ESSAYS ON ENVIRONMENTAL CHALLENGES AND REGULATIONS IN CHINA

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Lin Yang

May 2021

ESSAYS ON ENVIRONMENTAL CHALLENGES AND REGULATIONS IN CHINA

Lin Yang, Ph.D.
Cornell University 2021

As one of the most rapidly growing economies, China has been experiencing pressing environmental and urban challenges due to a dramatic increase in fossil fuel consumption, and a lack of stringent and well-enforced environmental regulations. To address problems such as air pollution, traffic congestion, and weak enforcement, China has issued extensive environmental and transportation regulations. My dissertation aims to empirically estimate the causal effects of environmental policies and public infrastructures on environmental outcomes. The dissertation is comprised of three chapters. Chapter 1, joint with Shanjun Li, Yanyan Liu, and Avralt Od-Purejav, estimates the impact of subway expansions on air quality by leveraging fine-scale air quality data and the rapid build-out of 14 new subway lines in Beijing from 2008 to 2016. Chapter 2 is a review article, joint with Shanjun Li, Jianwei Xing, and Fan Zhang, which reviews findings in the recent literature on the impacts of a host of urban transportation policies used in developed- and developing-country settings. Finally, Chapter 3 studies the role of accurate measurements in effective regulations. Using high-resolution satellite-based pollution measures, this chapter examines local governments' strategic pollution control behavior and its implications on dynamic representativeness based on the staggered roll-out of the air pollution monitoring system in China.

BIOGRAPHICAL SKETCH

Lin Yang was born in Mengzhou, Henan Province, P.R. China, in 1992. During her undergraduate at China Agricultural University, she transferred to the University of Maryland in 2012 through a 2+2 Dual Degree Program. She received her bachelor's degrees from both the Department of Agricultural and Resource Economics at the University of Maryland with honors and the Department of Economics and Management at China Agricultural University. In 2014, Lin entered the Master of Science program at the Dyson School of Applied Economics and Management at Cornell University. In her two years of graduate study as a master's student, she developed a research interest in environmental economics. She then continued to pursue a doctoral degree in Applied Economics at Cornell University, under the supervision of Professor Shanjun Li, conducting cutting-edge research on the environment, transportation, and urban challenges in China. During the five years of her Ph.D. study, Lin and her coauthors published their research in the top journal of environmental economics. With a strong passion for research and teaching, Lin decided to pursue an academic career after obtaining her doctoral degree. She will join the Hong Kong University of Science and Technology (Guangzhou) as an assistant professor in Fall 2021.

This dissertation is dedicated to my parents.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Shanjun Li, for his unwavering patience and support throughout my graduate study. I learned so much from him, not only in professional knowledge but also about the important qualities a researcher should have. I would also like to thank my committee members, Professor Cathy Kling, Professor Ariel Ortiz-Bobea, and Professor Ivan Rudik, for their invaluable comments and advice in completing this dissertation and pursuing an academic career.

Next, big thanks go out to my amazing cohort, my peers in the SEERE group, and the lifelong friends I have made here at Cornell. They have inspired me every day and made my life in Ithaca more wonderful than I could ever imagine. I look forward to our paths crossing again in the future.

Finally, I would like to thank my family for always being cheerleaders in my life and my husband for his unshakable love and companionship regardless of the distance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

DOES SUBWAY EXPANSION IMPROVE AIR QUALITY?

1.1 Introduction

Traffic congestion and air pollution pose pressing urban challenges in many developing and emerging countries. Based on real-time driving data in 2016, TomTom Traffic Index shows that all but one of the top 20 most congested cities are from developing and emerging economies, and eight of them were located in China. Meanwhile, East and South Asian countries, such as Bangladesh, China, India, and the Persian Gulf experienced the highest level of $PM_{2.5}$ concentration in 2015. Ambient $PM_{2.5}$ is the leading environmental factor for death, accounting for about 4.2 million deaths in 2015, nearly 40 percent of which occurred in China (Global Burden of Disease 2015).

The Beijing municipal government has been investing heavily in transportation infrastructures, such as buses, roads, and subway lines to combat traffic congestion and air pollution in the city. From 2007 to 2015, the government's total investment in transportation infrastructure amounted to over 430 billion Yuan (about USD 67 billion). During this period, Beijing rolled out14 new subway lines with a total length of 440 kilometers. The city's rapid subway expansion is still ongoing: another 12 subway lines with a total length of nearly 378 kilometers are under construction and scheduled to open before the end of 2020. Similar large-scale expansion of subway systems is taking place in major cities throughout China.

Despite the massive investment in subway infrastructure in Beijing and other major cities in China, rigorous evaluation of the impacts of subway expansion is lacking. This paper investigates the impact of subway expansion on air quality by exploiting the rapid expansion of Beijing's subway system from 2008 to 2016. The expansion of the subway network can create two countervailing forces that could affect air quality. First, the improved subway coverage could lead some commuters to switch from traveling using private cars to using subways (Mohring 1972). This traffic diversion effect or "Mohring Effect", should relieve traffic congestion and thus reduce air pollution. Second, the improvement in traffic conditions could make driving more attractive and induce additional travel demand using private cars, resulting in a traffic creation effect (Vickrey 1969; Duranton and Turner 2011). The net effect of subway expansion on air quality is ultimately an empirical question.

Our empirical analysis leverages rich spatial and temporal variation in air quality and subway coverage across Beijing from 2008 to 2016. The data on air quality come from daily air quality readings from 27 monitors throughout the city. During the data period, 252 new subway stations opened (out of totally 345 subway stations in operation by the end of 2016). The primary empirical strategy examines how air quality across different locations in the city is affected by the changes in the subway network density over time and across space. The main identification concern stems from the potential endogeneity in subway station location choices, in that the locations could be chosen based on the projections of some unobserved factors that could affect future traffic congestion and air pollution. For example, planners might locate subway stations in areas with projected growth in population or travel demand and hence deterioration in air quality. An endogenous location of this sort would lead to an underestimation of the real impact of subway expansion on air quality.

To address this endogeneity concern, we instrument the subway density measure by constructing an alternative density measure based on the historical subway planning map following Baum-Snow (2007). Many of Beijing's subway lines were planned more than 20 years ago, long before traffic congestion and air pollution were of concern. The identification assumption hinges on the fact that these lines were originally designed to facilitate national defense because Beijing, as the Chinese capital, is the home to the country's central government agencies. The subway lines that were originally proposed had similar coverage as the lines that were eventually built. Controlling for a rich set of temporal and spatial fixed effects, the IV results show that a one-standard-deviation increase in the subway density improves air quality by two percent.[1] The estimate implies that the city-wide average reduction in pollution ranges from 0.02 percent from the opening of Line 16 (with a length of 20km) to 0.24 percent from the opening of Line 6 (with a length of 78km).

This approach, based on a continuous measure of network density, allows for the spillover effect of subway expansion across the whole network/city, but relies on the assumption that the impact diminishes over distance. To further examine the robustness of our results, we use a distance-based difference-in-differences (DID) method based on the assumption that the impact of subway expansion on air quality is local. We define the locations (of air quality monitors) within 2km of a subway station as the treatment group and the locations

[1] Air quality is measured using Air Pollution Index (API) from 2008 to 2012 and Air Quality Index (AQI) from 2013. These indices are translated from the dominant pollutant of the day piece-wise linearly. From 2008 to 2012, the index accounts for sulfur dioxide (SO_2), nitrogen dioxide (NO_2), suspended particulates (PM_{10}). Starting from 2013, the index accounts for SO_2, NO_2, PM_{10}, $PM_{2.5}$ and O_3.

farther than 20km away from a subway station as the control group. The locations between 2km to 20km are used as a buffer zone and are dropped in the analysis to avoid misclassifying the treatment status. Between the treatment and the control group, we focus on changes in air quality 60 days before and after the opening of a subway line. This focus on on a shorter time window can better address the concern of unobservables, but it is limited by only being able to examine the short-term impacts.

The key identification assumption of DID is that in the absence of a subway opening, the air quality in the treatment and the control group would follow similar trends. Subway construction could potentially cause ground construction dust and worsen the traffic congestion, leading to an overestimation of the pollution reduction effect. However, this concern is mitigated because Beijing's safety regulations require a three-month trial running period before the opening of a new subway line, thus, physical construction has to end at least three months before the opening of the line (Gu et al. 2018). We use an event study analysis to show the parallel trends hold for pre-opening periods in general. We also show robustness of our findings by restricting the control group to the monitors that are located 20km farther from the new subway stations but within 2km distance of subway stations opened in the past and to be opened in the future.

The DID specification shows that subway expansion improves air quality in the vicinity (within 2km) of the new subway line by 7.7 percent, relative to the area outside of the 20km radius within the 60-day time window. Allowing the effects to vary over time, we show that the effect becomes the largest around 50-60 days after opening. The DID specification considering heterogeneity in subway density shows that air quality improves further as more new subway stations are opened near a monitoring station.

Our paper adds to the emerging literature on the impact of subway expansion on air quality. Chen and Whalley (2012) estimate the causal effect on air pollution from the opening of one subway line in Taipei based on a regression discontinuity (RD) framework. They find that the opening of the Taipei Metro reduced air pollution from carbon monoxide (CO), one key tailpipe pollutant, by 5 to 15 percent. Zheng et al. (2019) use the DID method to estimate the impacts of the opening of the first subway line in Changsha, China and find an 18 percent reduction in CO in the areas proximate to subway stations. Gendron-Carrier et al. (2018) examine 43 cities across the world that had a new subway system open from 2000 to 2014. Using the satellite data on Aerosol Optical Depth around city centers, the paper estimates that particulate concentrations drop by 4 percent following the opening of a new subway system and that the effect persists for up to eight years. Nevertheless, recent papers by Beaudoin and Lin-Lawell (2017) and Rivers et al. (2020) find no evidence of air quality improvement from the expansion of public transit.

This study leverages fine-scale air pollution data and multiple subway lines within the same city to examines the impact of subway expansion on air quality. Different from the RD or DID frameworks in the literature, we use a continuous density measure to characterize the expansion and employ an IV strategy based on the historical planning for identification in our main analysis. By using the continuous measure of subway network density that varies across locations in the city, our analysis focuses on the marginal impact of subway expansion, rather than the impact of building the first subway line.

Rapid urbanization is a global trend, especially in developing and fast-growing economies, and building subway lines constitutes a common supply-side strategy to address traffic congestion and air pollution from automobile usage. Subway construction requires substantial investment, so it is important to understand the benefits of this investment. Based on our empirical results using subway network density, we conduct a back-of-the-envelope calculation of the benefits of subway expansion through improved health outcomes and reduced traffic congestion. The health benefits include both mortality and morbidity impacts, while the benefit from traffic congestion relief stems from the value of reduced travel time of commuters. Our conservative analysis shows that the subway expansion observed during our sample period can provide a total discounted health benefit of $0.6-2.0 billion during a 10-year period and $1.0-3.1 billion during a 20-year period, accounting for only 1.1-3.6 percent and 1.4-4.4 percent of the total upfront construction cost and the total discounted operating cost during the same period. Our estimates suggest that although there does exist non-trivial health benefits from improved air quality, the benefit from the reduction of traffic congestion estimated from the literature is more than one order of magnitude larger.

We organize the remainder of the paper as follows. Section 2 discusses the background and related data sets. In Section 3, we describe the empirical strategy. In Section 4, we discuss the estimation results and policy implications. Section 5 concludes.

1.2 Background and Data

In this section, we discuss the challenges of air pollution and the rapid expansion of the Beijing subway system. We then present the main datasets.

1.2.1 Air Quality in Beijing

During the past several decades, China has experienced unprecedented economic growth. From 1980 to 2016, the country's per capita GDP increased significantly, from less than $200 to over $8,000 in nominal terms according to the World Bank national accounts data. Meanwhile, air quality in major cities such as Beijing is deteriorating. Figure 1.1 shows daily and annual $PM_{2.5}$ concentrations in Beijing from 2008 to 2017. The average level is about twice as high as the Chinese annual standard, and six to ten times the U.S. standard.[2]

A rich economic literature has shown robust evidence of the adverse impact of outdoor air pollution on premature mortality and contemporaneous adult health (Chay and Greenstone 2003; Currie and Neidell 2005; Greenstone and Hanna 2014; Lelieveld et al. 2015; Schlenker and Walker 2016; He et al. 2016). The epidemiology literature has linked chronic obstructive pulmonary disease (COPD), ischemic heart disease (IHD), and lung cancer (LC) to $PM_{2.5}$ (Burnett et al. 2014). According to the Global Burden of Diseases (Cohen et al. 2017), outdoor air pollution contributed to 4.2 million premature deaths in the world in 2015; 40 percent of those occurred in China.

[2] The U.S. Environmental Protection Agency (EPA) sets the U.S. standard as 12 $\mu g/m^3$ annually and 35 $\mu g/m^3$ daily while the China Ministry of Environmental Protection (MEP) sets the Chinese standard as 35 $\mu g/m^3$ annually and 75 $\mu g/m^3$ daily.

The major sources of outdoor air pollution such as $PM_{2.5}$ include power plants, automobiles, and industrial activities. The relative contribution of each source varies across locations. Quantifying the contribution of urban traffic to $PM_{2.5}$ is challenging because tailpipe emissions lead to secondary $PM_{2.5}$, whereby motor-vehicle emissions are transformed into ambient air pollution through complicated chemical processes. In practice, air quality modeling has yielded a wide range of results. In U.S. cities, the contribution of motor-vehicles to air pollution ranges from 5 percent in Pittsburgh, PA to 55 percent in Los Angeles, CA (Tager et al. 2010). Zhang et al. (2013) estimate the contribution of traffic and waste incineration to air pollution to be 4 percent while Lelieveld et al. (2015) find that motor-vehicle travel alone contributes 3 percen of air pollution in Beijing. However, due to different definitions of the toxic level of each pollutant (such as $PM_{2.5}$, NO, SO_2 and O_3), the level of air pollution from ground traffic remains uncertain.

1.2.2 Beijing Subway Expansion

During the past two decades, the Chinese automobile industry has grown to be by far the largest in the world, with a total output of around 29 million units including 24.8 million passenger vehicles, in 2017. Private vehicle ownership in China was uncommon before 2000 but the sales of new passenger vehicles in China increased dramatically after the turn of the century, growing from less than one million units in 2001 to nearly 25 million in 2017 and surpassing the U.S. market in 2009. Beijing has led the way in vehicle ownership growth, transitioning from a city on bikes to a city in cars during this period: Beijing's stock of passenger vehicles increased from about 1.1 million units in 2001 to nearly six million units in 2018. Beijing is now routinely ranked as one of the most congested cities in the world, with the average traffic speed during peak travel times often less than 15 miles per hour.

The Beijing municipal government has taken several measures in order to control the air pollution and the traffic congestion caused by the city's increasing vehicle ownership. One measure is the driving restriction policy started in 2008 whereby vehicles are banned from driving one day per week based on the last digit of the license plate. During important events such as the 2008 Olympic Games or when the air pollution is extremely hazardous (e.g., during the "red alert" days), half of all private vehicles are restricted from the road (with the restriction based on odd and even numbers).[3] Viard and Fu (2015) find that traffic restriction in Beijing led to a 19 percent decline of API during every-other-day restrictions and a seven percent decline during one-day-per-week restrictions. This is consistent with the findings of Chen et al. (2013), who examine the effectiveness of different environment measures that the Chinese government adopted to prepare for the 2008 Olympic Games.

Driving restriction policy, on the other hand, may have incentivized households to buy a second vehicle in order to bypass the driving restrictions.[4] In an additional attempt to curb the growth in vehicle ownership, the Beijing munic-

[3]The Emergency Management Division from the Beijing Environmental Protection Bureau issues air pollution alerts based on the four-tiered pollution warning system. *Blue*: AQI> 200 for one or more days; *Yellow*: AQI> 200 for 2 or more days; *Orange*: AQI> 200 for 3 or more days and AQI> 300 for 2 consecutive days; *Red*: AQI> 200 for 4 or more days and AQI> 300 for 2 consecutive days or AQI> 500 for any 24-hour period.

[4]Davis (2008) studies the effectiveness of driving restriction in Mexico City and finds that the driving restriction leads to worse air quality because more households buy a second vehicle (which tends to be old and release higher pollution). Zhang et al. (2017) find similar results for the driving restriction policies implemented in Bogota, Colombia.

ipal government adopted a quota system for new vehicles in 2012 by capping the monthly number of new vehicle sales. In addition, a limited number of vehicle licenses is allocated through a lottery system (Li 2018). The winning odds of the license plate lottery in Beijing have decreased from 1:10 in early 2012 to nearly 1:2000 in 2018 as the pool of lottery participants increases dramatically and the cap tightens over time.

Along with demand-side strategies to reduce traffic, the Beijing municipal government has also been investing heavily in transportation infrastructure such as buses, roads, and subway lines. From 2008 to 2016, 13 new subway lines and one airport expressway were constructed with a total length of 440 kilometers, making the Beijing subway system not only the most rapidly expanded but also the longest in the world.[5] Figure 1.2 shows the detailed timeline of Beijing subway expansion, which is still ongoing; another 12 subway lines are under construction and scheduled to open before the end of 2020 with a total length of nearly 378 kilometers. Many other cities in China are also rapidly expanding their subway systems.

1.2.3 Data Description

Table 1.1 describes the main variables of our analysis and they are constructed based on three major datasets. The first dataset contains daily air quality readings from all of the 27 monitors in Beijing. Figure 1.3 shows the spatial distri-

[5]Other four subway systems in the top five worldwide by length (2012): (i) the Shanghai subway is opened in 1995, with a total network length of 423km; (ii) the London subway is opened in 1863, with a total length of 402km; (iii) the New York City subway is first opened on Oct 1904 with a total length of 368km; and (iv) the Seoul subway is first opened in 1974, with a total length of 368km.

bution of the 27 air quality monitors; 11 of these are operated by the central government, and the rest are operated by the local government. Geographically, eight monitors lie within the 5th ring road, and the rest are outside the 5th ring road. Air pollution (*Air Pollution*) in Beijing is measured by two different indices: Air Pollution Index (API), available from January 1, 2008 to December 31, 2012, and Air Quality Index (AQI), available from January 1, 2013 to May 12, 2017. Both indices are measured at the monitoring station level on a daily basis. The API is based on three atmospheric pollutants, sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and suspended particulates (PM_{10}). In 2013, the Chinese government replaced API with AQI which considers $PM_{2.5}$ separately from PM_{10}, and includes ozone (O_3) and carbon monoxide (CO) as major pollutants. The API or AQI for a given day is calculated based on the level of the dominant pollutant during that day and the dominant pollutant is determined by a scoring system as shown in Table 1.2.[6]

The second dataset records the opening dates and the locations of subway lines. During the data period from 2008 to 2016, 13 new subway lines and one airport expressway with 252 new subway stations were opened. Figure 1.3 overlays air quality monitors with subway stations in Beijing as of 2016. Most of the subway stations are located in the central city.[7] Subway stations on the same line could be opened at different dates. For example, some subway stations on Line 8 were opened on the same day on Line 9. Our analysis is thus based on ten major opening dates during the sample period (Figure 1.2).[8]

[6]An alternative air quality measure is the Aerosol Optical Depth (AOD) data from satellites (Gendron-Carrier et al. 2018; Zou 2020). We do not use the AOD data due to the large number of missing observations caused by the cloud coverage at the daily level.

[7]On December 30, 2010, four subway lines (Line Daxing, Changping, Fangshan and Yizhuang) opened, targeting the suburban districts.

[8]The ten major opening dates are Jul 19, 2008; Sep 28, 2009; Dec 30, 2010; Dec 31, 2011;

Table 1.3 presents the opening dates of new subway lines with the lines' total length and number of new stations, as well as average measures of the subway density at the locations of air pollution monitoring stations, for each of the ten opening dates. The average standardized network density at the monitoring stations increases from 0.27 in 2008 to 0.96 in 2016. The construction of the subway network density is discussed in detail in the following section.

The third dataset contains daily weather variables: average temperature, average relative humidity, precipitation, and binary variables indicating rain, snow, storm, and fog. It also includes hourly wind direction (measured in degrees from 0° to 359°) and speed. Wind plays an important role in air pollution because it affects the movements of the fine particulates. Since our unit of observation is daily, we need to convert hourly wind speed and direction to the daily level. We calculate the daily wind direction and speed based on the vector summation of hourly wind direction and speed.[9] We then categorize the daily wind directions into 16 groups. Table 1.4 presents summary statistics for the main daily weather variables and the daily wind directions.[10]

Table 1.5 presents the sample averages of ln(*Air Pollution*) 60 days before and after the opening of each new subway line. The top panel shows the sim-

Dec 30, 2012; May 5, 2013; Dec 28, 2013; Dec 28, 2014; Dec 26, 2015; and Dec 31, 2016. The opening dates within 60 days apart from these major opening are combined with the closest major opening date.

[9]For example, wind at 8:00 am is 30 degree (angle from North) with wind speed 4 mph; wind at 9:00 am is 90 degree (E) with a speed 4 mph. The summation of the two wind vectors would be a 60 degree wind vector with a speed 7 mph.

[10]One may consider a simple average of the hourly wind direction indicators for a day, but this measurement could neglect the magnitude of wind speed and thus bias its impact. Wind directions at the subway station level can help us assess the subway expansion's impact at a particular monitoring station precisely. This approach will be practically feasible to implement as the satellite data (such as Aerosol Optical Depth data) and smart phones (or cards) data are becoming more available.

ple averages, while the bottom panel presents the average residuals after controlling for weather conditions and a rich set of time and location fixed effects (the same set of controls to be used in the regression analysis). The treatment group is defined as the monitoring stations within 2km of a new subway line, while the control group is defined as the monitoring stations more than 20km away from the new subway line. The top panel shows a 4 percent increase in air pollution level on average after the opening of a subway line. This counterintuitive result could be driven by seasonality: nine out of the 14 new lines were opened in December and air quality tends to be worse in January and February than in November and December due to winter heating. The bottom panel shows that after partialling out time and location fixed effects and weather conditions, the opening of a new subway line is associated with an 4.6 percent reduction in air pollution level on average.

Figure 1.4 depicts average residuals of $\ln(\textit{Air Pollution})$ from 60 days before to 60 days after the opening of each new subway line for the treatment group and the control group, after partialling out weather conditions and a rich set of time and location fixed effects. The treatment group appears to have a higher air pollution level than the control group (relative to their baseline levels) one month before the opening of the new lines but have a lower level of air pollution about 20 days after the opening. The difference between the two groups seems to increase over time after the opening with the treatment group exhibiting a lower level of air pollution.

1.2.4 Subway Network Density

The key explanatory variable in our main empirical specification is an inverse distance-weighted subway density:

$$Density_{it} = \sum_{j \in \mathcal{N}_t} \frac{1}{Distance_{ij}^2},$$

where i, j, and t index air pollution monitoring stations, subway stations, and days, respectively. $\mathcal{N}_t$ is the set of existing subway stations at time t. The subway network density for monitoring station i at time t is the weighted number of subway stations at time t, in which the weight is the inverse of squared distance from the monitor to a corresponding subway station in operation at time t. Following the density measure commonly adopted in the urban literature (Ewing and Cervero 2010), this measure can be considered as the number of subway stations per unit area centered around a given monitoring station. The density measure increases with the number of subway lines. However, a new subway line will change the density measure differently across monitoring stations. The density will increase more for the monitoring stations closer to the subway line.

This subway density measure, however, does not account for the heterogeneity across subway stations or subway lines in their contribution to the whole subway system. For example, major transfer stations that connect multiple subway lines or subway lines in the center of the system play more important roles in the connectivity of the system. To capture this heterogeneity, we generate an alternative density measure which takes into account the ridership of each subway line for robustness checks.[11] The following equation

[11] The ridership information at the subway station level would be ideal to be used as a

shows the ridership-weighted subway density measure ($\widetilde{Density}_{it}$):

$$\widetilde{Density}_{it} = \sum_{j \in \mathcal{N}_t} \frac{Weight_{j\ell}}{Distance_{ij}^2},$$

where $Weight_{j\ell}$ denotes the weight of subway station j on subway line ℓ, which equals the ridership share of line ℓ among all subway lines in operation at time t.

Table 1.3 reports the number of new stations at each opening and the average standardized density in the vicinity of air quality monitors at each opening.

1.3 Empirical Strategy

In this section, we discuss our empirical methods and the identification challenges. The main empirical framework employs subway network density as the key explanatory variable and uses the instrumental variable (IV) approach to address endogenous subway locations. We then present the difference-in-difference (DID) framework as an alternative strategy and discuss analysis of heterogenous treatment effects.

weight. Unfortunately, we could not find such data set at this point. To proxy the ridership at the station level, we use ridership data at the subway line level, treating that each station in a certain subway line has the same ridership.

1.3.1 Network Density and Air Quality

We estimate the following equation:

$$\begin{aligned}\ln(Air\,Pollution_{it}) =& \beta_1(Density_{it}/\sigma) + Monitor_i + Trend_{it} \\ &+ Weather_t\boldsymbol{\beta_2} + Monitor_i \times Driving_t \\ &+ Year_t + Season_t + DoW_t + Holiday_t + \varepsilon_{it}. \end{aligned} \tag{1.1}$$

The outcome variable, $\ln(Air\,Pollution_{it})$, is the logarithm of daily Air Pollution Index (API) during 2008-2012 and Air Quality Index (AQI) from 2013 onward. $i = 1, \ldots, 27$ is the index for monitoring stations and $t \in [\text{Jan } 1, 2008, \ \text{Dec } 31, 2017]$ is the index for day. The key explanatory variable is the standardized subway network density to facilitate interpretation, where $Density_{it}$ is defined above and σ is the standard deviation of the density. $Weather_t$ is a vector of weather variables including average temperature (^oC), relative humidity (%), wind speed (m/s), precipitation (mm), dummies for rain, snow, storm, and fog, and 16 wind direction dummies.

We include monitor fixed effects ($Monitor_i$) to control for unobserved location attributes that affect air quality. We also control for a set of temporal fixed effects including year fixed effects ($Year_t$), season fixed effects ($Season_t$), day of week fixed effects (DoW_t) and holiday fixed effects ($Holiday_t$). To control for other confounding factors that may vary across time but are not adequately controlled by the time fixed effects, we include a monitor-specific time trend, $Trend_{it}$, to allow the unobserved time trend to vary across monitors.[12]. We also interact monitor fixed effects with driving restriction policy ($Driving_t$) to allow

[12] $Trend_{it}$ is a vector of monitor-specific linear time trends (the interaction of the dummy for monitor i and the linear time trend t).

the effects of driving restrictions to vary by locations. Beijing's driving restriction policy bans some vehicles from driving on a given workday depending on the last digit of the license plate number. This policy follows a pre-set rotation schedule in terms of which pair of numbers (1 and 6, 2 and 7, 3 and 8, 4 and 9, or 5 and 0) is restricted on a given day, and it is not adjusted based on traffic conditions. Because the last digits of license plates are not evenly distributed and this policy thus changes the traffic conditions on the road ((Yang et al. 2020)), we construct $Driving_t$ as a vector of five dummies indicating the five pairs of the last digits of license plates. ε_{it} is the random error term.

The key identification challenge is the potential endogeneity of the density variable resulting from non-random placement of subway stations. City planners may place the subway lines and stations in anticipation of the future growth (e.g., population or commercial activities) of different parts of the city, which could have implications for the traffic congestion level. If the subway lines are more likely to be placed in areas with higher anticipated growth of economic activities (hence congestion), the framework using the network density as the key explanatory variable may underestimate the impact of subway expansion on air quality improvement.

To address the concern of non-random placement of subway stations, we use the historically planned subway network to construct an instrument for the density measure, following Baum-Snow (2007), which uses historical highway plans in the U.S. to instrument for observed highway routes. We obtain historical subway plans in 1957, 1983, 1999 and 2003, as shown in Figure 1.6. We use the 2003 plan to construct the instrument because it has the most lines and because many of the lines appear in earlier plans. The 1957 plan is the first known plan and provides the basis for the subsequent plans while the 1983 plan defines the "Horizontal+Vertical+Ring" framework of the Beijing subway system, which continues to be used. Because we do not observe the planned opening dates from the historical plans, we assign the actual opening dates to the planned lines. In order to introduce another layer of randomness, we also implement random opening dates within a window of the observed opening date as a robustness check.[13]

The exogeneity assumption of the IV hinges on the fact that the original subway plans were designed to facilitate national defense, with little or no regard for future travel demand or air quality. Many of the lines were planned several decades before the construction, long before air pollution and traffic congestion became a concern. During the first planning period of the subway system about 60 years ago, the population in Beijing was less than 3 million, with only 5,000 vehicles. Building a subway system requires huge investments and advanced technologies. The then-premier, Zhou Enlai, said, "Beijing is building the subway purely for defense reasons. If it was for transport, purchasing 200 buses would solve the problem." [14]

Beijing's vehicle stock was only 1.5 million in 2003, compared to nearly 6

[13]Following Faber (2014), we construct an alternative IV in the earlier version where we use the minimum spanning tree (MST) method to construct hypothetical subway lines with the origin and destination given by the historical subway plans. We straighten up all the historical subway lines and reallocate the observed subway stations to the nearest location on the hypothetical lines. We find similar results using the two different sets of IV.

[14]A quote from the article "The birth of the Beijing subway: Premier Zhou said that the preparation of the subway is to prepare for the battle" well explains the situation that China faced back in the 1950s, "In June 1950, the new China, which was just half a year after the founding of the People's Republic of China, was forced to become involved in the Korean War. At the same time, the US Seventh Fleet entered the Taiwan Strait. ... In such an international situation, war preparedness should be the first factor to be considered in Beijing's urban planning." http://discovery.cctv.com/20070926/100879.shtml.

million by 2018. The rapid increase in vehicle ownership after 2003 was unlikely to be predicted by policy makers and the historical plan is thus unlikely to be correlated with the spatial pattern of traffic congestion and air pollution within the city. The IV is correlated with the density measure because the constructed subway lines largely follow the historical plans, which contain a similar number of transferring stations and level of connectivity as the current subway system.

The empirical approach based on subway network density relies on the spatial and temporal variation of the network expansion. The subway density measure is not a city-wide measure but is local in nature. A new subway line would increase the density more for nearby monitoring stations than for those farther away from the line. The underlying assumption is that the impact of subway expansion on air quality is not uniform across the city but diminishes over distance. With this assumption, this approach allows for system-wide impact or the spatial spillover effect of subway expansion on air quality.

1.3.2 Difference-in-Differences Specification

As an alternative specification, we use the DID method which assumes the impact of subway expansion to be confined locally. This assumption allows us to define treatment and control groups. While this assumption may appear to be ad hoc, the advantage of the DID approach is that it can be easily adapted to examine the potential heterogeneity in impacts (e.g., the dynamic impact over time).

Our DID strategy compares the air quality 60 days before and 60 days after each of the 10 opening dates of subway stations between the treatment and the control group. Since the subway lines are designed to serve different areas of Beijing, the set of treated and control monitors vary across different opening dates. We choose the time windows to be 60 days before and after the opening dates to avoid the overlap between the pre-opening and post-opening periods of two consecutive lines. In DID regressions, we restrict our sample to the observations that fall in the 120-day windows around the opening dates.

We define the treatment group as the monitoring stations within 2km of a subway station and the control group the monitoring stations farther than 20km of a subway station. We treat the area in between as the buffer zone and drop the monitors in the buffer zone in the DID analysis to address the concern of misclassifying treatment status.

The choice of the treatment group is based on the radius of the impact on commuters' mode of travel to subway stations. The typical length of time that commuters take to travel to subway stations is between 5 and 15 minutes. Walking and biking are the two most common commuting modes to subway stations in Beijing. The typical walking distance is about 1km (or 12 minutes based on a walking speed of 5km/hour) while the typical biking distance is about 3km. We choose the average of the two as the radius of impact to define the treatment group.[15]

As the subway system is a network, the impact of the opening of a new

[15]The walking and biking distances are approximated based on the Guideline of Designing and Planning for Areas along Urban Rail from Ministry of Housing and Urban-Rural Development of the People's Republic of China, and Yang et al. (2018b). We also conduct a spatial lag analysis to determine the 2km cut-off for the treated and 20km cut-off for the control groups. The results are available upon request.

subway station on air quality could go beyond 2km. The DID provides estimates of *local* effects within 2km of subway stations, which is different from the estimates of city-wide effects in the density specification discussed earlier. The impact is likely to be larger in the areas closer to subway stations due to the stronger impact on travel mode choices. Therefore, we expect the estimates from the DID to be larger than the estimated impacts from the IV method using the density measure, which is confirmed by our empirical findings (to be discussed later).[16]

Following a general framework by Bertrand et al. (2004) and Hansen (2007) with multiple groups and time periods, the basic DID framework is specified as

$$\begin{aligned}\ln(Air\ Pollution_{it}) =& \theta Treated_{it} \times \mathbf{1}(Post_t) + Monitor_i + Trend_{it} \\ &+ Weather_t \boldsymbol{\beta} + Monitor_i \times Driving_t \\ &+ Year_t + Season_t + DoW_t + Holiday_t + \varepsilon_{it}, \end{aligned} \tag{1.2}$$

where $Treated_{it}$ is a treatment indicator that takes the value of 1 if monitor i is within 2km of any subway stations that were opened on date τ ($\tau - 60 \leq t \leq \tau + 60$). $\mathbf{1}(Post_t)$ is a dummy variable indicating whether an observation is within 60 days *after* opening of these new subway stations, that is, $\tau \leq t \leq \tau+60$. The parameter of interest is θ which captures the impact of the subway opening on air pollution for areas in the vicinity of the new subway stations within 60 days after the opening. Other control variables are defined as in Equation 1.1.

[16]To the extent that the opening of a subway station could impact the traffic flow of the whole city including areas 20km away, the DID approach confounds control with treatment and could underestimate the true impact. Indeed, when we define the control group as the monitoring stations 15km away from a subway station and shrink the buffer zone accordingly, we find a smaller impact, consistent with the intuition above. We choose 20km to reduce the potential bias.

The key assumption of the DID is that, in the absence of a new subway opening, air quality in the treatment and control groups follow parallel trends. Most monitoring stations in the control group are in the suburban districts of the city as shown in Figure 1.3. One may be concerned that those monitors in the control group may be too far away from the city center and thus would have different trends from those in the treatment group.

We take two strategies to address this concern. Our first strategy takes advantage of the staggered rollout design of the subway lines. We use the monitors that are located 20km farther from the new subway stations but within 2km distance of subway stations either opened in the past or to be opened in the future as the control group. Because both the treatment and control groups contain only monitoring stations that are close to subway stations, the two groups likely share similar (observed and unobserved) characteristics. The underlying assumption of this method is the randomness of the opening date.

Second, we use event study analysis to show the parallel trends hold for pre-opening periods in general. We divide the 120-day time window around opening dates into twelve 10-day intervals (six pre-opening periods $n = -5, -4, ..., 0$, and six post-opening periods $n = 1, 2, ..., 6$) and run the following regression:

$$\begin{aligned}\ln(Air\,Pollution_{it}) = &\sum_{n\neq 0}\delta_n P_t(n) \times Treated_{it} + Monitor_i + Trend_{it} \\ &+ Weather_t\boldsymbol{\beta} + Monitor_i \times Driving_t \\ &+ Year_t + Season_t + DoW_t + Holiday_t + \varepsilon_{it} \qquad (1.3)\end{aligned}$$

where $P_t(n) = \mathbf{1}\left[\tau + 10 \cdot (n-1) \leq t \leq \tau + 10 \cdot n\right]$, indicating interval n. The base interval is the 10-day intervals before the opening dates (i.e., $n = 0$).

Table 1.6 (and Figure 1.5) presents the coefficient estimates of δ_n. The results support the parallel trends assumption in general: compared with the base interval (10-day window before opening dates), the subsequent changes in air quality between the treatment and control groups are not significantly different for four out of the five pre-opening intervals in the specification exploiting staggered rollout design (Column 4). In the specification that does not exploit the staggered rollout design (Column 3), three out of the five pre-opening intervals show parallel trends, with the base interval and the parallel trends assumption only being marginally rejected in one of the remaining two intervals. In contrast, we find statistically significant effects of air pollution reduction in four out of six post-opening intervals for the same two specifications (Columns 3 and 4).

One additional identification concern may arise from air pollution induced by subway construction, which differs between the treatment and the control group. The construction of a subway station involves both underground and ground work, which may generate construction dust and worsen the air quality. If the construction leads to higher pollution levels close to new subway stations before opening dates, the DID framework could mistake the pollution reduction from the mere completion of the construction itself as the impact of the subway expansion and hence overestimate the true impact. However, this concern is mitigated because under the national standard of subway construction in China, every subway line is subject to an intensive trial run over a three-month period during which the subway train is tested after the ground work has been finished completely.[17] Since our DID analysis focuses on the

[17] The first phase of the trail run process has no passengers on board and during the second phase of the process, typically the last 20 days of the process, the subway with passengers (not

120-day window around opening dates during which the subway construction is already completed, we do not expect construction dust to confound our results.

We estimate two alternative specifications to relax the assumption of uniform effects of subway opening across opening dates and stations. First, we allow the impact to vary by number of days after the subway opening, as specified in Equation 1.4.

$$\begin{aligned}\ln(Air\,Pollution_{it}) =& \psi_1 Treated_{it} \times \mathbf{1}(Post_t) + \psi_2 Treated_{it} \times \mathbf{1}(Post_t) \times Days_t \\ &+ \psi_3 Treated_{it} \times \mathbf{1}(Post_t) \times Days_t^2 + Monitor_i + Trend_{it} \\ &+ Weather_t\boldsymbol{\beta} + Monitor_i \times Driving_t \\ &+ Year_t + Season_t + DoW_t + Holiday_t + \varepsilon_{it} \end{aligned} \tag{1.4}$$

where $Days_t$ is the number of days after the opening of the subway station. This specification allows the effect to occur gradually since it may take time for commuters to adjust their travel modes.

In the second specification, we examine the heterogeneity of treatment effects by allowing the impact to differ based on the number of new subway stations within the vicinity of the treated monitors as in Equation 1.5.

$$\begin{aligned}\ln(Air\,Polution_{it}) =& \eta N_{it} \times Treated_{it} \times \mathbf{1}(Post_t) + Monitor_i + Trend_{it} \\ &+ Weather_t\boldsymbol{\beta} + Monitor_i \times Driving_t \\ &+ Year_t + Season_t + DoW_t + Holiday_t + \varepsilon_{it} \end{aligned} \tag{1.5}$$

where N_{it} is the number of subway stations opened at date τ ($\tau - 60 \leq t \leq \tau + 60$) within the 2km distance of the monitor i. This specification captures

the public) will be tested following the scheduled time and route.

the notion that when more subway stations are located nearby, commuters are more likely to use the subway to reach their destinations and hence to reduce driving and air pollution more in the vicinity areas.

1.4 Empirical Results

In this section, we first present the estimated impacts of subway expansion on air quality using the IV method and the DID method in the first two subsections. We then present the results from a benefit-cost analysis based on back-of-the-envelope calculations.

1.4.1 Estimates Based on Network Density

Table 1.7 shows the OLS results using the continuous density measure shown in Equation 1.1. The key variable is the standardized subway network density. We sequentially add weather variables, wind conditions, a rich set of location and time fixed effects, and the driving restriction policy as control variables. Column (1) does not have monitoring station fixed effects, and the result shows a positive correlation between subway density and the level of air pollution. This result is likely driven by the fact that the city center, where the subway network is denser, tends to have higher pollution levels. Once monitor fixed effects are included, the results show that higher subway density is associated with a lower level of air pollution. This negative relationship is robust across columns (2) to (4). Column (3) adds monitor fixed effects interacting with the driving restriction policy, while column (4) further includes a monitor-specific time trend. Adding the monitor-specific time trend helps to alleviate the concern about the endogenous location of subway lines. Subway lines may tend to be placed in areas with faster projected growth in economic activities (and hence more air pollution); without controlling for this, the impact of subway expansion on air quality would be underestimated, as confirmed by the results in columns (3) and (4).[18]

The results from the full model (column (4) of Table 1.7) suggest that a one standard-deviation increase in subway density reduces the air pollution level by 1.5 percent. This estimation exploits the variation in network density and air pollution across space and locations. It can be interpreted as the longer term impact, when we compare it with estimates from the DID framework presented in the next section or from the literature, which typically relies on a shorter time window around the intervention to address confounding factors.

The weather variables have intuitive signs: high temperature and humidity are associated with a higher level of air pollution while rainfall/snow and wind are associated with a lower level of air pollution. High temperature can lead to faster formation of ground-level ozone and fine particulate matter while high humidity (without precipitation) makes it difficult for the natural air current to dissipate the pollutants. Precipitation in the form of rainfall or snow, as well as high wind, can help pollutants dissipate more quickly.

We address the potential endogeneity of network density measure using IV in Table 1.8. Column (1) is identical to column (4) in Table 1.7 to facilitate comparison. Column (2) instruments for the density variable with a hypothetical

[18]We have also tried two alternate monitor-specific time trends: time squared and time cubed. The OLS results are robust to the order of the time trend, we find similar estimates with monitor-specific squared time trend and monitor-specific cubed time trend.

density measure based on the 2003 subway planning map and uses the actual opening date of each line. The impact from 2SLS is slightly larger in magnitude than that from OLS. Column (3) uses a random opening date during a six-month window around the observed opening date to construct the IV. This helps to address the concern that policymakers may choose the opening date partly based on the projected pollution level. In practice, the opening of a new subway is often celebrated with a ceremony at which high-level government officials from both the Beijing municipal government and the central government are present. Seven out of the 10 opening dates in our sample fall in the last few days of a calendar year. In addition to the coincidence of celebrating a new subway line opening together with the beginning of a new year, this choice of dates is also likely due to the fact that it is easier to gather high-level government officials during the public holidays.

Columns (4) to (6) of Table 1.8 use the ridership-weighted density measures in which higher weights are assigned to subway lines with larger ridership in the network. Column (4) comes from OLS, while columns (5) and (6) come from 2SLS. Column (5) uses the observed opening date to construct the IV, while column (6) randomizes the opening date. Column (6) produces slightly larger estimates than columns (4) and (5), suggesting that a one standard-deviation increase in population-weighted density reduces the level of air pollution by 3.5 percent. In both specifications with different density measures, 2SLS results are slightly larger than OLS estimates.

Table 1.9 translates the parameter estimates of the IV regression with observed opening dates (column 5 of Table 1.8) into the impact for each subway line. To estimate average subway density in Beijing, we calculate the subway network at the Traffic Administration Zone (TAZ) level.[19] Figures 1.7 and 1.8 map the subway network density at the TAZ level at the end of 2007 (the year before our study period), 2009, 2011, 2013, and 2016. The subway network, which is denser at the city center, has been expanding rapidly with openings of new subway lines. For example, the opening of Line 6 (opened on December 30, 2012) increases the population-weighted density by 0.12 overall, which in turn leads to a 0.24 percent decrease in air pollution level. In the aggregate, the total 14 lines built from 2008 to 2016 result in a 1.01 percent decrease in air pollution in Beijing. Our estimates of the pollution reduction effect are smaller than that of Gendron-Carrier et al. (2018), who find a four percent reduction in air pollution after the opening of a new subway system. However, the majority of new subway systems considered in Gendron-Carrier et al. (2018) were the first subway lines in their corresponding cities, which could explain the larger estimated impacts than those in our case. In addition, studies using the DID or the regression discontinuity method tend to have larger estimates (Chen and Whalley 2012; Zheng et al. 2019), as these estimates may capture a shorter term and more local impact than ours. This is consistent with our analysis using the DID method below, which shows a larger impact than the estimate based on the continuous density measure.

[19]The city of Beijing is divided into 1911 Traffic Administration Zones (TAZs) for the purpose of city planning. Each TAZ has similar population size so the average subway density at the TAZ level is roughly equivalent to the population-weighted average of the density at the district level.

1.4.2 Difference-in-Differences Estimates

Table 1.10 presents the results from the basic DID model (Equation (1.2)). The results across columns exhibit similar patterns to those in Table 1.7. With the absence of monitoring station fixed effects in columns (1) to (3) of Table 1.10, air pollution level is positively associated with subway opening. After controlling for monitor fixed effects, Columns (4) to (6) provide similar estimates of the effects of subway opening on air pollution based on the DID model. The results from column (6) suggest that within a 60-day time window after a subway line's opening, the monitors in the vicinity (within 2km) of subway stations exhibit a 7.7 percent reduction in air quality compared to the monitors outside the 20km radius.

The DID specifications produce relatively larger impact estimates compared to those from the framework based on continuous density measures, likely for two reasons. First, the DID method focuses on a shorter-time window, while the method with density measures relies on variation during the whole data period. Thus, the DID estimates should be viewed as shorter-term impacts. Second, the DID method estimates the impacts of subway expansion on the areas within a 2km radius of new subway lines which are likely larger than the city-wide effects estimated by the method with network density measures.

Table 1.11 reports regression results using different time windows (from 10 to 180 days) before and after the opening dates. The estimates are not statistically different across 40- to 100-day windows (column 4 to 10). When we increase the window to 110 days and longer, however, the average effect seems to fade away. This is consistent with the notion that it may take some time for commuters to adjust their travel modes in the short term and hence for the impact on air pollution to be materialized. In the longer term, reduced traffic congestion could lead to additional driving demand, mitigating the initial reduction of air pollution. This dynamic is consistent with traffic diversion in the short-term and with induced traffic demand in the longer-term, as discussed in the introduction.

Table 1.12 shows the effect under a continuous measure of the time variables. We interact the treated group indicator with the linear and quadratic term of days post-opening, respectively. We also compare the specifications under two different time windows (60 days and 120 days). The results from our model specifications with the quadratic term of days post-opening (columns 2 and 4) suggest that the effect of subway opening on air pollution is non-linear. The subway opening begins to have a negative effect on air pollution after approximately 15-20 days; the magnitude of the effect then increases at a decreasing rate, with a turning point being around 50-60 days, after which the effect diminishes.

Table 1.13 presents the DID specification which accounts for the number of subway stations in the vicinity of treated monitors. The result shows that one additional subway station added to the vicinity of a monitor reduces air pollution by 2 to 4.1 percent, depending on model specifications. Compared to the IV method based on the network density measure, the DID method yields qualitatively the same results but considerably larger point estimates. Take the previous example of Line 6. The opening of Line 6 improves air quality in Beijing by 0.70 percent (assuming no effects on buffered locations) to 6.04

percent (assuming the buffered locations have the same impact as the treated locations). This comparison reflects the interplay of the two countervailing forces: the traffic division effect of public transit investment (the Mohring effect), and the induced demand effect. The second channel takes longer to occur and dampens the positive impact on air quality improvement observed in the short term. Nevertheless, our estimates suggest that the first channel is the dominant force in the longer run.

1.4.3 Cost-Benefit Analysis

This section presents a back-of-the-envelope analysis on the benefit of subway expansion through two channels. The first benefit is on human health including both mortality and morbidity from improved air quality. The second benefit comes from congestion relief and the value of saved travel time for commuters.

Our empirical analysis finds that subway expansion leads to statistically significant improvement in air quality. Table 1.9 shows the estimated air quality improvement due to each subway line based on the benchmark specification (based on the IV results in Table 1.8). The population weighted air quality improvement ranges from 0.02 percent by Line 16 opened on December 31, 2016 to 0.24 percent by Line 6 opened on December 30, 2012. Recent literature from both epidemiology and economics has shown that the long-term exposure to airborne particulates can lead to elevated mortality especially among infants and morbidity due to cardiorespiratory diseases (Chay and Greenstone 2003; Currie and Neidell 2005; Currie and Walker 2011; Knittel et al. 2016;

Greenstone and Hanna 2014; He et al. 2016; Ebenstein et al. 2017).

To calculate the mortality impact of subway expansion, we take the estimates from Ebenstein et al. (2017) that study the impact of long-term exposure to airborne particulate matter on mortality using a regression discontinuity design. They find that a 10-$\mu g/m^3$ increase in PM_{10} increases cardiorespiratory mortality by 8 percent; this impact varies across age cohorts but not across gender. Following the analysis in Barwick et al. (2019) to monetize the mortality impact, the mortality cost amounts to \$13.38 billion across the Chinese population from a 10-$\mu g/m^3$ increase in PM_{10}, or \$64.9 per household in Beijing when adjusted for the Beijing per capital income (in 2015 dollars). The morbidity cost of air pollution comes from Barwick et al. (2019), who provide the first comprehensive analysis of the morbidity cost in China based on the universe of credit and debit card spending. They find that the morbidity cost from a 10-$\mu g/m^3$ increase in $PM_{2.5}$ is \$20.2 (in 2015 dollars) per household for China.[20]

The congestion relief benefit comes from the value of the saved commuting time. Using a regression discontinuity design, Yang et al. (2018a) estimate that each new subway line reduces travel delay by an average of 15 percent based on the subway lines that opened between 2009 to 2015. The Beijing Annual Transportation Report shows that the average traffic delay time is around 20 minutes per hour. We assume that these delays occur during the peak hours (7am-9am and 5pm-7pm) on the weekdays and that approximately two mil-

[20] We follow the emerging literature on the morbidity costs of air pollution (Deschenes et al. 2017; Barwick et al. 2019), which estimate that the morbidity costs could amount to about two thirds of the mortality costs. Landrigan et al. (2018) summarized a series of studies that suggest that the morbidity costs resulting from pollution-related disease might conservatively increase mortality costs by 10-ÂĂÂŞ70%, and some individual country studies suggest that the increment might be even greater: 25% for Colombia, 22-78% for China, and 78% for Nicaragua.

lion commuters (who travel by cars and buses) are affected. The value of time (VOT) for automobile travel is often assumed to be half of the market wage (Parry and Small 2009), which is 62.98 Yuan per hour ($9.5 per hour) based on the monthly wage of 10,077 Yuan.

Panel (a) of Table 1.14 presents the cost-benefit calculations during a 10-year period after the opening of each subway line. The cost includes both the upfront construction cost and the operating cost (Column 1). We discount the operating cost and the benefit at a 5 percent annual discount rate. The total cost from all the subway lines during the sample period is $56.3 billion (with the construction cost being $46.7 billion). The health benefit amounts to $0.64 billion (Column 2), or 1.13 percent of the total cost (Column 4), while the benefit from congestion relief is $26.9 billion (Column 6), or 48 percent of the total cost (Column 8). Panel (b) of Table 1.14 presents the cost-benefit calculations during a 20-year period where the benefit from health and congestion relief accounts for 1.38 percent and 58 percent of the total cost, respectively. The analysis suggests that the health benefit from improved air quality is a relatively small portion compared to the overall benefit of subway expansion.

However, our benefit estimates in Columns (2), (4), (6), and (8), are conservative for three reasons. First, the mortality benefit is based on the Value of a Statistical Life (VSL) of $2.27 million (in 2015) from (Ashenfelter and Greenstone 2004), rather than the central estimate of $8.7 million figure recommended by the U.S. EPA. Second, the value of time is assumed to be 50 percent of the wage, rather than 100 percent of the hourly wage (Small 2012; Wolff 2014). Third, the benefit calculation includes neither the benefit from improved commute reliability nor the benefit from a larger choice set of travel modes (Small et al. 2005).

We then calculate an upper bound of the health benefit and congestion relief benefits in 10-year and 20-year respectively, presented in Columns (3), (5), (7), and (9). These estimates are based on the VSL of $8.7 million from the U.S. EPA and the VOT of 100 percent of hourly wage in Beijing. At the upper bound, the health benefit amounts to $2.01 billion or 3.57 percent of the total cost while the benefit from congestion relief is $53.71 billion or 95.34 percent of the total cost during a 10-year period. During a 20-year period, the upper bound of benefits from health and congestion relief accounts for 4.36 percent and 116.41 percent of the total cost respectively. Together, the total benefits from health and time saving alone exceed the costs during a 20-year timeframe, recognizing that subway systems could have a life span of at least several decades or over 100 years.[21]

Our analysis suggests that although the health benefit of subway expansion is nontrivial, it is much smaller than the benefits from congestion relief. Large sources of air pollution in Beijing include motor vehicle emissions, industrial activities, coal burning, and construction dust, as well as long-range transported pollution from nearby cities. According to Beijing Environmental Protection Bureau, automobiles are the largest source of $PM_{2.5}$, accounting for 22 percent in the whole city and about one third of the total in the urban core in 2012. The second largest source of $PM_{2.5}$ in 2012 is coal burning (17 percent), followed by construction site dusts (16 percent). Unless driving is substantially reduced, the impact on air quality improvement from infrastructure investment alone is likely to be small, especially when the road usage is

[21] London has the oldest subway system which started in 1890 and the New York City subway system began operation in 1904.

not priced.

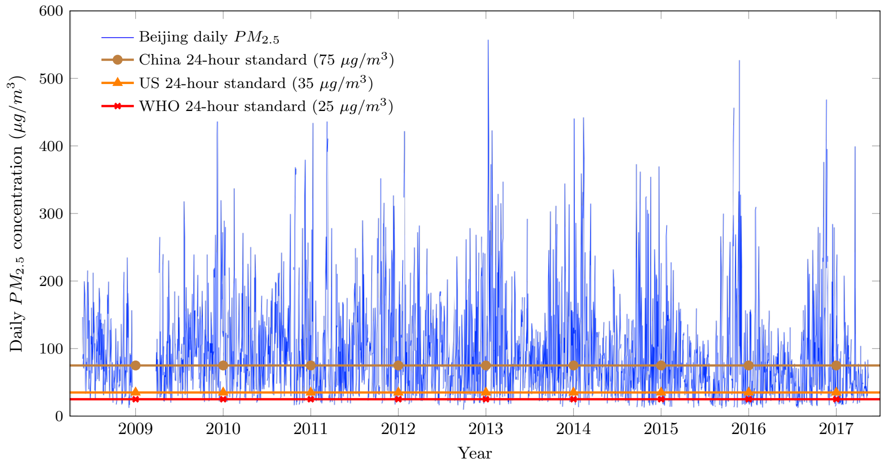
1.5 Conclusion

To address worsening air pollution and traffic congestion across urban areas in China, central and local governments are undertaking large investment in transportation infrastructure such as roads, rail, and subway systems. China's total investment in transportation infrastructure in 2014 amounted to 2.5 trillion yuan ($409 billion), about four percent of its GDP. Beijing has been leading the way among major cities in public transportation infrastructure by rapidly expanding its subway lines. Between 2002 and 2015, the Beijing municipal government invested nearly 300 billion Yuan (or USD 47 billion) on 16 new subway lines and Beijing now has the second longest subway network of 599km in the world, after Shanghai.

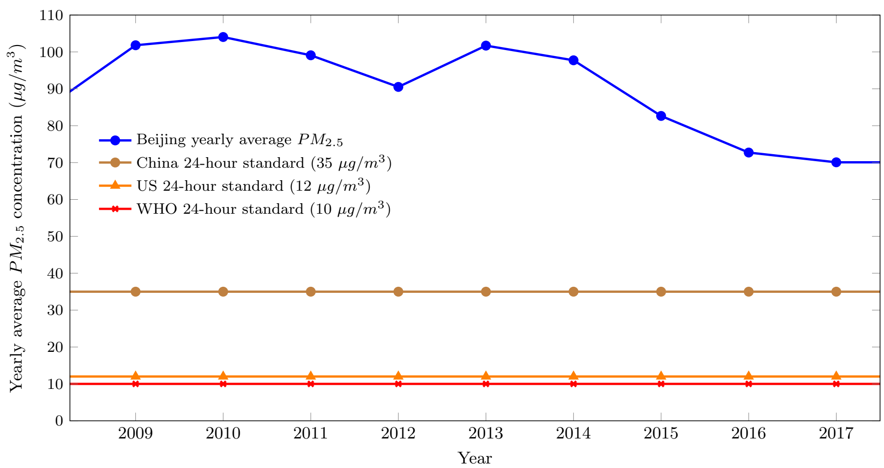
While previous literature has examined the congestion relief function of public transportation, there is limited evidence regarding the impact of subway expansion on air quality. By leveraging fine-scale air pollution data and the rapid rollout of 14 new lines from 2008 to 2016 in Beijing, we find that the opening of new subway stations improves air quality from a variety of empirical specifications. An IV analysis based on the network density measure shows that a one standard-deviation increase in the density improves air quality by two percent. The 20-year total discounted health benefits of the subway expansion amounts to $1.0-3.1 billion due to reduced mortality and morbidity from improved air quality. Nevertheless, the benefit would only account for 1.4 to 4.4 percent of the total cost, including both the construction and operating cost. Our findings suggest that most of the cost from subway expansion needs to be justified from traffic congestion relief and other economy-wide impacts. Future research could examine the impact of subway expansion on the location choices of households, labor participation decisions, and firm entry and exit, all of which could have important implications on the broader economy.

Figure 1.1: Beijing $PM_{2.5}$ Concentration ($\mu g/m^3$)

(a) Daily



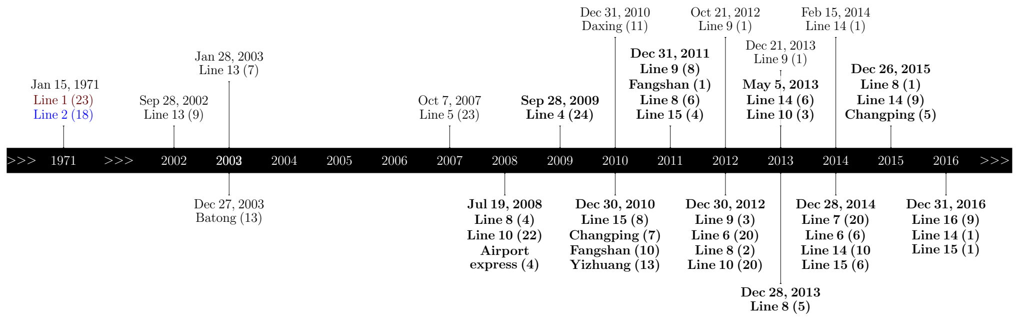
(b) Annual



Note: Panel (a) and (b) shows daily and annual average $PM_{2.5}$ concentrations in Beijing from 2008 to 2017 respectively. The average level is about twice as high as the Chinese annual standard, and six to ten times the U.S. standard. The U.S. EPA sets the U.S. standard as 12 $\mu g/m^3$ annually and 35 $\mu g/m^3$ daily whereas the China MEP sets the Chinese standard as 35 $\mu g/m^3$ annually and 75 $\mu g/m^3$ daily.

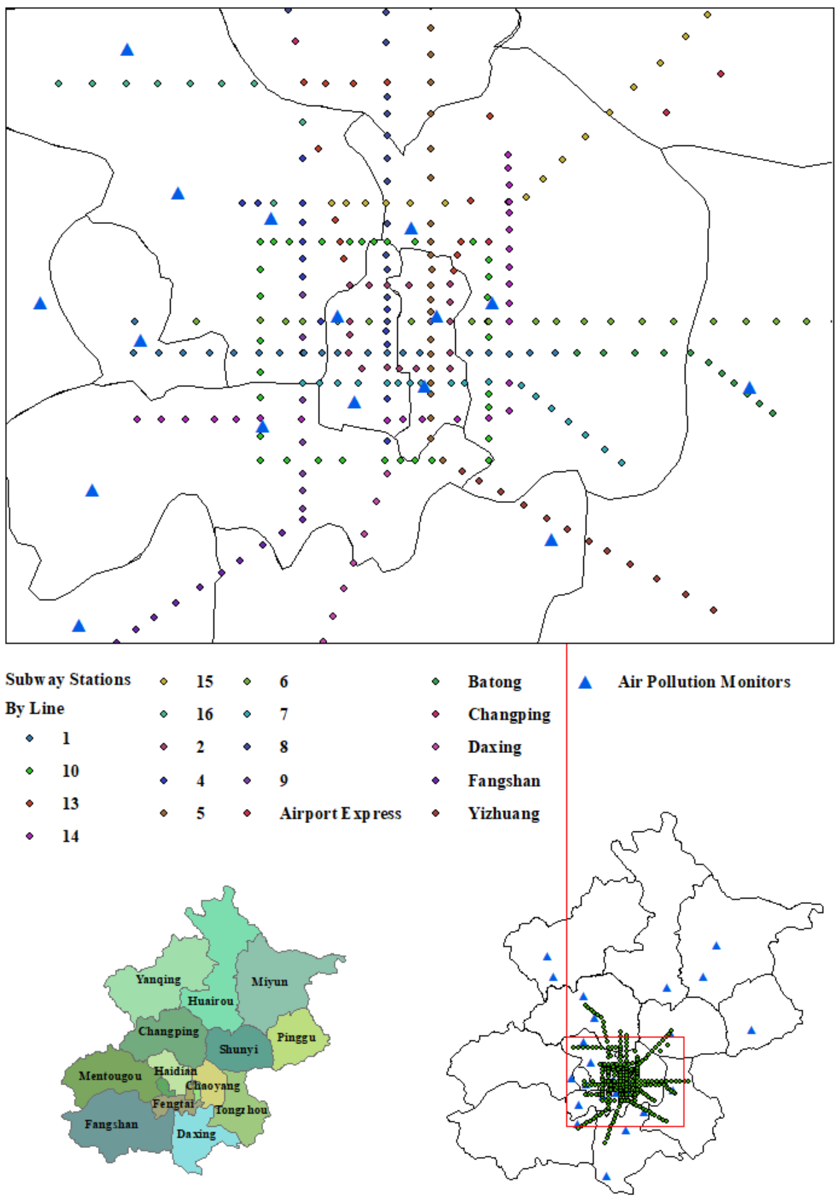
Source: U.S. Embassy and Consulates in China.

Figure 1.2: Beijing Subway Expansion Timeline



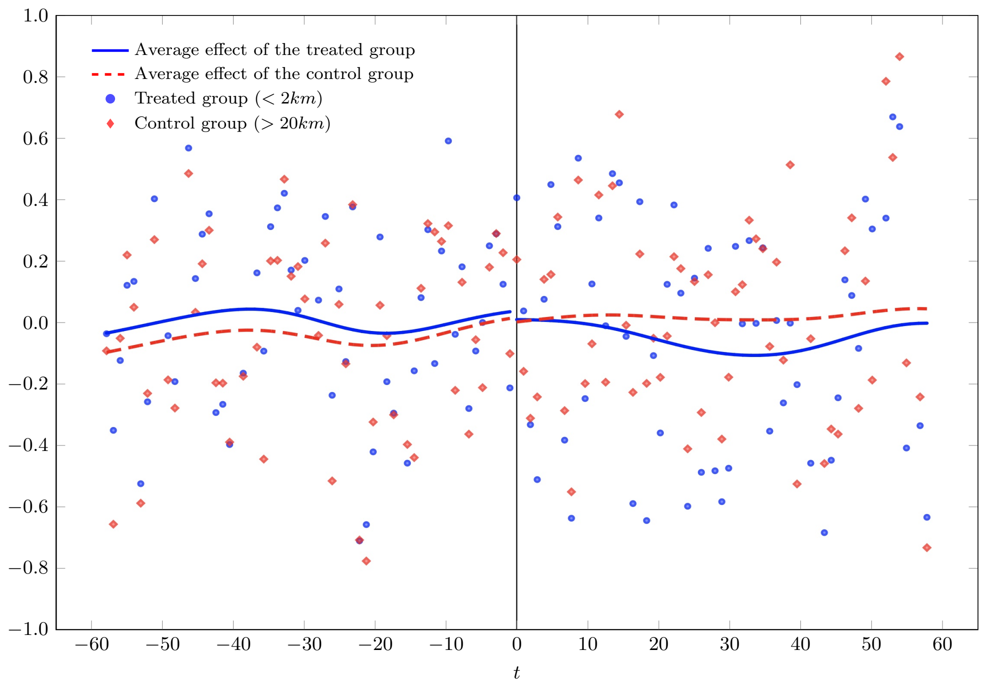
Note: Figure shows the timeline of Beijing subway expansion and the major openings. The number of new subway stations for each opening or expansion is shown in the parentheses and the major openings are shown in bold. We consider the following ten opening dates are major openings: Jul 19, 2008; Sep 28, 2009; Dec 30, 2010; Dec 31, 2011; Dec 30, 2012; May 5, 2013; Dec 28, 2013; Dec 28, 2014; Dec 26, 2015; and Dec 31, 2016. The opening dates within 60 days apart from these major opening are combined with the closest major opening date. From 2008 to 2016, 13 new subway lines and one airport expressway were constructed with a total length of 440 kilometers and 252 new subway stations opened, making the Beijing subway system not only the most rapidly expanded but also the longest in the world.
Source: www.bjstats.gov.cn/xwgb/tjgb/ndgb/201402/t20140213_267744.htm.

Figure 1.3: Air Quality Monitors and Subway Stations



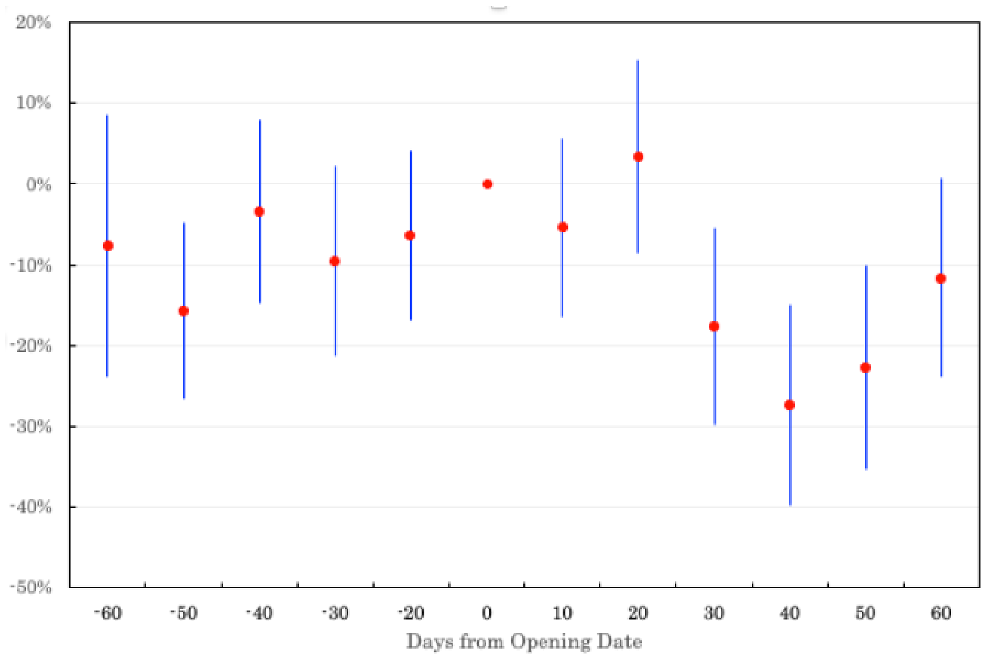
Source: www.bjstats.gov.cn/xwgb/tjgb/ndgb/201402/t20140213_267744.htm

Figure 1.4: Residualized ln(*Air Pollution*) for 60 Days Before and After the Opening



Note: Residualized plots of ln(*Air Pollution*) after controlling for weather conditions, monitor fixed effects, time fixed effects: year, season, day of week and holiday, and monitor-specific time trends.

Figure 1.5: Event Study Analysis of Subway Openings



Note: The estimates in this graph are based on the parallel trend testing analysis with the specification exploiting staggered rollout design (Column 4 from Table 1.6) and are compared with the base interval (10-day window before opening dates).

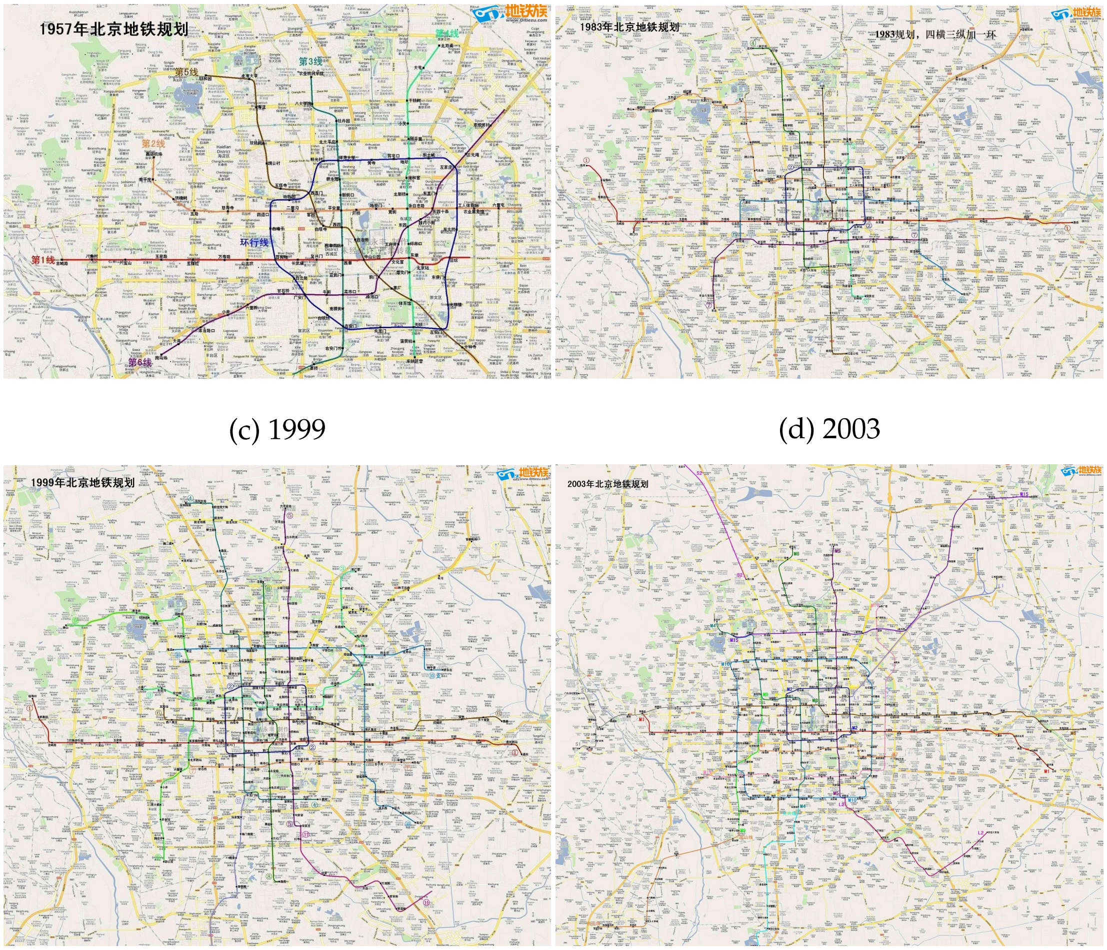
Figure 1.6: Historical Construction Plans of The Beijing Subway System

(a) 1957

(b) 1983

(c) 1999

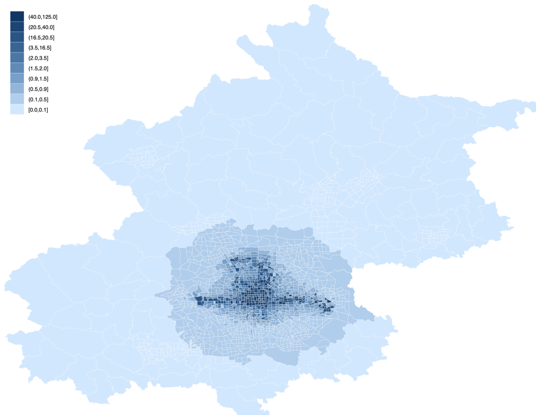
(d) 2003



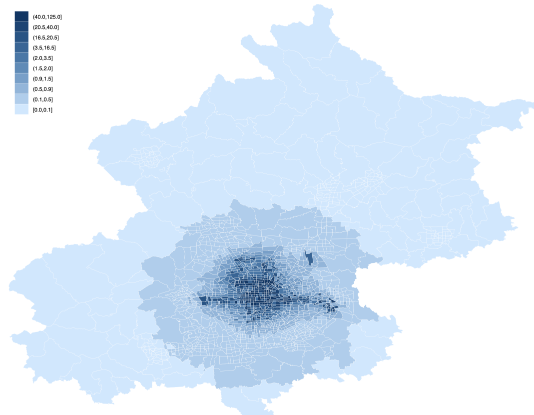
Source: www.ditiezu.com

Figure 1.7: Subway Expansion and Network Density at the TAZ level

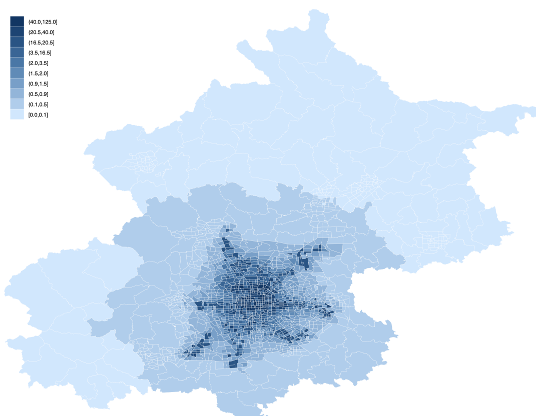
(a) 2007



(b) 2009



(c) 2011



(d) 2013

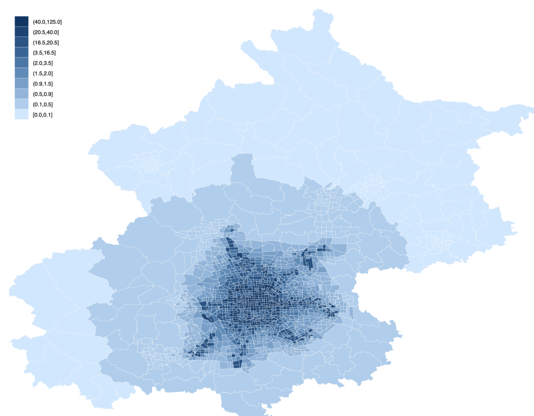


Figure 1.8: Beijing Subway Network Density at the TAZ level as of 2016

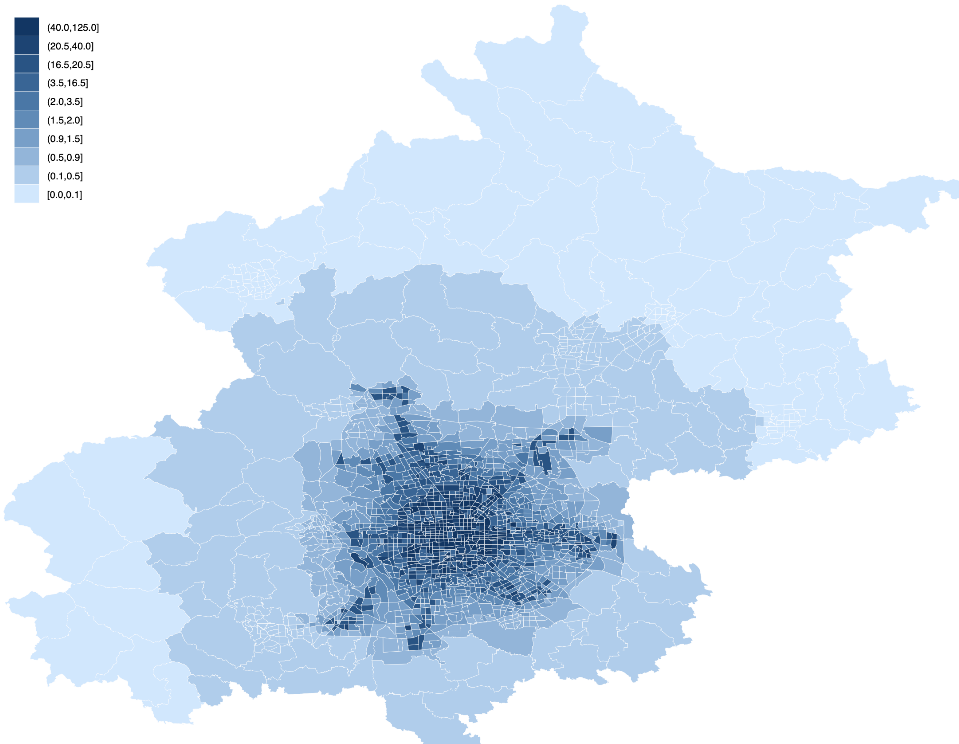


Table 1.1: Variable Descriptions

Variable	Definition
(a) Air Pollution Indicators, monitoring station (i) × daily (t)	
API_{it}	Air Pollution Index ranging from 0 to 500. This index measured between 2008 and 2012 and it accounts for sulfur dioxide (SO_2), nitrogen dioxide (NO_2), suspended particulates (PM_{10}).
AQI_{it}	Air Quality Index ranging from 0 to 500. This index has been measured since 2013 and it accounts for SO_2, NO_2, PM_{10}, $PM_{2.5}$ and O_3.
(b) Subway Density Measures, monitoring station (i) × daily (t)	
$Density_{it}$	Subway network density centered at monitoring station i, which is defined as the total number of stations at time t weighted by the inverse of squared distances from monitoring station i to each subway stations in Beijing.
$\widetilde{Density}_{it}$	Subway network density centered at monitoring station i, which is defined as the total number of stations at time t weighted by both the daily ridership of each subway line and the inverse of squared distances from monitoring station i to each subway stations in Beijing.
$Treated_{it}$	Treated group or treated air pollution monitoring station. 1 if it is treated, 0 otherwise. Air pollution monitoring station i is treated when there is at least one new subway station (j) opened within 2km distance and we keep it as treated for 60 days after the opening, defined as $\mathbf{1}(Post_t) \times \mathbf{1}(Distance_{ij} \leq 2km, j \in N_\tau)$.
$N_{it} \times Treated_{it}$	Heterogeneity of treated group or treated air pollution monitoring station, which counts total number of new subway stations opened within 2km distance and kept as treated for 60 days after the opening, defined as $N_{it} = \mathbf{1}(Post_t) \times \sum_{j \in N_t} \mathbf{1}(Distance_{ij} \leq 2km)$.
(c) Weather Variables, daily (t)	
Air temperature (^oC)	Average daily temperature.
Relative humidity (%)	Average daily relative humidity.
Precipitation (mm)	Total daily rainfall or snowmelt.
Wind speed (km/h)	Average daily wind speed.
Wind direction (cat.)	The vector summation of hourly wind direction with its speed as the the length of each vector.
Rain/Snow/Storm/Fog	Dummy: 1 if there was rain/snow/storm/fog , 0 otherwise.

Table 1.2: Conversion from Pollutants Concentration to API and AQI

Air Pollution Index (API)		Pollutants					
value	level	PM_{10} ($\mu g/m^3$)	$PM_{2.5}$ ($\mu g/m^3$)	O_3 ($\mu g/m^3$)	CO (mg/m^3)	NO_2 ($\mu g/m^3$)	SO_2 ($\mu g/m^3$)
0-50	Excellent	0-50				0-80	0-50
50-100	Good	50-150				80-120	50-150
100-200	Slightly polluted	150-350				120-280	150-800
200-300	Moderately polluted	350-420				280-565	800-1600
300-400	Severely polluted	420-500				565-750	1600-2100
400-500	Severely polluted	500-600				750-940	2100-2620

Air Quality Index (AQI)		Pollutants					
value	level	PM_{10} ($\mu g/m^3$)	$PM_{2.5}$ ($\mu g/m^3$)	O_3 ($\mu g/m^3$)	CO (mg/m^3)	NO_2 ($\mu g/m^3$)	SO_2 ($\mu g/m^3$)
0-50	Good	0-50	0-35	0-100	0-2	0-40	0-50
50-100	Moderate	50-150	35-75	100-160	2-4	40-80	50-150
101-150	Unhealthy for SG	150-250	75-115	160-215	4-14	80-180	150-475
151-200	Unhealthy	250-350	115-150	215-265	14-24	180-280	475-800
201-300	Very unhealthy	350-420	150-250	265-800	24-36	280-565	800-1600
>300	Hazardous	>420	>250	>800	>36	>565	2100-2620

Note: During 2008-2012, the Chinese government adopts the Air Pollution Index (API) which takes into account three pollutants. Starting from 2013, the Chinese government replaces API with Air Quality Index (AQI) which considers $PM_{2.5}$ separately from PM_{10} as a major pollutant, and also Ozone.

Table 1.3: Beijing Subway Expansion and Network Density

Opening date (τ)	Subway line (ℓ)	Subway length (km)	N. of Stations new (N_τ)	N. of Stations total ($\mathcal{N}_\tau$)	Standardized Density non-weighted ($Density_\tau/\sigma$)	Standardized Density ridership-weighted ($\widetilde{Density}_\tau/\widetilde{\sigma}$)
Before 2008	1, 2, 5, 13, BT	140	93	93	0.27	0.27
July 19, 2008	8, 10, AE	57	30	123	0.39	0.44
Sep 28, 2009	4	28	24	147	0.45	0.52
Dec 30, 2010	15, DX, CP, FS, YZ	108	49	196	0.57	0.54
Dec 31, 2011	9	36	19	215	0.62	0.56
Dec 30, 2012	6	70	46	261	0.80	0.75
May 5, 2013	14 (West)	14	9	270	0.82	0.76
Dec 28, 2013	8 (Extension)	7	7	277	0.84	0.77
Dec 28, 2014	7	62	42	319	0.93	0.81
Dec 26, 2015	14 (East)	11	15	334	0.94	0.82
Dec 31, 2016	16	20	11	345	0.96	0.82

Note: The names of suburban subway lines are shown as abbreviation: Airport Express (AE), Batong (BT), Daxing (DX), Changping (CP), Fangshan (FS) and Yizhuang (YZ). There were 93 subway stations operating before our data period. Network density centered at an air pollution monitoring station is defined as the weighted sum of subway weighted by the squared inverse distance from the monitoring station to each subway station operating in the network as of the opening date. It is standardized by dividing its standard deviation. The ridership-weighted density is the reweight of the density by ridership of subway line. Standard deviations of the both densities are $\sigma = 3.58$ and $\widetilde{\sigma} = 29.77$ respectively. All density measures are averaged across monitoring stations for each opening date.

Table 1.4: Summary Statistics

Main variables	Mean	S.D.	Min	Max	N
(a) Air Pollution					
API_{it}	82.84	48.56	5.00	500.00	49103
AQI_{it}	124.64	80.02	8.00	500.00	54939
(b) Subway Density					
$Density_{it}$ (non-weighted)	2.48	3.58	0.01	16.26	297
$\widetilde{Density}_{it}$ (ridership-weighted)	0.19	0.30	0.00	1.34	297
$N_{it} \times Treated_{it}$	0.16	0.69	0.00	6.00	297
(c) Weather variables					
Air temperature (^oC)	12.97	11.39	-15.04	33.05	3533
Wind speed (m/s)	1.97	1.58	0.02	10.21	3533
Precipitation (mm)	1.97	8.82	0.00	262.64	3339
Relative humidity (%)	54.64	20.20	6.97	97.83	3533
Wind direction ($cat.$)	7.95	4.94	1.00	16.00	3533

Note: The air quality panel summarizes the daily Air Pollution Index from 2008-2012 and Air Quality Index since 2013 from 27 air quality monitors in Beijing. The density panel summarizes the daily subway density measures at monitoring station level. The weather panel summarizes the daily, city-level weather conditions.

Table 1.5: Changes in Air Pollution Before and After Openings

	ln(*Air Pollution*)			
	Before	After	Diff.	Diff-in-Diff.
Control	4.428	4.437	0.009	
	(0.008)	(0.008)	(0.011)	
Treated	4.483	4.535	0.052	0.043
	(0.018)	(0.022)	(0.028)	(0.031)
	Residualized ln(*Air Pollution*)			
	Before	After	Diff.	Diff-in-Diff.
Control	0.005	−0.004	−0.009	
	(0.005)	(0.005)	(0.007)	
Treated	0.022	−0.033	−0.055	−0.046
	(0.014)	(0.015)	(0.021)	(0.022)

Note: The top panel shows the sample mean of ln(*Air Pollution*) 60 days before and after each subway line opens. The bottom panel shows the sample means of residualized ln(*Air Pollution*) after after controlling for weather conditions, monitor fixed effects, time fixed effects: year, season, day of week and holiday, and monitor-specific time trends. The treatment group is defined as the monitoring stations within 2km of a new subway line while the control group is defined as the monitoring stations more than 20km away from the new subway line. The standard errors are in parentheses.

Table 1.6: Parellel trend test

	Dependent variable: $\ln(Air\ Pollution_{it})$			
	(1)	(2)	(3)	(4)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau - 60 \leq t < \tau - 50)$	−0.080	−0.099	−0.067	−0.076
	(0.061)	(0.061)	(0.062)	(0.083)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau - 50 \leq t < \tau - 40)$	−0.147***	−0.158***	−0.147***	−0.157***
	(0.043)	(0.044)	(0.045)	(0.056)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau - 40 \leq t < \tau - 30)$	−0.010	−0.022	−0.022	−0.034
	(0.049)	(0.051)	(0.052)	(0.058)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau - 30 \leq t < \tau - 20)$	−0.065	−0.076	−0.088*	−0.095
	(0.050)	(0.050)	(0.052)	(0.060)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau - 20 \leq t < \tau - 10)$	−0.029	−0.044	−0.063	−0.064
	(0.045)	(0.047)	(0.047)	(0.054)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau < t \leq \tau + 10)$	−0.090*	−0.102**	−0.062	−0.054
	(0.048)	(0.049)	(0.045)	(0.056)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau + 10 < t \leq \tau + 20)$	0.016	0.000	0.041	0.034
	(0.053)	(0.053)	(0.051)	(0.061)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau + 20 < t \leq \tau + 30)$	−0.178***	−0.190***	−0.178***	−0.176***
	(0.052)	(0.052)	(0.052)	(0.062)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau + 30 < t \leq \tau + 40)$	−0.256***	−0.267***	−0.277***	−0.274***
	(0.053)	(0.053)	(0.054)	(0.063)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau + 40 < t \leq \tau + 50)$	−0.172***	−0.185***	−0.225***	−0.227***
	(0.057)	(0.057)	(0.056)	(0.064)
$\mathbf{1}(Distance_{ij} \leq 2km) \times \mathbf{1}(\tau + 50 < t \leq \tau + 60)$	−0.044	−0.054	−0.112**	−0.116*
	(0.051)	(0.051)	(0.053)	(0.063)
Time Window (days)	τ±60	τ±60	τ±60	τ±60
Weather Controls	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
Monitor FE	Y	Y	Y	Y
Monitor FE × Driving	N	Y	Y	Y
Monitor FE × Trend	N	N	Y	Y
Staggered Rollout	N	N	N	Y
N	17231	17231	17231	3314
R^2	0.53	0.53	0.54	0.56

Note: Each column reports results from an OLS regression where the dependent variable is ln(*Air Pollution*) and the key explanatory variables are the treatment dummies (the interaction of each 10 days within the 60-day time window around the opening dates and there is a new subway station within 2km from the monitoring station). The control group is the monitors outside 20km. The unit of observation is monitor-day. Column (4) relies on the staggered rollout. The weather controls include daily variables: temperature (C^0), relative humidity (%), precipitation (mm), wind speed (km/h), sets of dummies for wind direction and the interactions with the wind speed , dummies for rain, snow, storm, fog. The time fixed effects include year, season, day-of-week, holiday-of-sample dummies. Parentheses contain standard errors clustered at the day level. Significance: *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

Table 1.7: OLS: The Impact of Subway Network Density on Air Pollution

	Dependent variable: ln($Air\ Pollution_{it}$)			
	(1)	(2)	(3)	(4)
$Density_{it}/\sigma$	0.049***	−0.006***	−0.007***	−0.015***
	(0.001)	(0.002)	(0.002)	(0.003)
Temperature (^0C)	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Relative humidity (%)	0.009***	0.009***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Rainfall/snow (mm)	−0.002*	−0.002*	−0.002*	−0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Wind speed (m/s)	−0.071**	−0.071**	−0.072**	−0.071**
	(0.031)	(0.031)	(0.031)	(0.031)
Constant	4.027***	4.085***	4.086***	4.057***
	(0.073)	(0.074)	(0.079)	(0.080)
Monitor FE	N	Y	Y	Y
Monitor FE × Driving	N	N	Y	Y
Monitor FE × Trend	N	N	N	Y
N	86758	86758	86758	86758

Note: Each column reports results from an OLS regression where the dependent variable is ln($Air\ Pollution$) and the key explanatory variable is the standardized subway network density $Density_{it}/\sigma$. Subway network density in a given location is defined as the weighted sum of subway stations weighted by the squared inverse distance from the location to each subway station in the network. The unit of observation is monitor-day. The weather controls include dummies for daily rain, snow, storm, fog. All columns have controlled for weather, wind directions, and a set of time fixed effects (Year, Season, Day of Week and holidays). Parentheses contain standard errors clustered at the day level. Significance: $^{*}p < 0.1$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 1.8: IV: The Impact of Subway Network Density on Air Pollution

(a) Standardized non-weighted density			
	(1)	(2)	(3)
Dependent variable: ln(*Air Pollution*)	OLS	IV	IV
		Second Stage	
$Density_{it}/\sigma$	−0.015***	−0.020***	−0.028***
	(0.003)	(0.004)	(0.009)
Random Opening Dates		N	Y
		First Stage	
$Density_{it}/\sigma$ (2003 Planning)		0.789***	0.651***
		(0.004)	(0.012)
F-stat		48808	3160

(b) Standardized ridership-weighted density			
	(4)	(5)	(6)
Dependent variable: ln(*Air Pollution*)	OLS	IV	IV
		Second Stage	
$\widetilde{Density}_{it}/\sigma$ (ridership weighted)	−0.026***	−0.024***	−0.035***
	(0.007)	(0.005)	(0.011)
Random Opening Dates		N	Y
		First Stage	
$Density_{it}/\sigma$ (2003 Planning)		0.655***	0.520***
		(0.004)	(0.012)
F-stat		57069	2605

Note: The last two column report results from IV regression where the dependent variable is ln(*Air Pollution*) and the key explanatory variable for Panel (a) is the standardized subway network density, $Density_{it}/\sigma$. Panel (b) shows results with the key explanatory variable as density measure using line ridership as extra weights for the subway stations, $\widetilde{Density}_{it}/\sigma$. Column (2), (3), (5) & (6) report the result from IV regressions with different specifications. The instrument is the subway network density based on the 2003 subway plan map. Column (2) and (5) use the same opening dates for actual subway system and the IV. Column (3) and (6) assign random opening dates for lines in 2003 plan as the 3 months before or after the real opening dates. The unit of observation is monitor-day. All columns control for the daily weather variables: temperature (C^0), relative humidity (%), precipitation (mm), wind speed (km/h), sets of dummies for wind direction and the interactions with the wind speed , dummies for rain, snow, storm, fog; the time fixed effects: day-of-week, quarter-of-year, year, holiday-of-sample dummies; spatial fixed effects: dummies for air pollution monitoring stations and the interactions with the time trend and driving restriction policy dummies. Parentheses contain standard errors clustered at the day level. Significance: *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

Table 1.9: Marginal Impact of Subway Expansion on Air Pollution

Opening date	Cumulative Standardized Density		Marginal Increase in Density		Marginal Reduction in air pollution (%)	
	non-weighted (1)	ridership-weighted (2)	non-weighted (3)	ridership-weighted (4)	non-weighted (5)	ridership-weighted (6)
Before 2008	0.230	0.201	-	-	-	-
July 19, 2008	0.307	0.300	0.077	0.098	0.154	0.236
Sep 28, 2009	0.365	0.366	0.057	0.066	0.115	0.157
Dec 30, 2010	0.432	0.380	0.068	0.015	0.135	0.035
Dec 31, 2011	0.459	0.391	0.027	0.010	0.054	0.025
Dec 30, 2012	0.577	0.515	0.118	0.125	0.237	0.299
May 5, 2013	0.595	0.532	0.017	0.016	0.035	0.039
Dec 28, 2013	0.604	0.535	0.010	0.004	0.020	0.009
Dec 28, 2014	0.697	0.575	0.093	0.040	0.185	0.095
Dec 26, 2015	0.726	0.587	0.029	0.012	0.058	0.029
Dec 31, 2016	0.735	0.589	0.009	0.002	0.018	0.005
Total			0.505	0.387	1.009	0.930

Note: Network density centered at a TAZ is defined as the weighted sum of subway weighted by the squared inverse distance from the centroid of the TAZ to each subway station operating in the network as of the opening date. It is standardized by dividing its standard deviation. The ridership-weighted density is the reweight of the density by ridership of subway line. Standard deviations of the both densities are $\sigma = 16.38$ and $\widetilde{\sigma} = 19.62$ respectively. All density measures are averaged over TAZs for each opening date.

Table 1.10: Difference-in-Difference Estimates with a Fixed Time Window

	Dependent variable: ln(*Air Pollution*)					
	Without Monitor FE			DID		
	(1)	(2)	(3)	(4)	(5)	(6)
$Treated_{it} \times \mathbf{1}(Post_t)$	0.105***	0.082***	0.099***	−0.073***	−0.075***	−0.077***
	(0.026)	(0.020)	(0.013)	(0.019)	(0.019)	(0.018)
Temperature (^0C)		−0.011***	0.010***	0.010***	0.010***	0.012***
		(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Relative humididy (%)		0.008***	0.015***	0.015***	0.015***	0.015***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Precipitation (mm)		−0.007*	−0.006*	−0.006*	−0.007*	−0.006
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Wind speed (m/s)		−0.078*	−0.105***	−0.104***	−0.106***	−0.103***
		(0.042)	(0.034)	(0.034)	(0.034)	(0.034)
Constant	4.434***	4.144***	3.727***	3.846***	3.854***	3.765***
	(0.018)	(0.104)	(0.140)	(0.141)	(0.153)	(0.153)
Time Window (days)	$\tau\pm60$	$\tau\pm60$	$\tau\pm60$	$\tau\pm60$	$\tau\pm60$	$\tau\pm60$
Weather controls	N	Y	Y	Y	Y	Y
Wind Directions	N	Y	Y	Y	Y	Y
Wind Directions × Speed	N	Y	Y	Y	Y	Y
Year FE	N	N	Y	Y	Y	Y
Season FE	N	N	Y	Y	Y	Y
Day of Week FE	N	N	Y	Y	Y	Y
Monitor FE	N	N	N	Y	Y	Y
Monitor FE × Driving	N	N	N	N	Y	Y
Monitor FE × Trend	N	N	N	N	N	Y
N	18214	17231	17231	17231	17231	17231
R^2	0.00	0.29	0.45	0.52	0.53	0.54

Note: Each column reports results from an OLS regression where the dependent variable is ln(*Air Pollution*) and the key explanatory variable the interaction of treatment and post-opening. Columns (4) to (6) show the DID estimates with different sets of controls. The treatment group is defined as the monitoring stations within 2km of a new subway line while the control group is defined as the monitoring stations more than 20km away from the new subway line. The unit of observation is monitor-day. The weather controls include dummies for rain, snow, storm, fog. Parentheses contain standard errors clustered at the day level. Significance: *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

Table 1.11: Difference-in-Difference Estimates with Varying Time Windows

	Dependent variable: $\ln AQI$					
	(1)	(2)	(3)	(4)	(5)	(6)
$Treated_{it} \times \mathbf{1}(Post_t)$	−0.046	−0.031	−0.029	−0.052***	−0.057***	−0.052***
	(0.038)	(0.025)	(0.020)	(0.018)	(0.016)	(0.015)
Time Window (days)	$\tau\pm10$	$\tau\pm20$	$\tau\pm30$	$\tau\pm40$	$\tau\pm50$	$\tau\pm60$
	(7)	(8)	(9)	(10)	(11)	(12)
	−0.066***	−0.062***	−0.075***	−0.047***	−0.022	−0.015
	(0.014)	(0.013)	(0.013)	(0.016)	(0.016)	(0.016)
Time Window (days)	$\tau\pm70$	$\tau\pm80$	$\tau\pm90$	$\tau\pm100$	$\tau\pm110$	$\tau\pm120$
	(13)	(14)	(15)	(16)	(17)	(18)
	−0.008	−0.007	−0.009	−0.009	−0.019	−0.023
	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Time Window (days)	$\tau\pm130$	$\tau\pm140$	$\tau\pm150$	$\tau\pm160$	$\tau\pm170$	$\tau\pm180$

Note: Each column reports results from an OLS regression using different time windows ((1) to (18): $Open_t = \tau\pm10, \tau\pm20, \ldots, \tau\pm60, \ldots, \tau\pm180$-day) where the dependent variable is $\ln(Air\ Pollution)$ and the key explanatory variable is the treatment indicator (the interaction of the time window dummy and the treated group indicator), $Treated_{it} = \mathbf{1}(Post_t) \times \mathbf{1}(Distance_{ij} \leq 2km)$. The May 5th, 2013 opening is dropped from the sample to avoid overlapping events and to extend the time window. The treatment group is defined as the monitoring stations within 2km of a new subway line while the control group is defined as the monitoring stations more than 20km away from the new subway line. The unit of observation is monitor-day. All columns control for the daily weather variables: temperature (C^0), relative humidity (%), precipitation (mm), wind speed (km/h), sets of dummies for wind direction and the interactions with the wind speed , dummies for rain, snow, storm, fog; the time fixed effects: day-of-week, quarter-of-year, year, holiday-of-sample dummies; spatial fixed effects: dummies for air pollution monitoring stations and the interactions with the time trend and driving restriction policy dummies. Parentheses contain standard errors clustered at date level. Significance: $^*p < 0.1$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 1.12: Difference-in-Difference Estimates with Continuous Time Measurement

	Dependent variable: $\ln AQI$			
	(1)	(2)	(3)	(4)
$Treated_{it} \times \mathbf{1}(Post_t)$	0.075**	0.151***	−0.001	0.110***
	(0.036)	(0.056)	(0.026)	(0.041)
$Treated_{it} \times \mathbf{1}(Post_t)$ $\times Days_{it}$	−0.004***	−0.012***	−0.000	−0.006***
	(0.001)	(0.004)	(0.000)	(0.002)
$Treated_{it} \times \mathbf{1}(Post_t)$ $\times Days^2_{it}/100$		0.013*		0.005***
		(0.007)		(0.001)
Time Window (days)	$\tau\pm60$	$\tau\pm60$	$\tau\pm120$	$\tau\pm120$
N	15467	15467	30933	30933
R^2	0.56	0.56	0.47	0.47

Note: Each column reports results from an OLS regression where the dependent variable is ln(*Air Pollution*). The treatment group is defined as the monitoring stations within 2km of a new subway line while the control group is defined as the monitoring stations more than 20km away from the new subway line. The unit of observation is station-day. All columns control for the daily weather variables: temperature (C^0), relative humidity (%), precipitation (mm), wind speed (km/h), sets of dummies for wind direction and the interactions with the wind speed , dummies for rain, snow, storm, fog; the time fixed effects: day-of-week, quarter-of-year, year, holiday-of-sample dummies; spatial fixed effects: dummies for air pollution monitoring stations and the interactions with the time trend and driving restriction policy dummies. Parentheses contain standard errors clustered at date level. Significance: *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

Table 1.13: Difference-in-Differences Estimates with Heterogenous Effect

	Dependent variable: ln(*Air Pollution*)			
	(1)	(2)	(3)	(4)
$N_{it} \times Treated_{it} \times \mathbf{1}(Post_t)$	−0.020***	−0.024***	−0.032***	−0.041**
	(0.007)	(0.007)	(0.007)	(0.018)
Temperature (^0C)	0.010***	0.010***	0.012***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)
Relative humididy (%)	0.015***	0.015***	0.015***	0.017***
	(0.001)	(0.001)	(0.001)	(0.001)
Precipitation (mm)	−0.006*	−0.007*	−0.006	−0.009**
	(0.004)	(0.004)	(0.004)	(0.004)
Wind speed (m/s)	−0.104***	−0.106***	−0.103***	−0.130***
	(0.034)	(0.034)	(0.034)	(0.038)
Constant	3.815***	3.826***	3.738***	3.258***
	(0.140)	(0.152)	(0.152)	(0.271)
Time Window (days)	$\tau\pm60$	$\tau\pm60$	$\tau\pm60$	$\tau\pm60$
Weather Controls	Y	Y	Y	Y
Wind Directions	Y	Y	Y	Y
Wind Directions × Speed	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Season FE	Y	Y	Y	Y
Day of Week FE	Y	Y	Y	Y
Monitor FE	Y	Y	Y	Y
Monitor FE × Driving	N	Y	Y	Y
Monitor FE × Trend	N	N	Y	Y
Staggered Rollout	N	N	N	Y
N	17231	17231	17231	3314
R^2	0.52	0.53	0.54	0.55

Note: Each column reports results from an OLS regression where the dependent variable is ln(*Air Pollution*) and the key explanatory variable is the interaction of treatment, post-opening, and number of new subway stations within 2km of each monitor. The control group is defined as the monitoring stations more than 20km away from the new subway line. The unit of observation is station-day. Column (4) relies on the staggered rollout. The weather controls include dummies for rain, snow, storm, fog. Parentheses contain standard errors clustered at date level. Significance: $^*p < 0.1$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 1.14: Cost-Benefit Analysis of Subway Expansion

Opening Date	Total Cost Billion $	Health Benefit Billion $ lower VSL=2.3	Health Benefit Billion $ upper VSL=8.7	Health Benefit % of Cost lower VSL=2.3	Health Benefit % of Cost upper VSL=8.7	Congestion Benefit Billion $ lower VOT=0.5	Congestion Benefit Billion $ upper VOT=1.0	Congestion Benefit % of Cost lower VOT=0.5	Congestion Benefit % of Cost upper VOT=1.0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(a) 10 Years of Operation									
Jul 19, 2008	5.69	0.08	0.26	1.45	4.58	2.69	5.37	47.28	94.46
Sep 28, 2009	3.61	0.06	0.20	1.79	5.65	2.69	5.37	74.52	148.66
Sep 30, 2010	7.05	0.08	0.25	1.14	3.61	2.69	5.37	38.16	76.23
Sep 31, 2011	5.19	0.03	0.10	0.60	1.90	2.69	5.37	51.83	103.56
Sep 30, 2012	10.37	0.13	0.42	1.28	4.04	2.69	5.37	25.94	51.80
May 5, 2013	3.15	0.03	0.08	0.84	2.66	2.69	5.37	85.40	170.51
Sep 28, 2013	1.96	0.02	0.05	0.77	2.43	2.69	5.37	137.24	274.73
Sep 28, 2014	11.58	0.15	0.47	1.28	4.04	2.69	5.37	23.23	46.39
Sep 26, 2015	2.94	0.04	0.14	1.49	4.70	2.69	5.37	91.50	182.43
Sep 31, 2016	4.81	0.01	0.04	0.26	0.81	2.69	5.37	55.93	111.73
Total	56.34	0.64	2.01	1.13	3.57	26.90	53.70	63.10	95.34
(b) 20 Years of Operation									
Jul 19, 2008	6.21	0.13	0.40	2.05	6.47	4.15	8.29	66.83	133.50
Sep 28, 2009	4.14	0.10	0.32	2.41	7.62	4.15	8.29	100.24	200.39
Dec 30, 2010	7.57	0.12	0.39	1.64	5.18	4.15	8.29	54.82	109.51
Dec 31, 2011	5.71	0.05	0.15	0.84	2.67	4.15	8.29	72.68	145.18
Dec 30, 2012	10.89	0.20	0.65	1.88	5.94	4.15	8.29	38.11	76.10
May 5, 2013	3.67	0.04	0.13	1.11	3.52	4.15	8.29	113.08	225.65
Dec 28, 2013	2.48	0.02	0.07	0.94	2.96	4.15	8.29	167.34	334.42
Dec 28, 2014	12.10	0.23	0.72	1.89	5.96	4.15	8.29	34.30	68.51
Dec 26, 2015	3.47	0.07	0.21	1.95	6.16	4.15	8.29	119.60	239.04
Dec 31, 2016	5.33	0.02	0.06	0.36	1.13	4.15	8.29	77.86	155.51
Total	71.22	0.98	3.11	1.38	4.36	41.50	82.90	84.49	116.41

Note: All the monetary terms are in 2015 dollars and discounted by an annual discount rate of 5%. The total cost includes both the construction cost and the operating cost. The construction cost accounts for 82.9% of the total cost during a 10-year period for the lines in the sample period and 65.6% for the period of 20 years. The health benefit includes the saving from mortality and morbidity costs. The lower bound health benefit calculations are based on the Value of a Statistical Life (VSL) of $2.3 million (in 2015) as in Ashenfelter and Greenstone (2004). The upper bound health benefits are based on the central estimate of $8.7 million as recommended by U.S. EPA. The savings from congestion relief is calculated based on the reduced time delay by subway opening using estimates from Yang et al. (2018). The lower bound of congestion cost saving assumes the value of time (VOT) to be 50% of the wage, and the upper bound assumes 100% of wage as the VOT.

CHAPTER 2

TRANSPORTATION AND THE ENVIRONMENT IN DEVELOPING COUNTRIES

2.1 Introduction

Air pollution and climate change represent serious threats to human health. In 2016, air pollution was responsible for approximately 7 million deaths from various life-shortening diseases, including heart disease, lung cancer, and stroke, according to the World Health Organization (WHO). [1] The public health and economics literatures have established that air pollution increases mortality, especially among the most vulnerable groups, including infants and older adults, and that it leads to large morbidity costs. [2] Climate change is expected to cause far-reaching and sweeping economic and societal changes that are likely to affect agriculture, biodiversity, economic growth, geopolitics, human health, and world peace. [3]

Many developing countries, especially rapidly growing countries, are experiencing pressing environmental challenges as a result of the dramatic increase in fossil fuel consumption to meet the need for consumption and production, limited access to clean technologies, and the lack of stringent and

[1] WHO Global Ambient Air Quality Database: https://www.who.int/airpollution/data/cities/en/.

[2] Studies on the mortality impact of air pollution include Chay and Greenstone (2003), Currie and Neidell (2005), Currie and Walker (2011), Knittel et al. (2016), and Deryugina et al. (2019). Studies on morbidity costs include citetMoretti2011, Deschenes et al. (2017), Barwick et al. (2019), and Williams and Phaneuf (2019).

[3] Recent papers on the impacts of climate change include Nordhaus (2006) and Dell et al. (2012) on economic growth; Mendelsohn et al. (1994), Schlenker et al. (2005), DeschÃłnes and Greenstone (2007), and Burke and Emerick (2016) on agriculture; DeschÃłnes and Moretti (2009), DeschÃłnes and Greenstone (2011), and Barreca et al. (2016) on mortality; and Miguel et al. (2004),Hsiang et al. (2011), and Jia (2014) on social conflict.

well-enforced environmental regulations. The populations in these countries are particularly vulnerable to adverse environmental conditions because of the lack of effective government interventions and the costs of and limits on options available to individuals to prevent or mitigate the effects of pollution.

Figure 2.1a depicts the level of fine particulate matter ($PM_2.5$) across the globe, illustrating that concentrations tend to be higher in low- and middle-income countries. The United States and Japan were historically the worldâĂŹs major contributors of carbon dioxide (CO_2) emissions. Figure 2.1b shows that over the past two decades, emissions from developing countries such as China and India surged to catch up. In 2006, China surpassed the United States to become the worldâĂŹs largest emitter of CO_2.

Rapid urbanization in developing countries presents both challenges and opportunities in addressing environmental challenges (Kahn 2006). The worldâĂŹs urban population increased from less than 1.4 billion (or 36%) in 1960 to nearly 4.2 billion (or 55%) in 2018. By 2050, over two-thirds of the worldâĂŹs population are projected to live in urban areas, and the rural to urban migration during this process will mostly occur in developing countries. On the one hand, the high concentration of people and activities in cities could lead to severe traffic congestion and exacerbate air pollution, especially with the rise in vehicle ownership in emerging economies. On the other hand, cities have the potential to organize economic activities spatially to reduce energy consumption and environmental impacts and to better take advantage of the economies of scale in public transit. Understand the role of transportation in addressing urban environmental challenges has important implications for policy design to foster the emergence of green cities.

The transportation sector, which relies heavily on fossil fuels, is a major source of air pollution and greenhouse gas (GHG) emissions. The WHO estimates that road transport contributes 30% of particulate emissions in European cities and up to 50% in member countries of the Organisation for Economic Co-operation and Development (OECD). (Carter 2019) According to the US Environmental Protection Agency (EPA), the transportation sector is responsible for about 10% of PM2.5, more than 55% of nitrogen oxide (NOx), and about 10% of volatile organic compound (VOC) emissions in the United States.

Because of the rapid rise in private vehicle ownership and travel demand, as well as the relatively low fuel efficiency in developing countries, the transportation sector plays an increasingly significant role in local air quality. Figure 2.2 shows the increase in new passenger vehicle registrations in selected countries from 2005 to 2017. Among developed countries, new vehicle sales were stable or declined slightly during this period. By contrast, China and India experienced dramatic increases in vehicle ownership, with total new passenger vehicles in China increasing fivefold. Although per capita vehicle ownership is still relatively low in developing countries, and total gasoline consumption is still far behind that in the United States (Figure 2.3), the upward trend in these countries is substantial.

As household income rises in developing countries, the need for travel and the desire for automobile ownership grow, and willingness to pay for a cleaner environment increases as well. In order to strike a balance between these competing incentives, policy makers need to recognize that (a) automobile usage generates several types of externalities (including pollution, congestion, noise, accidents, and road damage) that need to be addressed by policy interventions(Parry et al. 2007), and (b) automobiles and transportation infrastructure are durable goods. Short-term decisions on vehicle purchase and transportation network design can have far-reaching implications for emissions trajectory for decades to come. Government policies need to be forward looking and take long-run household behavioral responses into account.

A suite of policy tools has been used to reduce urban air pollution from automobiles. These tools include demand- and supply-side policies that aim to encourage travel mode shifts, reduce emission intensity levels via fuel-economy and emission standards, and promote alternative fuels or zero-emission technologies. These policies can also be distinguished as command-and-control or market-based instruments. This article reviews recent research on each of these policy tools, with a focus on their application in developing countries. It discusses the cost-effectiveness of each policy in addressing the pollution challenge, giving special attention to the empirical challenges in identifying the causal impacts of policies. [4]

2.2 Policies to Promote Modal Shifts

2.2.1 Expansion of Public Transit

Faced with increasing air pollution and traffic congestion, local governments use the expansion of road and public transit networks as the first line of de-

[4]Because the choice of travel modes is also tied to housing and job location choices, smart urban planning can play an important role in curbing car-related emissions by reducing travel distances or eliminating the need to travel altogether. This review does not examine the cost-effectiveness of urban policies. Section 5 briefly discusses actions that can help improve system-wide efficiency, such as ridesharing and autonomous driving.

fense. In Beijing, the government invested more than $67 billion in transportation infrastructure between 2007 and 2015, greatly expanding the public transit network by adding 14 new subway lines and more than 200 bus routes (3,300 new buses).[5] Similar expansions are happening in India, Mexico, and many other emerging economies. [6] Although it requires massive funding, this type of investment can also stimulate economic activities and facilitate trade (Redding and Turner 2015).

With the goal of reducing traffic-related air pollution and traffic congestion, supply-side policies such as expanding the public transit network can create two countervailing forces on air quality. Improving the public transit network can divert commuters from driving private vehicles to public transport (Mohring 1972). This traffic diversion effect (the Mohring effect) could potentially reduce traffic congestion and vehicle emissions. However, improving transportation infrastructure (by increasing road capacity or enhancing public transit) can reduce the cost of travel and driving, resulting in an increase in travel demand and driving (Vickrey 1969) and leading to more pollution. Duranton and Turner (2011) find that traffic volume increased as a result of the expansion of highway capacity in US cities between 1983 and 2003. Although the expansion of road capacity can initially reduce traffic congestion and air pollution, it increases travel demand in the long run, eroding the initial improvement in traffic conditions.

Given the high cost of transportation infrastructure and potential countervailing forces at play, empirically estimating the impact of supply-side policies

[5]Beijing Transport Annual Report (Beijing Transport Institute).

[6]India Road Investment. (India Brand Equity Foundation); Mexico Transport Infrastructure Development. (Oxford Business Group)

on air quality is important. The central challenge lies in finding exogenous variation in public transit infrastructure, which could be confounded with other unobserved factors. For example, urban planners may situate public transit (such as the subway) in areas where population and economic activities are projected to grow. In this case, air quality in those areas may have deteriorated in the counterfactual scenario of no expansion of public transit. The issue of endogenous location could bias the true impact of subway expansion in the empirical analysis.

To tackle the identification challenge, researchers have used the regression discontinuity (RD)-in-time, difference-in-differences (DID), and event study (ES) approaches, which rely on different identification assumptions. The key assumption behind the RD-in-time and ES approaches is that no unobservables exhibit discrete changes at the time of treatment (e.g., subway opening), so as not to confound the impact of the treatment. [7] The key identification assumption behind DID is the parallel trend assumption, i.e., unobservables do not affect the treatment and control groups differently in the absence of the treatment.

Chen and Whalley (2012) estimate the effects of opening one subway line in Taipei on air pollution based on the RD-in-time framework. They find that opening the subway line reduced carbon monoxide (CO) emissions by 5âĂŞ15%. Employing an approach similar to that used by Chen and Whalley

[7] The RD-in-time method uses time as the running variable and assumes that the impact of unobservables on air quality can be captured by flexible functions of the time trend. The identification relies mainly on time-series variation, different from the traditional RD in cross-sectional settings. The RD-in-time, in essence, is the same as the event study method which explicitly uses pre- and post- event data for identification. Hausman and Rapson (2018) discuss the pitfalls and recommendations for addressing them when applying the RD-in-time method.

(2012), Goel and Gupta (2017) use the RD-in-time method to examine the impact of the Delhi Metro expansion on air quality. They find a 34% localized reduction in CO in the short run. Using an ES design, Gendron-Carrier et al. (2018) examine 43 cities across the world that opened new subway systems between 2000 and 2014. They find that particulate concentrations dropped by 4% on average following the opening of a new subway system. Zheng et al. (2019) use the DID method to estimate the impacts of the opening of the first subway line in Changsha, China. They find an 18% reduction in CO in areas close to subway stations.

A strategy to deal with the endogenous location concern of the public transit is the instrumental variable (IV) method. The instrument should provide variation that affects location choices but is exogenous to contemporaneous shocks to air pollution and other outcome variables. Baum-Snow (2007) uses planned routes (many of which were not built) as the IV for US highways to examine the trend of suburbanization, and Faber (2014) constructs a hypothetical highway system in China based on historical planning maps using a minimum spanning tree (MST) approach to examine trade integration and industrialization. Li et al. (2019) follow a similar strategy to examine the impact of BeijingâĂŹs subway system, using the original planning routes as the IV.

An additional empirical challenge in this literature is accounting for the spillover effect of the transportation network. Local changes in road or subway networks could have a system-wide impact, making it difficult to find a valid control group in the DID framework. Li et al. (2019) employ a continuous measure of subway network density as the key regressor to estimate the citywide effect. The network density varies across space within a city and over time for a given location as a result of the expansion of the subway network. A new subway line would more sharply increase network density in adjacent areas than in areas that are farther away. Using the predetermined planning map as an IV, the authors estimate the effect of subway expansion on air pollution for the rapid build-out of 14 subway lines in Beijing from 2008 to 2016. They contrast the estimates based on this approach with those from a distance-based DID approach. The DID approach focuses on the local effect and provides a larger estimate; the network density approach allows for the spillover effect across the whole network and relies on the assumption that the impact diminishes over distance.

Several other studies examine the impact of expanding bus and railway services on air quality or traffic congestion. Anas and Timilsina (2009) use a simulation model to study the lock-in effects of transportation infrastructure in Beijing. They find that increasing bus services in the city center would reduce overall CO2 emissions and that expanding suburban roads would increase them. Lalive et al. (2013) and Bel and Holst (2018) show that increasing rail services in Germany and expanding bus rapid transit (BRT) services in Mexico City reduced emissions of pollutants such as CO, NOx, and PM2.5. Three studiesâĂŤby Silva et al. (2012) in Brazil, Anderson (2014) in Los Angeles, and Bauernschuster et al. (2017) in GermanyâĂŤtake advantage of exogenous variations in public transit supply created by strikes of public transit workers to show that decreased public transit use led to more air pollution and traffic congestion.

Although improving transportation infrastructure is necessary to address traffic congestion and promote economic activities, it is unlikely to be a cost-effective way to improve environmental quality. Findings from the literature suggest that expanding subways has at best a modest effect on reducing air pollution in the short run and that these effects may erode over time as travelers adjust their travel behavior. Beaudoin and Lin-Lawell (2017) and Rivers et al. (2020) find no evidence of air quality improvement from the expansion of public transit. In fact, Beaudoin and Lin-Lawell find that the increase in US public transit supply between 1991 and 2011 led to a small deterioration in overall air quality, especially for NO2 and PM10. Li et al. (2019) estimate that the benefit from pollution reduction generated by the rapid subway expansion in Beijing represents only a small fraction of the overall construction and operating costs; the benefit from congestion relief is much larger and of the same order of magnitude as the costs.

2.2.2 Restrictions on Driving and Vehicle Purchase

Governments can use a variety of demand-side policies to incentivize commuters to change their travel behavior (switching from driving to public transit, for example, or driving less during congested hours). This subsection discusses command-and-control approaches (Subsection 2.3 discusses market-based policies).

The command-and-control approach has been widely adopted, especially in developing countries. This approach includes driving and vehicle-purchasing restrictions. The driving restriction (or road space rationing) policy was first introduced in Athens, Greece in 1982; Santiago, Chile was the second city to adopt it in 1986. In 1989, Mexico City started perhaps the longest-running and best-known license-based driving restriction policy. Based on the last digit of the vehicleâĂŹs license plate, the policy restricts about 20% of vehicles from driving on each workday. In 2008, BeijingâĂŹs municipal government adopted the driving restriction policy to prepare for the 2008 Olympic Games. Initially, the government adopted an even-day/odd-day policy, whereby a vehicle could be driven only on an odd or even day, based on its license plate. After the Olympics, the restriction was relaxed so that the license number-based ban applied on only one designated weekday, a policy also used in Mexico City. In recent years, Paris, Rome, Milan, Oslo, and New Delhi imposed temporary driving restrictions to address congestion and air pollution. Many German cities implemented low emission zone policies, which ban high-polluting vehicles from driving in certain areas (Wolff 2014). [8]

Several studies examine the impact of these policies on traffic congestion and the environment. Like the literature on supply-side policies, these studies commonly adopt quasi-experimental strategies, such as the RD-in-time, DID, and ES methods, using identification from both spatial and temporal variations. The empirical findings are mixed, highlighting the importance of understanding competing forces and consumer responses in policy design.

Using the RD-in-time method, Davis (2008) finds that the driving restriction led to worse air quality in Mexico City because the policy incentivized drivers to circumvent the restriction by purchasing a second vehicle, which tended to be older and dirtier. In contrast, the evidence on the environmental impact of BeijingâĂŹs driving restriction policy has been largely positive. Using both RD-in-time and DID methods, Viard and Fu (2015) show that the every-other-day driving restrictions in Beijing led to a 19% reduction in air pol-

[8]See Wolff and Perry (2010) for a review of low emission zone policies in European cities.

lution and that the one-day-a-week restrictions led to a 7% reduction. Zhong et al. (2017) confirm that the driving restriction policy in Beijing reduced both traffic congestion and air pollution and, as a result, emergency room visits also declined. [9]

The difference in the environmental outcomes of the license-based driving restriction between Mexico City and Beijing suggests that an effective policy design needs to pay attention to the broad operating environment, which affects consumer responses to the policy and, ultimately, the effectiveness of the policy. In response to the driving restriction policy, commuters in Beijing have mainly resorted to public transit instead of the purchase of a second vehicle to meet their travel demand (Xu et al. 2015). There are two important institutional differences between Mexico City and Beijing. First, as previously discussed, Beijing has been investing heavily in improving the public transit system since 2007. The expansion of public transit, including subway and buses, provides residents alternative travel modes. Second, the Beijing municipal government adopted two policies that limited householdsâĂŹ ability to purchase a second vehicle. At the time of the driving restriction policy, the Beijing government also implemented a policy to restrict sales and registrations of used vehicles from other cities that did not meet BeijingâĂŹs tailpipe emission standards. In addition, Beijing implemented a quota system on vehicle purchases from 2011 that limited householdsâĂŹ ability to purchase a second (new or used) vehicle. The theoretical model developed by Zhang et al. (2017) highlights the uncertainty of the effects on air quality that result from license plated-based driving

[9]These results are consistent with the findings of Chen et al. (2013), who examine short-term environmental measures, including the driving restriction policy and other policies the Chinese government adopted in preparing for the 2008 Olympic Games. They find a positive but temporary impact of the measures on air quality.

restrictions. They show both theoretically and empirically that the same policy could lead to different outcomes depending on the substitution among travel modes, the purchase of second vehicles, and atmospheric chemistry, which could result in differential impacts across pollutants.

Another command-and-control policy to curb the growth in travel demand is a vehicle quota system. In 1990 Singapore adopted such a policy, allocating licenses (known as certificates of entitlement, COEs) through a monthly auction system. The cap is defined over different categories of vehicles based on engine power. The COE price ran as high as SGD 50,000 (about US$35,000) for large passenger vehicles. In 1994 Shanghai started an auction system to allocate limited vehicle licenses; it switched to an online system with a reservation price in 2008. The monthly cap has been about 10,000 units. The number of bidders per month is about 150,000 to 200,000; the average winning bid was about CNY 90,000 (about $14,000) in recent years.

In 2011, the Beijing municipal government implemented a vehicle quota policy to reduce air pollution and traffic congestion. It uses a lottery system to allocate limited vehicle licenses. The lottery was initially held monthly; since 2014, it has been held every second month. The quota was reduced over time, and the odds of winning decreased substantially. [10] Five other cities in China now have vehicle quota systems based on various allocation mechanisms: Guiyang and Guangzhou adopted license lotteries in July 2011 and August 2012, respectively; Tianjin, Hangzhou, and Shenzhen started to implement a hybrid system in January, March, and December 2014, respectively.

[10] The odds of winning the license plate lottery in Beijing decreased from 1:10 in early 2011 to nearly 1:2,000 in 2018 because the cap tightened, and the pool of lottery participants increased dramatically.

Li (2018) uses a structural econometric model to compare the allocative efficiency and environmental outcomes of auction and lottery systems. The analysis suggests that the lottery system leads to a large welfare loss from misallocation, although it has an advantage over an auction in terms of reducing externalities such as air pollution from automobile usage. [11] The environmental impacts of the purchase restriction policy warrant future research. The short-run impact is likely small; hence, given that the first-order effect of the policy is on the flow rather than the vehicle stock, the short-term impact is hard to empirically detect. The long-run impact could be more significant, but it is harder to identify because there is more room for confounding factors to be at play in a longer time horizon.

2.2.3 Congestion Pricing

The command-and-control approaches that restrict driving or vehicle ownership are not the first-best policies to address environmental and congestion externalities; such policies can lead to unintended consequences (Davis 2008). Market-based policies have gained more traction from policy makers in recent years. This subsection discusses congestion pricing as a market-based policy tool to affect travel behaviors such as travel time, distance, frequency, and modes.

Congestion pricing was first proposed by (Vickrey 1959), who recognized congestion as a classic externality and identified the mispricing of transport resources as its root cause. To maximize the efficiency gain, congestion pricing

[11] Chin and Smith (1997), Koh (2003), Chen and Zhao (2013), and Xiao et al. (2017) have examined the impact of the quota policy on vehicle purchases and consumer welfare in Singapore and China.

can be designed to allow charges to vary by location and time based on the spatial and temporal variation of the congestion externality. To address distributional concerns and further promote the use of public transit, the revenue raised can be used to improve access to and the quality of public transit.

Singapore adopted the first congestion pricing scheme in 1975. Some European cities have adopted area-based congestion pricing (London in 2003, Stockholm in 2006, Milan in 2008, Gothenburg in 2013). In the United States, several area-based schemes were proposed but failed to be implemented over the years. New York stateâĂŹs 2019 budget proposes congestion pricing on vehicles that enter Manhattan below 60th Street. If adopted, New York City would become the first US city to use congestion pricing.[12] Real-time congestion pricing is now technically feasible. Singapore is slated to become the first city to use a GPS-based system in 2020. The more flexible congestion charges raise privacy concerns, however, as they rely on collecting commutersâĂŹ travel information (Parry et al. 2007). [13]

Several studies examine the effectiveness of congestion pricing designs. They include Olszewski and Xie (2005) (Singapore), Beevers and Carslaw (2005) (London), Simeonova et al. (2018) (Stockholm), and Gibson and Carnovale (2015) (Milan). These studies find that the schemes reduce congestion by approximately 10âĂŞ30% and provide significant environmental ben-

[12]Several dozen high-occupancy toll (HOT) lanes with variable or dynamic tolls are operating or planned in the United States. For example, the express lanes on Interstate 66 near Washington, DC charge single-occupancy vehicles a fee that fluctuates according to traffic conditions.

[13]SingaporeâĂŹs electronic road pricing system features about 80 entry points that record passing vehicles around the city. The charges are not based on distance traveled, and they vary only infrequently. The system is being upgraded to a GPS-based system with the ability to incorporate time-varying and location-specific charges.

efits for the priced area. [14] (For a review of studies on the impacts of existing congestion pricing schemes, see Anas and Lindsey (2011) [15])

Congestion pricing has not been implemented in any developing country. Two recent studies attempt to study the potential benefits of adopting such a policy in these countries. Using real-time, fine-scale traffic data from Beijing, Yang et al. (2020) analyze the relationship between traffic density and speed. They estimate the optimal time-varying and location-specific congestion charges to be between CNY 0.05 and CNY 0.39 per kilometer; they conclude that the pricing scheme could help relieve peak-hour traffic congestion and lead to annual welfare gains of CNY 1.5 billion.

With a similar focus but a different method, Kreindler (2018) uses GPS data on more than 100,000 commuter trips in Bangalore, India to conduct a randomized experiment for the morning commute. He compares two congestion charge policies that impose fees for driving through certain areas during peak hours. Based on the experimental price variation, he estimates commutersâĂŹ preference for scheduling flexibility relative to their value of time. He concludes that the costs of rescheduling travel to inconvenient times will almost entirely offset the benefits of the saved travel time, resulting in only a small consumer welfare gain.

[14]Drivers may respond to the charges by driving around the priced area. Such behaviors may lead to more traffic and emission outside the area, as Gibson and Carnovale (2015) suggest.

[15]Daniel and Bekka (2000), Bigazzi and Figliozzi (2013), and Fu and Gu (2017) study highway tolls and their environmental impacts. Using data from 98 Chinese cities, and both RD-in-time and DID methods, Fu and Gu (2017) show that eliminating highway tolls increases air pollution by 20% and decreases visibility by 1 km.

2.3 Policies to Promote Alternative Fuel Vehicles

The past two decades have witnessed the rapid development and diffusion of alternative fuel vehicle (AFV) technologies amid heightened concern over energy security and transportation-related air pollution and GHG emissions. AFV technologies include flexible-fuel vehicles (FFVs), hybrid vehicles (HEVs), battery electric vehicles (BEVs), plug-in hybrid electric vehicles (PHEVs), fuel cell vehicles (FCVs), and natural gas vehicles (NGVs). By reducing the consumption of gasoline and running on cleaner fuels, AFV technologies provide potential pathways to mitigate or even eliminate the environmental externalities associated with petroleum consumption.

AFV technologies face common adoption barriers in the early stage of diffusion, such as higher upfront costs, limited model choices, consumersâĂŹ lack of familiarity with the new technology, and the potential undervaluation of future fuel-cost savings. To help speed the diffusion of AFVs, governments in both developed and developing countries have provided various incentives to consumers and automakers. They include both monetary and nonmonetary incentives for purchasing AFVs and mandates and regulations that require automakers to produce AFV vehicles.

This section investigates the effectiveness of various policy tools as evidenced by the findings of recent studies of AFVs. Most of the empirical studies focus on markets in developed countries, but their conclusions and policy implications could be generalized to developing countries.

2.3.1 Subsidies for Adoption of Alternative Fuel Vehicles

The most widely used policies to stimulate the adoption of AFVs rely on tax credits and rebates. Tax credits are usually claimed on a tax return; rebates are either provided after mailing in a proof of purchase or directly deducted upon purchasing an AFV. Governments can also implement vehicle scrappage schemes, in which buyers of energy-efficient vehicles receive rebates if they trade in their old emission-intensive vehicles (Li et al. 2013; Jacobsen and van Benthem 2015).

A large body of empirical studies estimates the effects of subsidies on consumer adoption of AFVs. The stated preference approach was especially popular during the early stage of the diffusion of alternative fuel technology because of the lack of data. A challenge of the stated preference analysis is that the hypothetical purchase environment is often different from the real world, and the choices that respondents make in a survey may not reflect their true preferences in a real vehicle purchase situation, biasing the elasticity estimates. With the increasing availability of sales data and the adoption of real policies, recent studies have used the revealed preference method by exploiting the spatial and temporal variation in market sales of AFVs and incentive policies while controlling for vehicle model characteristics and consumer demographic variables.

The effectiveness and efficiency of subsidy programs hinge on several factors. The first is the lack of additionality: the notion that incentives do not always result in additional AFV sales, because many buyers who receive the subsidy might still have purchased AFVs without it. This problem may be pronounced during the early deployment stage because early adopters of AFVs are consumers who embrace new technologies, and who have the strongest environmental awareness and, usually, higher incomes. Therefore, their purchase decisions do not heavily rely on the provision of subsidies; these people would probably have purchased AFVs without the subsidies.

Various empirical studies document the challenge of additionality. Chandra et al. (2010) argue that the HEV tax rebates offered by Canadian provinces subsidized many consumers who would have bought HEVs in any case. Beresteanu and Li (2011) find that HEV sales in the United States would still be growing rapidly, even without tax incentives. Huse and Lucinda (2014) find that a substantial share of FFV consumers in Sweden would have purchased FFVs in the absence of the cash rebates because of the vehiclesâĂŹ lower operational costs. Xing et al. (2019) find that federal income tax credits for purchasing plug-in electric vehicles (PEVs) [16] in the United States resulted in a 29% increase in PEV sales, but 70% of the credits went to households that would have purchased PEVs without the credits. [17]

Replacing a one-size-fits-all policy with one that targets marginal buyers who are more responsive to the subsidy and would purchase AFVs only with the subsidy could improve effectiveness. Marginal buyers are those who consider the higher upfront cost the only obstacle to the adoption of AFVs or those who view the subsidy amount as sufficient compensation for the utility loss from the other drawbacks of AFVs. Using the EV subsidy receipts data and vehicle transaction prices, Muehlegger and Rapson (2018) show that

[16] PEVs include both battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs).

[17] See DeShazo (2016) for a literature review on the effectiveness of US subsidy programs for PEVs.

EV demand by low- and middle-income households is price elastic and that the pass-through of the subsidy is complete among these consumers. Xing et al. (2019) find that cost-effectiveness is greater for policies that that eliminate or reduce subsidies for high-income households and provide more generous subsidies for low-income households. Improving the targeting of subsidy policies is an important issue and an active area of research in other energy- and poverty-reduction programs (Allcott et al. 2015; Kitagawa and Tetenov 2018).

The second important factor is the need to design a subsidy that pays attention to the heterogeneity in benefits across locations and vehicles. Although PEVs produce little or zero tailpipe emissions on the road, substantial heterogeneity of environmental impacts could exist when factoring in upstream emissions. The environmental advantage of PEVs over conventional vehicles is lower in locations where electricity is generated through fossil fuels. Holland et al. (2016) find considerable heterogeneity in the environmental benefits of PEV adoption in the United States, depending on the location. They therefore argue for a regionally differentiated PEV policy. The environmental benefit of PEVs is the largest in California, where the damage from gasoline vehicles is great, and the electric grid is relatively clean. In contrast, PEVs cause more harm than gasoline vehicles in places such as North Dakota, where electricity is generated mostly from coal. For the many developing countries that rely on coal for electricity generation, the environmental benefit of PEVs is an important empirical question.

A third factor to consider in policy design is that the environmental benefits of AFVs hinge on the amount of gasoline replaced by alternative fuels. This figure is challenging to estimate because consumers who choose to buy

AFVs may be different from others in driving demand. Consumers who purchase AFVs may have greater environmental awareness; they may thus have purchased another fuel-efficient vehicle had they not purchased an AFV. As a result, the reduction in emissions may be small (Xing et al. 2019)). In addition, the ability of AFVs to reduce pollution depends on how many miles AFVs are driven and how many miles would have been driven by gasoline vehicles. Because they have a shorter range and charging is inconvenient, BEVs may not be driven as much as conventional vehicles. Davis (2019) finds that both BEVs and PHEVs are driven considerably fewer miles per year than gasoline vehicles and suggests that PEVs may therefore imply smaller environmental benefits than previously believed.

For AFVs, such as FFVs and PHEVs, that piggyback on gasoline vehicles and can run on both gasoline and alternative fuels, âĂIJfuel arbitrageâĂİ could also weaken the effectiveness of subsidies in reducing emissions. With relatively low gasoline prices, FFV and PHEV drivers are more incentivized to choose gasoline over ethanol or electricity, given the lack of ethanol and electric fueling infrastructure and the inconvenience of refueling. Huse and Lucinda (2014) estimate that CO2 savings would fall by 14% if gasoline usage among FFV drivers increased to 50% and by 18% if such gasoline usage increased to 75%. Salvo and Huse (2013) find imperfect substitutability between gasoline and ethanol among flexible-fuel motorists in Brazil because consumers discriminate among fuel options based on characteristics other than price, including engine performance, the station-stopping cost, and the origin of the fuel. They suggest that substantial investments in consumer education on less-established alternative fuel technologies are required because consumer demand for the âĂIJincumbentâĂİ gasoline is sticky.

To summarize, when designing AFV subsidy policies, policy makers need to account for fuel-switching behaviors and alternative fuel usage. Both factors affect the effectiveness of such policies in reducing emissions. One possible solution is to adjust the AFV subsidy amount based on the frequency of alternative fuel usage when such data are available. Providing valuable fuel price information and accessible price comparison could increase the usage of alternative fuels (Salvo 2018).

2.3.2 Subsidies for Alternative Fueling Infrastructure

FFVs and PHEVs can be fueled at any gasoline station. The diffusion of other AFVs relies heavily on alternative fueling infrastructure, which is limited during the early deployment stage. The interdependence between the building of fueling stations and the adoption of AFVs gives rise to the chicken and egg problem: Consumers are reluctant to adopt AFVs unless there are sufficient alternative fueling stations, but governments and private companies are reluctant to build such stations when few AFVs are on the road. The installation of home charging for PEVs could reduce dependence on public fueling stations; the fueling of FCVs depends entirely on public hydrogen stations. [18]

In addition to providing subsidies to AFV buyers, many governments have been subsiding construction of AFV fueling stations. It is important to under-

[18]FCVs are powered by hydrogen and fueled with pure hydrogen gas from hydrogen fueling stations. They can fuel in less than 10 minutes and have a driving range of about 300 miles. As of October 2019, there were only 41 hydrogen stations in the United States. The FCV market will not witness significant penetration unless the mass deployment of hydrogen stations occurs.

stand whether subsidizing one side of the market is more efficient than subsidizing the other side. Dimitropoulos et al. (2016) find that early adopters of PHEVs are sensitive to changes in the detour time to reach a fast-charging station. They argue that policies that expand fast-charging stations could be an effective stimulus for the early adoption of BEVs, potentially saving public spending for the stimulation of the adoption of electric vehicles.

Li et al. (2017) and Springel (2019) quantify the indirect network effects in the PEV market in the United States and Norway, respectively. Both studies find that the network effects of charging stations on PEV adoption are larger than the effects of PEV stock on investment in charging stations; they therefore suggest that subsidizing charging stations is more effective in speeding PEV diffusion at the initial rollout stage. This finding is likely driven by the fact that early adopters are less price sensitive and more concerned about whether they can conveniently refuel wherever they drive.

At the early stage of a technology deployment, the existence of multiple standards of the complementary service may lead to efficiency loss. Li (2019) finds that unifying the three incompatible standards for charging EVs in the United States would have increased consumer surplus by US$500 million between 2011 and 2015 and allowed car manufacturers to sell 20.8% more EVs.

2.4 Fuel Standards and Emissions Regulation

Instead of directly providing incentives to alter consumer vehicle purchase and driving behavior, governments can impose mandates and regulations on vehicle producers to reduce pollution. This section discusses the main mandates on manufacturers: fuel-economy standards, fuel-content regulations, and tailpipe emission standards.

2.4.1 Fuel-Economy Standards

Many countries adopted fuel-economy standards that require vehicle manufacturers to improve fleet-wide fuel efficiency and provide a minimum level of alternative fuel vehicles. Nine governments, including the United States, Japan, the European Union, and China, have established fuel-economy and GHG emission standards for passenger vehicles. The standard an automaker needs to meet is usually a (sales-) weighted average of the target for each vehicle model in the automakerâĂŹs fleet. Automakers who fail to meet the requirement either pay the penalty or buy regulatory credits from the market under the credit-trading regime. [19]

One argument that supports fuel-economy regulations is that consumers may undervalue fuel economy and fail to adopt fuel-saving technologies. The empirical literature has mixed evidence on the extent to which consumers discount future fuel-cost savings (Busse et al. 2013; Allcott and Wozny 2014; Sallee et al. 2016; Grigolon et al. 2018). Studies that evaluate the efficiency of fuel-economy regulations in reducing gasoline consumption consistently find that gasoline taxes can achieve the same goal at a much lower cost (Goldberg 1998; Austin and Dinan 2005; Jacobsen 2013; Anderson and Sallee 2016).

However, due to the political challenge of increasing taxes and the diffi-

[19] In addition to establishing fuel-economy mandates, some governments require that a certain share of the entire fleet each automaker sells be zero emission vehicles (ZEVs). California, for example, requires that 4.5% of vehicles produced be ZEVs in 2018 and 22% by 2025.

culty of quantifying the marginal social harms, the external cost of gasoline consumption in many countries around the world is not properly reflected by the gasoline tax (Parry and Small 2005). If the regulator decides to implement fuel-economy mandates, there are several lessons from the literature that are relevant for this situation in developing countries. First, with a binding fuel-economy mandate, providing additional AFV subsidy may have little impact on reducing energy consumption or GHG emissions. Fuel-economy mandates essentially increase the cost of producing vehicles that are less fuel efficient and encourage automakers to sell more AFVs. However, when the additional AFV subsidy induces extra AFV sales, the mandate stringency is relaxed, and automakers can thus sell more gasoline vehicles and still maintain compliance. Therefore, the AFV subsidy implicitly subsidizes gas guzzlers, as it makes it easier to sell them. One possible solution is to exclude AFVs from the average fleet fuel-economy calculation so that the mandate takes only gasoline vehicles into account. Second, the fuel-economy mandates in many countries are now attribute based; the stringency of the regulation depends on the vehicleâĂŹs weight or size, and larger and heavier vehicles are subject to a less-stringent requirement. However, this policy design provides an incentive for automakers to increase vehicle size, which could undermine the gains from fuel economy (Whitefoot and Skerlos 2012; Ito and Sallee 2018). Policy makers should be aware of the potential for vehicle substitution across sizes to occur when assigning fuel economy targets for different vehicle segments.

2.4.2 Fuel-Content Regulations and Tailpipe Emission Standards

Implementing fuel-content regulations that restrict the chemical composition of the fuel is another strategy to reduce the harmful pollutants from fuel consumption. Most developed countries enforce the European UnionâĂŹs Fuel Quality Directive, which requires sulfur levels below 10 ppm for both gasoline and diesel vehicles. In contrast, many developing countries still set sulfur limits above the level recommended by the United Nations.

When designing these policies, regulators should be mindful of firmsâĂŹ responses and unintended consequences. Auffhammer and Kellogg (2011) find that the US federal gasoline content regulation, which allowed refiners flexibility in choosing a compliance mechanism, did not improve air quality because refiners lacked incentives to reduce the emission components that are most closely related to ozone formation. By contrast, the standards used in California that better target harmful components are more effective in improving air quality.

Emissions-control systems can be installed that reduce tailpipe emissions per gallon of fuel combusted. Tailpipe emission standards set the maximum amount (grams per mile) of targeted pollutants allowed in exhaust emissions from a fuel combustion engine. A number of countries, including the United States, Canada, Japan, members of the European Union, China, and India, have implemented this type of regulation. Tests conducted at specified intervals measure vehicle emissions, typically Particulate Matter (PM), NOx, CO, and hydrocarbons. Under such regulations, manufacturers may only sell vehicles that comply with standards. In the United States, the EPA manages and implements emission standards. California is allowed to implement more stringent emission standards, which are set by the California Air Resources Board (CARB).

In 2000, both China and India introduced their first emission standards, based on European regulations of that time. Since then, both countries have tightened their standards several times to address serious air pollution in urban areas. Based on the most more stringent European standards (Euro 6) already imposed in the EU, the new standards are slated to go into effect in 2020 in both China and India. The effectiveness and efficiency of these policies remain to be studied.

2.5 Future Research Areas

Many developing economies, especially rapidly growing ones, are facing pressing environmental challenges due to the increased use of fossil fuels for energy and the lack of effective and stringent regulations. As income rises in these countries, demand for environmental quality increases, putting pressure on governments to reevaluate their positions on economic growth and environmental quality. This article reviews recent studies on policies related to road transportation, their impacts on the environment, and the implications for developing countries.

In theory, market-based policies such as congestion pricing and credit trading have efficiency advantages over command-and-control approaches in addressing the externalities associated with transportation. Their applications have been very limited, however, especially in developing countries. Future research could shed light on the pros and cons of different policy instruments while paying attention not only to the efficiency and distributional impacts of existing policies, but also to the emerging opportunities afforded by new technologies and practices. The following questions warrant future research.

First, understanding how consumers in developing countries value fuel economy could help policy makers estimate the efficiency and effectiveness of policy tools such as gasoline taxes and fuel-economy standards to incentivize the purchase of fuel-efficient vehicles. This question is especially important for developing countries, where many vehicle buyers are first-time buyers, and information on fuel economy may not be well understood. If consumers are not well informed, or if they do not pay attention to fuel cost, they may choose vehicles that are less fuel efficient than optimal, providing a justification for government intervention through fuel-economy regulations. [20] As discussed in Section 4.1, there is no consensus on the extent to which consumers undervalue fuel economy in developed countries. Even less evidence is available on this issue in developing countries (Greene 2010). Chugh et al. (2011) find no strong evidence that consumers undervalue fuel economy in India. Comparing the vehicle consumption tax and fuel tax in China, Xiao and Ju (2014) find that increases in the fuel tax decrease total car sales but do not effectively

[20]Even when fully informed about the fuel economy of each vehicle model, consumers may not have the correct perception of the monetarized value of fuel economy, because of âĂIJMPG Illusion,âĂİ in which consumers mistakenly think that fuel costs scale linearly in miles per gallon rather than gallons per mile (Larrick and Soll 2008). The estimated welfare cost of the MPG Illusion is negligible, however, and is not sufficient to justify the current fuel economy regulations (Allcott 2013). Using data from experiments that provide fuel economy information to new vehicle buyers in the United States, Allcott and Knittel (2019) find no impact of the information intervention on consumersâĂŹ vehicle choices; the authors suggest that US fuel economy standards are more stringent than necessary in addressing imperfect information.

encourage consumers to choose fuel-efficient vehicles. When choosing vehicles, consumers are more sensitive to changes in upfront costs than fuel costs. Further studies are needed to understand consumer preferences, information access, and awareness of fuel economy in developing countries.

Second, new technologies related to transportation could offer opportunities as well as challenges for developing countries in addressing environmental issues. Autonomous vehicles have the potential to profoundly transform the transportation sector and the economy in general (Winston and Karpilow 2019). However, a decade may pass before these vehicles comprise a significant share of the overall vehicle market. Thus, related empirical data may not be available for many years. [21] By contrast, ridesharing services are already prevalent in many parts of the world. As travel demand and vehicle ownership increase in developing countries, encouraging ridesharing could potentially help combat severe air pollution and congestion. By increasing the flexibility of travel and providing a new travel mode, ridesharing could increase consumer surplus (Cohen et al. 2016). Its impact on consumer travel behavior and the environment is not well understood. The environmental impact hinges critically on the emissions of the substituted travel modes and on total travel demand.[22]

By exploiting the spatial and temporal variation of Uber entry and Uber

[21] Researchers would need to use projections and simulations to conduct forward-looking analysis on how autonomous vehicles would affect travel behavior, vehicle choices, housing locations, and the broad economy.

[22] If ridesharing mainly encourages carpooling and reduces private car driving, it may reduce overall on-road emissions. If the introduction of this new travel mode increases travel demand, if ridesharing replaces walking or public transit trips, or if deadheading to search for customers (out-of-service movement) accounts for a significant component of vehicle miles, ridesharing may increase on-road pollution and impose new challenges to pollution reduction in the transportation sector.

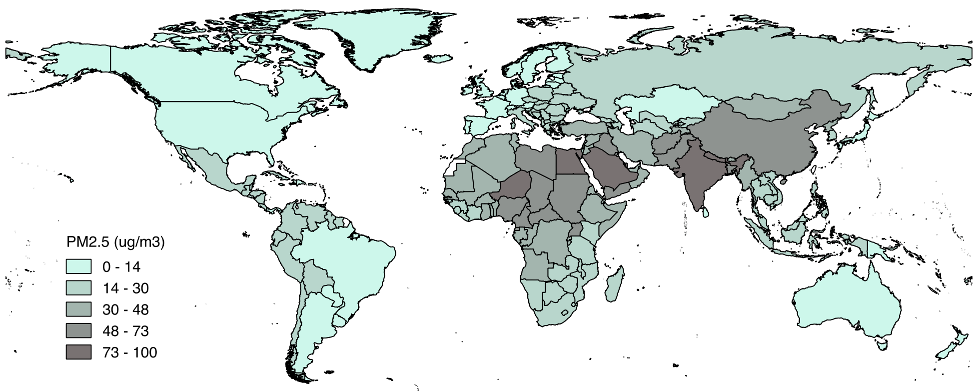
penetration in the US market, Hall et al. (2018) estimate the impact of ridesharing on public transit ridership. They find that Uber complements public transit by solving the last-mile problem of public transit. They suggest that ridesharing could worsen pollution and congestion by increasing the number of car trips without taking the substitution between ridesharing and private vehicle driving into account. However, if ridesharing complements public transit and encourages more consumers to switch from driving private cars to using public transits for the entire trip, ridesharing may help reduce the overall on-road tailpipe emissions and congestion.

In addition, ridesharing may help mitigate air pollution and congestion through other channels. By better matching consumers and drivers, it reduces the time taxi drivers spend finding consumers. Reduced search time on the road could potentially reduce congestion and fuel consumption from the combined taxi and ridesharing market (Hahn and Metcalfe 2017). In addition, people who switch from driving private cars to ridesharing can save time and fuel wasted when finding parking spaces, a problem in more populated cities (Winston 2013). The extent to which ridesharing reduces air pollution and congestion through these channels depends on the substitutability between ridesharing and private driving and taxi riding. Understanding the full impact requires estimating consumer travel mode choice incorporating ridesharing.

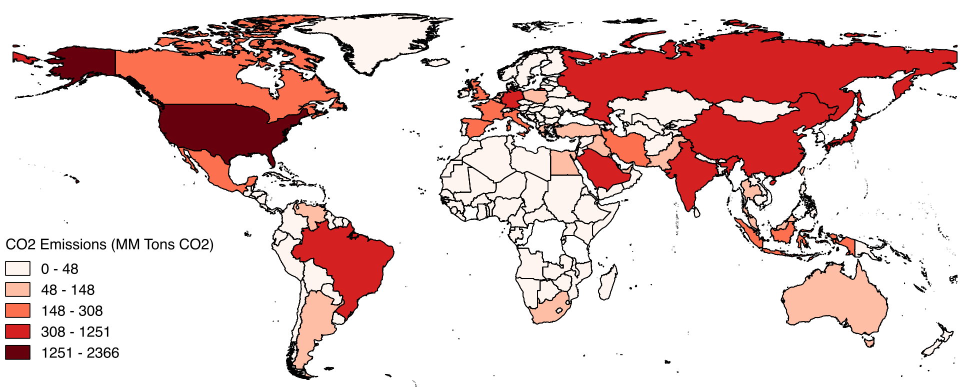
Third, transportation policies could have broad social and economic impacts by changing a variety of household choices. Previous studies have shown that changes in commuting cost can affect labor participation decisions, fertility, and productivity (Duranton and Turner 2012; Black et al. 2014; Liu et al. 2018). Few studies have examined the broad impacts of transportation-related policy beyond immediate goals, especially in developing countries. Understanding household location choices and the general equilibrium impacts of transportation policies could help policy makers better understand their impacts on the environment and urban structure, as well as the distributional consequences. Transportation innovations and policies such as infrastructure expansion and congestion pricing affect the commuting cost and household location choices, as predicted by classical urban models (LeRoy and Sonstelie 1983). The spatial pattern of household locations in turn affects travel choices (travel mode and distance) and the environment. To examine these questions, researchers can employ equilibrium sorting models that incorporate consumer heterogeneity and allow for general equilibrium feedback between economic agents and the environment (Epple and Sieg 1999; Sieg et al. 2004; Kuminoff et al. 2013). These types of models can shed light on the interactions between transportation policies and housing markets, and provide a unified framework with which to analyze and compare the effectiveness of different transportation policies to address the environment, traffic congestion, and social welfare.

Figure 2.1: Fine particulate matter (PM2.5) and greenhouse gas emissions by country

(a) PM2.5 concentration in 2017

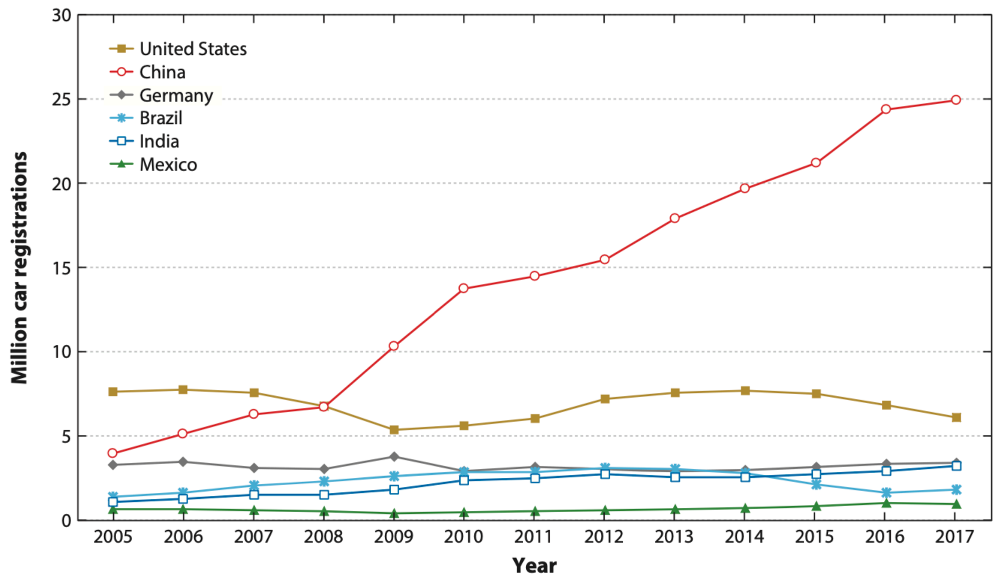


(b) Average CO2 emissions between 2005 and 2016



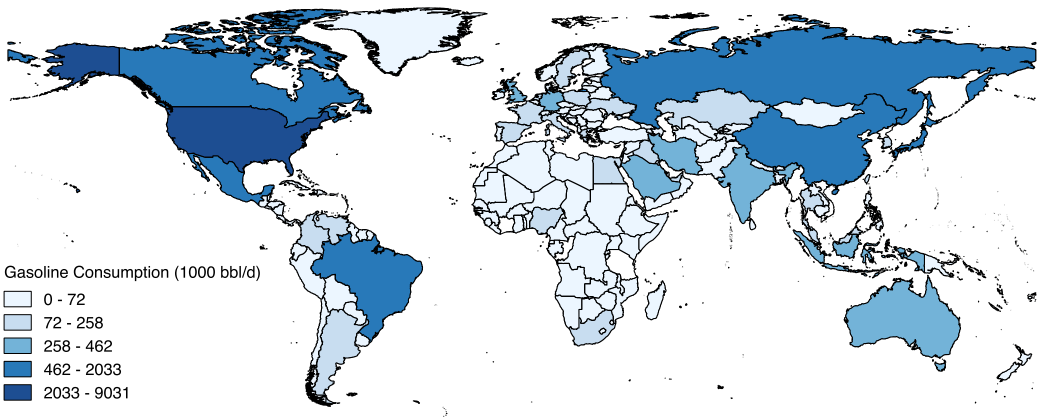
Note: Reconstructed based on data from the World Bank (https://data.worldbank.org/indicator/EN.ATM.PM25. MC.M3) and the US Energy Information Administration (https://www.eia.gov/international/data/world).

Figure 2.2: New passenger car registrations by country and changes from 2005 to 2017



Note: Data from the International Organization of Motor Vehicle Manufacturers (http://www.oica.net).

Figure 2.3: The global distribution of gasoline consumption between 2005 and 2016



Note: Data from the US Energy Information Administration (https://www.eia.gov/international/data/world).

CHAPTER 3

POLLUTION MONITORING, STRATEGIC BEHAVIOR, AND DYNAMIC REPRESENTATIVENESS

3.1 Introduction

Enforcement of and compliance with regulations hinge on accurate measurements of implementation and outcomes.[1] Imperfect monitoring of national regulations can lead to strategic compliance at the local level, which will further bias measurements and cause policy failures. Implementation of national policies at local levels under fiscal and political incentives is a principal-agent problem inherent in the delegation of authority by governments to bureaucratic officials (Aghion and Tirole 1997).[2] Given the ubiquitous information asymmetry between central and local governments, local regulators are likely to implement targeted strategies to meet national policy goals. In the field of environmental regulation, studies have found firms and local governments responding to different regulation stringencies in ways that result in unintended consequences such as pollution spillover (Kahn 2004; Kahn and Mansur 2013; Kahn et al. 2015; Chen et al. 2018; Karplus et al. 2018). For example, Auffhammer et al. (2009) find targeted regulatory efforts in response to nonattainment designations under the Clean Air Act in the U.S., and He et al. (2020) find that Chinese local officials enforce tighter water quality regulations on polluters

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

[2]There exists a rich theoretical literature outlining contracts that align the principal's and agent's incentives (Laffont and Tirole 1993; BÃl'nabou and Tirole 2006). In the political contract between central and local governments, the incentives include monetary incentives such as subsidies and fines, as well as political incentives such as hierarchical assignments of duties and promotions.

immediately upstream of monitoring stations. Thus, an accurate measure of environmental quality that accounts for local regulators' strategic behavior is critical for decentralized regulation enforcement.

Air quality evaluation in major countries around the world is mainly based on stationary, in situ monitors that aim to provide a representative measure of local air quality. China launched a nation-wide, real-time air quality monitoring and disclosure program in 2013. Over 1400 monitors in three staggered waves of cities were quickly built, and air quality in China has greatly improved in the past few years. However, the monitors do not cover the entirety of China. The central government intends to use national policy goals to achieve better air quality but only observe the air pollution at monitored areas. Consequently, the local regulation enforcement tends to target "monitor readings" instead of the actual air quality. Studies find data manipulation issues in China's air quality data before this real-time monitoring was introduced, indicating the importance of "monitor readings" to local regulators. (Andrews 2008; Chen et al. 2012; Ghanem and Zhang 2014) Although better monitoring technologies help improve data quality significantly (Greenstone et al. Forthcoming), strategic responses at local levels can still exist. Previous studies by Zou (2020) and Grainger et al. (2019) have shown firms' and local regulators' strategic behaviors in responding to either the intermittent monitoring schedule or choices of new monitor sites for the monitoring system in the U.S. However, there is a lack of empirical analysis of strategic responses to spatial gaps in monitored areas at the local level. Moreover, previous studies have not examined the monitors' spatial representativeness from a dynamic perspective. Even if the monitor siting was representative ex-ante, strategic responses could invalidate the representativeness ex-post.

In this paper, I leverage high-resolution satellite-based air pollution measures to examine local officials' strategic behaviors in pollution reduction and the implications on dynamic spatial representativeness of ground monitors in China. I use a distance-based Difference-in-Differences analysis with treatment intensity to study the strategic behaviors. The staggered roll-out of the new monitoring system allows cities that joined in different waves to serve as treated and control cities for each other. I then examine the strategic pollution reductions by defining a treatment intensity indicator. The areas near monitors are classified as "monitored" areas, and areas far away from monitors are "unmonitored" areas. I then compare the pollution changes before and after a monitor is opened. In order to learn if such strategic behaviors would change regulatory effectiveness, I examine the spatial representativeness of ground air pollution monitors by comparing population-weighted average pollution levels of an entire city to the city's average pollution based on monitored locations. In doing this, I find that most of the monitors represent the city's average air quality well at the years of monitors roll-out. However, the spatial representativeness is changing over years, indicating spatially differentiated pollution changes within a city.

My paper fills the spatial gaps of ground-level monitoring data by using fine-scale grids data to study the pollution changes over space. The satellite images include annual PM2.5 (fine inhalable particles, with diameters that are generally 2.5 micrometers and smaller) grids at the 1km by 1km resolution (over nine million grids for all of China) from 2000 to 2017.[3] Using the annual

[3]By combining satellite-based measures of AOD with chemical-transport modeling and

level data, I avoid concerns about the missing data in most monthly and daily satellite data. Moreover, the fine-scale grids provide rich spatial variations. This satellite-based PM2.5 data is becoming popular in economic studies because it fills the gaps in ground monitoring networks and validates the data quality at the ground level. (Sullivan and Krupnick 2018; Fowlie et al. 2019) To provide evidence supporting the political incentives behind strategic pollution reductions, I collect data on city characteristics such as population, GDP, etc., as well as information about local officials from the China Political Elite data, which records the local officials' career path, age, and education.

The main finding of this paper is that areas adjacent to monitors experience 6.5% lower PM2.5 concentrations than those farther away, and the results are robust to alternative definitions of monitored and unmonitored groups. [4] The baseline impact of monitoring on overall air pollution is positive (pollution increases), showing that the strategic pollution reduction may lead to pollution leakages to unmonitored areas. I use an event study analysis to show that the parallel trends hold for pre-opening periods in general. Moreover, by including post-opening periods, I find that the difference in pollution becomes larger as the final assessment deadline approaches. [5] My results are robust to placebo tests of random monitor locations and random monitor opening dates. To eliminate the concerns about measurement errors in the satellite-derived PM2.5 data, which may correlate with ground monitors spatially, I also run the same analysis using raw daily satellite Aerosol Optical Depth (AOD) read-

land characteristics, van Donkelaar et al. (2019) derive ground-level concentrations of PM2.5 at high levels of spatial disaggregation.

[4]In the main finding, cells within 3km of a monitor are defied as the monitored area, and cells outside 3km are in the unmonitored group.

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

ings and find robust results.

One additional identification concern may arise from the fact that most monitors are placed in urban centers with poor air quality, so the political interpretation of the results may not be appropriate. Thus, the difference in pollution reduction patterns between the monitored and unmonitored groups may not necessarily be caused by local regulators' strategic responses to stringent environmental targets. Instead, the results could be driven by pollution transported from polluted areas to the cleaner area. Another possibility is that regulators choose to prioritize more polluted areas first instead of gaming the evaluations. I eliminate this type of concern by conducting a heterogeneity analysis in which I compare the strategic pollution reductions for monitors located in dirtier areas to monitors located in cleaner areas of a city. I find no significant impact of monitors being in a polluted area on strategic reductions.

I have conducted heterogeneity analyses to support the political interpretation of strategic pollution reduction. First, I find strong heterogeneity across cities according to the timing of entering the new monitoring program. The later a city joins the monitoring program, the larger strategic responses that are observed. Second, I have also conducted a heterogeneity analysis by cities' pollution compliance levels, where I find a larger strategic reduction in cleaner cities and cities with pollution levels approaching the national standard. Third, cities with younger mayors who have greater promotion chances have larger strategic responses. Lastly, I find that having an economic recession in the previous year shifts local officials' regulation focus from environmental performance to economic growth, and leads to smaller strategic reductions. Taken together, these findings consistently confirm the existence of local officials' strategic pollution reduction, which arises from the misalignment between the national policy goal and local bureaucratic incentives.

Local officials employ a few strategies to reduce pollution near monitors strategically. The next part of the paper discusses the channels through which the spatial differences in pollution reductions occur. The potential channels could include local measures such as directly cleaning the air near monitors or shutting down restaurants and small workshops near monitors, and non-local measures such as relocating polluting sources away from monitors or implementing traffic control. Local pollution reduction measures reduce air pollution in areas adjacent to monitors without increasing pollutions elsewhere, whereas non-local measures will lead to pollution leakages to unmonitored areas. My results suggest that non-local measures dominate, and pollution leaks to areas more than 60km away from monitors. Although there is no data available to test for the mechanisms directly, the political incentives behind the strategic behaviors are strongly supported by government reports, media news, and multiple heterogeneity analyses.

I provide policy suggestions for a better air pollution monitoring system. My analysis of spatial representativeness suggests that most of the monitors are good representations of a city's average air quality at the beginning of monitors roll-out. However, given local officials' strategic responses and the fact that monitor locations are unlikely to change once sited, my simulation of future monitors' representativeness shows that the ground monitoring system will not be representative in the long run. Since ground monitors are costly to build, and the observed strategic response may still exist even with new monitors, it is important for the central government to combine ground monitor readings with external sources of pollution measurements such as satellite, mobile monitors, and public supervision.

This paper makes the following contributions. First, my results highlight the importance of accounting for local regulators' strategic responses when the central government designs national policies. By documenting the gap in pollution reductions for monitored and unmonitored areas, I provide evidence that policies that are ex-ante efficient will not necessarily be efficient with the existence of strategic local responses. My paper is the first empirical study which links the local official's strategic behaviors with the dynamic change in monitor representativeness and examines the underlying political incentives.

My paper adds to the growing literature on the political economy of environmental regulation by highlighting the implementation of national regulations at the local level. (Kahn 2004; Kahn et al. 2015; Jia and Nie 2017; Chen et al. 2018; He et al. 2020) A few of these studies focus on the upstream-downstream gap in China's water pollution regulation. A recent study by He et al. (2020) discusses how imperfect performance monitoring of water pollution in China can break down the central-local alignment. In my paper, I show that the gaps in ground monitoring networks can lead to significant deviation in the local air pollution regulations from what the central government observes.

Second, I contribute to the growing literature on the environmental monitoring regulation and enforcement (Gray and Shimshack 2011; Duflo et al. 2013; Shimshack 2014). While existing literature mainly focuses on the air pollution monitoring system in the U.S. (Grainger et al. 2019; Zou 2020), where they look at either the intermittent monitoring schedule or monitor siting from a static spatial point of view. My paper adds to the limited studies looking at the new air quality monitoring program in China and particularly examines the dynamic changes in monitors' spatial representativeness due to local officials' strategic responses to gaps in monitor coverages. My paper relates closely to two of the concurrent studies. Greenstone et al. (2020) show the improvement of data quality with the help of the new monitoring system, and Barwick et al. (2020) focus on the relationship between information disclosure in the new program and people's avoidance behaviors. My study complements the previous two in that I reveal the heterogeneous impact of the system on air quality caused by local regulators' strategic responses to gaps in monitoring coverages. With the strategic responses, the information disclosed to the public would be inaccurate, and people's avoidance behavior may be biased (especially for rural households). My study is also widely applicable to monitoring regulation in other countries in both the developed and developing world because they either have monitoring networks that were built decades ago or need to design a new monitoring system.

Third, this paper adds to the literature on the value of satellite data in environmental regulations. Taking advantages of the high-resolution satellite images of air pollution, I am able to fill the gaps in ground monitoring and examine the pollution changes across different regions. In particular, I use satellite measures to evaluate the population-weighted pollution levels in each city and the representativeness of the ground monitoring system. Similar studies in the U.S. context also prove the value of satellite data and show the bias in attainment and non-attainment designations using only ground monitor's readings and the resultant welfare losses (Sullivan and Krupnick 2018; Fowlie et al. 2019). In addition to air pollution regulations, the value of satellite data in fields like climate change, wildfire surveillance (Ruminski et al. 2007), forest land cover (Hansen et al. 2013), and biodiversity (Turner et al. 2015) has been increasingly recognized by regulators and researchers.

Finally, I provide policy implications for an improved air pollution monitoring and enforcement. The central government should use auxiliary pollution information from remote-sensing data and public supervisions, together with the ground-level monitoring data, to evaluate pollution conditions. Although it is difficult to directly test the mechanism of local regulators' strategic pollution reductions due to data limitations, I provide indirect evidence for the role of economic development pressure, local regulators' characteristics, and public pressure. My results support the political incentives behind local officials' strategic behaviors and show the importance of an incentive-compatible enforcement from the central government.

The remainder of the paper is organized as follows. Section 2 provides a brief background on environmental regulations and the monitoring system in China. Section 3 describes the main data sources. Section 4 presents the main identification of local officials' strategic pollution reductions. Section 5 explores channels and mechanisms underlying the strategic behavior. Section 6 discusses policy implications for the air pollution monitoring system. Section 7 concludes.

3.2 Institutional Background

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

3.2.1 Political System in China

Political incentives are one of the internal mechanisms of both economic development and environmental protection, especially in China. A salient feature in China's political system is that the central government sets targets and links the local officials' promotion to their performance in these targets. Local of-

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

ficials, in turn, are highly incentivized and are given great flexibility in local regulatory plans to meet the national targets. Studies in political economics have examined the principal-agent problem lies in China's economic development. The incentive-based strategic responses by local governments have led to many unintended consequences such as inequality, collusion, corruption, and cheating, which may undermine the policy goals. (Li and Zhou 2005; Fisman and Wang 2015; Oliva 2015; Jia and Nie 2017; Jia 2017)

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

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

3.2.2 Environmental Regulations in China

Air pollution regulation has been a top priority of the central government of China in the past decade. It declared "war on air pollution," implementing a series of mitigation actions, such as the "Air Ten" action plan that was announced in 2013, (the Air Pollution Prevention and Control Action Plan). The action plans add detailed pollution control requirements to the 12^{th} FYPs in terms of targets, standards, measures, and technologies. In addition to the plans, a raft of new environmental protection laws and guidance are enacted, which are claimed to be the "strictest ever" environmental policies regulations to show the central government's determination to win this "war".

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

Stringent central regulations have helped improving air quality in China, according to the ground monitor readings. For example, Greenstone et al. (2020) estimate the air pollution trend since 2013 ("Air Ten") and show that all of the air pollutant concentrations dropped sharply, except for O3, which saw a modest increase. PM2.5 levels dropped by 27.7 $\mu g/m^3$, or about 41 percent from the 2013 level. However, the sharp reduction in air pollution is based on the ground monitor readings, which may be subject to bias due to gaps in spatial coverages. My paper aims to dig deeper into this pollution reduction trend and study the local governments' strategic pollution reduction behaviors using the newly disclosed monitoring system.

3.2.3 Monitoring Systems for Ambient Pollutants

Evaluating a city's air quality and local officials' performance is mainly based on the stationary, in situ monitors. Along with the evolution of China's environmental regulation and policies, the monitoring system for ambient pollutants evolves significantly. The data quality in China has been criticized a lot, especially for air pollution data before 2013: only 74 major cities had monitors, the data was reported by local governments as a daily air pollution index, and not available to the public. Obviously, local governments have great power to manipulate the reported air pollution data. As shown in Ghanem and Zhang (2014), when the policy goal is the number of "blue sky days" in a year, that is when the air pollution index is less than 100, the air pollution data reported by local governments is bunching at the cut-off. [8]

To win the "war against pollution" after 2013, China launched a nationwide, real-time air quality monitoring and disclosure program, which quickly built-out over 1400 monitors. Several major improvements have been made in this new monitoring program. Firstly, PM2.5 is listed as a major pollutant. Secondly, the monitored data are uploaded to the cloud automatically, which

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

significantly eliminates the data manipulation issue in the pre-automation self-reported pollution data. A recent study by Greenstone et al. (Forthcoming) shows the improvement in data quality with the new monitoring system, and the increased public awareness of pollution prevention.

There are three types of monitors in China: 1. Monitors controlled by the central government; 2. Monitors controlled by local government; 3. Micro Monitors for specific polluting sources. The central government control monitors are the first group of monitors set up before air pollution becomes a society-wide concern. Also, the local government has a relatively low involvement in the central monitors. Most importantly, the performance of local officials in eliminating air pollutions is based on the readings of central monitors. To help better control for polluting sources, the local officials build many local government control monitors, which are not included in evaluating a city's average pollution. [9]

In order to regulate the siting and operation of the monitors, the central government issued guidelines for air quality monitoring. The guidelines include the monitors' location choices, monitoring techniques, management of the monitoring data, and penalties for data manipulation and other human intervention of the monitors. The central government state that only central monitors will be counted into the evaluation of cities' average air quality conditions and local official's performance in pollution reduction. Local monitors, although built under the same guidelines, will only be helping local officials in detecting polluting sources and designing for local regulatory plans.

[9] As of 2016, there were more than 2000 monitors in China, including both central and local monitors.

Three waves of prefectural cities entered the monitoring system successively in each year between 2012 and 2014. Major development regions such as the Jing-Jin-Ji region, the Yangtze River Delta region, and the Pearl River Delta region, as well as a few large cities such as provincial capitals, are the first wave to enter the new monitoring network. In these cities, many of the monitors were built and operated long before the new monitoring system was introduced. Entering the program means upgrading the existing monitors to automation, as well as adding new monitors. By the end of 2012, 496 monitors in 74 cities started to work. The second wave and the third wave then added around 450 and 550 monitors into the network each year. Cities in three waves vary largely in terms of their hierarchy level and overall environmental performances. Figure 3.1 shows the three roll-out waves of monitors in China. The national monitoring network with 1499 central monitors is designed to serve for urban areas of 336 cities. The number of monitors in each city is based on the population density and a city's pollution level in the past three years.

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

There is a lack of empirical evidence for local officials' pollution reduction strategy facing the new monitoring system. Since the scattered monitors lead to gaps in measuring the pollution exposure for unmonitored areas, bias may still exist due to local officials' strategic responses in having spatially differentiated pollution control measurements in monitored areas and unmonitored areas. The issue is not unique in China. Fowlie et al. (2019) and Sullivan and Krupnick (2018) examine the misclassification of attainment and non-attainment designation of counties due to the gaps in ground air pollution monitors in the U.S, and the potential welfare loss using the satellite-based pollutant data as references. Grainger et al. (2019) also use satellite NOx data to check the strategic siting behaviors of attainment and non-attainment counties. They find avoidance behaviors of local officials in attainment counties near the non-attainment threshold, where they strategically place new monitors at a relatively clean area of the county. Inspired by these studies, I use the remote-sensing data to fill the gaps in ground air pollution monitoring system and find evidence for local officials spatially differentiated pollution control strategies.

3.3 Data

3.3.1 Remote Sensing Data

In order to examine the spatial difference in air pollution regulations, this paper fills the gap in the ground monitoring system using high-resolution images of the major air pollutant, PM2.5, which are derived from the original satellite measures of Aerosol Optical Depth (AOD). The satellite AOD data comes from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) algorithm. AOD measures the total vertical distribution of particles and gases within a grid according to the light extinction coefficient. It indicates how much direct sunlight is prevented from reaching the ground by aerosol particles and can be used to infer ground-level pollution, particularly for fine particles such as PM2.5 and PM10. Atmospheric science literature has shown a strong correlation between satellite measure and ground-level pollution data.[10] Since the satellite measures are largely affected by cloud coverages over an area, missing data is a big issue when using remote sensing data with fine spatial and temporal resolutions. Studies of the remote sensing techniques find better correlations between AOD and ground-level PM with coarser spatial and temporal resolutions by month or year (Hoff and Christopher 2009).

The satellite images this paper uses include annual PM2.5 grids (1km by 1km resolution, nine million grids for whole China) from 2000 to 2017. By combining satellite-based measures of AOD with chemical-transport modeling and land characteristics, van Donkelaar et al. (2019) derive ground-level concentrations of PM2.5 at high levels of spatial disaggregation. One concern with the satellite-derived ground-level pollution measure is the measurement errors caused by the calibration of the satellite data using ground monitoring data. Even though that van Donkelaar et al. (2019) use geographical weighting method to give smaller weights to cells further away from ground monitors, and larger weights to cells closed to ground monitors, one may be worried

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

about different measurement errors may occur at cells with different distances to monitors. To address this concern, the authors conducted cross-validation tests, where they remove part to all of the ground monitors from the calibration. The derived PM2.5 data still performs well. [11]

3.3.2 Spatial Representativeness of Ground Monitors

With the fine-scale pollution data and spatial information of the new ground monitoring network, I examine the spatial representativeness of these monitors. First, I apply the kernel density estimation to compare the pollution distribution of monitored cells with that of unmonitored cells, following the methodology from Grainger et al. (2019), which define a z-score for each grid in each city to measure the within-county variation. [12] I also compare the spatial distributions of different types of ground monitors: central vs. local. The kernel density estimation result in Figure 3.2(a) shows that monitors are mostly placed in a relatively more polluted area in a city. This is consistent with the intuition that most monitors are placed in urban areas to cover more population. Figure 3.2(b) shows that local monitors are placed in a slightly cleaner area comparing to central monitors. This is intriguing because one would expect the local officials to put local monitors nearer to polluting sources in order to regulate air pollution directly.

One thing to notice is that almost all monitors are located in urban areas, and the sparse central monitors are the only base in evaluating the air pollu-

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

[12] Z-score is calculated by taking the observed value in grid cell i in city c and year t, subtract the average for that city, and scale it by the city level standard deviation.

tion condition of a city. The gaps in the ground monitoring network might cause the regulation focus to bias toward urban citizens. Instead, the less-monitored places, i.e., the rural areas' pollution, will not be considered in evaluating the local officials' environmental performance. Contrarily, the satellite-based measurements give a highly spatial resolved coverage of the air pollution in the entire city area. To examine the difference between monitor-based and satellite-based city average PM2.5, I use the 1km by 1km gridded population count from 2015 Census to weigh each cell and calculate the weighted average PM2.5 for each city. Taking this as the "true" city-level PM2.5, I then compare it with the monitor-based population-weighted average PM2.5. The map in Figure 3.3 shows the monitors representation errors in the years that cities joined the system. I regard the cities with errors within ±10% as having well-representative monitors. The warm colors are cities where monitors over-represent the "true" city-level PM2.5, and the cool colors are cities with under-representative monitors. The representation errors exhibit large spatial variations, where two-thirds of cities have over-representative monitors, consistent with the kernel density figures. I have also included a set of interesting correlates in Appendix A.2 and A.3 to check if the leaders' characteristics, the GDP per capita, or industrial type matters for the "representation errors".

The representation errors in Figure 3.3 are static at the moment of their openings. If the pollution reduction patterns are even across space, then the representativeness of monitors would not change as long as the monitors' locations do not change. However, though monitors are unlikely to move for a long period, local regulators' strategic responses to the static monitor locations would change the monitor's spatial representativeness overtime. From the representation error maps in each year (Appendix A1), this is indeed the case. Monitors' spatial representativeness exhibits dynamic changes in years after cities joining the program, which greatly motivates my study of local regulators' strategic pollution reduction behaviors.

3.3.3 Other Data and Summary Statistics

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

The summary statistics are presented in Table 3.1 and 3.2. Table 3.1 presents satellite PM2.5 summary statistics by calendar year. Over the period of study, the PM2.5 level increased significantly before 2013, and then declined. After the declaration of the "war against pollution", there is an overall improvement in air quality. In Table 3.2 , I present a summary statistic by different waves of cities, where I summarize the population-weighted PM2.5 density using the 2015 population in each grid cell as the weight. I also summarize the population-weighted PM2.5 density at cells containing monitors, which are in general higher than the city average PM2.5 in all three waves of cities. In addition to the PM2.5 densities, I also include a summary of city characteristics such as the population and GDP by the three waves. Comparing the three waves, I find that cities in earlier waves tend to be dirtier and have more population. In terms of GDP, and the GDP in each industry, cities joining the program earlier tend to be more economically developed. The difference in city characteristics among waves may lead to different environmental strategies and regulation outcomes. Because cities in earlier waves are high in the hierarchy rank, city officials' characteristics could be different. From the summary statistics of city mayors' age and education, I find that wave one cities have slightly older mayors and more mayors with PhD degrees. Most mayors in wave two cities own master's degrees, and most city mayors in wave three have bachelor's degrees.

3.4 Strategic Pollution Reduction After Monitoring

3.4.1 Empirical Framework

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

$$\ln(PM2.5_{iwt}) = \alpha Open_{wt} + \beta Near_i \times Open_{wt} + Cell_i + Year_t + Trend_{wt} + \varepsilon_{it} \quad (3.1)$$

The outcome variable, $\ln(PM2.5_{iwt})$, is the logarithm of annual PM2.5 concentration at the 1km×1km grid cell. i is the index for grid cells within cities opened in wave w at year t. In my study, there are over nine million cells' annual PM2.5 from 2000 to 2017 in the raw data. $Open_{wt}$ is the treatment indicator that takes the value of 1 if cell i is in a wave w city after joining the new monitoring program. The treatment intensity is defined by $Near_i$, which equals 1 if the grid cell i is in an area adjacent to a ground monitor (monitored area), and 0 if the cell i is in areas far away from monitors (unmonitored area). In most cases, I am less interested in the causal effect of the monitoring program per se (α), but rather more in the difference in the causal effect in monitored vs. unmonitored areas (β) after monitoring. Due to the large spatial and temporal variations in air pollution, there may be confounders that would bias β from identifying the difference in pollution reductions across space. Especially, cells in monitored and unmonitored areas could have different location attributes that affect air quality. To address these concerns, I report results of estimations with a rich array of controls, including cell fixed effect and year fixed effect. I also include a wave-specific time trend to allow the unobserved time trend in pollution to vary across waves. The identification variation is then from comparing cells in monitored vs. unmonitored areas before vs. after new waves of monitor roll-out. Since pollution observed at a cell is likely driven by emissions elsewhere that also affect nearby cells, I cluster standard errors at the city level.

Cities selected into the program in different waves may be due to wave-specific unobservables that are time-variant. Cities in earlier waves tend to be larger cities with more population, higher GDP per capita, higher levels of air pollution and industrial emissions, etc. I include wave by year fixed effects in Equation (3.2). Although the fixed effects absorb the baseline impact of monitoring on overall pollution (α from Equation (3.1)), Equation (3.2) provides a clearer identification of changes in treatment effect by treatment intensity (β). It also has more flexible controls than the wave-specific time trend. The identification variation now is from comparing monitored vs. unmonitored cells in the same wave cities, before and after monitoring. The key explanatory variable is $Near_{it}$ which is an interaction of $Near_i$ and $Open_{wt}$.

$$\ln(PM2.5_{it}) = \beta Near_{it} + Cell_i + Wave_w \times Year_t + \varepsilon_{it} \tag{3.2}$$

3.4.2 Baseline Results

In the baseline results, I estimate equation (3.1) and (3.2) using my preferred sample from 2007 to 2017, which includes five years prior to monitoring and all post monitoring years to have a relatively balanced panel. [13] The monitored area is defined as grid cells within 3km of a monitor. The results are robust to alternative definitions of monitored areas such as 2km, 5km, and 10km, and unmonitored areas such as outside 3km, 30km, and 50km where I drop the cells in between to address the concern of misclassifying monitoring status. The DID with treatment intensity provides estimates of local effects within the choice of the treatment intensity groups, where results using different monitored areas could represent different pollution control strategies that local officials adopt. I will discuss more in the next section.

Table 3.3 presents the baseline DID result by estimating Equation (3.1) and adding controls sequentially. In the first four columns, *Open* captures the base-

[13] For cities in the first wave, the sample period is [2007, 2017] with five years pre and post monitoring; for cities in the second wave, the sample period is [2008, 2017] with five years pre and four years post monitoring; for cities in the third wave, the sample period is [2009, 2017] with five years pre and three years post monitoring.

line impact of joining the monitoring program on air pollution, comparing to control cities. The baseline DID estimates of the causal impact of monitoring on air pollution are positive (pollution increases) and significant across the controls. In Figure 3.4, I conduct an event study for the causal impact of monitoring, where I replace the treatment indicator $Open_{wt}$ with opening dummies for each year pre and post monitoring. The event study figure shows no pretrend, and significant increases in pollution after cities joined the program. [14]

I then include the treatment intensity indicator $\mathbf{1}$(0-3km) in column (5) & (6) to capture the heterogeneous treatment effects of the monitoring program on pollution in monitored (cells within 3km) vs. unmonitored (cells outside 3km) areas.[15] The results from column (6) show that pollution in monitored areas is decreased after monitoring by 2%. Unmonitored areas exhibit 4% higher pollution after a city joins the program, indicating the potential pollution leakages.

In Table 3.4, I show the baseline DID results are robust to alternative definitions of treatment intensity groups. The first three columns present results for monitored areas defined as cells within 3km of monitors, and compare to different unmonitored areas such as cells outside 30km and 50km of monitors.

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

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

The areas in between are dropped to have a clearer defination of treatment intensity groups. Column (4)-(6) expand the monitored areas to five distance bins to show how the treatment effect varies over space. Consistent with intuition, the difference in pollution changes between monitored and unmonitored groups are larger when two groups are more apart from each other, and the differences are smaller when monitored areas are further from monitors.

In Table 3.5, I use wave specific year fixed effects to absorb the baseline causal effect of monitoring and show the relative changes between treatment intensity groups (Equation (3.2)). Column (1) presents the main finding of my paper. Pollution in monitored areas is 6.5% less than that in unmonitored areas after monitors roll-out. Similar to Table 3.4, the results are robust to alternative treatment intensity groups. The first three columns present results for monitored areas as cells within 3km to monitors, and compare to different unmonitored areas such as outside 30km and 50km of monitors. The areas in between are dropped to have a clearer definition of treatment intensity groups. Column (4)-(6) expand the monitored areas to five distance bins to show how the treatment effect varies over space. Consistent with intuition, the difference in pollution changes between monitored and unmonitored groups is larger when two groups are more apart from each other, and the differences are smaller when monitored areas are further from monitors.

3.4.3 Identification

The key assumption is that in the absence of a monitor opening or switching to automation, air quality in the monitored and unmonitored areas follow parallel trends. In other words, I assume that the only reason that ambient air quality might show a significant difference between areas nearby monitors and areas far away from monitors is because that local officials strategically put more efforts into reducing “local” air pollution. As directed by the central government, most monitoring stations are placed in urban centers to cover populated areas. One may be concerned that cells in the unmonitored areas are too far away from the city center and thus would have different pollution trends from those in the monitored areas. While the parallel trend assumption is not directly testable, I conduct a “placebo” test and an event study analysis to support the assumption. To address the identification concern of endogenous monitor locations, I conduct another “placebo” test with random monitor placements.

Placebo Tests

First, I conducted a “placebo” test using only pre-program periods and randomly assign opening years for all monitors at the same locations. The rationale behind the placebo test is that cells in “monitored” and “unmonitored” areas should not be significantly different over a false-opening year in the absence of the monitoring program. For each monitor, I randomly assign an opening year between 2007 to 2011 for 500 times. I then conduct 500 estimations of equation (3.2) and plot the distribution of the coefficients in Figure 3.5. Comparing with the observed coefficient, I find that the observed coefficient lies outside of the 99% confidence interval of the coefficients from 500 placebo tests, which center around 0.016. This result shows that before the monitoring program, a false opening would not lead to larger pollution reductions in monitored areas than unmonitored areas.

Second, in order to show that my findings indeed a result of local pollution reductions in monitored areas, I conduct a placebo test with random monitor locations. Keeping the number of monitors and the year of joining the program unchanged, I randomly relocate all the monitors within each city 500 times. The underlying idea is that if local officials only conduct strategic reductions in areas very closed to monitors, then no significantly different pollution reduction should be observed in areas with a false monitor opening compare to other areas in the city. After matching the 500 groups of placebo monitors with the satellite grid cells, I estimate equation (3.2) and plot the distribution of the coefficients in Figure 3.6. The observed coefficient lies outside of the 99% confidence interval, suggesting that local pollution reductions happened only at the observed monitored areas.

Event Study

I use event study analysis to show the parallel trends between monitored and unmonitored groups hold for pre-opening periods in general. I divide the years around opening dates into five pre-opening periods $n = -5, -4, ..., -1$, and six post-opening periods $n = 0, 1, ..., 5$ and run the following regression:

$$\ln(PM2.5_{it}) = \sum_{n=-5}^{5} \beta_n \phi(n) \times Near_{it} + Cell_i + Wave_w \times Year_t + \varepsilon_{it} \tag{3.3}$$

where $\phi(n) = \mathbf{1}[n \leq t \leq n+1]$, indicating interval n. The base interval is the year before the opening year (i.e., $n = -1$). I expand the dataset used in main DID analysis (PM2.5 in 2009-2017) to year of 2007 which allows wave 1 and wave 2 cities to have the same number of five pre-opening periods. However, the number of post-opening periods for cities in different waves would not be the same due to data availability.

Figure 3.7 (and Column 1 in Table 3.6) presents the coefficient estimates of $\phi(n)$. The results support the parallel trends assumption in general: compared with the base interval (1-year before opening years), the subsequent changes in air pollution between the monitored and unmonitored areas are not significantly different for the four pre-opening intervals in the specification. In contrast, I find statistically significant different air pollution reduction between the monitored and unmonitored groups in the post-opening intervals for the same specification. The fifth year prior to monitoring exhibits a significant difference, which could be due to more unobserved policy changes in years further before monitoring. Column (2)-(5) in Table 3.6 show the event study estimation results are robust to alternative definitions of $Near_{it}$.

Eliminate Alternative Explanations

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

A similar concern lies in the political interpretation of local officials' strategic behaviors. One may argue that local officials choosing to reduce more pollution in monitored areas is not a strategy that they play to gaming the performance evaluation. Instead, they choose a more cost-effective way to reduce pollution in a relatively more polluted area, which happens to be the area adjacent to a monitor. To address this type of concerns of monitored areas being coincident with polluted areas, I examine the heterogeneity of treatment effects where I allow the impact to differ based on the relative pollution levels of cells within the vicinity of monitors as in Equation (3.4),

$$\ln(PM2.5_{it}) = \beta Near_{it} + \eta Near_{it} \times Dirty_{it} + Cell_i + Wave_w \times Year_t + \varepsilon_{it} \tag{3.4}$$

where $Dirty_i$ is a dummy variable which equals 1 if the PM2.5 of a cell i is higher than the city average PM2.5 level in year t. This specification examines the potential concern of monitors locating in the dirty area of a city. The coefficient η will show the different pollution gaps between monitors in a dirty area and clean area. The results are reported in Table 3.7, where I include alternative definitions of monitored and unmonitored areas to show robust results. Cells within 3 km of a monitor are the monitored cells in the first three columns, and I then expand the monitored areas to include more distance bins. From columns (1) to (6), I show that no matter which monitored groups, being in the dirtier area of a city does not lead to large pollution reductions as concerned. In fact, the magnitude of the interaction terms with *Dirtier* is almost zero comparing to the strategic pollution reductions in monitored areas.

Another possibility that may generate similar results is the measurement errors from satellite-derived pollution measures. The PM2.5 data I use is derived from the raw satellite images, which require information from monitor-based sources. The Geographical Weighted Regression method used when deriving PM2.5 from satellite images assigns larger weights to areas closer to ground monitors, and smaller weights to farther areas. One may be concerned that the resulted measurement errors from the data generating process will be correlated with the distances to monitors and also varied over time when more ground monitors are opened. If this is the case, then the spatially different pollution patterns could simply because of the spatially differentiated measurement errors. Although van Donkelaar et al. (2016) have conducted several out-of-sample cross-validation tests to justify their satellite-derived PM2.5 data, I conduct a robustness check using the raw satellite images to further eliminate this possible explanation. Using the raw AOD data from the NASA MODIS product, I manually aggregate the daily AOD images at 3km by 3km resolution into annual AOD, and match with the ground monitors. The grid cells containing monitors are monitored cell, and those do not contain any monitors are unmonitored cells. Estimating Equation (3.2) using the AOD data shows a similar result. After monitoring, pollution in monitored cells decreases comparing to unmonitored cells. (Table 3.8)

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

3.5 Heterogeneous Effects and Potential Mechanisms

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

3.5.1 Channels for Strategic Reduction

I present the spatial distribution of the impact of monitoring and discuss abundant qualitative evidence of the local officials' pollution control strategies to support the political interpretation of my findings. By replacing the binary indicator of one monitored group and one unmonitored group used in Eq (3.1) with fifteen treatment intensity groups, I show the spatial distribution of the treatment effect by distances from monitors in Table 3.11. The changes in the impact of monitoring over space also indicate the potential channels of strategic pollution reductions. The coefficient estimates of *Open* represents the impact of monitoring on air pollution in the base group, which includes cells more than 300km away from the closest monitors. Combining with the interaction terms, the strategic pollution reductions exist within 70km ranges of monitors and are robust in magnitudes. Beyond 70km, the overall impact of monitoring turns positive and continues to increase for cells further away. With more distance bins in the unmonitored groups, Table 3.11 represents the potential pollution migration patterns across space after monitoring. Note that most of these central monitors are placed in population-dense (urban) areas.

Column (2) in Table 3.11 summarizes the population in each distance bin. Although the monitoring enforcement seems to divert air pollution away from areas near monitors, this does not necessarily lead to policy failure when considering the population exposed to air pollution. However, this could exacerbate inequality issues if pollutions are leaking to rural areas. I provide more discussion in Section 6 on the dynamic changes of monitors' representativeness in population-weighted pollution exposure.

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

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

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

A more effective strategy in the longer term would be relocating polluting sources from small-scale workshops, restaurants to large industrial plants to suburban or rural areas that are commonly unmonitored. This type of strategy would be preferred considering either economic development or environmental performance (improving monitor readings). However, it would impose the largest environmental damages and bias of central regulations. Based on the baseline DID results in Table 3.3, relocation of polluting sources seems to be the most common strategy given that unmonitored areas become more polluted after monitoring.

The strategies that local officials use to achieve better monitor readings are hard to test empirically due to data limitations. For example, traffic controls and water spraying in monitored areas are short-term actions which may only be caught by constantly observing the abnormal phenomenon near monitors. Instead, I use several heterogeneity analyses to indirectly support the findings of local regulators' strategic responses.

3.5.2 Heterogeneity in Strategic Pollution Reductions

I present evidence from heterogeneous analysis to show that the political incentives of local politicians are indeed the driving forces behind my main findings.

a) Roll-out Waves of Entering the Program

In addition to the annual assessment, the local officials face a final assessment of air pollution reductions at the end of 2017. They may use more aggressive strategies to reduce monitor readings when the final assessment approaches. On the contrary, major cities in earlier waves, especially those in the key development regions, face more stringent PM2.5 reduction goals. It is unclear which incentivizes local officials more in taking more aggressive strategic responses, the stringent target or the approaching deadline. In Figure 3.8, I investigate heterogeneity in the impact of treatment intensity on pollution reductions by roll-out waves. I find the cities in later waves show larger strategic pollution reductions in monitored areas, indicating more aggressive strategies as the deadline approaches. Another possible explanation is that cities in wave one and two cities are those with monitor readings upgraded from manual to automation in the new system, rather than having new monitors opened. Thus, the strategic pollution reduction might exist before the cities join the new monitoring system.

b) Compliance Levels

Local official's pollution control strategy could be varying with the existing pollution conditions. In Table 3.9 and Figure 3.9, I explore the heterogeneity by cities' average pollution levels, using the national annual PM2.5 standard 35 ug/m3 as a reference. I use the population-weighted city average pollution at the monitored cells at the years of monitors roll-out. Cities with average pollution levels below the annual standard are defined as clean cities. I find that clean cities tend to have larger strategic pollution reductions in monitored areas after monitoring. Restricting the sample to cities with average PM2.5 from 30-40 ug/m3 shows similar results. In order to see if the heterogeneity by compliance level varies with roll-out waves, I include additional analysis using subsamples in each wave. Clean cities in wave one tend to have more aggressive strategies. This could be due to the fact that wave one cities are in general dirtier than other cities. Thus, dirtier cities in wave one are the most polluted cities in China and under strict supervision by the central government. To see how strategic response varies by the closeness to the national standard, I include another layer of interaction, *Compliance*, which is the difference between city PM2.5 and 35 ug/m3. For a clean city, when its pollution level approaches the national standard, I find larger strategic pollution reductions. The heterogeneous effect by cities' compliance levels indicates that local officials facing different compliance status choose different strategies to meet the environmental targets.

c) Leader Characteristics

City mayors play an essential role in policy regulation and implementations. I investigate whether a city mayor's characteristics have an impact on the strategic pollution reductions after monitoring. Figure 3.11 shows the heterogeneity analysis by city mayor's age, where I separate the sample into two subsamples by city mayors' age. A mayor has better chances to be promoted to a higher position at an age younger than 57. Thus, a younger mayor may have larger incentives to perform well in the environmental evaluation and adopt more strategic pollution reduction methods in monitored areas. For mayors older than 57, which means they have little to no promotion opportunities, they would be less incentivized to achieve policy targets. Figure 3.11 shows such results that cities with mayors younger than 57 tend to have larger strategic reductions in monitored areas. On the other hand, I do not find any significant impact of a mayor's educational background on their strategic behaviors. This may suggest the strategic reduction methods are common knowledge for leaders across education levels and do not require elite training.

d) Economic Growth

In general, there are tradeoffs between economic development and pollution abatement for local regulators. Prioritizing environmental regulations may hurt the local GDP growth and local officials may have different strategic behaviors in pollution control when facing different economic conditions. To examine the role of economic growth pressure, I generate a dummy variable indicating the growth or recession of a city's GDP in the previous year and interact with the DID treatment intensity term. Table 3.10 shows the results for all cities, and for each wave of cities. I find that no matter in which roll-out wave, when a local official faces downward pressure on economic growth, they tend to reduce strategic measures that improve monitor readings. This set of heterogeneity results suggest that local regulators are balancing both their efforts and performance in economic growth and pollution control. The gap in pollution changes between monitored and unmonitored areas is indeed a result of local regulators' strategic pollution reductions.

e) Information Transparency (Public Pressure)

Local official's strategic behaviors can potentially be captured by residents if they have full information about air pollution monitors, such as locations and readings. With the new monitoring system, information about the central monitors are publicly available through multiple sources, including the MEP's website and third-party online platforms. In addition, a few provinces have launched their own online air pollution disclosure platform. They provide detailed information about the monitor locations, including both central and local monitors. In China, eleven provinces have an online platform, which shows their effort in improving information transparency. Moreover, local residents can perform additional supervision on air pollution monitors and check consistency with online information. In Figure 3.12, I investigate such heterogeneity and find that provinces with online pollution disclosure do not show significant strategic pollution reductions in monitored areas after monitoring, which suggests the importance of information transparency and public pressure could potentially reduce local official's strategic behaviors.

3.6 Policy Implications and Suggestions

3.6.1 A Well-representative Monitoring System

One would expect to see local regulators to have very different strategic behaviors facing the new monitoring system, because monitors' siting could over-represent, well-represent, or under-represent the average city pollution levels. Even though the central government intended to place the monitors in populated areas to improve the representativeness, the over-representing monitors (monitor-based pollution larger than city average pollution) could exacerbate local government's strategic pollution reductions. I conduct a heterogeneity analysis to show that it is necessary to build a monitoring system that well-represents the average city pollution level. I split the sample into three groups: "over-represent" cities with representation errors greater than 10%, "well-represent" cities with errors between -10% to 10%, and "under-represent" cities with smaller than -10% errors.

Figure 3.13 shows the event study on three subsamples. I find that cities with over-representing monitors tend to have more aggressive strategic reductions after monitoring, comparing to well-representing cities. It is hard to find a clear trend for cities with under-represented monitored pollution due to the few numbers of "under-represent" cities. The heterogeneous results are intuitive because if the central government places monitors in the dirtiest area of a city, local officials will be more incentivized to reduce the pollution only in the monitored area. However, unmonitored areas could still have more pollution than the national standard due to pollution leakages. Thus, it would be necessary for the central government to evaluate the cities' average pollution thoroughly and use the population-weighted average pollution as references for monitoring sites. The well-represent cities still have a slightly downward trend after monitoring. This indicates that a well-representative monitoring system could, to some extent, reduce the strategic responses at the local level but would not prevent the behaviors from happening. In fact, the strategic responses may change the spatial representativeness in the long run.

3.6.2 Dynamic Monitors Representativeness

From the representation error map in Figure 3.3, the current monitoring system in most cities shows good representativeness when the cities first joined the program. However, similar to the monitoring systems in developed countries, monitor locations are unlikely to change once the monitors were placed. For example, the current air quality monitors in the U.S. were built two decades ago, and covered populated areas following federal guidelines. Other than adding new monitors to nonattainment counties, the existing monitor locations have not changed ever since. Thus, even though monitors were sited to be well-representing counties' overall air quality in the 90s, the representativeness can be dynamic due to human interventions in monitored areas. Using my estimates for the relative pollution reductions in monitored areas (cells within 3km have 6.5% more pollution reductions), and the last observed year of pollution in my data in 2017, I calculate the projected pollution levels for five years from 2018 to 2022. I do not conduct simulations for a longer period into the future because there could be large uncertainty and new regulations. I find that in the near future, the over-representative monitors seem to become more representative of a city's overall air quality. However, there are also more cities exhibiting negative representation errors, 42 cities at years of monitoring vs. 52 cities in 2022. Even though the monitoring system works fine in my projected years, it is possible that with the strategic responses, monitors would become less representative in the long term. Moreover, the pollution leakages to unmonitored areas, mostly rural regions, could cause large health impacts and biased evaluation of policy goals.

3.6.3 The Remote Sensing Data and Other Pollution Information

The key to eliminating or preventing local official's strategic responses to the ground monitoring system is to add referencing data sources into the evaluation. In an ideal world with ground monitors everywhere, local officials are impossible to predict which sets of monitors would be used to evaluate their environmental performance. Thus, the only strategy left is to improve air quality city-wide. This seems unrealistic because ground monitors are large in size and costly to build and maintain. The satellite-based pollution measures can be a good source to fill the gap in ground monitor coverages. As shown in (Sullivan and Krupnick 2018) and Fowlie et al. (2019), remote-sensing data has helped the authors to assess the extent to which the existing U.S. ground monitor-based measurements over- or under-estimate true exposure to PM2.5 pollution. In my context, I have used the satellite-based data to re-evaluate the policy goals set by the "Air Ten" action plan for the end of 2017. Unlike the monitor-based pollution patterns estimated in Greenstone et al. (2020), PM2.5 decreases by 40% from 2013 to 2018, my estimates find an overall increase in the city-wide pollution level. This suggests that monitor-based evaluation would overstate the environmental performance and distort future policy design.

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

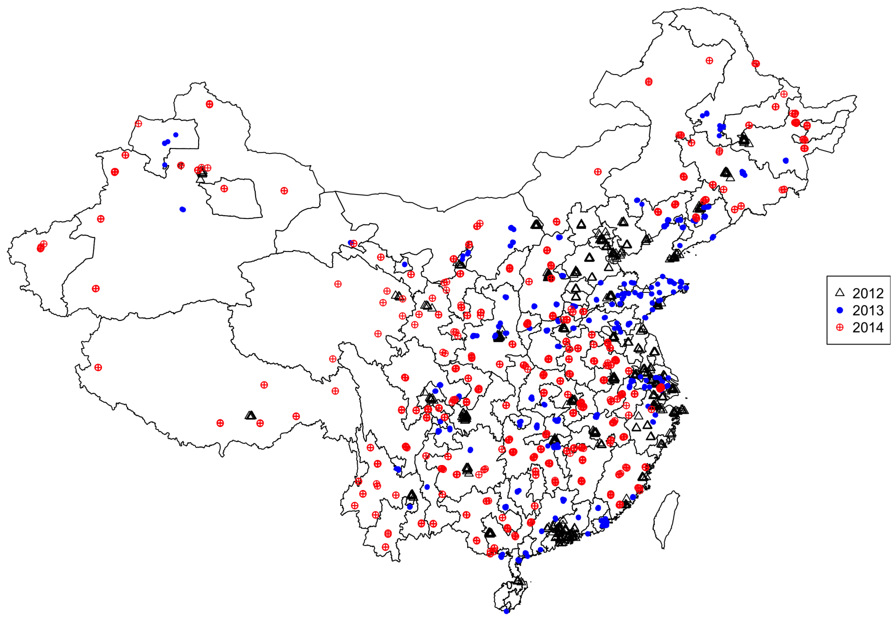
3.7 Conclusion

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

I find that there exists a significant difference between pollution changes in areas adjacent to monitors and areas far away from monitors after monitoring. This result is robust to different definitions of monitored and unmonitored areas. Although the new ground monitoring network has improved data quality significantly, the gaps in monitor coverages lead to pollution leakages from monitored areas to unmonitored areas. The baseline DID result shows that pollution in unmonitored areas increases after monitors roll-out, which indicates that the underlying mechanism of such strategic reduction is non-local, relocating polluting sources away from monitors. By studying the heterogeneous impact of cities' pollution levels, the characteristics of local leaders, the role of public pressure, and the role of economic growth, I provide evidence supporting the political interpretation of the strategic pollution reductions. Overall, my results are consistent with the expectation that strategic pollution reductions are more likely to arise with larger incentives to improve monitor readings, such as in cities with younger mayors and cities with approaching assessment deadlines.

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

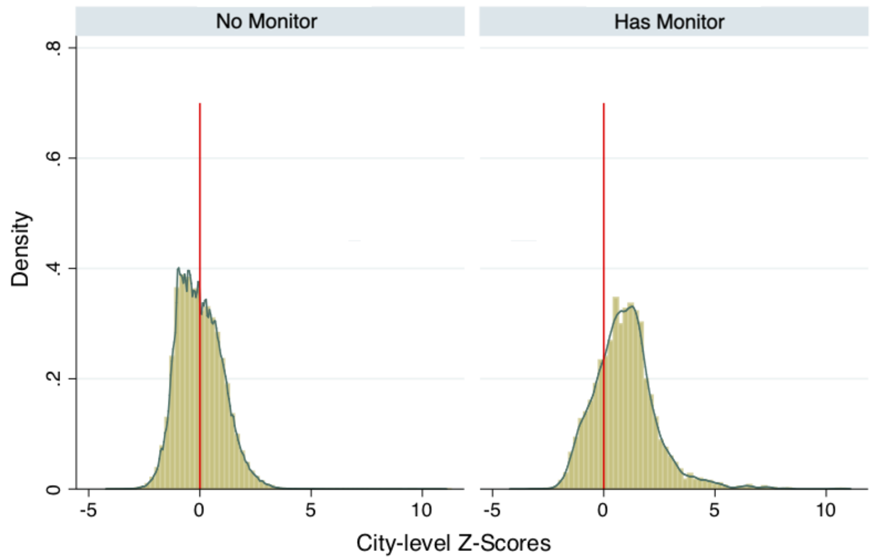
Figure 3.1: Roll-out of Monitoring Stations in China



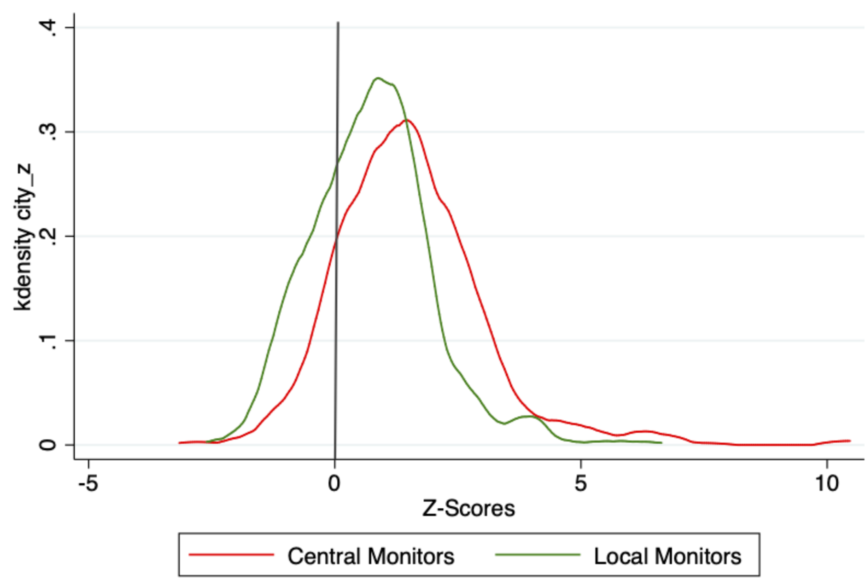
Note: This figure shows the roll-out of air pollution monitoring stations in China by three waves from 2012 to 2014. All monitors on the map are central government-controlled monitors.

Figure 3.2: Kernel Densities for PM2.5 Z-Scores

(a) Kernel Densities for PM2.5 Z-Scores: Central monitors vs. No monitor

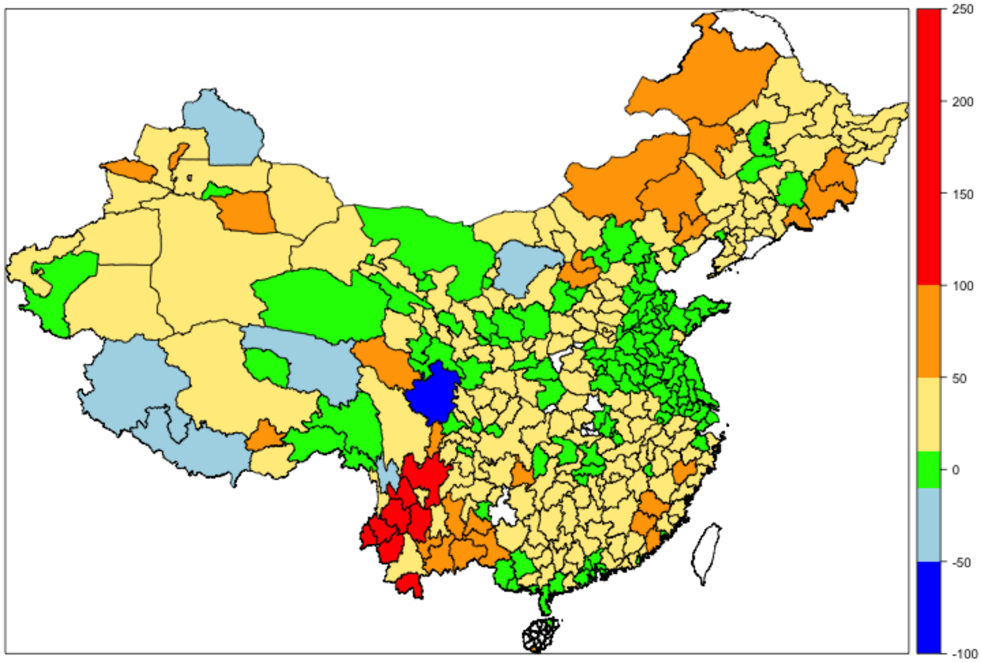


(b) Kernel Densities for PM2.5 Z-Scores: Central monitors vs. Local monitors



Note: Each figure shows the kernel density estimate for the distribution of city-level z-scores. Z-score is calculated by taking the observed value in grid cell i in city c and year t, subtract the average for that city, and scale it by the city level standard deviation. Figure (a) compares the distribution of city-level z-scores at cells containing central monitors to cells without monitors using data from 2009 to 2017. Figure (b) compares the distribution of city-level z-scores at cells containing central monitors to cells containing local monitors in 2016.

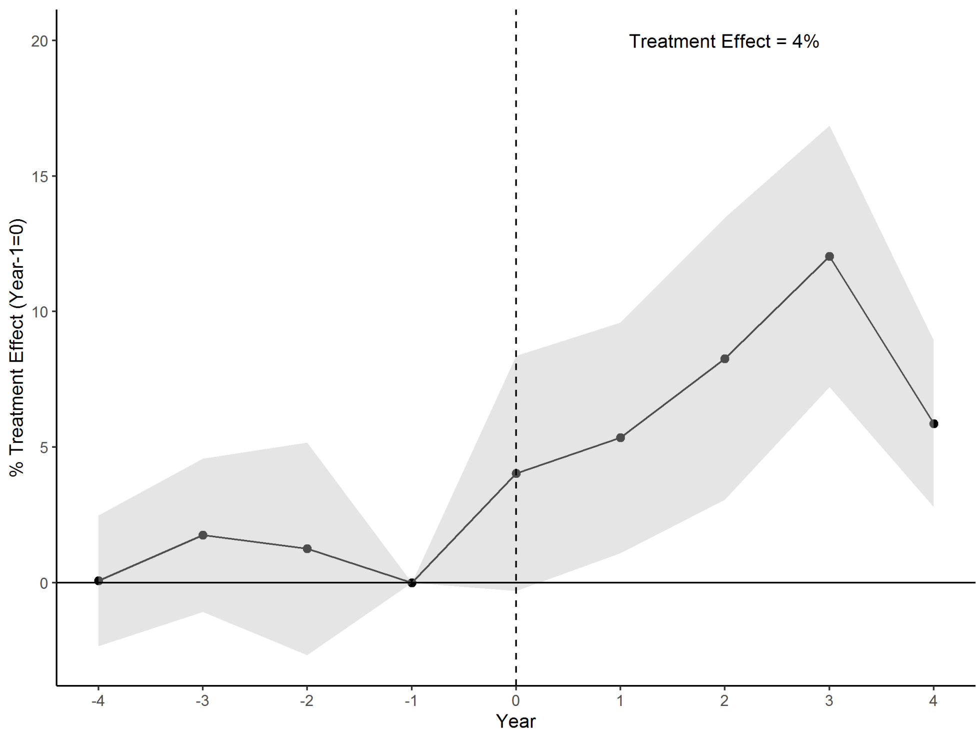
Figure 3.3: Monitor Representation Errors at Opening Years: All Cells vs. Monitored Cells



Error=(monitor avg-city avg)%
Error > 0: Monitors Over-Estimate City Avg Pollution;
Error < 0: Monitors Under-Estimate City Avg Pollution

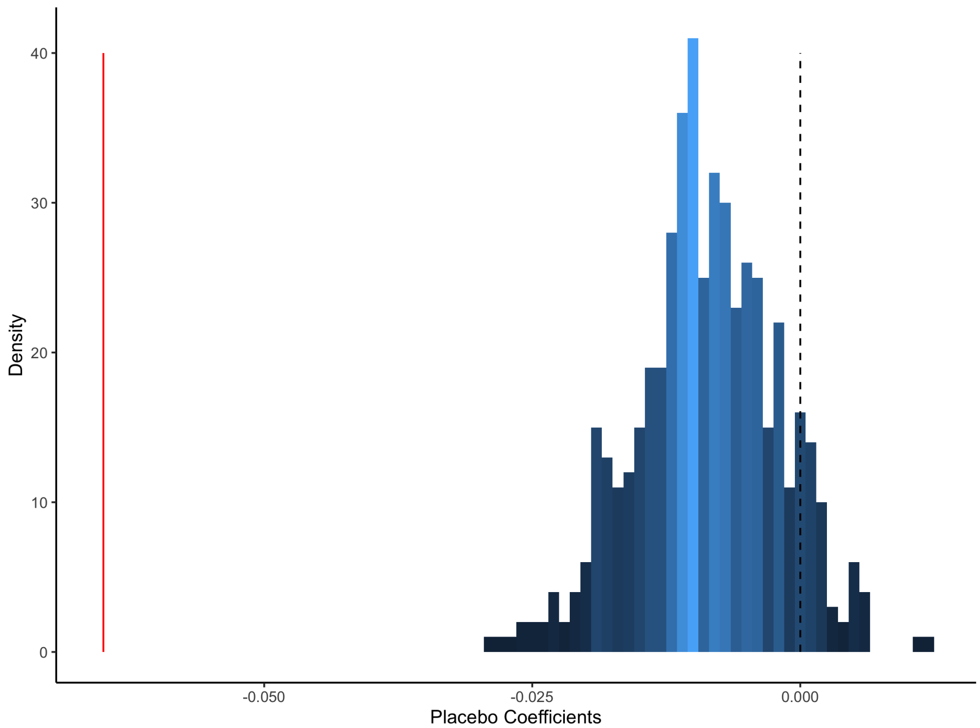
Note: This figure shows the monitors representation errors in the years of joining the new monitoring program. The representation error is defined as the percentage difference between city average pollution level calculated based on only monitored cells and city average pollution based on all cells. All the pollution levels are weighted by the 2015 grid-level population count. Cities in green means the monitors well-represent city average PM2.5, with representation errors in $[-10\%, 10\%]$. Cities in warm colors (error $> 10\%$) have monitors over-representing the city average pollution, and those in cool colors (error $< -10\%$) meansthat the monitors under-present city average pollution level. The map is based on raw data and presented at the city level. Representation error maps for each year from 2012 to 2017 are in Appendix A.1.

Figure 3.4: Event Study of Monitor Opening on Air Pollution



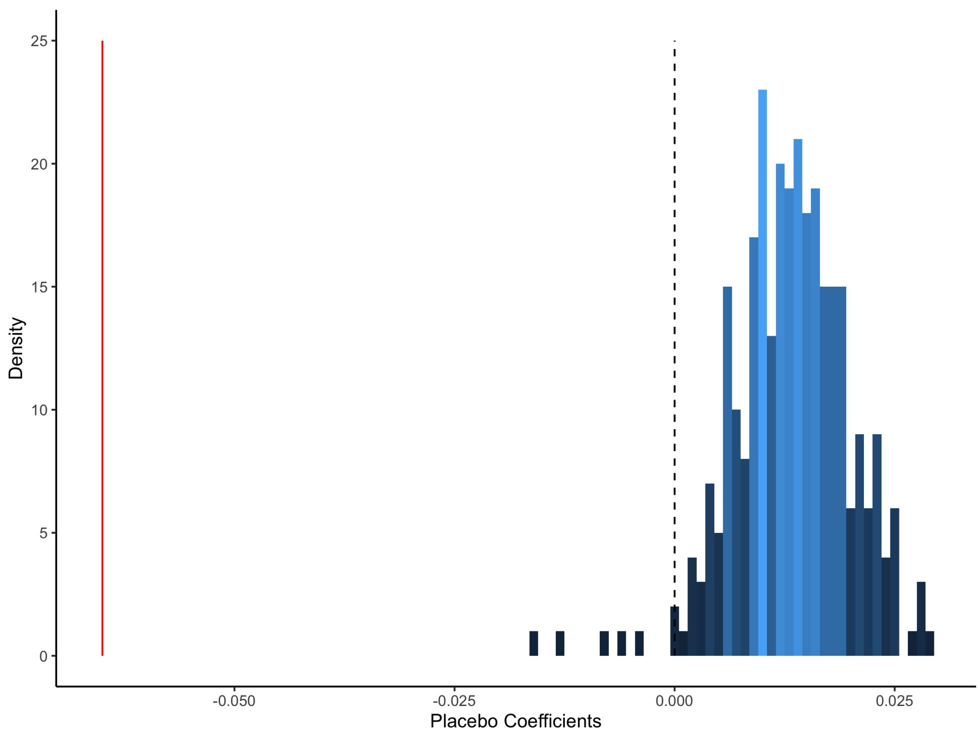
Note: This figure shows the event study results of monitor opening on air pollution controlling for cell fixed effects, year fixed effects, and wave-specific time trend. I regress the PM2.5 on four pre-opening indicators and four post-opening indicators. The year before monitoring is the base interval. Standard errors are clustered at city level.

Figure 3.5: Placebo Test with Random Opening Years in Pre-Monitoring Periods



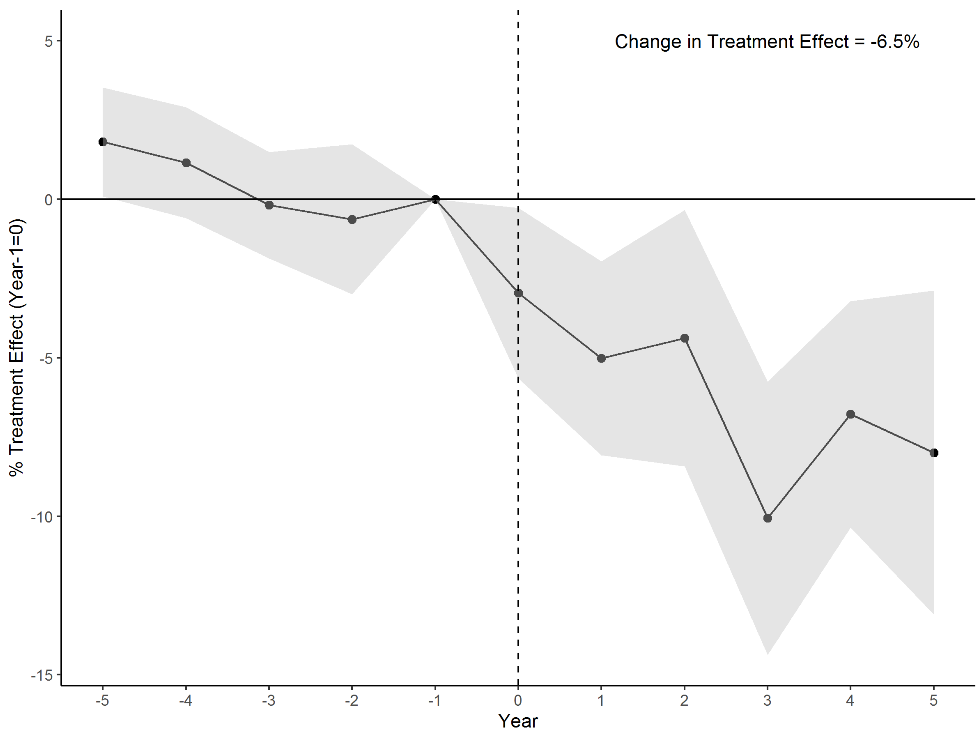
Note: This figure shows the results of a âĂIJplaceboâĂİ test using only pre-program periods and randomly assign each monitor an opening year. I conduct 500 estimations of the treatment intensity analysis and plot the distribution of the 500 placebo coefficients and compare them with the observed effect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

Figure 3.6: Placebo Test with Random Monitor Locations



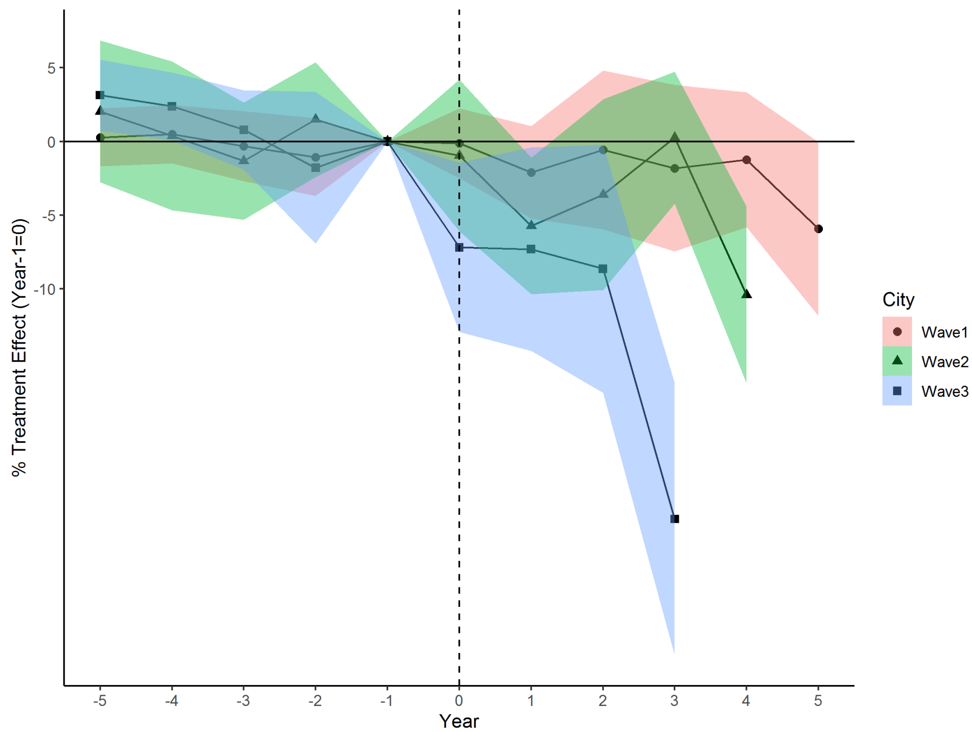
Note: This figure shows the results of a âĂIJplaceboâĂİ test that conducts 500 randomly relocations of all monitors within a city and keep the opening year unchanged. I conduct 500 estimations of equation (2) and plot the distribution of the 500 placebo coefficients, and compare them with the observed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

Figure 3.7: Event study: Change in Impact of Monitoring on Air Pollution in Monitored vs. Unmonitored Areas



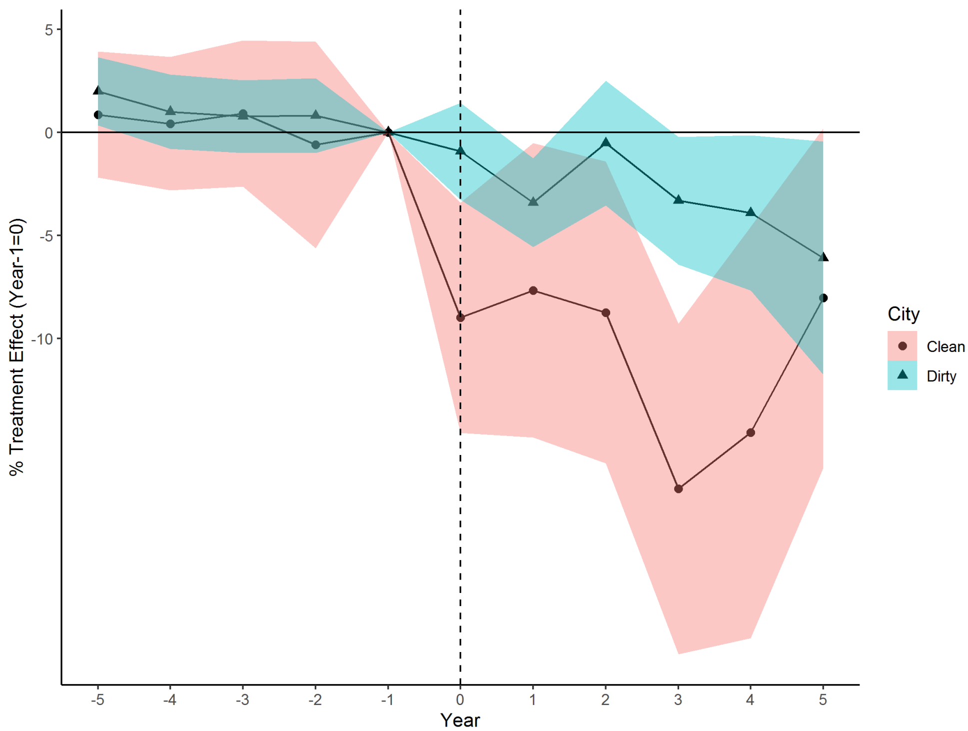
Note: This figure shows the event study results of monitor opening with treatment intensity on air pollution. (Column 1 from Table 3.4), where I regress the PM2.5 on interactions of treatment intensity indicator Near, and five pre-opening indicators and six post-opening indicators, controlling for cell fixed effects, and wave by year fixed effects. The year before monitoring is the base interval. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Figure 3.8: Heterogeneity Analysis by Waves: Change in Impact of Monitoring on Air pollution in Monitored vs. Unmonitored Areas



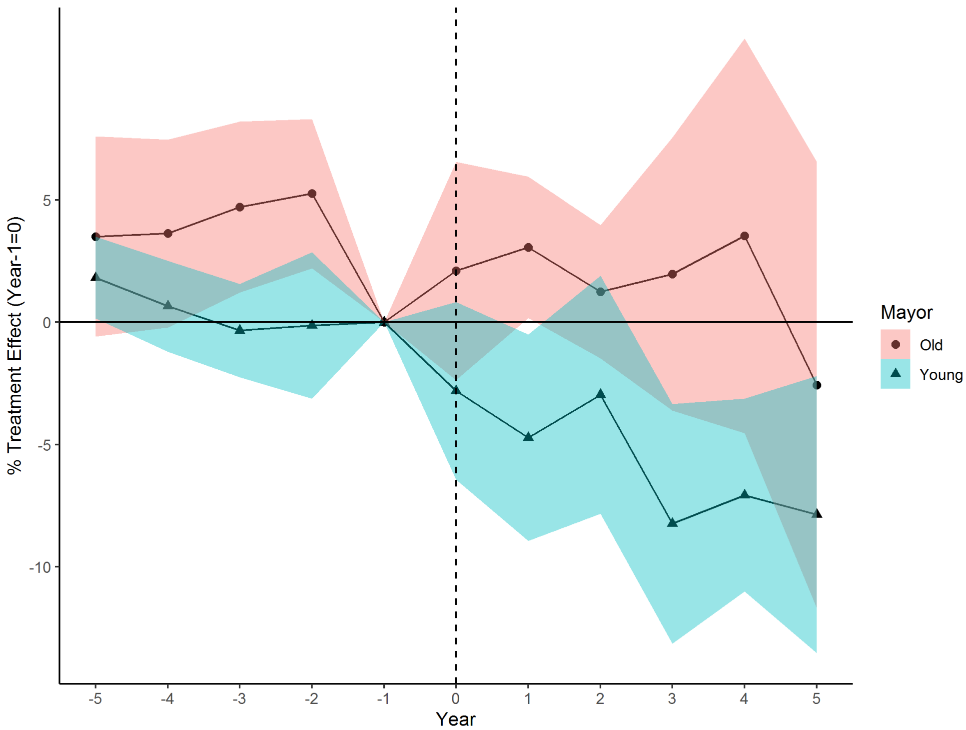
Note: This figure shows the event study results of monitor opening with treatment intensity on air pollution for three subsamples divided by roll-out waves. Using each city group, I regress the PM2.5 on interactions of treatment intensity indicator Near, and pre-opening and post-opening indicators, controlling for cell and year fixed effects. The year before monitoring is the base interval. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Figure 3.9: Heterogeneity Analysis by City Average Pollution Level: Change in Impact of Monitoring on Air Pollution in Monitored vs. Unmonitored Areas



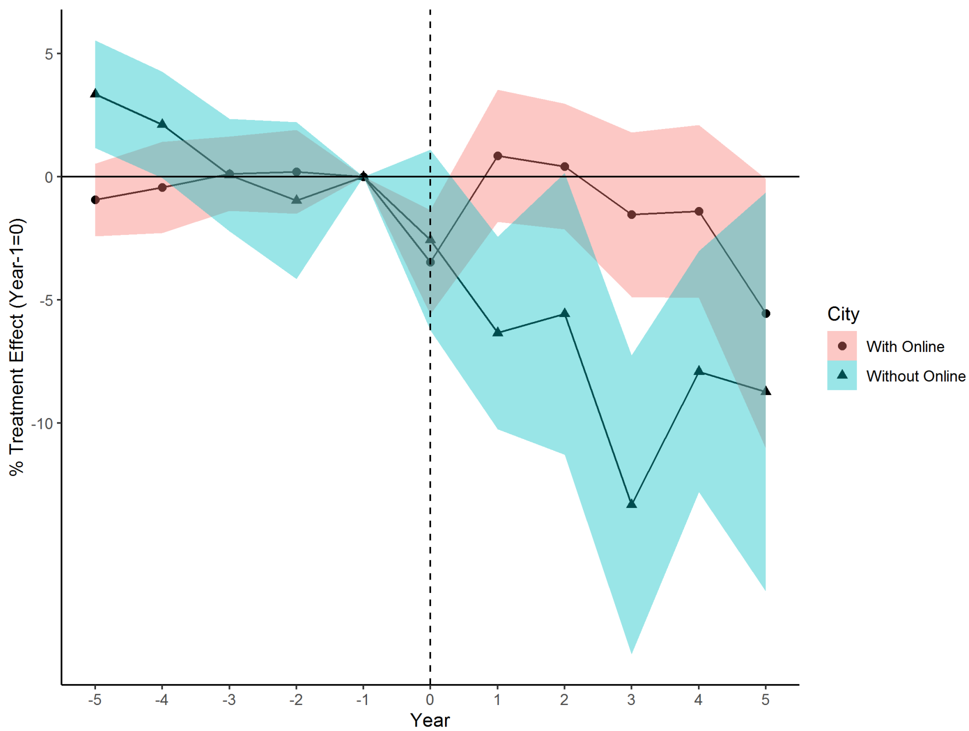
Note: This figure shows the event study results of monitor opening with treatment intensity on air pollution, for two subsamples of cities classified by comparing the citiesâĂŹ average PM2.5 level with national annual standard, 35 $\mu g/m^3$. Using each city group, I regress the PM2.5 on interactions of treatment intensity indicator Near, and five pre-opening indicators and six post-opening indicators, controlling for cell fixed effects, and wave by year fixed effects. The year before monitoring is the base interval. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Figure 3.10: Heterogeneity Analysis by City MayorsâĂŹ Age: Change in Impact of Monitoring on Air Pollution in Monitored vs. Unmonitored Areas



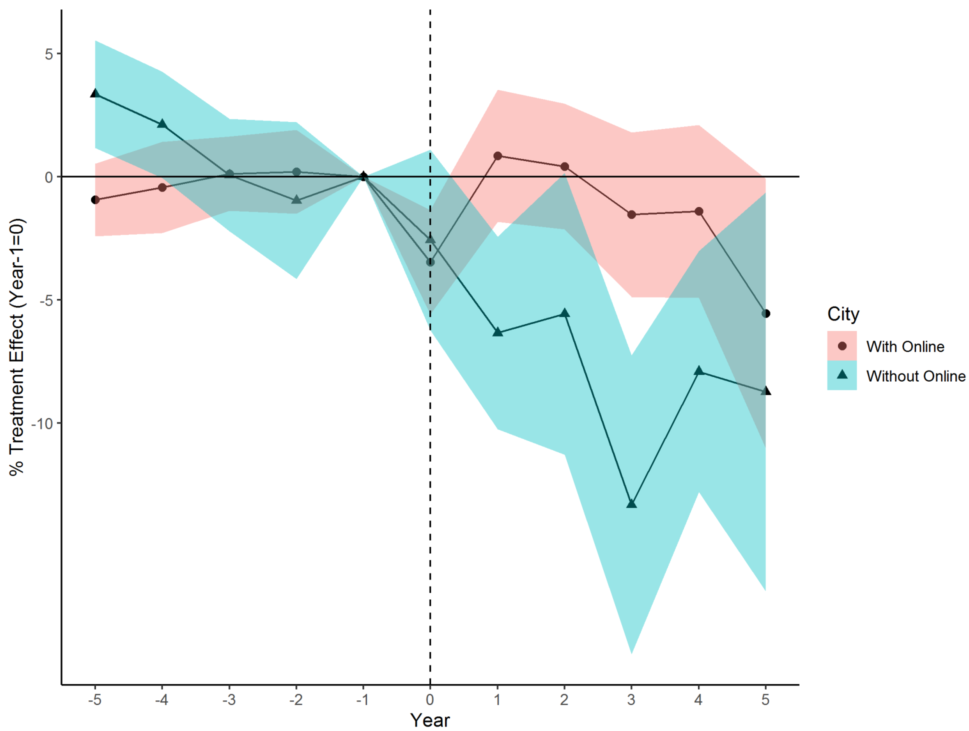
Note: This figure shows the event study results of monitor opening with treatment intensity on air pollution, for two subsamples of cities classified by city mayorsâĂŹ age. The cutoff point for mayorâĂŹs age is 57 because this is the ceiling threshold for a mayor to get promoted. Using each city group, I regress the PM2.5 on interactions of treatment intensity indicator Near, and five pre-opening indicators and six post-opening indicators, controlling for cell fixed effects, and wave by year fixed effects. The year before monitoring is the base interval. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Figure 3.11: Heterogeneity Analysis by Province Online Disclosure: Change in Impact of Monitoring on Air Pollution in Monitored vs. Unmonitored Areas



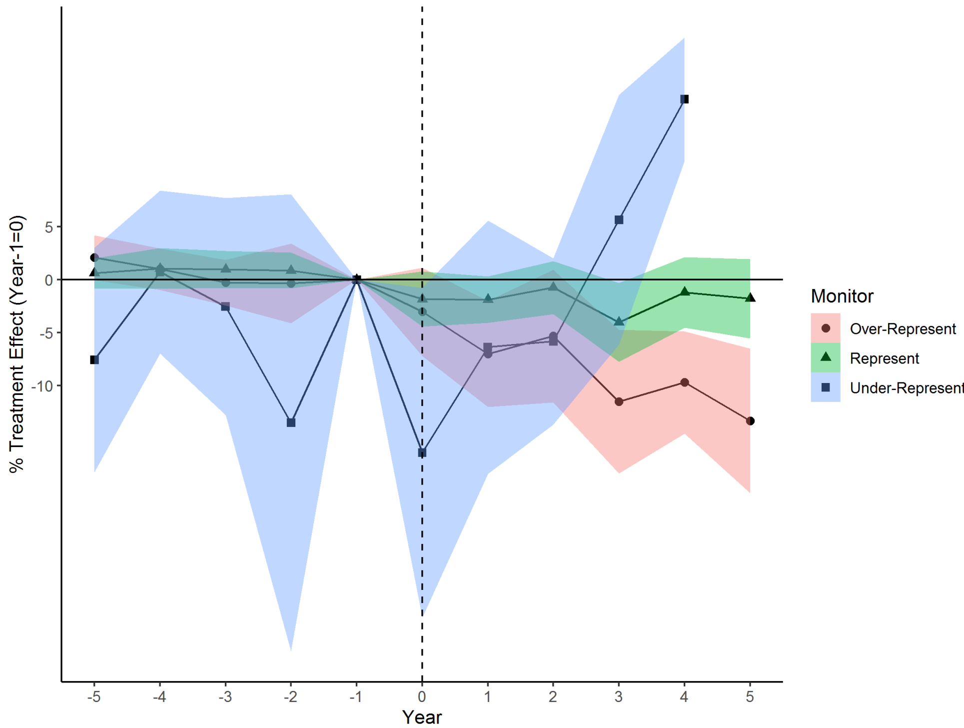
Note: This figure shows the event study results of monitor opening with treatment intensity on air pollution, for two subsamples divided by whether a province has its own online pollution disclosure platform or not. Using each city group, I regress the PM2.5 on interactions of treatment intensity indicator Near, and five pre-opening indicators and six post-opening indicators, controlling for cell fixed effects, and wave by year fixed effects. The year before monitoring is the base interval. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Figure 3.12: Heterogeneity Analysis by Province Online Disclosure: Change in Impact of Monitoring on Air Pollution in Monitored vs. Unmonitored Areas



Note: This figure shows the event study results of monitor opening with treatment intensity on air pollution, for two subsamples divided by whether a province has its own online pollution disclosure platform or not. Using each city group, I regress the PM2.5 on interactions of treatment intensity indicator Near, and five pre-opening indicators and six post-opening indicators, controlling for cell fixed effects, and wave by year fixed effects. The year before monitoring is the base interval. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Figure 3.13: Heterogeneity Analysis by Monitor Representativeness: the change in impact of monitoring on air pollution in monitored vs. unmonitored areas



Note: This figure shows the event study results of monitor opening with treatment intensity on air pollution, for three subsamples divided by monitors spatial representativeness at the years of opening. Representation errors are defined as the difference between population-weighted city average PM at monitored cells and at all cells. Over-represent cities have representation errors greater than 10%. Well-represent cities have error between -10% to 10%. Under-represent cities are with errors less than -10%. Using each city group, I regress the PM2.5 on interactions of treatment intensity indicator Near, and five pre-opening indicators and six post-opening indicators, controlling for cell fixed effects, and wave by year fixed effects. The year before monitoring is the base interval. Each estimate represents the difference in PM2.5 between monitored areas (cells within 3km of monitors) and unmonitored areas (cells outside 3km) at a given period. Standard errors are clustered at city level.

Table 3.1: Summar Statistics: Satellite-based Air Pollution (PM2.5, $\mu g/m^3$)

	Wave 1		Wave 2		Wave 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	city_avg	city_monavg	city_avg	city_monavg	city_avg	city_monavg
2009	54.25	59.26	49.23	56.48	40.02	48.10
	(17.24)	(18.35)	(17.23)	(17.21)	(19.98)	(21.31)
2010	53.10	57.98	50.23	57.65	41.34	49.22
	(17.44)	(18.46)	(19.93)	(19.88)	(22.90)	(23.84)
2011	50.51	55.35	47.15	54.34	38.15	45.99
	(17.54)	(18.66)	(18.26)	(18.57)	(20.40)	(21.47)
2012	46.91	51.45	44.97	52.08	36.40	44.13
	(16.20)	(17.56)	(18.35)	(18.73)	(19.51)	(20.78)
2013	54.66	59.73	51.55	58.92	41.30	49.34
	(20.81)	(22.16)	(21.38)	(21.57)	(22.56)	(23.74)
2014	55.31	60.29	50.58	57.97	41.83	50.48
	(18.73)	(19.81)	(19.14)	(18.89)	(22.85)	(24.03)
2015	51.67	56.54	47.92	54.45	37.85	44.87
	(18.51)	(19.68)	(18.25)	(18.37)	(19.63)	(20.25)
2016	46.48	51.33	43.37	51.23	34.32	42.41
	(17.71)	(18.59)	(17.33)	(17.83)	(18.65)	(20.53)
2017	52.48	56.34	47.98	54.73	40.57	47.45
	(15.51)	(15.81)	(15.77)	(15.84)	(19.01)	(20.14)

Notes: The underlying observations are at the city level. Standard deviations are in parentheses. Column (1), (3), (5) show population-weighted PM2.5, column (2), (4), (6) show population-weighted PM2.5 level at monitored cells are average of post-monitoring period. Population data is in 2015.

Table 3.2: Summar Statistics: Other Variables

	(1)	(2)	(3)
Variable	Wave1	Wave2	Wave3
City Pollution, GDP, Population			
Population Weighted PM2.5	51.71	48.11	39.09
	(17.95)	(18.55)	(20.76)
Population Weighted PM2.5 at Monitored Cells	55.87	55.48	46.55
	(19.27)	(18.67)	(21.60)
GDP Per Capita	63944	48187	30641
	(30064)	(31222)	(16396)
GDP in 3rd Industry	236.72	67.06	31.62
	(300.22)	(76.67)	(21.17)
GDP in 2nd Industry	204.54	94.08	44.45
	(173.30)	(89.73)	(30.89)
GDP in 1st Industry	20.52	16.20	14.79
	(15.71)	(10.53)	(9.88)
Population in 2015	4857550	2716721	1940899
	(3499542)	(1531311)	(1341832)
Leader's Characteristics			
Age	51.88	50.09	50.00
	(4.66)	(3.65)	(3.57)
Young (Age<57)	.930	.995	.984
	(.256)	(.070)	(.127)
Master	.497	.572	.540
	(.501)	(.495)	(.499)
PhD	.269	.218	.158
	(.444)	(.414)	(.364)
Bachelor	.2104121	.193	.261
	(.408)	(.395)	(.440)
Number of Cities	74	98	176

Notes: The underlying observations are at the city level. Standard deviations are in parentheses. Population-weighted PM2.5 are measured by 2009-2017 average, PM2.5 level at monitored cells are average of post-monitoring period. GDP data is from 2001-2017 for 281 cities. Leader's characteristics data ranges from 2009-2015.

Table 3.3: Baseline Difference in Differences Estimation Results

	Dependent variable: $\ln(PM2.5_{it})$					
	(1)	(2)	(3)	(4)	(5)	(6)
Open	0.188***	0.106***	0.048**	0.040**	0.049**	0.041**
	(0.038)	(0.026)	(0.022)	(0.020)	(0.022)	(0.020)
(0-3km)*Open					-0.100***	-0.062***
					(0.024)	(0.014)
Controls	No	Cell FE	Cell FE Year FE	Cell FE Year FE Wave×T	Cell FE Year FE	Cell FE Year FE Wave×T
Observations	84,349,384	84,349,384	84,349,384	84,349,384	83,293,774	83,293,774
R^2	0.009	0.958	0.965	0.966	0.965	0.966

Note: Column (1)-(6) show DID estimation results with different fixed effects. The first three columns represent baseline DID results, where *Open* is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. Column (4)-(6) show DID estimation results with treatment intensity defined by distances to monitors. Cells within 3km to the monitor are in the monitored group and cells outside 3km are unmonitored cells. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 3.4: Baseline DID Estimation Results, Alternative Unmonitored Areas

	Dependent variable: $\ln(PM2.5_{it})$					
Unmonitored Area:	>3km	>30km	>50km	>15km	>30km	>50km
	(1)	(2)	(3)	(4)	(5)	(6)
Open	0.041**	0.043*	0.043*	0.044**	0.046**	0.049**
	(0.020)	(0.022)	(0.025)	(0.021)	(0.022)	(0.024)
(0-3km)*Open	-0.062***	-0.068***	-0.076***	-0.066***	-0.070***	-0.079***
	(0.014)	(0.016)	(0.017)	(0.015)	(0.016)	(0.018)
(3-6km)*Open				-0.065***	-0.069***	-0.078***
				(0.014)	(0.015)	(0.017)
(6-9km)*Open				-0.066***	-0.071***	-0.079***
				(0.014)	(0.015)	(0.017)
(9-12km)*Open				-0.066***	-0.071***	-0.080***
				(0.014)	(0.015)	(0.017)
(12-15km)*Open				-0.067***	-0.072***	-0.081***
				(0.014)	(0.015)	(0.017)
Observations	83,293,774	74,496,330	65,280,883	83,293,774	77,485,464	68,270,017
R^2	0.966	0.966	0.966	0.966	0.966	0.966

Note: Column (1)-(6) show DID estimation results. *Open* is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. The first three columns use cells within 3km to the monitor as the monitored group and compare different unmonitored groups: cells outside 3km, 30km or 50km of the monitors. Column (4)-(6) add four more distance bins to the monitored group. All columns include cell fixed effects, year fixed effects and a wave specific time trend. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 3.5: Difference in Differences with Alternative Treatment Intensity Bins Estimation Results

	Dependent variable: $\ln(PM2.5_{it})$					
Unmonitored Area:	>3km	>30km	>50km	>15km	>30km	>50km
	(1)	(2)	(3)	(4)	(5)	(6)
(0-3km)*Open	-0.065***	-0.072***	-0.079***	-0.069***	-0.074***	-0.083***
	(0.013)	(0.015)	(0.016)	(0.014)	(0.015)	(0.017)
(3-6km)*Open				-0.068***	-0.073***	-0.082***
				(0.014)	(0.015)	(0.016)
(6-9km)*Open				-0.069***	-0.074***	-0.083***
				(0.014)	(0.014)	(0.016)
(9-12km)*Open				-0.069***	-0.074***	-0.084***
				(0.014)	(0.015)	(0.016)
(12-15km)*Open				-0.070***	-0.075***	-0.084***
				(0.014)	(0.015)	(0.016)
Observations	83,293,774	74,496,330	65,280,883	83,293,774	77,485,464	68,270,017
R^2	0.967	0.966	0.966	0.967	0.967	0.967

Note: Column (1)-(6) show DID estimation results with treatment intensity defined by distances to monitors. *Open* is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. The first three columns use cells within 3km to the monitor as the monitored group and compare different unmonitored groups: cells outside 3km, 30km or 50km of the monitors. Column (4)-(6) add four more distance bins to the monitored group. All columns include both the cell FE and Wave×Year FE. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 3.6: Event Study with Alternative Treatment Intensity Groups Estimation Results

	Dependent variable: $\ln(PM2.5_{it})$					
Monitored Area:	≤3km			≤10km		
Unmonitored Area:	>3km	>30km	>50km	>10km	>30km	>50km
	(1)	(2)	(3)	(4)	(5)	(6)
Near*(y-5)	0.018**	0.021**	0.025**	0.022**	0.024**	0.028***
	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.010)
Near*(y-4)	0.012	0.013	0.016	0.014	0.015	0.017
	(0.009)	(0.010)	(0.011)	(0.009)	(0.010)	(0.011)
Near*(y-3)	-0.002	-0.003	-0.007	-0.001	-0.003	-0.006
	(0.009)	(0.010)	(0.012)	(0.009)	(0.010)	(0.011)
Near*(y-2)	-0.006	-0.005	-0.008	-0.010	-0.008	-0.012
	(0.012)	(0.013)	(0.015)	(0.012)	(0.013)	(0.015)
Near*(y0)	-0.030**	-0.030**	-0.033*	-0.033**	-0.033**	-0.037**
	(0.014)	(0.015)	(0.017)	(0.014)	(0.015)	(0.017)
Near*(y+1)	-0.050***	-0.054***	-0.061***	-0.049***	-0.052***	-0.059***
	(0.016)	(0.017)	(0.019)	(0.016)	(0.017)	(0.019)
Near*(y+2)	-0.044**	-0.044*	-0.051*	-0.049**	-0.049**	-0.056**
	(0.021)	(0.024)	(0.028)	(0.021)	(0.023)	(0.027)
Near*(y+3)	-0.101***	-0.108***	-0.121***	-0.105***	-0.111***	-0.126***
	(0.022)	(0.025)	(0.029)	(0.022)	(0.025)	(0.028)
Near*(y+4)	-0.068***	-0.080***	-0.095***	-0.074***	-0.084***	-0.100***
	(0.018)	(0.022)	(0.026)	(0.017)	(0.020)	(0.024)
Near*(y+5)	-0.080***	-0.098***	-0.118**	-0.080***	-0.095***	-0.116***
	(0.026)	(0.034)	(0.046)	(0.026)	(0.033)	(0.045)
Observations	87,843,991	77,211,912	67,374,686	87,843,991	78,853,329	69,016,103
R^2	0.966	0.966	0.967	0.966	0.967	0.967

Note: Column (1)-(6) show event study results with different treatment intensity groups. *Near* is the monitored area indicator which equals one for cells within 3km from monitors in column (1)-(3), and 10km from monitors in column (4)-(6). y-5, y-4,...,y+5 represent each year within the 5-year time window around monitor openings. For each monitored group, the three columns compare different unmonitored groups: cells outside 3km (10km), 30km or 50km of the monitors. All columns include both the cell FE and Wave×Year FE. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 3.7: Heterogeneity Analysis: Difference in Differences with Treatment Intensity Estimation Results

	Dependent variable: $\ln(PM2.5_{it})$					
Unmonitored Area:	>3km	>30km	>50km	>15km	>30km	>50km
	(1)	(2)	(3)	(4)	(5)	(6)
(0-3km)*Open	-0.073***	-0.079***	-0.087***	-0.077***	-0.082***	-0.091***
	(0.015)	(0.017)	(0.019)	(0.016)	(0.017)	(0.019)
(3-6km)*Open				-0.071***	-0.076***	-0.085***
				(0.015)	(0.016)	(0.017)
(6-9km)*Open				-0.069***	-0.074***	-0.083***
				(0.014)	(0.014)	(0.016)
(9-12km)*Open				-0.070***	-0.075***	-0.085***
				(0.013)	(0.014)	(0.016)
(12-15km)*Open				-0.072***	-0.077***	-0.086***
				(0.013)	(0.014)	(0.016)
(0-3km)*Open*Dirtier	0.010	0.010	0.010	0.010	0.010	0.010
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
(3-6km)*Open*Dirtier				0.004	0.005	0.005
				(0.010)	(0.010)	(0.010)
(6-9km)*Open*Dirtier				0.0001	0.0002	0.0003
				(0.012)	(0.012)	(0.012)
(9-12km)*Open*Dirtier				0.002	0.002	0.002
				(0.012)	(0.012)	(0.012)
(12-15km)*Open*Dirtier				0.004	0.004	0.004
				(0.012)	(0.012)	(0.012)
Observations	83,293,774	74,496,330	65,280,883	83,293,774	77,485,464	68,270,017
R^2	0.967	0.966	0.966	0.967	0.967	0.967

Note: Column (1)-(6) show DID estimation results with treatment intensity defined by distances to monitors. *Open* is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. *Dirtier* is a dummy variable indicating if the pollution in a cell is above the average city PM2.5. The first three columns use cells within 3km to the monitor as the monitored group and compare different unmonitored groups: cells outside 3km, 30km or 50km of the monitors. Column (4)-(6) add four more distance bins to the monitored group. All columns include both the cell FE and Wave×Year FE. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 3.8: Robustness Check Using Raw AOD Data

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

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

Table 3.9: Hereogeneous Analysis: Clean vs. Dirty Cities by Roll-out Waves

	Dependent variable: $\ln(PM2.5_{it})$			
	All	Wave1	Wave2	Wave3
	(1)	(2)	(3)	(4)
(0-3km)*Open	-0.062***	-0.015	-0.052***	-0.133***
	(0.013)	(0.015)	(0.018)	(0.035)
(0-3km)*Open*1(Clean City)	-0.030	-0.085***	0.003	-0.031
	(0.025)	(0.024)	(0.056)	(0.034)
(0-3km)*Open*1(Clean City)*Compliance	-0.004**	-0.008**	-0.009	-0.006***
	(0.002)	(0.003)	(0.012)	(0.002)
Observations	86,844,613	9,856,110	16,784,304	60,204,199
R^2	0.967	0.954	0.954	0.963

Note: Column (1)-(4) show DID estimation results with heterogeneity in treatment effect, separated by roll-out waves. Open is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. $\mathbf{1}(CleanCity)$ is a dummy variable indicating cells inside a city with average pollution level (based on monitored cells at the opening years) below the national standard, 35 μ/g^3. *Compliance* represents the closeness to the national standard. All columns use cells within 3km to the monitor as the monitored group and cells outside 3km as unmonitored groups. Column (2)-(4) show the estimation using subsamples of cities in three waves. All columns include both the cell FE and year FE (Wave×Year FE for column (1)). Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 3.10: Hereogeneous Analysis: GDP Growth Pressure by Roll-out Waves

	Dependent variable: $\ln(PM2.5_{it})$			
	All	Wave1	Wave2	Wave3
	(1)	(2)	(3)	(4)
(0-3km)*Open	-0.064**	-0.017	-0.045**	-0.139*
	(0.025)	(0.013)	(0.022)	(0.072)
(0-3km)*Open*1(Economic Decline)	0.079***	0.031	0.092***	0.116***
	(0.021)	(0.047)	(0.027)	(0.029)
Observations	41,532,584	7,983,534	10,647,527	22,901,523
R^2	0.945	0.956	0.933	0.940

Note: Column (1)-(4) show DID estimation results with heterogeneity in treatment effect, separated by roll-out waves. Open is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. $\mathbf{1}(EconomicDecline)$ is a dummy variable indicating cells inside a city that experienced an economic recession in the previous year (decreased GDP). All columns use cells within 3km to the monitor as the monitored group and cells outside 3km as unmonitored groups. Column (2)-(4) show the estimation using subsamples of cities in three waves. All columns include both the cell FE and Wave×Year FE. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

Table 3.11: Impact of Monitoring on Air Pollution with respect to Distances from Monitors

	Dependent variable: ln(PM2.5)		
	(1)	Population (Million)	(2)
Open	0.151***	Outside 300km	1.914
	(0.048)		
(0-10km)*Open	-0.172***	0-10km	261.386
	(0.041)		
(10-20km)*Open	-0.172***	10-20km	108.383
	(0.042)		
(20-30km)*Open	-0.173***	20-30km	97.286
	(0.042)		
(30-40km)*Open	-0.170***	30-40km	88.179
	(0.043)		
(40-50km)*Open	-0.165***	40-50km	80.283
	(0.043)		
(50-60km)*Open	-0.160***	50-60km	70.712
	(0.044)		
(60-70km)*Open	-0.153***	60-70km	55.526
	(0.045)		
(70-80km)*Open	-0.145***	70-80km	43.907
	(0.046)		
(80-90km)*Open	-0.135***	80-90km	32.314
	(0.047)		
(90-100km)*Open	-0.126***	90-100km	24.369
	(0.048)		
(100-150km)*Open	-0.092*	100-150km	51.003
	(0.053)		
(150-200km)*Open	-0.052	150-200km	11.966
	(0.058)		
(200-300km)*Open	-0.051	200-300km	8.113
	(0.050)		
Observations	83,293,774		
R^2	0.967		

Note: This table shows DID estimation results with treatment intensity bins. Open is the treatment indicator that equals one if a cell is in a city that has joined the new monitoring program. The coefficient estimates of Open represents the impact of monitoring on air pollution in the base group, which includes cells outside of 300km of monitors. The interactions represent the effect in each treatment intensity group. Column (2) shows the total population in each distance bin using 2015 population data. Controls include both the cell FE and Wave×Year FE. Standard errors are clustered at city level. Significance: *p<0.1; **p<0.05; ***p<0.01

BIBLIOGRAPHY

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

Allcott, Hunt, "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market," *American Economic Journal: Economic Policy*, August 2013, *5* (3), 30–66.

_ **and Christopher Knittel**, "Are Consumers Poorly Informed about Fuel Economy? Evidence from Two Experiments," *American Economic Journal: Economic Policy*, February 2019, *11* (1), 1–37.

_ **and Nathan Wozny**, "Gasoline Prices, Fuel Economy, and the Energy Paradox," *Review of Economics and Statistics*, December 2014, *96* (5), 779–795.

_ **, Christopher Knittel, and Dmitry Taubinsky**, "Tagging and Targeting of Energy Efficiency Subsidies," *American Economic Review*, May 2015, *105* (5), 187–191.

Anas, A. and R. Lindsey, "Reducing Urban Road Transportation Externalities: Road Pricing in Theory and in Practice," *Review of Environmental Economics and Policy*, December 2011, *5* (1), 66–88.

Anas, Alex and Govinda R. Timilsina, "Lock-in Effects of Road Expansion on Co2 Emissions: Results from a Core-Periphery Model of Beijing," *Policy Research Working Papers*, August 2009, pp. The World Bank, Washington, D.C.

Anderson, Michael L., "Subways, Strikes, and Slowdowns: The Impacts of Public Transportation on Traffic Congestion," *American Economic Review*, 2014, *104(9)*, 2763–2796.

Anderson, Soren T. and James M. Sallee, "Designing Policies to Make Cars Greener," *Annual Review of Resource Economics*, October 2016, *8* (1), 157–180.

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

Ashenfelter, Orley and Michael Greenstone, "Using Mandated Speed Limits to Measure the Value of a Statistical Life," *Journal of Political Economy*, 2004, *112* (S226-67).

Auffhammer, Maximilian and Ryan Kellogg, "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality," *American Economic Review*, October 2011, *101* (6), 2687–2722.

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

Austin, David and Terry Dinan, "Clearing the Air: The Costs and Consequences of Higher Cafe Standards and Increased Gasoline Taxes," *Journal of Environmental Economics and Management*, November 2005, *50* (3), 562–582.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S. Shapiro, "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century," *Journal of Political Economy*, February 2016, *124* (1), 105–159.

Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur, "The Morbidity Cost of Air Pollution: Evidence from Consumer Spending in China," *Working Paper*, 2019, pp. Cornell University. Ithaca, NY.

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

Bauernschuster, Stefan, Timo Hener, and Helmut Rainer, "When Labor Disputes Bring Cities to a Standstill: The Impact of Public Transit Strikes on Traffic, Accidents, Air Pollution, and Health," *American Economic Journal: Economic Policy*, February 2017, *9* (1), 1–37.

Baum-Snow, Nathaniel, "Did Highways Cause Suburbanization?," *Quarterly Journal of Economics*, 2007, *122* (2), 775–805.

Beaudoin, Justin and C.-Y. Cynthia Lin-Lawell, "Is Public TransitâĂŹs âĂŸGreenâĂŹ Reputation Deserved? Evaluating the EïňĂects of Transit Supply on Air Quality," *Working paper*, 2017, pp. University of Washington Tacoma and Cornell University, Ithaca, NY.

Beevers, S and D Carslaw, "The Impact of Congestion Charging on Vehicle Emissions in London," *Atmospheric Environment*, January 2005, *39* (1), 1–5.

Bel, GermÃă and Maximilian Holst, "Evaluation of the Impact of Bus Rapid Transit on Air Pollution," *Transport Policy*, 2018, *63*, 209–220.

Beresteanu, Arie and Shanjun Li, "Gasoline Prices, Government Support, and the Demand for Hybrid Vehicles in the United States: Hybrid Vehicle Demand," *International Economic Review*, February 2011, *52* (1), 161–182.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Bigazzi, Alexander Y. and Miguel A. Figliozzi, "Marginal Costs of Freeway Traffic Congestion with on-Road Pollution Exposure Externality," *Transportation Research Part A: Policy and Practice*, November 2013, *57*, 12–24.

Black, Dan A., Natalia Kolesnikova, and Lowell J. Taylor, "Why Do so Few Women Work in New York (and so Many in Minneapolis)? Labor Supply of Married Women Across Us Cities," *Journal of Urban Economics*, January 2014, *79*, 59–71.

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

Burke, Marshall and Kyle Emerick, "Adaptation to Climate Change: Evidence from US Agriculture," *American Economic Journal: Economic Policy*, August 2016, *8* (3), 106–140.

Burnett, Richard T., C. Arden Pope, Majid Ezzati, Casey Olives, Stephen S. Lim, Sumi Mehta, Hwashin H. Shin, Gitanjali Singh, Bryan Hubbell, Michael Brauer, H. Ross Anderson, Kirk R. Smith, John R. Balmes, Nigel G. Bruce, Haidong Kan, Francine Laden, Annette PrÃijss-UstÃijn, Michelle C. Turner, Susan M. Gapstur, W. Ryan Diver, and Aaron Cohen, "An Integrated Risk Function for Estimating the Global Burden of Disease

Attributable to Ambient Fine Particulate Matter Exposure," *Environmental Health Perspectives*, April 2014, *122* (4), 397–403.

Busse, Meghan R, Christopher R Knittel, and Florian Zettelmeyer, "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review*, February 2013, *103* (1), 220–256.

BÃl'nabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, 2006, *96* (5), 1652–1678.

Carter, Charles, "The War on Smog.," 2019.

Chandra, Ambarish, Sumeet Gulati, and Milind Kandlikar, "Green Drivers or Free Riders? An Analysis of Tax Rebates for Hybrid Vehicles," *Journal of Environmental Economics and Management*, September 2010, *60* (2), 78–93.

Chay, Kenneth Y. and Michael Greenstone, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *Quarterly Journal of Economics*, 2003, *118*, 1121–1167.

Chen, Xiaojie and Jinhua Zhao, "Bidding to Drive: Car License Auction Policy in Shanghai and Its Public Acceptance," *Transport Policy*, May 2013, *27*, 39–52.

Chen, Yihsu and Alexander Whalley, "Green Infrastructure: The Effects of Urban Rail Transit on Air Quality," *American Economic Journal: Economic Policy*, 2012, *4(1)*, 58–97.

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

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

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

Chin, Anthony and Peter Smith, "Automobile Ownership and Government Policy: The Economics of Singapore's Vehicle Quota Scheme," *Transportation Research Part A: Policy and Practice*, March 1997, *31* (2), 129–140.

Chugh, Randy, Maureen Cropper, and Urvashi Narain, "The Cost of Fuel Economy in the Indian Passenger Vehicle Market," *Energy Policy*, November 2011, *39* (11), 7174–7183.

Cohen, Aaron J, Michael Brauer, Richard Burnett, H Ross Anderson, Joseph Frostad, Kara Estep, Kalpana Balakrishnan, Bert Brunekreef, Lalit Dandona, Rakhi Dandona, Valery Feigin, Greg Freedman, Bryan Hubbell, Amelia Jobling, Haidong Kan, Luke Knibbs, Yang Liu, Randall Martin, Lidia Morawska, C Arden Pope, Hwashin Shin, Kurt Straif, Gavin Shaddick, Matthew Thomas, Rita van Dingenen, Aaron van Donkelaar, Theo Vos, Christopher J L Murray, and Mohammad H Forouzanfar, "Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015," *The Lancet*, 2017, *389* (10082), 1907–1918.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe, "Using Big Data to Estimate Consumer Surplus: The Case of Uber," *NBER Working Paper*, 2016, p. w22627.

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

_ **and Reed Walker**, "Traffic Congestion and Infant Health: Evidence from E-ZPass," *American Economic Journal: Applied Economics*, 2011, *3(1)*, 65–90.

da Silva, Cacilda Bastos Pereira, Paulo HilÃąrio Nascimento Saldiva, Luis Fernando Amato-LourenÃğo, Fernando Rodrigues-Silva, and Simone Georges El Khouri Miraglia, "Evaluation of the Air Quality Benefits of the Subway System in sÃčo Paulo, Brazil," *Journal of Environmental Management*, June 2012, *101*, 191–196.

Daniel, Joseph I. and Khalid Bekka, "The Environmental Impact of Highway Congestion Pricing," *Journal of Urban Economics*, March 2000, *47* (2), 180–215.

Davis, Lucas W., "The Effect of Driving Restrictions on Air Quality in Mexico City," *Journal of Political Economy*, 2008, *116*, 38–81.

_ , "How Much Are Electric Vehicles Driven?," *Applied Economics Letters*, October 2019, *26* (18), 1497–1502.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken, "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, July 2012, *4* (3), 66–95.

Deryugina, Tatyana, Garth Heutel, Nolan Miller, David Molitor, and Julian Reif, "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction," *NBER Working Paper*, 2019, p. w22796.

Deschenes, Olivier, Michael Greenstone, and Joseph S. Shapiro, "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program," *American Economic Review*, October 2017, *107* (10), 2958–89.

DeschÃłnes, Olivier and Enrico Moretti, "Extreme Weather Events, Mortality, and Migration," *The Review of Economics and Statistics*, 2009, *91*, 659–681.

_ **and Michael Greenstone**, "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *The American Economic Review*, 2007, *97* (1), 354–385.

_ **and** _ , "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US," *American Economic Journal: Applied Economics*, 2011, *3* (4), 152–185.

DeShazo, J. R., "Improving Incentives for Clean Vehicle Purchases in the United States: Challenges and Opportunities," *Review of Environmental Economics and Policy*, January 2016, *10* (1), 149–165.

Dimitropoulos, Alexandros, Jos N. van Ommeren, Paul Koster, and Piet Rietveld, "Not fully charged: Welfare effects of tax incentives for employer-provided electric cars," *Journal of Environmental Economics and Management*, July 2016, *78*, 1–19.

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

Duranton, Gilles and Matthew A Turner, "The Fundamental Law of Road Congestion: Evidence from US Cities," *American Economic Review*, October 2011, *101* (6), 2616–2652.

_ **and Matthew A. Turner**, "Urban Growth and Transportation," *The Review of Economic Studies*, October 2012, *79* (4), 1407–1440.

Ebenstein, Avraham, Maoyong Fan, Guojun He, Michael Greenstone, and Maigeng Zhou, "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*, 2017, *114* (39), 10384–10389.

Epple, Dennis and Holger Sieg, "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, August 1999, *107* (4), 645–681.

Ewing, Reid and Robert Cervero, "Travel and the built environment," *Journal of the American Planning Association*, 2010, *76(3)*, 265–294.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *The Review of Economic Studies*, 2014, *81* (3), 1046–1070.

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

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

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

Fu, Shihe and Yizhen Gu, "Highway Toll and Air Pollution: Evidence from Chinese Cities," *Journal of Environmental Economics and Management*, May 2017, *83*, 32–49.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew Turner, "Subways and Urban Air Pollution," *NBER Working Paper*, January 2018, p. w24183.

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

Gibson, Matthew and Maria Carnovale, "The Effects of Road Pricing on Driver Behavior and Air Pollution," *Journal of Urban Economics*, September 2015, *89*, 62–73.

Goel, Deepti and Sonam Gupta, "The Effect of Metro Expansions on Air Pollution in Delhi," *The World Bank Economic Review*, February 2017, *31* (1), 271–294.

Goldberg, Pinelopi Koujianou, "The Effects of the Corporate Average Fuel Efficiency Standards in the US," *The Journal of Industrial Economics*, 1998, *46* (1), 1–33.

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

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

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

Greene, David, "How Consumers Value Fuel Economy: A Literature Review.," 2010.

Greenstone, Michael and Rema Hanna, "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India," *American Economic Review*, 2014, *104*, 3038–3072.

_ , Guojun He, Ruixue Jia, and Tong Liu, "Can Technology Solve the Principal-Agent Problem? Evidence from ChinaâĂŹs War on Air Pollution," *American Economic Review: Insights*, Forthcoming.

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

Grigolon, Laura, Mathias Reynaert, and Frank Verboven, "Consumer Valuation of Fuel Costs and Tax Policy: Evidence from the European Car Market," *American Economic Journal: Economic Policy*, August 2018, *10* (3), 193–225.

Gu, Yizhen, Junfu Zhang, and Ben Zou, "Subways and Road Congestion," *SSRN Electronic Journal*, 2018.

Hahn, Robert and Robert Metcalfe, "The Ridesharing Revolution: Economic Survey and Synthesis.," *Working Paper*, 2017, pp. Brookings Institution, Washington, DC.

Hall, Jonathan D., Craig Palsson, and Joseph Price, "Is Uber a Substitute or Complement for Public Transit?," *Journal of Urban Economics*, November 2018, *108*, 36–50.

Hansen, Christian B., "Generalized least squares inference in panel and multi-level models with serial correlation and fixed effects," *Journal of Econometrics*, October 2007, *140* (2), 670–694.

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

Hausman, Catherine and David S Rapson, "Regression Discontinuity in Time: Considerations for Empirical Applications," *Annual Review of Resource Economics*, 2018, *10*, 533–52.

He, Guojun, Maoyong Fan, and Meigong Zhou, "The Effect of Air Pollution on Mortality in China: Evidence from the 2008 Beijing Olympic Games," *Journal of Environmental Economics and Management*, 2016, *79*, 18–39.

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

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

Holland, Stephen P., Erin T. Mansur, Nicholas Z. Muller, and Andrew J. Yates, "Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors," *American Economic Review*, December 2016, *106* (12), 3700–3729.

Hsiang, Solomon M., Kyle C. Meng, and Mark A. Cane, "Civil Conflicts Are Associated with the Global Climate," *Nature*, August 2011, *476* (7361), 438–441.

Huse, Cristian and Claudio Lucinda, "The Market Impact and the Cost of Environmental Policy: Evidence from the Swedish Green Car Rebate," *The Economic Journal*, August 2014, *124* (578), F393–F419.

Ito, Koichiro and James M. Sallee, "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel Economy Standards," *The Review of Economics and Statistics*, May 2018, *100* (2), 319–336.

Jacobsen, Mark R, "Evaluating US Fuel Economy Standards in a Model with Producer and Household Heterogeneity," *American Economic Journal: Economic Policy*, May 2013, *5* (2), 148–187.

Jacobsen, Mark R. and Arthur A. van Benthem, "Vehicle Scrappage and Gasoline Policy," *American Economic Review*, March 2015, *105* (3), 1312–1338.

Jia, Ruixue, "Weather Shocks, Sweet Potatoes and Peasant Revolts in Historical China," *The Economic Journal*, March 2014, *124* (575), 92–118.

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

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

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

__ , *Green Cities:Urban Growth and the Environment*, Brookings Institution Press, 2006.

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

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

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

Kitagawa, Toru and Aleksey Tetenov, "Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice," *Econometrica*, 2018, *86* (2), 591–616.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders, "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health," *Review of Economics and Statistics*, May 2016, *98* (2), 350–366.

Koh, Winston T.H., "Control of Vehicle Ownership and Market Competition: Theory and SingaporeâĂŹs Experience with the Vehicle Quota System," *Transportation Research Part A: Policy and Practice*, November 2003, *37* (9), 749–770.

Kreindler, Gabriel E, "The Welfare EïňĂect of Road Congestion Pricing: Experimental Evidence and Equilibrium Implications," *Working Paper*, 2018, pp. MIT Department of Economics, Cambridge, MA.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins, "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets," *Journal of Economic Literature*, December 2013, *51* (4), 1007–1062.

Laffont, Jean-Jacques and Jean Tirole, "A theory of incentives in procurement and regulation.," *MIT press.*, 1993.

Lalive, Rafael, Simon Luechinger, and Armin Schmutzler, "Does Supporting Passenger Railways Reduce Road Traffic Externalities?," *CEPR Discussion Paper No. DP9335*, 2013.

Landrigan, Philip J, Richard Fuller, Nereus J R Acosta, Olusoji Adeyi, Robert Arnold, Niladri (Nil) Basu, ..., and Ma Zhong, "The Lancet Commission on pollution and health," *The Lancet*, 2018, *391* (10119), 462 – 512.

Larrick, R. P. and J. B. Soll, "ECONOMICS: The MPG Illusion," *Science*, June 2008, *320* (5883), 1593–1594.

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

Lelieveld, J., J. S. Evans, M. Fnais, D. Giannadaki, and A. Pozzer, "The Contribution of Outdoor Air Pollution Sources to Premature Mortality on a Global Scale," *Nature*, September 2015, *525* (7569), 367–371.

LeRoy, Stephen F. and Jon Sonstelie, "Paradise Lost and Regained: Transportation Innovation, Income, and Residential Location," *Journal of Urban Economics*, January 1983, *13* (1), 67–89.

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

Li, Jing, "Compatibility and Investment in the U.S. Electric Vehicle Market," *Working Paper*, 2019, pp. MIT Sloan School of Management, Cambridge, MA.

Li, Shanjun, "Better Lucky Than Rich? Welfare Analysis of Automobile Licence Allocations in Beijing and Shanghai," *The Review of Economic Studies*, October 2018, *85* (4), 2389–2428.

_, Joshua Linn, and Elisheba Spiller, "Evaluating âĂIJCash-for-ClunkersâĂİ: Program effects on auto sales and the environment," *Journal of Environmental Economics and Management*, March 2013, *65* (2), 175–193.

__, **Lang Tong, Jianwei Xing, and Yiyi Zhou**, "The Market for Electric Vehicles: Indirect Network Effects and Policy Design," *Journal of the Association of Environmental and Resource Economists*, March 2017, *4* (1), 89–133.

__, **Yanyan Liu, Avralt-Od Purevjav, and Lin Yang**, "Does Subway Expansion Improve Air Quality?," *Journal of Environmental Economics and Management*, July 2019, *96*, 213–235.

Liu, Antung A., Joshua Linn, Ping Qin, and Jun Yang, "Vehicle Ownership Restrictions and Fertility in Beijing," *Journal of Development Economics*, November 2018, *135*, 85–96.

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

Mendelsohn, Robert, William D Nordhaus, and Daigee Shaw, "The Impact of Global Warming on Agriculture: A Ricardian Analysis," *The American Economic Review*, 1994, *84* (4), 753–771.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti, "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, August 2004, *112* (4), 725–753.

Mohring, Herbert, "Optimization and Scale Economies in Urban Bus Transportation," *American Economic Review*, 1972, *62* (4), 591–604.

Muehlegger, Erich and David Rapson, "Subsidizing Mass Adoption of Electric Vehicles: Quasi-Experimental Evidence from California," *NBER Working Paper*, December 2018, p. w25359.

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

Nordhaus, W. D., "Geography and Macroeconomics: New Data and New Findings," *Proceedings of the National Academy of Sciences*, March 2006, *103* (10), 3510–3517.

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

Olszewski, Piotr and Litian Xie, "Modelling the Effects of Road Pricing on Traffic in Singapore," *Transportation Research Part A: Policy and Practice*, August 2005, *39*, 755–772.

Parry, Ian W H and Kenneth A Small, "Does Britain or the United States Have the Right Gasoline Tax?," *The American Economic Review*, 2005, *95* (4), 31.

Parry, Ian W. H and Kenneth A Small, "Should Urban Transit Subsidies Be Reduced?," *American Economic Review*, May 2009, *99* (3), 700–724.

Parry, Ian W. H., Margaret Walls, and Winston Harrington, "Automobile Externalities and Policies," *Journal of Economic Literature*, 2007, *45* (2), 373–399.

Redding, Stephen J. and Matthew A. Turner, *Transportation Costs and the Spatial Organization of Economic Activity, Elsevier Handbook of Urban and Regional Economics*, Vol. 5, Elsevier, 2015.

Rivers, Nicholas, Soodeh Saberian, and Brandon Schaufele, "Public transit and air pollution: Evidence from Canadian transit strikes," *Canadian Journal of Economics*, 2020.

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

Sallee, James M., Sarah E. West, and Wei Fan, "Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations," *Journal of Public Economics*, March 2016, *135*, 61–73.

Salvo, Alberto, "Flexible Fuel Vehicles, Less Flexible Minded Consumers: Price Information Experiments at the Pump," *Journal of Environmental Economics and Management*, November 2018, *92*, 194–221.

— **and Cristian Huse**, "Build It, but Will They Come? Evidence from Consumer Choice Between Gasoline and Sugarcane Ethanol," *Journal of Environmental Economics and Management*, September 2013, *66* (2), 251–279.

Schlenker, Wolfram and W. Reed Walker, "Airports, Air Pollution, and Contemporaneous Health," *The Review of Economic Studies*, April 2016, *83* (2), 768–809.

— **, W Michael Hanemann, and Anthony C Fisher**, "Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach," *American Economic Review*, February 2005, *95* (1), 395–406.

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

Sieg, Holger, V Kerry Smith, H Spencer Banzhaf, and Randy Walsh, "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods," *International Economic Review*, 2004, *45* (4), 1047–1077.

Simeonova, Emilia, Janet Currie, Peter Nilsson, and Reed Walker, "Congestion Pricing, Air Pollution and ChildrenâĂŹs Health," *NBER Working Paper*, 2018, p. w24410.

Small, Kenneth A, "Valuation of travel time," *Economics of transportation*, 2012, *1* (1), 2–14.

Small, Kenneth A., Clifford Winston, and Jia Yan, "Uncovering the distribution of motorists' preferences for travel time and reliability," *Econometrica*, 2005, *73* (4), 1367–1382.

Springel, Katalin, "Network Externality and Subsidy Structure in Two-Sided Markets: Evidence from Electric Vehicle Incentives," *Working Paper*, 2019, pp. Georgetown University. Washington, D.C.

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

Tager, Ira, Kenneth Demerjian, Mark Frampton, and et al., "Traffic-related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects," Special Report 17, Health Effects Institute, Boston 2010.

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

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

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

Viard, V. Brian and Shihe Fu, "The effect of Beijing's driving restrictions on pollution and economic activity," *Journal of Public Economics*, 2015, *125*, 98–115.

Vickrey, William S., "Statement on the Pricing of Urban Street Use. Hearings: U.S. Congress, Joint Committee on Metropolitan Washington D.C. Problems," 1959, pp. 466–477.

_ , "Congestion Theory and Transport Investment," *The American Economic Review*, 1969, *59(2).*

Whitefoot, Kate S. and Steven J. Skerlos, "Design incentives to increase vehicle size created from the U.S. footprint-based fuel economy standards," *Energy Policy*, February 2012, *41*, 402–411.

Williams, Austin M. and Daniel J. Phaneuf, "The Morbidity Costs of Air Pollution: Evidence from Spending on Chronic Respiratory Conditions," *Environmental and Resource Economics*, October 2019, *74* (2), 571–603.

Winston, Clifford, "On the Performance of the U.S. Transportation System: Caution Ahead," *Journal of Economic Literature*, September 2013, *51* (3), 773–824.

_ **and Quentin Karpilow**, *Autonomous Vehicles: The Road to Economic Growth or Broken Dreams?* Book Manuscript 2019.

Wolff, H. and L. Perry, "Policy Monitor: Trends in Clean Air Legislation in Europe: Particulate Matter and Low Emission Zones," *Review of Environmental Economics and Policy*, June 2010, *4* (2), 293–308.

Wolff, Hendrik, "Keep Your Clunker in the Suburb: LowâĂŘEmission Zones and Adoption of Green Vehicles," *The Economic Journal*, August 2014, *124* (578), F481–F512.

Xiao, Junji and Heng Ju, "Market Equilibrium and the Environmental Effects of Tax Adjustments in China's Automobile Industry," *Review of Economics and Statistics*, May 2014, *96* (2), 306–317.

_ **, Xiaolan Zhou, and Wei-Min Hu**, "Welfare Analysis of the Vehicle Quota System in China," *International Economic Review*, May 2017, *58* (2), 617–650.

Xing, Jianwei, Leard Benjamin, and Shanjun Li, "What Does an Electric Vehicle Replace?," *NBER Working Paper*, 2019, p. w25771.

Xu, Yangfei, Qinghua Zhang, and Siqi Zheng, "The Rising Demand for Subway After Private Driving Restriction: Evidence from Beijing's Housing Market," *Regional Science and Urban Economics*, September 2015, *54*, 28–37.

Yang, Jun, Purevjav Avralt-Od, and Shanjun Li, "The Marginal Cost of Traffic Congestion and Road Pricing: Evidence from a Natural Experiment in Beijing," *American Economic Journal: Economic Policy*, 2020, *12* (1), 418–453.

_, Shuai Chen, Ping Qin, Fangwen Lu, and Antung A. Liu, "The Effect of Subway Expansions on Vehicle Congestion: Evidence from Beijing," *Journal of Environmental Economics and Management*, 2018.

Yang, Yongchong, Ying Liu, and Liang Li, "Urban Cycling Hot Spot Extraction Based on Sharing-bikes' Big Data," *Bulletin of Surveying and Mapping*, 2018, *8*, 68–73.

Zhang, R., J. Jing, J. Tao, S.-C. Hsu, G. Wang, J. Cao, C. S. L. Lee, L. Zhu, Z. Chen, Y. Zhao, and Z. Shen, "Chemical characterization and source apportionment of PM2.5 in Beijing: seasonal perspective," *Atmospheric Chemistry and Physics*, 2013, *13*, 7053–7074.

Zhang, Wei, C.-Y. Cynthia Lin Lawell, and Victoria I. Umanskaya, "The Effects of License Plate-Based Driving Restrictions on Air Quality: Theory and Empirical Evidence," *Journal of Environmental Economics and Management*, March 2017, *82*, 181–220.

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

Zheng, Siqi, Xiaonan Zhang, Weizeng Sun, and Jianghao Wang, "The Effect of a New Subway Line on Local Air Quality: A Case Study in Changsha," *Transportation Research Part D: Transport and Environment*, March 2019, *68*, 26–38.

Zhong, Nan, Jing Cao, and Yuzhu Wang, "Traffic Congestion, Ambient Air Pollution, and Health: Evidence from Driving Restrictions in Beijing," *Journal of the Association of Environmental and Resource Economists*, September 2017, *4* (3), 821–856.

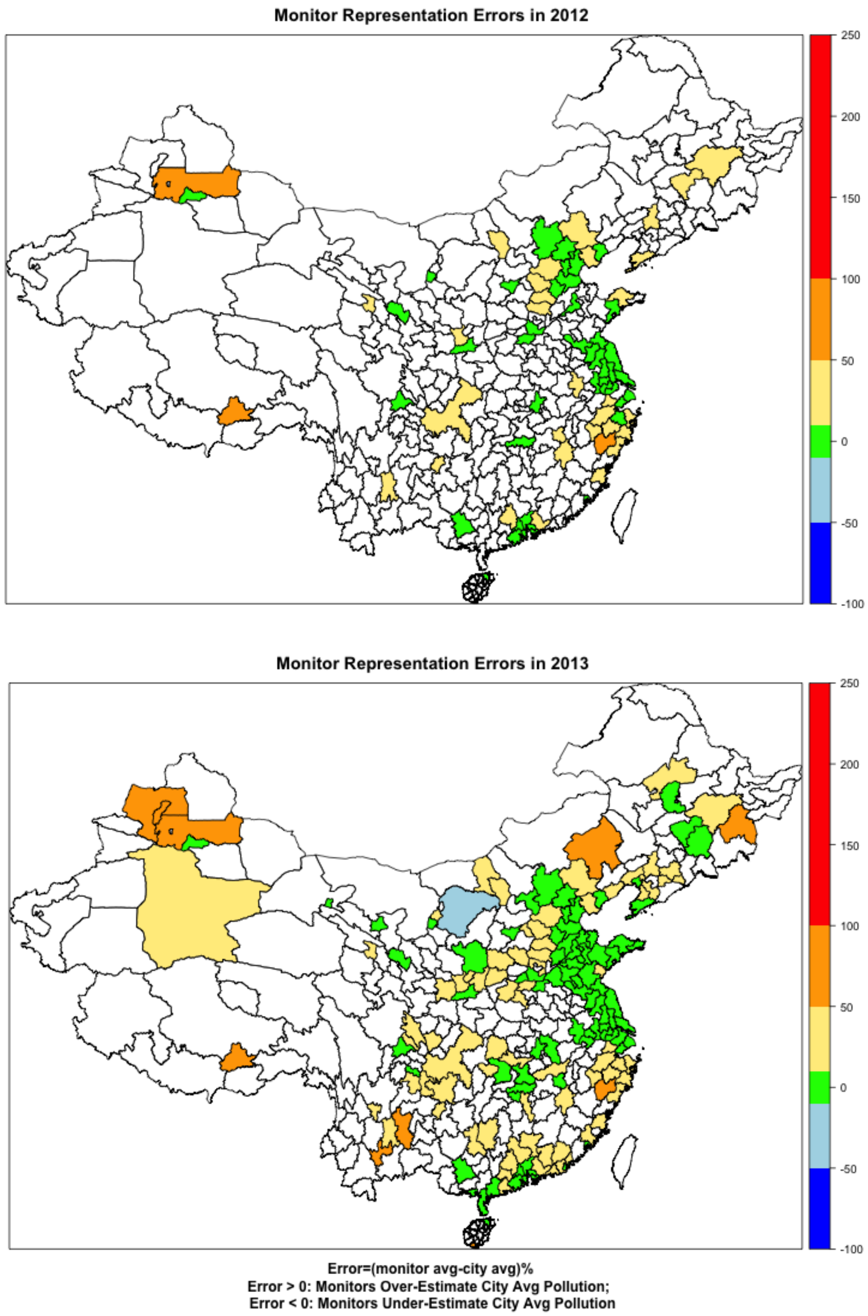
Zou, Eric, "Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality," *American Economic Review*, 2020.

APPENDIX A

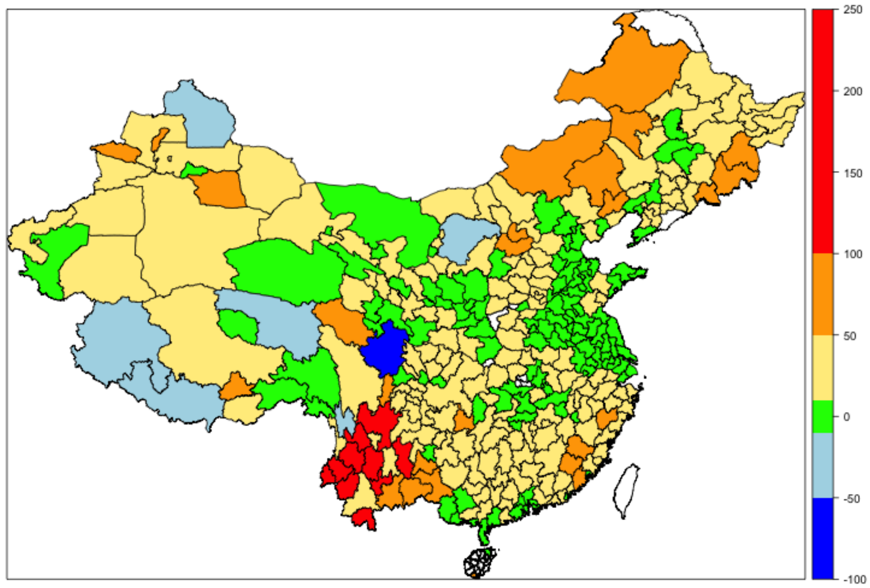
CHAPTER 3 OF APPENDIX

A.1 Appendix

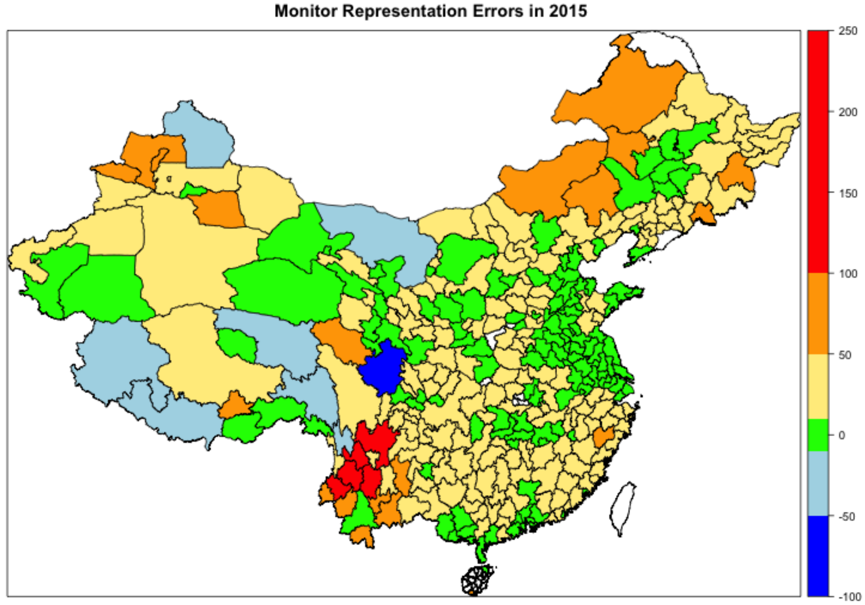
Figure A.1: Monitor Representation Errors by Year: all cells vs. monitored cells



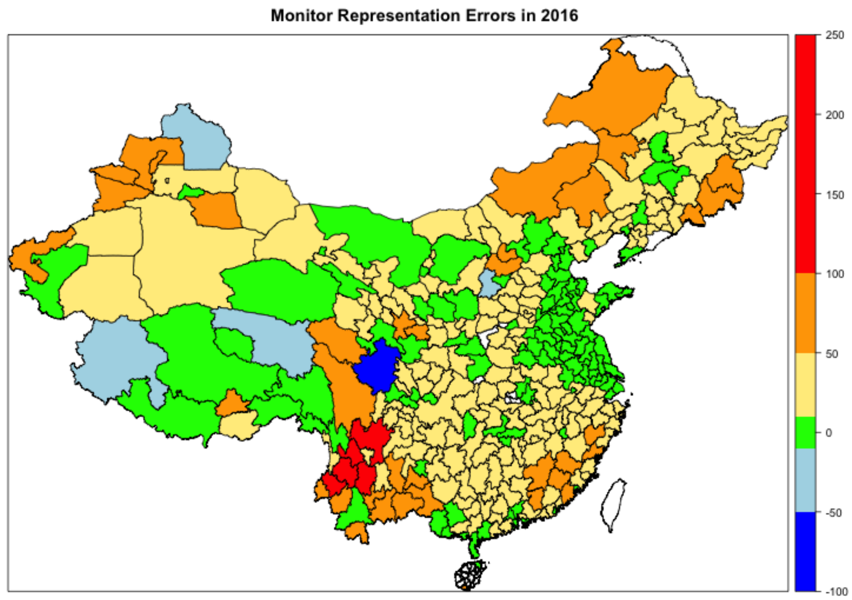
Monitor Representation Errors in 2014

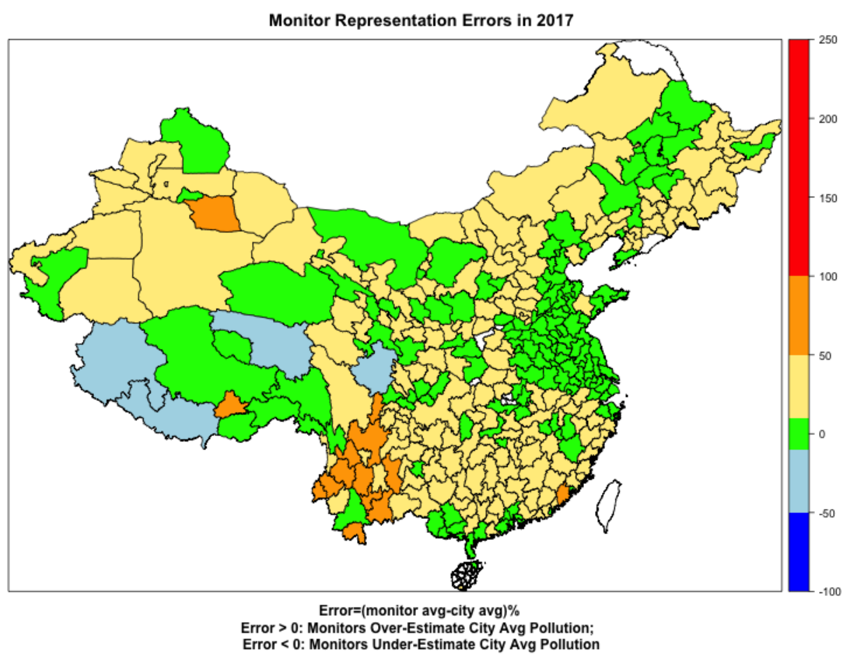


Monitor Representation Errors in 2015



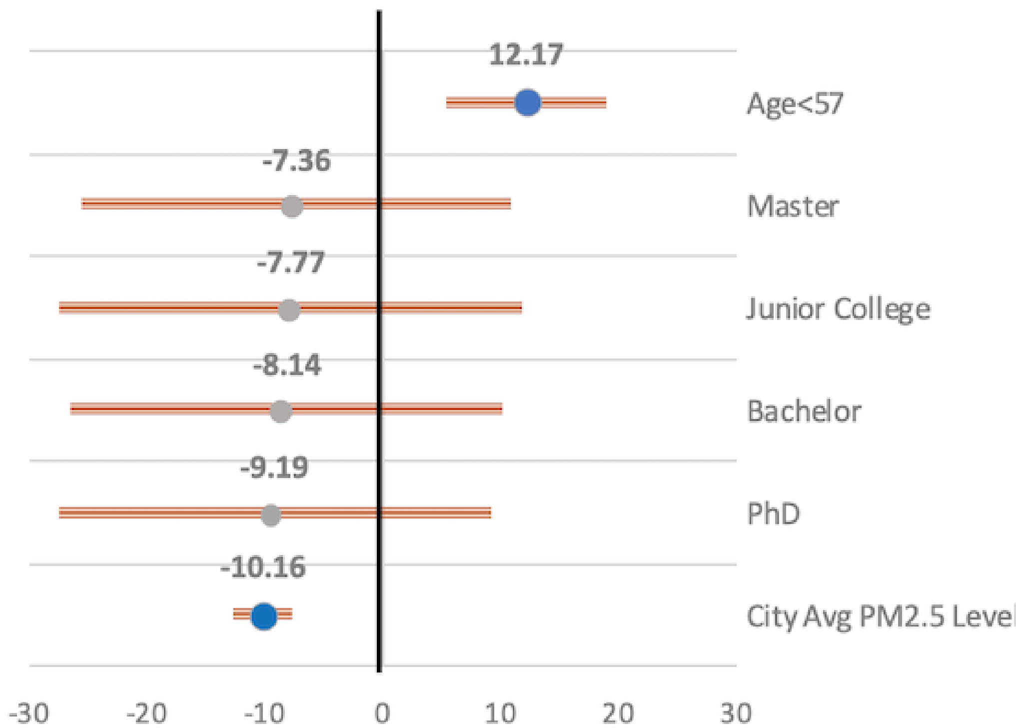
Error=(monitor avg-city avg)%
Error > 0: Monitors Over-Estimate City Avg Pollution;
Error < 0: Monitors Under-Estimate City Avg Pollution





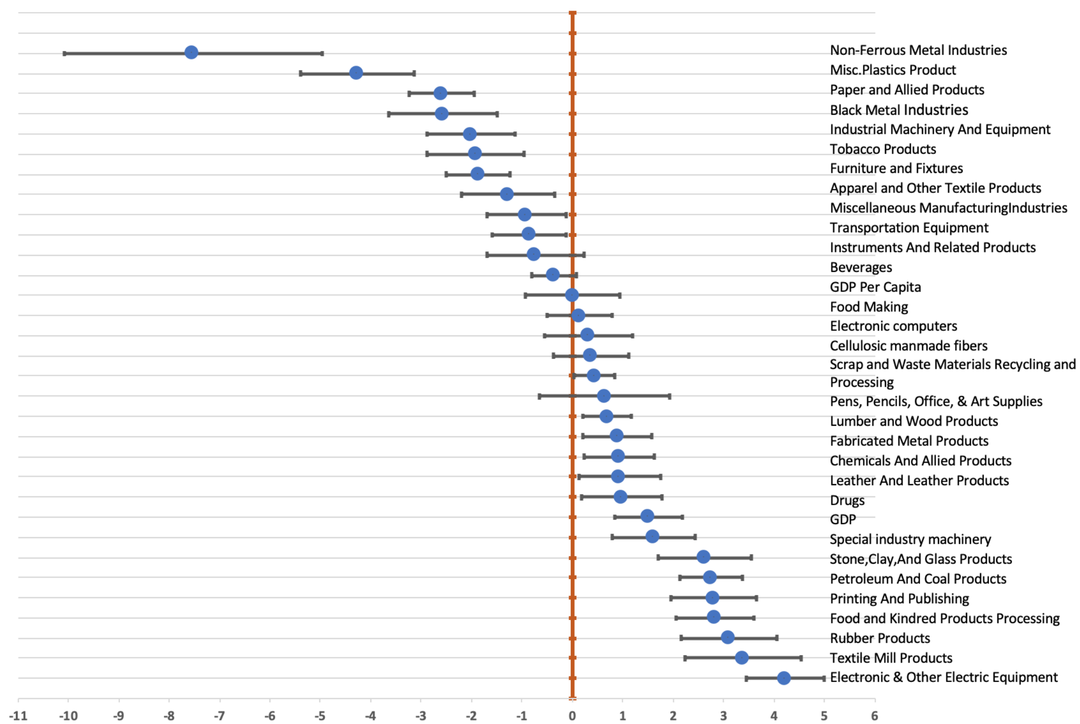
Note: The figures show the monitors representation errors from 2012 to 2017. The representation error is defined as the percentage difference between city average pollution level calculated based on only monitored cells and city average pollution based on all cells. All the pollution levels are weighted by the grid level population in 2015. Cities in green means the monitors well-represent city average PM2.5, with representation errors in [-10% , 10%]. Cities in warm colors (error > 10%) means the monitors over-represent city average pollution, and those in cool colors (error < -10%) means the monitors under-present city average pollution level. The map is based on raw data and presented at city level.

Figure A.2: Correlates of Monitor Representation Errors



Notes: This graph reports coefficient estimates with 95% Confidence intervals from a single panel regression of measurement errors on city characteristics. Year and Province Fixed Effects are included.

Figure A.3: Correlates of Monitor Representation Errors



Notes: This graph reports coefficient estimates with 95% Confidence intervals from a single panel regression of measurement errors on industrial concentrations (based on 2011 data). Year and Province Fixed Effects are included.