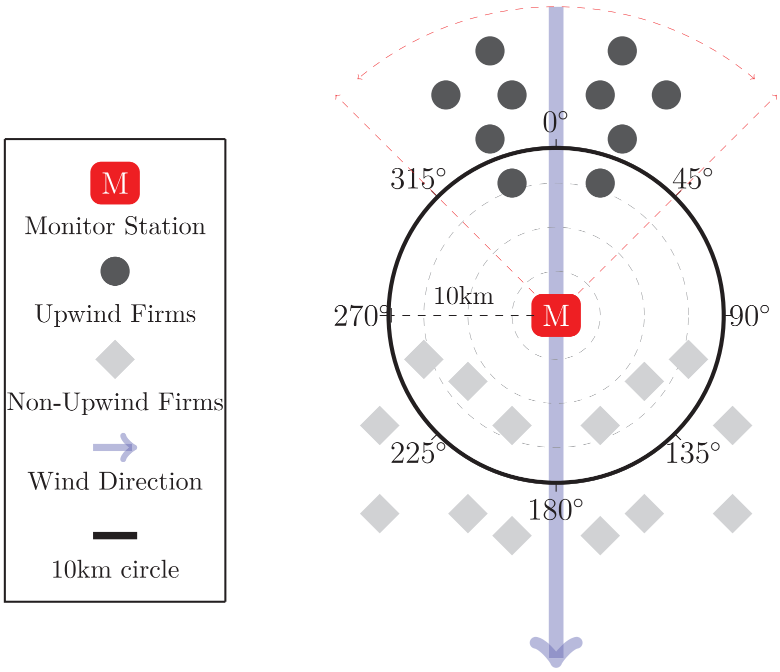


Notes: This figure shows the relative increase in enforcement for each distance bin after 2015. Error spikes represent 95 percent confidence intervals. Formally, we estimate the following equation:

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

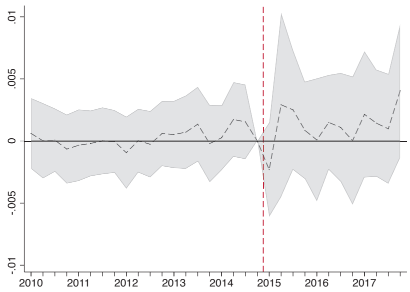
where m_{it}^d is an indicator for there being a monitor within distance d from firm i in quarter t; and all other variables are the same as in Equation 1. Hence, we are here estimating the average change in enforcement in the post-period relative to the pre-period.

Figure D4. Classifying Upwind Firms

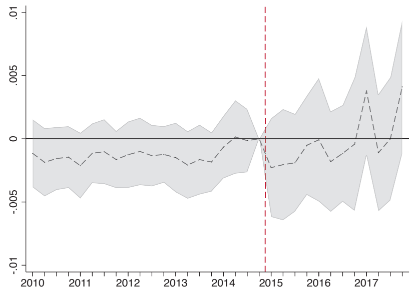


Notes: This figure illustrates our procedure for classifying whether a firm is upwind or not (i.e. whether the wind moves emissions towards the monitor or not). The thick blue arrow illustrates the dominant wind direction in a quarter. We follow previous work (Freeman et al., 2019) and define all firms that are within 45 degrees of the wind vector that passes through the monitor (i.e. the area confined by the dashed red lines) as upwind. Upwind firms are identified by black dots in the figure, while non-upwind firms are identified as grey diamonds. The 10km solid black circle illustrates the criteria used in the baseline specification to identify firms close to a monitor.

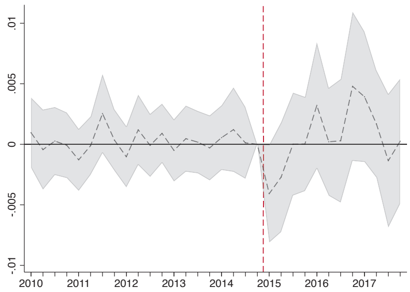
Figure D5. Firm-level: Placebo Nonparametric Event Study



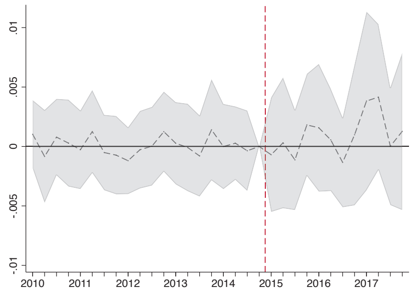
(a) 0–5 km (envir. bureau)



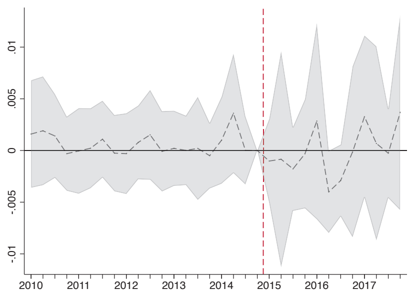
(b) 5–10 km (envir. bureau)



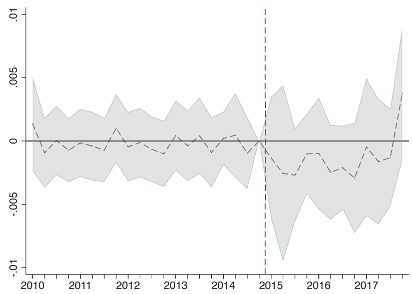
(c) 10–15 km (envir. bureau)



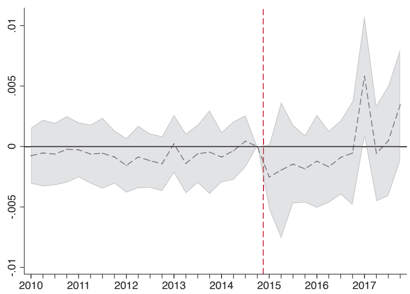
(d) 15–20 km (envir. bureau)



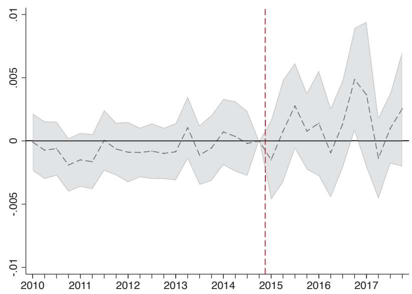
(e) 0–5 km (firm centroid)



(f) 5–10 km (firm centroid)



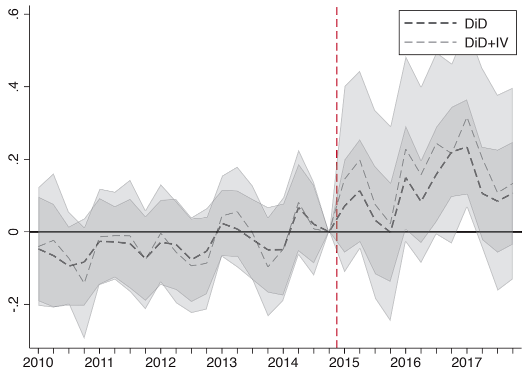
(g) 10–15 km (firm centroid)



(h) 15–20 km (firm centroid)

Notes: This figure shows the estimates of the nonparametric event study using Equation 1 for two placebo firm distances: kilometers to the local environmental bureau (figures a-d) or the kilometers to the city's firm centroid (figures e-h). The shaded area represents 95 percent confidence intervals calculated using robust standard errors clustered at the city level.

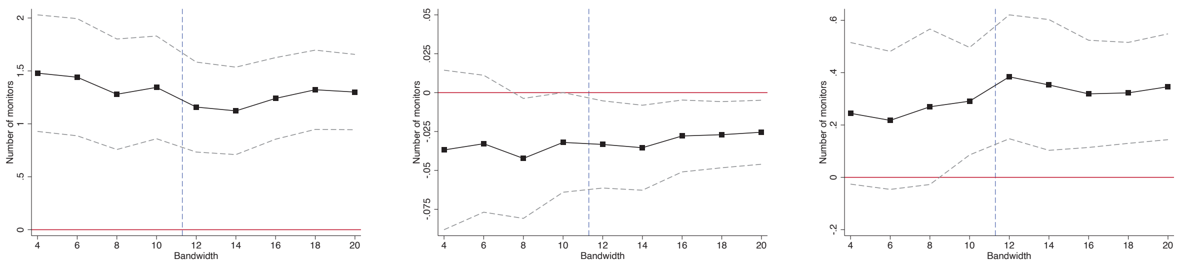
Figure D6. City-level: Enforcement Event Study



(a) DiD

Notes: This figure present the estimates from Equation 3 of city-level enforcement (log(# firms)) using two different specifications (DiD, DiD+IV). The shaded area represents 95 percent confidence intervals based on standard errors clustered on the city.

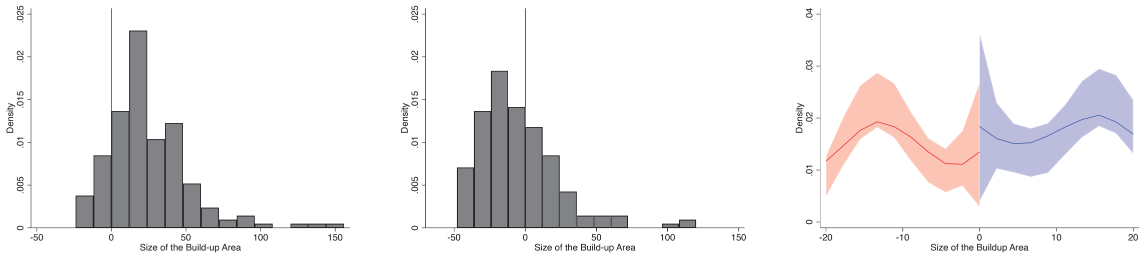
Figure D7. City-level: Alternative RD Bandwidths



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

Notes: These figures report the sensitivity of the RD coefficients to alternative bandwidths. The vertical axis shows the RD coefficients, while the horizontal axis shows the bandwidth used to estimate the respective coefficient. The blue dashed line marks the optimal bandwidth (11.3) using the approach suggested by Calonico, Cattaneo and Titiunik (2014).

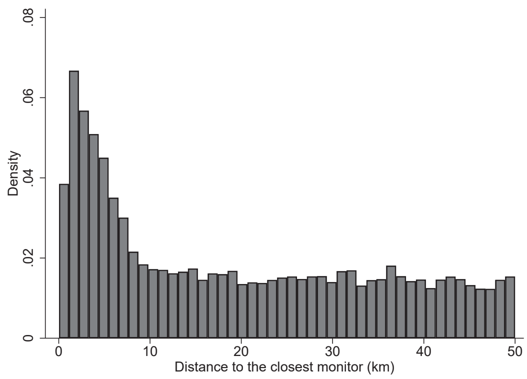
Figure D8. Histogram of Running Variables



(a) Cutoff 1: 20 (km^2) **(b)** Cutoff 2: 50 (km^2) **(c)** Estimated Density

Notes: The figures provide histograms and estimated densities of the size of the built-up area for our sample over the two cutoffs we use in the analysis. The p-value for the null hypothesis that the density of the size of the built-up area is continuous at the threshold is 0.642.

Figure D9. Distance to the Closest Monitor



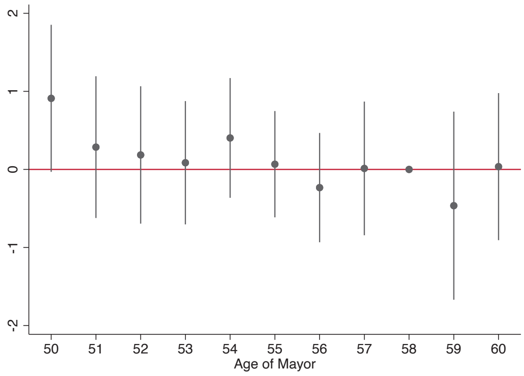
Notes: This figure shows the distribution of the distance between ASIF firms and the closest monitor. The sample is restricted to firms that are located within 50 km from a monitor.

Figure D10. Media Reporting on Enforcement Around Monitors

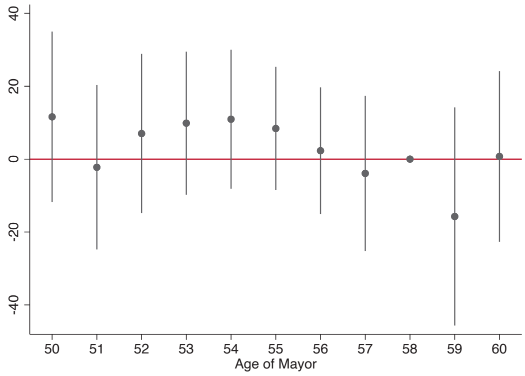
(a) Search Results in Chinese **(b)** Translation

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 step-up 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**: www.baidu.com

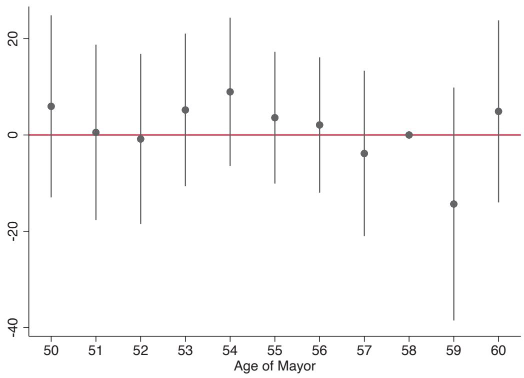
Figure D11. Balance Graphs: Mayor's Age and City Characteristics



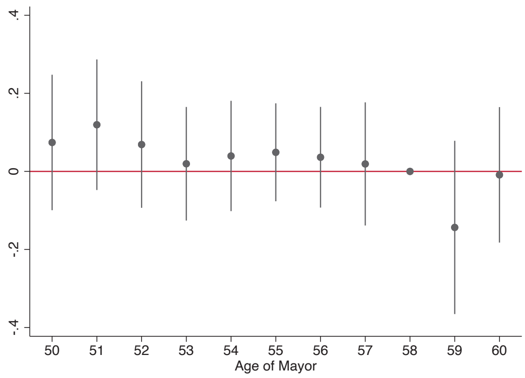
(a) Age vs # Monitors



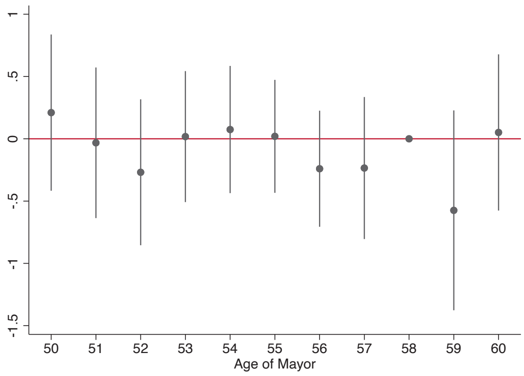
(b) Age vs Size of Buildup Area



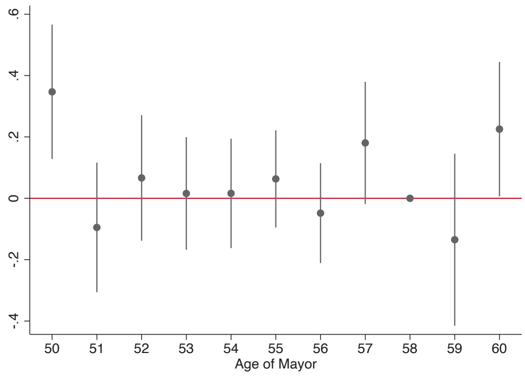
(c) Age vs Urban Population



(d) Age vs PM before 2015



(e) Age vs Night light before 2015



(f) Age vs Enforcement before 2015

Notes: These figures report the balance of cities' baseline characteristics by mayor's age at the time of the NPC, using the same approach as in Figure 6. Reported coefficients are relative to the effect for mayors who would be 58 years old at the time of the NPC. Error spikes represent 95 percent confidence intervals.

Issued in the series Discussion Papers 2020

2020

01/20 January, **Laura Khoury**, Clément Brébion and Simon Briole. "Entitled to Leave: the impact of Unenployment Insurance Eligibility on Employment Duration and Job Quality"

02/20 January, Thomas Buser, **Alexander Cappelen**, Uri Gneezy, Moshe Hoffman and **Bertil Tungodden**. "Competitiveness, gender and handedness: a large-sample intercultural study"

03/20 February, **Patrick Bennett**, Chiara Ravetti and Po Yin Wong. "Losing in a Boom: Long-term Consequences of a Local Economic Shock for Female Labour Market Outcomes"

04/20 April, **Øivind A. Nilsen**. "The Labor Market in Norway: 2000-2018"

05/20 April, **Simen A. Ulsaker**. "Exclusionary contracts and incentives to innovate"

06/20 May, **Alexander W. Cappelen**, **Ranveig Falch**, **Erik Ø. Sørensen** and **Bertil Tungodden**. "Solidarity and Fairness in Times of Crisis"

07/20 May, Gozde Corekcioglu, Marco Francesconi and **Astrid Kunze**. "Do Generous Parental Leave Policies Help Top Female Earners?"

08/20 June, **Ola Honningdal Grytten**. "Weber revisited: A literature review on the possible Link between Protestantism, Entrepreneurship and Economic Growth"

09/20 June, Eva M. Berger , Ernst Fehr, **Henning Hermes** , Daniel Schunk and Kirsten Winkel. "The Impact of Working Memory Training on Children's Cognitive and Noncognitive Skills"

10/20 June, **Ola Honningdal Grytten**. "Two centuries of economic growth: Norwegian GDP 1816-2020"

11/20 July, **Ola Honningdal Grytten**, Magnus Lindmark and Kjell Bjørn Minde. "Energy Intensity and the Environmental Kuznets Curve"

12/20 August, **Ola Honningdal Grytten**. “Puritan Motivation for Serial Entrepreneurship: The Haugean Example”

13/20 August, Julian Johnsen, Hyejin Ku and **Kjell G. Salvanes**. “Competition and Career Advancement: The Hidden Costs of Paid Leave”

14/20 August, Patrick Bennett, Richard Blundell and **Kjell G. Salvanes**. “A Second Chance? Labor Market Returns to Adult Education Using School Reforms”

15/20 August, Paul Brandily, Clément Brébion, Simon Briole and **Laura Khoury**. “A Poorly Understood Disease? The Unequal Distribution of Excess Mortality Due to COVID-19 Across French Municipalities”

16/20 September, **Ingvild Almås**, **Vincent Somville** and Lore Vandewalle. “The Effect of Gender-Targeted Transfers: Experimental Evidence From India”

17/20 September, **Ola Honningdal Grytten.** “The Wealth of a Nation: Norways Road to Prosperity”

18/20 September, Asbjørn G. Andersen, Simon Franklin, Tigabu Getahun, Andreas Kotsadam, **Vincent Somville** and Espen Villanger. “Does Wealth Reduce Support for Redistribution? Evidence from an Ethiopian Housing Lottery”

19/20 September, **Ingvild Almås**, Lars Ivar Berge, **Kjetil Bjorvatn**, **Vincent Somville** and **Bertil Tungodden**. “Adverse selection into competition: Evidence from a large-scale field experiment in Tanzania”

20/20 September, Julian Vedeler Johnsen, Kjell Vaage and **Alexander Willén**. “Interactions in Public Policies: Spousal Responses and Program Spillovers of Welfare Reforms”

21/20 October, **Aline Bütikofer**, Rita Ginja, **Fanny Landaud** and **Katrine Løken**. “School Selectivity, Peers, and Mental Health”

22/20 November, Barton Willage and **Alexander Willén.** “Postpartum Job Loss: Transitory Effect on Mothers, Long-run Damage to Children”

2021

01/21 January, Sebastian Axbard and **Zichen Deng**. "Informed Enforcement. Lessons from Pollution Monitoring in China"

NHH

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no