From Fog to Smog: the Value of Pollution Information*

Panle Jia Barwick Shanjun Li Liguo Lin Eric Zou

August 2019

Abstract

During 2013-2014, China launched a nation-wide real-time air quality monitoring and disclosure program, a watershed moment in the history of its environmental regulations. We present the first empirical analysis of this natural experiment by exploiting its staggered introduction across cities. The program has transformed the landscape of China's environmental protection, substantially expanded public access to pollution information, and dramatically increased households' awareness about pollution issues. These transformations in turn triggered a cascade of behavioral changes in household activities such as online searches, day-to-day shopping, and housing demand when pollution was elevated. As a result, air pollution's mortality cost was reduced by nearly 7% post the program, amounting to an annual benefit of RMB 120 billion. The resulting benefit is an order of magnitude larger than the cost of the program and the associated avoidance behavior. Our findings highlight considerable benefits from improving access to pollution information in developing countries, many of which are experiencing the world's worst air pollution but do not systematically collect or disseminate pollution information.

JEL Classification: D80, I10, Q53, Q58
Keywords: Pollution monitoring, Information disclosure, Smog, Avoidance

*Barwick: Department of Economics, Cornell University and NBER (email: panle.barwick@cornell.edu); Li: Dyson School of Applied Economics and Management, Cornell University, NBER and RFF (email: sl2448@cornell.edu); Lin: School of Economics, Shanghai University of Finance and Economics (email: lin.liguo@mail.shufe.edu.cn); Zou: Department of Economics, University of Oregon (email: ericzou@uoregon.edu). We thank Antonio Bento, Trudy Cameron, Lucas Davis, Todd Gerarden, Jiming Hao, Joshua Graff Zivin, Matt Khan, Grant McDermott, Ed Rubin, Ivan Rudik, Joe Shapiro, Jeff Shrader, Shuang Zhang for helpful comments. We thank Jing Wu and Ziye Zhang for generous help with data. Luming Chen, Deyu Rao, Binglin Wang, Tianli Xia, and Nahim bin Zahur provided outstanding research assistance.

1 Introduction

Economists have long emphasized the importance of information in decision making. In almost any decision environment, perfect information is necessary to ensure individually optimal choices and general market efficiency (e.g., Stigler, 1961; Hirshleifer, 1971; Grossman and Stiglitz, 1976). However, information as an input to decision making is often imperfect in real-world settings, especially for information with public good properties (such as forecasts on weather and pollution and disease prevention). The difficulties in appropriating private returns for this type of information call for government intervention. Understanding the value of providing such information is crucial for the optimal level of government investment in information gathering and reporting (Nelson and Winter, 1964; Craft, 1998).

There is little research on the value of providing pollution-related information in developing countries despite them experiencing the worst pollution in the world, largely because pollution information is either not collected or deliberately withheld by the government.[1] Consequently, questions like whether citizens can engage in effective pollution avoidance, what is the value of information, and how much public support is optimal remain largely unanswered. These issues are pressing since public funding for improving information infrastructure often competes with meeting basic needs in health care, nutrition, and education for the poor.

China provides a perfect setting for studying the role of pollution information. During the 2000s, its daily average concentration of fine particulate matter ($PM_{2.5}$) exceeded 50 ug/m^3, five times over the World Health Organization guideline. Despite the hazardous level of exposure, a comprehensive monitoring network was non-existent. Dissemination of the scant data that were collected was politically controlled and, in many cases, forbidden. In 2013, amid the social outcry on the lack of transparency and a dramatic shift in government position on air pollution, China launched a nation-wide real-time air-quality monitoring and disclosure program (henceforth, the information program), a watershed moment in the history of its environmental regulations. The emergence of the information program provides a unique opportunity to study changes in household behavior upon a sharp and permanent increase in the availability of pollution information. We present the first empirical analysis of this natural experiment by exploiting its staggered introduction across cities, using the most comprehensive database ever complied on social awareness, local air pollution levels, economic activities, and health outcomes that covers the period both before and after the information program.

[1] Among the 20 countries with the worst $PM_{2.5}$ level in 2018 (annual median $>$ 46 ug/m^3), only four (Nepal, Saudi Arabia, India, and China) installed a pollution monitoring system.

We first document that the information program has profoundly transformed the landscape of public access to pollution information and dramatically increased households' awareness about pollution issues. The frequency of air-pollution related articles in *People's Daily*, the government's official newspaper, rises from less than once-per-week to daily. The number of mobile phone applications ("apps") that stream air pollution data to users surges by 500%, four times faster than the growth of other apps. The term "smog" ("wu mai" in Chinese) became for the first time a buzzword in social media immediately after the program was launched. Purchases of air purifiers more than double one year after the information program is implemented in a city.

In turn, these changes in information access and public awareness have triggered a cascade of short-run and long-run behavioral changes in household activities such as day-to-day shopping and housing demand when pollution is elevated. In our short-run analysis, we exploit the universe of credit and debit card transactions in China from 2011 to 2015 to build a measure of outdoor purchase trips. Linking purchase activities to ambient air pollution, we show that the information program has boosted pollution avoidance by triggering a negative purchase-pollution elasticity of 3%. As expected, avoidance concentrates in plausibly "deferrable" consumption categories, such as supermarket shopping, outdoor dining, and entertainment, rather than in "scheduled" trips such as bill-pays, business-to-business wholesales, and cancer treatment sessions.

Our long-run analysis focuses on the housing market. Leveraging geo-location information from the near-universe of new home sales in Beijing during 2006-2014, we examine changes in the relationship between housing prices and local pollution levels induced by the information program using two different research designs. First, we employ the pixel-averaging technique ("oversampling") to enhance the original satellite data's spatial resolution from 10-by-10 km to 1-by-1 km (e.g., Fioletov et al., 2011; Streets et al., 2013). The high-resolution pollution measure allows us to conduct cross-sectional comparison within fine geographic units, such as community (geographically close to a Census block-group in the U.S.). We estimate a home value-pollution elasticity of -0.6 to -0.8 post disclosure. In contrast, the elasticity is small and statistically insignificant (-0.10 to 0.09) before the information program.

Second, we link China's emission inventory database with business registries to identify locations of major polluters in Beijing: the 10% of facilities that account for 90% of total industrial air emissions. These big polluters tend to be visible and well-known landmarks in the city. This allows us to estimate separate "distance gradient" curves (e.g., Currie et al., 2015) that express the home value as a function of proximity to the nearest major polluter before and after the information program. While there is no correlation between housing prices and proximity to polluters prior to the program, houses within 3 km of a major polluter depreciate 27% afterward, which corresponds to 42% of the inter-quartile range of the housing price dispersion. Thus, the information program facilitates the capitalization of air quality in the housing market, potentially improving residential sorting and social welfare.

These behavior changes could greatly mitigate the devastating consequences of China's elevated pollution. Our last set of empirical analyses examines changes in the mortality-pollution relationship as access to information improves. Using nationally representative mortality data from the Chinese Center for Disease Control and Prevention (CDC), we find a 5 percentage-point reduction in the mortality-pollution elasticity (especially for cardio-respiratory causes) post monitoring. Assuming a linear dose-response function and combining our findings with existing estimates on the causal effect of pollution on mortality in China (e.g., Ebenstein et al., 2017), access to pollution information has reduced premature deaths attributable to air pollution exposure by nearly 7%. It generates a health benefit that is equivalent to a 10 ug/m^3 reduction in PM_{10}, with an associated social Willingness-to-Pay in the order of RMB 120 billion annually based on recent estimates in the literature (Ito and Zhang, 2018). By our calculation, such social benefits outweigh the costs of defensive investments (such as air purifier purchases) and administrative costs of deploying and maintaining the program by at least an order of magnitude, making the information program one of the most successful environmental policies in a developing country.

We make three main contributions to the literature. First, our study provides to our knowledge the first empirical estimate of the value of a nation-wide program on pollution monitoring and disclosure.[2] Our empirical findings highlight the considerable benefits in collecting and disseminating pollution information in developing countries, many of which are experiencing the worst mortality damage from pollution exposure in the world (Landrigan et al., 2018). The success of China's program provides a benchmark for policy discussions (e.g., the cost-benefit analysis) on building information infrastructure in these countries.

Second, our study shows that information is a key determinant of avoidance behavior and defensive spending. Consumer activities (online searches, day-to-day shopping, and housing demand) exhibit little response to pollution until such information becomes widely available. This contrasts with the implicit assumption of perfect information on pollution exposure in the existing literature that uses revealed-preference to estimate the value of non-marketed environmental goods. To the extent that access to information is lacking in developing countries, this assumption underestimates consumers' true willingness-to-pay for

[2] A similar literature quantifies the value of weather forecasts, another type of government-provided information, as an important input to production decisions (Lave, 1963; Craft, 1998; Shrader, 2018; Jagnani et al., 2018).

environmental goods. Our findings provide a potential explanation for why environmental quality is severely undervalued in developing countries (Greenstone and Jack, 2015) and why the dose relationship between pollution and mortality can be different across developed and developing countries (Arceo, Hanna and Oliva, 2015).

Third, this study contributes to the broad empirical literature on the role of information in consumer choices. Growing evidence suggests that consumers misperceive product attributes in a wide range of contexts such as food nutritional contents (Bollinger, Leslie and Sorensen, 2011), insurance policy costs (Kling et al., 2012), vehicle fuel economy (Allcott, 2013), retirement savings (Bernheim, Fradkin and Popov, 2015), taxation (Chetty, Looney and Kroft, 2009), and energy prices (Shin, 1985; Ito, 2014). Information provision programs can improve consumers' perception of product attributes (Smith and Johnson, 1988; Oberholzer-Gee and Mitsunari, 2006), change consumer choices (Hastings and Weinstein, 2008; Dranove and Jin, 2010; Jessoe and Rapson, 2014; Wichman, 2017), and drive up average product quality (Jin and Leslie, 2003; Bai, 2018). In the context of air quality, recent studies have documented behavioral responses to pollution exposure in both the short- and long-terms. Our analysis shows that these behavioral responses could lead to improved health conditions and we use the associated benefits in dollar terms to provide a lower bound estimate of the value of pollution information.[3]

The rest of this paper is organized as follows. Section 2 reviews institutional details on the information program and describes data sources. Section 3 presents the theoretical framework. Section 4 documents the dramatic changes in information access and awareness after the program. Section 5 employs a unified framework to examine the effect of the program on short- and long-term avoidance behavior and mortality. Section 6 calculates the value of information. Section 7 concludes.

2 Institutional Background and Data

2.1 Environmental Regulations

The real-time $PM_{2.5}$ monitoring and disclosure program started in 2013 is a watershed moment in the history of China's environmental regulations. The program brought about a sharp and sudden change in the access of pollution information for the average residents and

[3] Cutter and Neidell (2009); Graff Zivin and Neidell (2009); Sun, Kahn and Zheng (2017); Zhang and Mu (2018) document changes in short-run avoidance and defensive spending while Chay and Greenstone (2005); Banzhaf and Walsh (2008); Bayer, Keohane and Timmins (2009); Mastromonaco (2015); Chen, Oliva and Zhang (2017); Freeman et al. (2019) show housing and migration decisions in the long-term in response to pollution information.

drastically enhanced the public awareness of the health impact of $PM_{2.5}$. To help understand this change, we provide a brief history of China's environmental regulations.

Environmental Regulations Prior to 2012 China established its first national ambient air quality standards (NAAQS) in 1982 which set limits for six air pollutants including Total Suspended Particulate (TSP), coarse particulate matter (PM_{10}), sulfur dioxide (SO_2), nitrogen oxides (NO_x), carbon monoxide (CO), and ozone (O_3). The standards were subsequently amended in 1996, 2000, and 2012. The 1996 amendment strengthened and expanded the standards to reflect the improvement in abatement capabilities while the 2000 amendment removed NO_x from the list and relaxed the standards for NO_2 and O_3 in response to non-compliance due to the increase in automobile usage.

Throughout much of the 1980's to early 2000's, the major threat of air quality was considered to be SO_2 due to coal burning. As acid rain caused widespread and visible damages to crops, forest, and the aquatic environment, the control of acid rain and SO_2 emissions was the focus of the environmental regulations (Yi, Hao and Tang, 2007). The prominent regulation is the two-control zone policy (TCZ) implemented from 1998 where prefectures with high PH values of precipitation or SO_2 concentration were designated as either the acid rain control zone (located in the south) or the SO_2 control zone (mostly in the north). A series of measures were imposed in these zones such as mandating the installation of flue gas desulfurization in coal-fired power plants and closing down the small coal-fired power plants (Tanaka, 2015). As a result of aggressive emissions control and clean energy policies, the average SO_2 concentration was reduced by nearly 45% from 1990 to 2002, with the majority of the cities achieving the national standard by 1998 (Hao and Wang, 2005).[4]

Starting from the early 2000, the source of air pollution shifted from coal burning to mixed sources, and particulate matter (PM) rather than SO_2 became the major pollutant. This shift was driven by the fact that while the emissions from coal-fired power plants have reduced significantly, the emissions from automobiles, industrial facilities, and construction have increased due to the dramatic growth in vehicle ownership, industrial activities (after China's WTO accession in 2001) and rapid urbanization. The regulatory focus was shifted to reducing urban air pollution through city-level efforts (Ghanem and Zhang, 2014), which proved to be ineffective due to the strong competing incentives for economic growth at the local level together with the weak monitoring and enforcement from the central government. Episodes of extreme air pollution were common especially during winters in many urban centers. U.S. Embassy in Beijing and consulates in Guangzhou and Shanghai started to

[4] The fraction of the acid rain zone in China's total terrain has decreased from the peak level of about 30% to 8.8% in 2015.

report hourly $PM_{2.5}$ in 2008 based on monitoring stations installed on-site. The $PM_{2.5}$ readings from these sites were often inconsistent with the official pollution reports and became sources of diplomatic tensions.[5]

Limited Pollution Awareness Prior to 2013 While air pollution has been a long-standing issue, public access to the daily pollution exposure was almost absent prior to 2013. Although the Ministry of Environmental Protection (MEP) publishes the daily Air Pollution Index (API) data for major cities starting from 2000, the reported API prior to the information program only partially reflected true air quality because it did not incorporate $PM_{2.5}$, which was the major air pollutant in many Chinese cities since the 2000s.[6] In addition, the API index was not incorporated into the mass media publications or broadcasts. Finally, the API data was gathered and reported by local environmental bureaus whose leaders were appointed by the local governments. The MEP did not control the monitoring stations and had limited ability to monitor the data quality. Recent research has found evidence of widespread manipulation of the API data (Andrews, 2008; Chen et al., 2012; Ghanem and Zhang, 2014; Greenstone et al., 2019).

While the dominant pollutant had shifted from SO_2 to particulate matter in the 2000's, there was no systematic collection of $PM_{2.5}$ data. As a result, consumer awareness of $PM_{2.5}$ was extremely limited prior to 2013. Poor visibility due to high levels of $PM_{2.5}$ was often characterized as *fog* rather than *smog* by both government agencies and the media. For example, newspaper headlines as well as China Meteorological Administration characterized flight delays and cancellations as being caused by widespread and dense fog in Beijing and Northern China in November 27, 2011. In fact, this was a major pollution event as shown in Figure 1 that displays the NASA satellite view of China and the high AOD measure from the NASA MODIS algorithm. A similar pollution event occurred in December 4-6, in 2011 when it was again covered as dense fog by major news media including China Central Television, the predominant state television broadcaster in Mainland China, and popular website such as sina.com.

The lack of awareness of $PM_{2.5}$ and fog-smog confusion among the public and the media were reflected upon by the prominent journalist-turned-environmentalist Chai Jing in her

[5] The then-vice minister of the Ministry of Environmental Protection (MEP), Wu, Xiaoqing, openly requested U.S. embassy and consulates to stop releasing $PM_{2.5}$ data from their monitoring stations during the press conference on the World Environment Day in 2012. He stated that the public release of air-quality data by the consulates "not only doesn't abide by the spirits of the Vienna Convention on Diplomatic Relations and Vienna Convention on Consular Relations, but also violates relevant provisions of environmental protection." (New York Times, June 5, 2012).

[6] API coverts the concentration of PM_{10}, SO_2, and NO_2 into a single index through a set of piece-wise linear transformations. The dominant pollutant on each day determines the level of API.

high-profile documentary on China's air pollution titled *Under the Dome* released in February 2015: "... I go back and check the headline from that day's newspaper (on December 1st, 2004): 'Fog at Beijing Capital Airport Causes Worst Flight Delays in Recent Years'. We all believed that was fog back then. That's what we called it.... as a former journalist, I started to blame myself because for all those years I had been reporting stories on pollution all across the country, I always thought pollution was about mining sites and those places near factories spewing smoke plumes. I never thought the skies that we saw every day could be polluted."[7]

The Information Program and Environmental Regulation Post 2012 In 2012, the MEP revised the NAAQS and for the first time in China's history set the national standards for $PM_{2.5}$. The new standards were slated to take effect nationwide in 2016 but some cities and regions were required to implement the standards earlier.[8] To help achieve the standards, China's State Council released the Action Plan on Air Pollution Prevention and Control in September 2013, which set specific targets for $PM_{2.5}$ reduction from 2013 to 2017 and outlined ten concrete policies such as promoting the role of market-based mechanisms and establishing monitoring and warning systems to cope with severe air pollution events.[9] In addition to this action plan, for the first time in the history of national five-year plans, the 13th Five-year Plan required prefecture-level cities or higher to reduce the $PM_{2.5}$ concentration by 18% from 2015 to 2020.

The recognition of $PM_{2.5}$ as a major pollutant and the aggressive policies to reduce $PM_{2.5}$ concentration marked an important shift of the China's long-standing strategy of prioritizing economic growth over environmental concerns and happened under the backdrop of China's 12th and 13th Five-Year Plans that set pollution reduction as one of the bureaucratic hard targets in the cadre evaluation system (Wang, 2017).[10] To effectively monitor local air pollution levels and to address the pitfalls of the previous reporting system of API (Greenstone

[7]The documentary has been compared with Al Gore's *An Inconvenient Truth* in terms of its style and impact. The film was viewed over 150 million times on popular website tencent.com within three days of its release, and viewed a further 150 million times by the time it was taken offline by the government four days later.

[8]The cities in Beijing-Tianjin-Hebei region, the Yangtze River Delta, and the Pearl River Delta as well as provincial capitals are required to implement the standards in 2012 while all prefecture-level cities are required to implement the standards by 2015.

[9]The plan may have been China's most influential and successful environmental policy during the past decade. Under this plan, $PM_{2.5}$ reduced by over 37% in Beijing-Tianjin-Hebei Region, 35% in the Yangtze River Delta, 26% in the Pearl River Delta, and over 30% on average in over 70 major cities (Huang et al., 2018).

[10]The mandate to reduce air pollution comes from the highest level of government officials. Premier Li, Keqiang described smog as "nature's red-light warning against inefficient and blind development" and declared war against pollution at the opening of the annual meeting of People's congress in March 2014. The phrase, *war on pollution*, has been quoted by President Xi, Jinping in national meetings since then.

et al., 2019), the MEP implemented a nationwide monitoring and disclosure program starting from 2013 with the focus of building a scientific and efficient system to monitor air quality and disclose publicly the real-time data.

The program contained two major provisions. First, it initiated continuous monitoring of major air pollutants, including $PM_{2.5}$, PM_{10}, O_3, CO, NO_2, and SO_2. This led to the installation of a comprehensive network of monitors which were built in three waves. In the first wave, monitoring networks were built between May and December 2012 in 74 major cities that represented the country's key population and economic centers (the Beijing-Tianjin-Hebei Metropolitan Region, the Yangtze River Delta, the Pearl River Delta, Direct-administered municipalities, and provincial capitals). Real-time readings on all major air pollutants were posted online and ready for streaming by December 31, 2012. By October 31, 2013, the second wave was completed, adding 116 cities from the list of the Environmental Improvement Priority Cities, and the National Environmental Protection Exemplary Cities.[11] In the final wave, achieved by November 20, 2014, the program reached the remaining 177 cities. The roll-out of the program is plotted in Figure 2. By the end of the third wave, the program had built more than 1,400 monitoring stations in 337 cities covering an estimated 98% of the country's population. A key feature of the monitoring roll-out is that it is based on cities' administrative hierarchy and well-known groupings that were designated long before the information program was initiated. The pre-determined nature of these groupings indicates that there is little scope of selecting cities into different roll-out waves based on unobservable characteristics or future projections of outcome variables.[12]

Second, the information program established a dissemination system to report a real-time Air Quality Index (AQI) that is on a single scale of 0-500. Monitoring results are displayed in real-time on MEP's website. Different from the old-generation monitoring stations used by the local environmental bureaus to report API, the new monitors are under the direct control of MEP. Data that are collected are directly transmitted to MEP's information center in real-time to avoid manipulation by the local government. Both hourly and daily AQIs, as

[11] The Environmental Improvement Priority Cities were designated in 2007 during the "Eleventh-Five-Year" period. The list has 77 cities including the four direct-administered municipalities, provincial capitals, cities in the Beijing-Tianjin-Hebei Metropolitan Region, the Yangtze River Delta, the Pearl River Delta, as well as other cities that are important regional economic centers and/or face significant environmental challenges. The National Environmental Protection Exemplary Cities program was created during the "Ninth-Five-Year" period; 68 cites were awarded the title from 1997 to 2007 based on a host of environmental quality criteria. Appendix Figure D.1 tabulates cities by waves and their associated city clusters.

[12] Appendix Table D.1 tabulates economics attributes for cities in each wave. Cities in earlier-waves tend to have a larger population, higher GDP per capita, higher levels of air pollution and industrial emissions, etc. On the other hand, as shown in Appendix Table D.2, these economic and environmental conditions do not change systematically after the program roll-out. Together, these evidence suggest that the choice of cities included in each wave is based on permanent differences in city characteristics, rather than based on city-level unobservables that correlate with the timing of the program roll-out.

well as concentrations of $PM_{2.5}$, PM_{10}, O_3, CO, NO_2, and SO_2, are available at individual station- and city-levels, with an interactive map showing the location of each monitoring station. Appendix Figure D.2 provides a screen-shot of the website interface. Importantly, the government allows private parties to access and stream data directly from the web. This functionality has spurred a surge in private websites and mobile phone applications that report real-time air quality information. We provide more details on how the public awareness on $PM_{2.5}$ and smog has been affected by the program in section 4.

2.2 Data

This section documents the primary data sources for our empirical analysis.

Mass Media, Phone Apps, and Web Search We draw on several digital sources to illustrate the evolution of public access to pollution information. First, we look at the publication trends by *People's Daily*, the government's official newspaper, and pull articles whose title or content contains the word "smog" from *People's Daily*'s digital archive. For each article that mentions "smog", we also identify a list of relevant cities.

Second, we scrape Apple's App Store to obtain Chinese mobile apps that contain air pollution information, using keywords including "air pollution", "atmospheric pollution", and "smog".[13] These apps typically display current hourly pollution levels; some also provide health related guidelines (e.g., avoid outdoor activities) when pollution levels are high. Appendix Figure D.3 is a screenshot from a typical pollution app. We also obtain apps in other major categories such as gaming, reading, and shopping and use them as a control group.

Third, the most widely used search engine in China, Baidu, publishes a search index that summarizes the number of queries for certain words in a city and day among both desktop and mobile users since 2011. We focus on the search index for "smog", the buzzword for air pollution. The search index is generated using an algorithm similar to Google Trends that is based on the underlying search traffic and reflects search intensity.

The prevalent usage of Internet and mobile phones among the Chinese population provides a unique opportunity to investigate pollution awareness using digital sources. Data from the China Internet Network Information Center (CINIC) show that, by the end of 2012, China had about 0.56 billion (or about 40% of population) Internet users, more than 80% of whom were active search engine users.[14] Mobile phone prevalence rose from 73.5 per

[13] The API returns the 200 most relevant apps for a given keyword.

[14] A CINIC 2013 survey of more than 2,800 Chinese phone respondents shows that more than 99% of Internet users have heard of Baidu, the most popular search engine (seconded by Google, 87%), and 98% have used it in the past six months (seconded by 360.cn, 43%).

100 people in 2011 to 95.6 per 100 people in 2016 (National Bureau of Statistics), with a smart-phone penetration rate of 72% in 2013 (Nielsen).

In addition to data on mass media coverage, mobile apps, and internet queries on air pollution, we have compiled a rich set of data on air purifier sales, bank card transactions, housing transactions, mortality rates, as well as the location of major polluters in Beijing and satellite measures of air pollution concentration, which we describe below.

Air Purifier Sales Air purifier sales data come from a leading market research firm and report the total units of air purifiers sold for both residential and institutional purposes at the monthly frequency for fifty cities from 2012-2015.[15] Among these fifty cities, thirty-four, eleven, and five are in the first, second, and third waves of program roll-out, respectively.

Bank-Card Transactions Data Households' shopping trips are constructed using the universe of debit and credit card transactions during 2011-2015 from UnionPay, the only inter-bank payment clearinghouse in China and the largest such network in the world. The database covers 59% of China's national consumption and 22% of its GDP in 2015 and is the most comprehensive data with fine spatial and temporal resolution on consumption activities for China (Appendix Figures D.4, D.5, D.6 provide summary statistics). For each transaction, we observe the merchant name and location, transaction amount and time, currency, etc.

Our key outcome variable is *purchase rate*, defined as the ratio between (1)the total number of transactions occurred in a city$\times$week, and (2) the total number of active cards with positive transactions in a city$\times$year. We focus on all transactions of a 1% random sample of cards, with an average of 18.3 million active cards at any given point in time.

Two features of the data are worth mentioning. First, our data contains a small fraction (about 3%) of transactions that are made online. We drop online transactions as it is difficult to trace these buyers' physical locations. Second, we do not observe items purchased in each transaction. Fortunately, UnionPay classifies merchants into 300 plus categories, such as department stores, supermarkets, etc. We use merchant category information in our analysis below.

Housing Transactions Data (Beijing) Our housing data contains a total of over 660,000 new home transactions in about 1,300 apartment complexes in Beijing from January 2006 to April 2014, with a near-universe coverage. Variables recorded include the transaction date and price, housing unit characteristics (floor, size of the unit, etc.), as well as attributes

[15]The firm name is withheld per our data use agreement.

and geo-location of the apartment complex. Beijing's housing market is fluid. Over 84% of transactions occur in complexes that are on the market for less than a year. Among the 1,300 apartment complexes, 64% are sold out in three years.

Polluter Data (Beijing) MEP conducts an annual survey of all major industrial polluters and compiles the Chinese Environmental Statistics (CES) database, the most comprehensive coverage of emissions in China and the source of the annual Environmental Yearbook (Liu, Shadbegian and Zhang, 2017; Zhang, Chen and Guo, 2018). We have access to the 2007 CES, which reports total industrial emissions across all pollutants for 587 polluters in Beijing. We obtain firm address and operation status by linking CES with firm registration records from Qixin (www.qixin.com) and geocode addresses using Baidu's Map API. Our study focuses on 407 polluters that operated throughout 2006-2014.

Mortality Data Chinese Center for Disease Control and Prevention (CDC) operates a Disease Surveillance Points (DSP) system that covers 161 counties and city districts and 73 million individuals during 2011-2016, a 5% representative sample of China's population.[16] DSP's mortality database, drawn from hospital records and surveys of the deceased's household, is one of the highest-quality health databases that have been used in recent medical and economic research (Zhou et al., 2016; Ebenstein et al., 2017). We observe the number of persons and total deaths by each county × week × gender × age-group and separately for the following six categories: chronic obstructive pulmonary diseases, heart diseases, cerebrovascular diseases, respiratory infections, digestive diseases, and traffic accidents. The first four groups are closely related to cardiovascular diseases, which are affected by air pollution exposure, while the latter two causes serve as placebo-style outcomes. To use the same geographic unit of analysis throughout the paper, we aggregate the county mortality data to the city level for a total of 131 cities. Among these 131 cities, 38 implemented the information program in the first wave, 38 in the second wave, and 55 in the last wave.

Satellite Data To overcome the challenge that reliable pollution data are only available after the information program, we obtain ambient air quality measures from Aerosol Optical Depth (AOD) via NASA's MODIS algorithm installed on satellite Terra's platform. The original data has a geographic resolution of 10 km × 10 km and a scanning frequency of 30 minutes, which we average to the city × day level from 2006-2015. MODIS records the degree to which sunlight is scattered or absorbed in the entire atmospheric column corresponding to the overpassed area under cloud-clear condition. As such, AOD captures concentration

[16] In China, counties are comparable to city districts and are smaller geo-units than cities.

of particle pollution such as sulfates, nitrates, black carbons, and sea salts, and serve as a proxy for outdoor particulate matter pollution (Van Donkelaar, Martin and Park, 2006). Appendix Figure D.7 documents a strong correspondence between AOD and PM2.5 post the information program.

We favor the MODIS AOD measure over alternatives (such as satellite-based ground-level $PM_{2.5}$ predictions) for several reasons. First, MODIS data can be easily aggregated from daily to weekly or quarterly levels. This allows us to use the same pollution measure throughout our analysis. In contrast, processed satellite-based $PM_{2.5}$ data are only distributed at certain temporal intervals (e.g., annual) and cannot be dis-aggregated in a straightforward manner. Second, MODIS AOD allows us to observe overlapping 10 km $\times$ 10 km grid cells, which is essential for the oversampling exercise in Section 5.3. Finally, while MODIS AOD is a common input in predicted ground-level $PM_{2.5}$, there is no consensus on the precise relationship between AOD and PM2.5 in the atmospheric science literature.

3 Theoretical Model

Classical economic theory argues that the value of information stems from the fact that information as an input to the decision process can help economic agents make better decisions, for example by resolving market uncertainty in demand and supply conditions (Stigler, 1961, 1962) or technological uncertainty in investment and production decisions (Lave, 1963; Hirshleifer, 1971). Pollution information affects the behavior of informed individuals who could take measures to reduce the harm from pollution. In this section, we present a stylized model to illustrate how the information program affects individual behavior and utility by incorporating the elements of information economics (Hirshleifer, 1971; Hilton, 1981) into the classical model of health demand and production (Grossman, 1972).

recognize the negative health impact of pollution

3.1 Model Setup

Individuals derive utility from the consumption of a numeraire good x, whose price is normalized to one, and health stock h: $U(x, h)$. Health stock depends on both the pollution level c and the extent of avoidance a (individuals' actions that mitigate the negative impact of pollution): $h = h(c, a)$.

Individuals face a budget constraint that is given by: $I + w * g(h) \geq x + p_a * a$, where I is non-labor income and w is the wage rate. Hours worked is denoted by $g(h)$ and is a function of the health stock. Individuals allocate their wage and non-wage income between consumption and engaging in the adaptation or avoidance behavior a, where p_a is the associated price (e.g., the cost of an air purifier or medication). We assume away dynamics and savings in this model to ease exposition. In addition, we use a to include broadly-defined (costly) adaptation behavior.[17]

Under imperfect information on pollution, consumers may or may not know the real pollution level c. They maximize utility by choosing the optimal consumption x and defensive investment a based on the *perceived* pollution level c_0:

$$\begin{aligned} &\max_{x,a} U(x, h) \\ \text{s.t. } & I + w * g(h) \geq x + p_a * a \\ & h = h(c_0, a) \end{aligned}$$

The health function $h = h(c_0, a)$ in the optimization can be viewed as an ex-ante health function which consumers relies on for decisions. It is different from the ex-post health outcome $h = h(c, a)$ experienced by consumers. This difference gives rise to the discrepancy between the (ex-ante) decision utility and the (ex-post) experience utility as described in (Bernheim and Rangel, 2009; Allcott, 2013).

Let avoidance under the perceived pollution c_0 be denoted by $a(c_0)$. Individuals' wage income is determined by the actual pollution level c and avoidance $a(c_0)$: $w * g[h(c, a(c_0))]$. The consumption of the numeraire good is denoted by:

$$x(c, c_0) = I + w * g[h(c, a(c_0))] - p_a * a(c_0)$$

which makes it explicit that x depends on the actual pollution c and perceived pollution c_0.[18] Individuals' (experience) utility based on the perceived pollution prior to the information program is:

$$U[X(c, c_0), h(c, a(c_0))] \equiv V(c, c_0)$$

where $V(\cdot, \cdot)$ denotes the indirect utility, where the first argument is the actual pollution c and the second argument is the perceived pollution level. To examine the behavioral changes and the welfare impacts of the information program, we make the following assumptions:

[17]Examples include reducing outdoor activities (Zivin and Neidell, 2009; Saberian, Heyes and Rivers, 2017), engaging in defensive spending (e.g., face masks and air purifiers) (Ito and Zhang, 2018; Zhang and Mu, 2018), and location choices and migration (Chay and Greenstone, 2005; Banzhaf and Walsh, 2008; Bayer, Keohane and Timmins, 2009; Chen, Oliva and Zhang, 2017).

[18]Individuals' actual income is determined by the ex-post health stock: $I + w * g[h(c, a(c_0))]$. Consumption of the numeraire good is a residual of the budget constraint, after subtracting the cost of avoidance.

Assumption (A1) Health stock is bounded and decreases in pollution and increases in avoidance: $\frac{\partial h}{\partial c} \leq 0$, and $\frac{\partial h}{\partial a} \geq 0$. In addition, the marginal health benefit of avoidance is decreasing: $\frac{\partial^2 h}{\partial a^2} \leq 0$. This assumption ensures that people don't engage in infinite amount of avoidance behavior. Finally, the worse the pollution, the larger the marginal health benefit of avoidance: $\frac{\partial^2 h}{\partial a \partial c} \geq 0$. The health benefit of avoidance is likely much higher when pollution is severe than when it is modest. Similarly, we assume that hours worked increases in health, but at a decreasing rate: $\frac{dg}{dh} \geq 0, \frac{d^2 g}{dh^2} \leq 0$.

We focus on interior solutions for the optimal level of avoidance behavior a. A necessary condition for an interior solution is $w * \frac{dg}{dh} * \frac{\partial h}{\partial a} |_{a=0} > p_a$. At low levels of avoidance, the marginal health benefit $\frac{\partial h}{\partial a}$ is likely to be large. In addition, there are many choices of different defensive mechanisms, some of which have low costs. For example, avoiding outdoor activities at times of high $PM_{2.5}$, wearing facial masks, or purchasing air purifiers are all cheap and effective defensive mechanisms. When avoidance is cheap, individuals will engage in an appropriate amount of avoidance to increase wage income (via improved health stocks), relax the budget constraint, and enjoy a higher consumption of the numeraire good.

Assumption (A2) Utility is quasi-linear $U(x, h) = x + u(h)$ and increases in health at a decreasing rate: $\frac{\partial U}{\partial h} \geq 0, \frac{\partial^2 U}{\partial h^2} \leq 0$. Quasi-linear utility functions are commonly used in the literature and help to simplify the exposition.

Assumption (A3) Let c_0 denote individuals' perception of air pollution prior to the information program. We assume that $c_0 < c$, that is, the perceived level of pollution is lower than the actual level.[19] In addition, pollution concentration c is perfectly observed post the program.

Proposition 1. *Under assumptions (A1)-(A3), the information program is predicted to result in the following impacts:*

1. *Avoidance behavior increases:* $a(c) > a(c_0)$
2. *Health improves and the (downward slopping) health-pollution response curve flattens:*

$$h(c, a(c)) > h(c, a(c_0)), \frac{dh}{dc}|_{c_0=c} \geq \frac{dh}{dc}|_{c_0<c}$$

3. *Indirect utility increases:* $V(c, c) > V(c, c_0)$

[19]Another interpretation of Assumption 3 is that people underestimate the negative health impact of pollution.

Appendix A provides the proof. The theoretical model predicts that after the information program, individuals engage in more pollution avoidance, which in turn reduces the health damages from pollution and increases consumer welfare. Our empirical analysis provides empirical tests on the first two predictions and uses the third prediction to quantify the value of the information program.

3.2 Value of information

To derive the value of information (VOI), recall that:

$$V(c,c) = U[x, h(c, a(c))] + \lambda\{I + w * g\left[h(c, a(c))\right] - x - p_a(a(c))\}$$

where $V(c,c)$ denotes the indirect utility when individuals correctly perceive pollution, and avoidance is chosen optimally according to the following condition:[20]

$$\left[U_h(c,a) + \lambda * w * g_h(h(c,a))\right] \frac{\partial h(c,a)}{\partial a} - \lambda p_a = 0 \tag{1}$$

The indirect utility before the information program is:

$$V(c,c_0) = U[x, h(c, a(c_0))] + \lambda\{I + w * g\left[h(c, a(c_0))\right] - x - p_a(a(c_0))\}$$

The key difference between $V(c,c)$ and $V(c,c_0)$ is in the choice of avoidance: $a(c)$ is determined by equation (1) rather than equation (A.1). To derive the value of information, we apply the Tailor's expansion to the indirect utility function $V(c,c)$ at the second argument $c = c_0$: $V(c,c) = V(c,c_0) + \frac{\partial V}{\partial c_0}(c - c_0) + Op(c - c_0)$. The value of information is therefore:

$$\begin{aligned} VOI &= V(c,c) - V(c,c_0) \\ &= \{U_h * \frac{\partial h}{\partial a} * \frac{\partial a}{\partial c_0} + \lambda * w * g_h * \frac{\partial h}{\partial a} * \frac{\partial a}{\partial c_0} - \lambda * p_a * \frac{\partial a}{\partial c_0}\}(c - c_0) + Op(c - c_0) \end{aligned}$$

where $O_p(c - c_0)$ denotes higher order terms of $(c - c_0)$. There are three terms in the curly bracket. The first refers changes in utility due to health improvement from the avoidance behavior. The second denotes changes in wage income. The third includes changes in the avoidance cost. The benefit of the program, or the value of information, is bounded below by the first and third terms, which we measure in our empirical analysis.

[20] x drops out from the indirect utility function since $U(x,h)$ is quasi-linear.

4 The Sea Change in Information Access and Awareness

4.1 Information Access: News and Mobile Apps

The government's official newspaper, *People's Daily*, and mobile phone apps are among the primary venues for the general public to access pollution information. In Figure 3a, we count the number of days in each month when *People's Daily* mentions "smog" in any articles. "Smog" is rarely mentioned before 2013. Almost immediately following the information program's initial roll-out, the frequency of "smog" appearing in People's Daily jumped to roughly 15 days per month.

One might be concerned that the sharp increase in "smog" mentions post 2013 is confounded by changes in the general environment (shifts in government policies etc.) instead of driven by the information program, which is gradually rolled out across cities. To examine this, we scan each smog article in *People's Daily* to determine the list of cities that are mentioned. This allows us to construct a time series measure of "smog" for each city. Figure 3b plots standardized "smog" (mean 0, standard deviation 1) as a function of time since the roll-out of the information program, spanning a year before and a year after. We estimate an event study controlling for month-of-year dummies (12 indicators) and year dummies (5 indicators). In other words, we test for a discontinuous change in "smog" mentions after a city begins to monitor pollution, conditional on general within-year seasonality as well as year-by-year changes in pollution.

The graphical pattern features a discrete increase exactly on the roll-out date (event month t=0). By one year after the roll-out, "smog" mentions in cities with the monitoring stations have increased by 50% of a standard deviation. In other words, there is a substantial increase in the chance that *People's Daily* mentions smog in a city after the city begins to monitor pollution. Assuming unobserved changes in the overall environment do not correlate exactly with the timing of the monitoring roll-out, the difference between pre ($t < 0$) vs. post ($t \geq 0$) coefficients identifies the causal impact of the information program. We have repeated this analysis using other keywords including "air pollution" and "atmospheric pollution", with very similar results.

We then examine the availability of pollution-related mobile phone apps. Unlike newspapers which provide pollution information at a daily frequency, information from apps are more readily accessible in real time. Given the high mobile phone penetration in China, pollution apps serve as a significant venue through which the public learn about their pollution exposure at the moment. We compare the distribution of release time for pollution apps with apps from other popular categories including gaming, music, video, reading, finance, sports, education, shopping, and navigation, which capture the majority of commonly-used apps.

Figure 4 presents the distribution of release time for pollution apps against apps in the control group. There is a clear surge in the density of pollution apps released after the information program, relative to non-pollution apps. The largest increase in the probability of releasing a pollution app occurs one quarter after the initial monitoring roll-out. In total, about 82% of pollution apps are released post January 2013, vs. 62% for non-pollution apps. Implicitly, this means that the availability of pollution apps have grown nearly 500% post 2013, which is four times faster than then growth of other apps.[21]

4.2 Awareness: Web Searches and Air Purifier Sales

We examine changes in the public awareness of air pollution issues in two ways. First, we measure the demand for pollution-related information by internet searches on Baidu that are related to "smog". This analysis is analogous to the examination on "smog" news in section 4.1. Figure 5a plots the time-series pattern of the search index at the national level. As in section 4.1, the smog search index increased sharply starting January 2013, the month of the initial roll-out. Post-2013 searches exhibit a strong secular pattern where the index is highest in winter seasons, as smog is more severe in winter partly due to coal-fueled heating.

Leveraging the search index at the city × daily level for over 300 cities, figure 5b plots the mean of the standardized search indexes in the year before and the year after the roll-out at a local city. Echoing results in section 4.1, the index is flat and near zero prior to the information program and rises rapidly when monitoring starts. By one year after the information program, smog searches have increased by 75% of a standard deviation. Examining other pollution-related search phrases such as "mask" and "air purifier" deliver very similar results.

Second, we examine changes in public and private investment in defensive equipment. We repeat the exact same analysis using data on monthly air purifier sales for fifty cities. Air purifier sales more than double, rising from 11,000 units sold per month in 2012 to over 25,000 units per month after 2013 (Figure 6a). Similar to web searches, air purifier sales also exhibit a strong seasonality with more sales in winter. Finally, the increase in sales coincides with the timing of the roll-out at a city (Figure 6b). For a typical city in our sample, air purifier sales increase by over 100% after monitoring begins.

[21] There was a mass of pollution apps released before 2013. These apps typically stream weather information and later incorporated real-time air quality contents post 2013. These apps are therefore categorized as pollution apps by the time we queried for the Appstore data.

4.3 Changes in Social and Economic Conditions

As we elaborate in Section 5.1, our research designs rely on the assumption that there are no confounding factors that systematically coincide with the timing of the information roll-out. China is experiencing rapid social and economic changes during the sample period. While our statistical analysis could control for general as well as city-specific time trends using fixed effects, one might be concerned about differential trends that correlate systematically with the timing of the roll-out.

Appendix Table D.2 presents a series of tests on differential shifts in city-level observables before and after the program. We focus on four classes of social and economic conditions: pollution levels (using satellite-based AOD), political and regulatory environment (the number of downfall local officials during the anti-corruption campaign, demographics of local political leaders, news mentions of regulation policies), and healthcare access (the number of medical facilities). Overall, we find no evidence that the general social and economic conditions or the air quality correlates with the information roll-out.

5 Pollution Disclosure, Behavior, and Health

5.1 Empirical Framework

As shown above, the information program has substantially expanded public access to pollution information and dramatically increased households' awareness about pollution issues. In turn, these changes have triggered a cascade of short-run and long-run behavioral changes in household activities and health outcomes, including avoidance behavior, housing choice and prices, and mortality. We use the following empirical framework to examine the change in the relationship between pollution exposure and the outcomes (i.e. the "slope") before versus after the program:

$$\mathrm{Outcome}_{ct} = \alpha \times \mathrm{Pollution}_{ct} + \beta \times \mathrm{Pollution}_{ct} \times \mathrm{d}_{ct} + \mathrm{x}'_{ct}\gamma + \varepsilon_{ct}, \tag{2}$$

where c denotes a city and t denotes time (e.g., day or week). $\mathrm{Pollution}_{ct}$ is the AOD measure of the ambient air quality. Dummy d_{ct} represents the information treatment and takes the value one for all periods after city t implements the information program based on the staggered roll-out schedule. Vector x_{ct} includes weather conditions and rich spatial and temporal fixed effects such as city fixed effects and time fixed effects. The last term ε_{ct} denotes remaining unexplained shocks.

All analyses below include full interactions between $\mathrm{Pollution}_{ct}$ and the treatment dummy

d_{ct}. Hence α represents the outcome-pollution gradient before the information program, and β represents changes in the gradient after the program, as denoted by β in the empirical framework.

Equation (2) highlights the difference between our study and the previous literature that estimates the causal effect of air pollution exposure. Conventionally, the key threat to identification arises because pollution exposure is likely to be correlated with the error term: $E(\text{Pollution}_{ct} \times \varepsilon_{ct}) \neq 0$. Such endogeneity could be due to omitted variables or errors in the measurement of pollution exposure.[22]Addressing endogeneity in air pollution is challenging and has been the subject of recent research on understanding the morbidity and mortality cost of air pollution (e.g., Bayer, Keohane and Timmins, 2009; Chen et al., 2013; Arceo, Hanna and Oliva, 2015; Deschenes, Greenstone and Shapiro, 2017; Ito and Zhang, 2018; Barwick et al., 2018). In contrast, the scope of our empirical analysis differs in two ways. First, in most cases we are not interested in the causal effect of pollution per se (which is α), but rather in the *change* in the causal effect before versus after the information program (which is β). Second, in our analysis, Pollution_{ct} is intended to be a direct measure of ambient pollution, rather than population exposure which is determined by the ambient air quality, avoidance behavior, and population distribution. In fact, in the analysis below we directly examine how avoidance and residential sorting respond to ambient air pollution with versus without readily available pollution information.

The key insight of our empirical framework is that, under reasonable assumptions, one can consistently estimate the change in pollution's causal effects (β) using OLS, without having to consistently estimate the level of the effect (α). Intuitively, β measures the change in the slope of the pollution-outcome relationship before and after the treatment. If we were to separately estimate the slope using data before and after the treatment, the endogeneity in pollution would lead to inconsistency in both estimates. However, if the nature of the endogeneity is not affected by the treatment, the inconsistency in the slope estimates would cancel out, leaving the OLS estimate of β to be consistent. The following two assumptions formalize this intuition:

Assumption (B1): $E(\varepsilon_{ct} | d_{ct}, x_{ct}) = 0$. This assumption implies that conditioning on city attributes and other controls x_{ct} such as city and week fixed effects and weather conditions, the treatment d_{ct} is exogenous. As discussed in section 2.1, the information program was implemented against the backdrop of MEP's promulgation of the national $PM_{2.5}$ standard, which marked a sudden and drastic change in the government's stance regarding the im-

[22]For example, satellite-based AOD captures particulate concentration in the entire air column above a ground spot, which might differ from ground-level exposure. In addition, ambient pollution might differ from actual exposure due to the outdoor-indoor difference in the pollution level.

portance of environmental quality. The roll-out schedule of the monitoring stations in three waves was largely based on the pre-determined city designations (e.g., provincial capitals or the list of environmental improvement priority cities designated in 2007), as shown in Figure 2 and Appendix Figure D.1.

One might be concerned about confounding factors that are correlated with these pre-determined city designations. For example, the enforcement of the national $PM_{2.5}$ standards established in 2012 might be systematically correlated with the roll-out schedule. To examine whether this assumption of conditional exogeneity is reasonable, we first compare a host of city attributes before and after the program, including variables that reflect the political and regulatory environment (the number of *People's Daily* news articles that report smog in a city for a given week, the number of anti-corruption cases, the age of the city mayor, whether the city mayor has a Ph.D. degree), healthcare access (the number of hospitals per 1,000 people), and most importantly, the pollution levels (both the weekly average and the maximum pollution reading in a city and week). If changes in the (implementation of the) environmental regulations are systematically correlated with the program roll-out, then we should expect pollution levels as well as proxies for the regulatory environment to change before and after the program. Results for the seven measures discussed above across four different specifications with an increasingly demanding set of controls indicate no discernible differences in any of these twenty-eight regressions (Appendix Table D.2), suggesting that the role of unobserved confounding factors is likely limited (Altonji, Elder, and Taber 2008, Emily and Andrews 2019).

In Section 5.2 below, we conduct two additional robustness analyses using a triple difference framework and randomized assignment. Our results are robust to confounders that are correlated to our treatment either spacially or over time.

Assumption (B2): $\text{Pollution}_{ct} = \mathrm{x}_{ct}'\theta + \nu_{ct}$, where $E\left(v_{ct}|\mathrm{x}_{ct}\right) = 0$ and $E\left(\varepsilon_{ct} \times \nu_{ct}\right) = \sigma_{\varepsilon\nu}$. This assumption implies that the endogeneity of Pollution_{ct}, which arises from the correlation of ν_{ct} and ε_{ct}, is not affected by the information program. This assumption is analogous to the "parallel trend" assumption in the Difference-in-Difference framework. In the analyses below, we plot the event study of the α coefficients, which are flat and stable in all estimations we have conducted. This contrasts sharply with a sizeable break at the time of the information treatment that is both economically and statistically significant and stable post the treatment. These patterns suggest that the outcome-pollution relationship is stable except for the information program. In other words, our assumption of a 'stable correlation' between the endogenous variable (Pollution) and the residuals is tenable in our setting.

Proposition 2. *Under Assumptions (B1) and (B2), the OLS estimate of β in equation (2) is consistent.*

The proof is provided in Appendix C. There are two sources of inconsistency in the OLS estimate of β: one from the endogeneity of the interaction term $\text{Pollution}_{ct} \times \text{d}_{ct}$, and the other from smearing due to the endogeneity of Pollution_{ct}. Under Assumptions (B1) and (B2), the inconsistency from these two sources cancels out. Based on Proposition 2, our subsequent analysis focuses on the OLS estimate of β. In Section 5.4 where we are interested in the baseline impact of pollution on mortality (α), we use both the regression discontinuity design and the IV strategy from the literature to address the endogeneity of Pollution_{ct}.

5.2 Pollution Disclosure and Avoidance

With access to reliable pollution information, households can take different measures to avoid or mitigate pollution exposure. Some low-cost and effective solutions include staying indoors, wearing facial masks, or using air purifiers when pollution is elevated. We first examine how the relationship between outdoor purchase trips and ambient pollution levels changes after the information program is implemented in a city via an event study:

$$\text{PurchaseRate}_{ct} = \sum_{k=-24}^{15} \beta_k \times \ln\,\text{Pollution}_{ct} \times \mathbb{1}(t=k) + \sum_{k=-24}^{15} \eta_k \times \mathbb{1}(t=k) + \text{x}'_{ct}\gamma + \varepsilon_{ct} \quad (3)$$

where c denotes city and t denotes week. The outcome variable "PurchaseRate_{ct}" is the number of card transactions in city c at week t per 10,000 active cards in the city in the corresponding year (Section 2.2). The pollution measure "$\ln\,\text{Pollution}_{ct}$" is logged average AOD. The key parameters of interest are the β's, which represent changes in purchase rate for one percent increase in AOD. To examine changes in the purchase-pollution relationship before versus after the program, we allow β's to vary over time relative to the roll-out month. Cities in different waves have different numbers of available pre and post periods. We examine an event window that spans 39 months (24 months before and 15 months after the program) and dummy out the remaining sample periods. This guarantees that there are nearly identical city $\times$ week observations underlying each event month.

We identify β's using week-to-week variations in air pollution net of a flexible set of geographic and time controls (x_{ct}) that include prefectiry-city FEs, week-of-year FEs, and year FEs. Standard errors are clustered at the city level to allow for arbitrary serial-correlations among the sample periods (weekly observations over five years). In order for the β's estimates to be representative of the population impact, we weight the regression using the number of active cards in a city and year as cities differ greatly in size.

Figure 7 summarizes the estimates of β_k coefficients. We restrict β_k to vary at the quarterly (3-month) level to average out noises in time trends. Two patterns emerge. First, before the program, the β_k estimates are flat and statistically indistinguishable from zero, suggesting a lack of behavioral responses to pollution when individuals have limited access to information. Second, β_k estimates exhibit a level-shift and become strongly negative after the program.

To examine the robustness of these patterns, we repeat the analysis across a range of specification choices (Table 1), modifying equation (3) in two ways. First, instead of the event dummies, we include the full interactions between the pollution term and the post-treatment dummy as in equation (2). Second, we increasingly tighten the fixed effects to exploit finer variation in the data. Column 1 uses city, week-of-year, and year fixed effects, which corresponds to the specification of Figure 7. Column 2 uses city and week-of-sample fixed effects (fixed effects for all weeks in 2011-2015), exploiting variation in pollution across cities in the same week-in-time. Column 3 further adds region × year fixed effects, allowing for common trends in transactions and pollution that are specific to each region.[23] Column 4 is our most stringent specification, controlling for city and region × week-of-sample fixed effects. We obtain similar estimation results across the board.

Consistent with the evidence in Figure 7, outdoor consumption trips are invariant to pollution before the information program, suggesting that households are unlikely to be engaging in any mitigating measures when pollution is elevated. In contrast, after a city implements the pollution monitoring and disclosure program, purchase trips become responsive to pollution levels: a doubling of the pollution level reduces purchase trips by 3 percentage points, according to our preferred specification in Column 4. This is not a trivial change given that the average week-to-week variation in AOD is 76% during our sample period. In addition, our analysis is at the weekly level, which by construction has already incorporated within-week inter-temporal substitution. The estimate reflects to a large extent permanently displaced outdoor trips as households seek to mitigate pollution exposure.

As a point of reference, Cutter and Neidell (2009) find that when 'Spare the Air' alert is issued in San Francisco Bay Area, *daily* traffic is reduced by 2.5-3.5% with the largest effect during and just after the morning commuting period.[24] Graff Zivin and Neidell (2009) estimate that *1-day* smog alerts issued in Southern California lead to a 8-15% reduction in attendance at two major outdoor facilities (the Los Angeles Zoo and the Griffith Park Obser-

[23]"Region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Centralsouth (81 cities), Southwest (54 cities), Northwest (52 cities).

[24]The 'Spare the Air' (STA) advisories, designed to elicit voluntary reductions in vehicle usage and encourage the usage of public transit and ride-sharing, are issued on days when ground-level ozone is predicted to exceed National Ambient Air Quality Standards.

vatory). These two studies focus on immediate (daily) behavioral changes after government-issued air quality warnings while our elasticity estimates are with respect to marginal changes of air quality over the course of a week post the information program.

In Appendix D, we report three sets of additional analyses that support our main findings. First, we examine heterogeneity by the "deferrable" vs. "scheduled" nature of the consumption. Deferrable categories include supermarkets, dining, and entertainment. These shopping trips are more likely subject to pollution avoidance. These categories experience a 5 to 7 percentage point increase in purchase-pollution elasticity with the most stringent set of controls and explain over 75% of the change in overall purchase-pollution gradient. On the other hand, we conduct placebo-style tests looking at the impact of information roll-out on "scheduled" consumption including billings (bills in utilities, insurance, telecommunication, and cable services), government services (court costs, fines, taxes), business-to-business wholesales, as well as cancer treatment centers. There is no statistical evidence that information availability changes "scheduled" consumption's responses to air pollution.

Second, we test the robustness of our results across a range of specification choices. To highlight a few examples, we find that the inclusion of flexible weather controls are not consequential to our estimation, that online transactions cannot explain away our findings, and that our conclusion holds for cities without U.S. Embassy or Consulates Offices (these Offices have independent $PM_{2.5}$ monitoring and so residents in the cities might have better information on air quality). Our findings are also robust to a more saturated research design where, for each wave of "treatment" cities, we introduce a group of "control" cities that neighbor the treatment cities, but have not yet experienced monitoring. We then estimate *differential* change in the transaction-pollution slope across the treatment vs. control cities. The logic of this test is similar to a triple difference design comparing purchase behavior before vs. after the information program, in treatment vs. control cities, on high vs. low pollution days.

Finally, we perform randomized inference to address the concern that we have a small number (three) of roll-out waves which is less than ideal for a staggered event study design. Appendix Figure D.8a compares our true t-test statistic with an "empirical null distribution" of test statistics obtained through 500 repetitions of random assignment of cities into monitoring roll-out waves.[25]Tested against the empirical null (rather than the theoretical null of a t distribution), our effect estimate is statistically significant at the 5% level.

[25]Conceptually, we take the list in Appendix Figure D.1 and shuffle cities across waves. We do this for 500 times, and each shuffle yields a coefficient estimate. Notice that, because we have three roll-out waves, each city has a nearly 1/3 chance of "landing" in its true wave. Our distribution of placebo estimates are therefore expected to center around a negative number.

5.3 Pollution Disclosure and Housing Choices

We now turn to assess housing market responses to pollution disclosure using the near-universe transactions of new homes sold in Beijing from January 2006 to April 2014. We observe the following features for every unit sold: transaction price, transaction date, apartment complex name, address, and attributes, floor level, and unit size.[26] Analogous to the previous analysis in section 5.2, we study the housing price-pollution relationship across neighborhoods with vary degrees of air pollution and examine the degree to which this relationship shifts before and after the information program was implemented in Beijing in January 2013. Because we only have 16 months of transactions post the treatment, we skip the event study and simply estimate the change in the price-pollution gradient.

Housing purchase decisions are likely to be affected by the long-run pollution level rather than day-to-day variations. As a result, we focus on year-to-year changes in housing prices. To do so, we first take all housing transactions and estimate the following equation:

$$\ln \text{TransactionPrice}_{ict} = \text{w}'_{ict}\gamma + \eta_{cy} + \varepsilon_{ict}, \tag{4}$$

where $\ln \text{TransactionPrice}_{ict}$ is the log transaction price of unit i in apartment-complex c on date t. The vector of unit characteristics w_{ict} includes floor fixed effects, sale month-of-year fixed effects, unit size and its quadratic term. Our variable of interest is η_{cy}, which are apartment-complex $\times$ year level averages of housing prices after controlling for observable attributes. There are on average 153 underlying housing transactions for each apartment-complex and year.

Once we obtain the estimated quality-adjusted housing price index at the apartment-complex $\times$year level, $\hat{\eta}_{cy}$, we examine the relationship between housing price and pollution using a framework similar to equation (2). We use two different measures of pollution at the sub-city level: fine-scale ambient air quality (AOD) at 1km-by-1km $\times$ year resolution, and distance to major polluters as a proxy for local pollution.

The regression equation is:

$$\hat{\eta}_{cy} = \alpha \cdot \ln \text{Pollution}_{cy} + \beta \cdot \ln \text{Pollution}_{cy} \times \mathbb{1}(\text{after monitoring}) + \text{x}'_{cy}\gamma + \varepsilon_{cy} \tag{5}$$

where $\ln \text{Pollution}_{cy}$ is one of the two pollution measures. β captures the change in pollution-

[26]There are three types of geographical units in this analysis: district, community (or "jiedao"), and an apartment complex. The municipality of the Metropolitan Beijing area is divided into sixteen districts, which is further divided into 180 communities and 1,200 apartment complexes. A community is comparable to a zip-code in the U.S. in terms of geographical coverage while an apartment complex is similar to a census tract.

housing gradient before and after the program. We discuss each of the these two sets of analysis below.

Fine-scale AOD and Housing Prices. To obtain a pollution measure with a high level of spatial resolution, we employ a frontier method in atmospheric science called "oversampling" that re-processes the original AOD data to increase its spatial resolution from 10-by-10 km to 1-by-1 km, while sacrificing the temporal resolution from daily to annual. Oversampling takes advantage of the fact that MODIS scans a slightly different, but overlapping, set of pixels at a given location on each of the satellite's overpass. When the researcher is not interested in the high temporal dimension (as in our case where we only need the annual pollution), it is possible to average across the overlapping overpasses to enhance the geo-spatial resolution of the AOD measure.[27] Figure 8 presents the pre- and post-oversampling average AOD concentration for the city of Beijing. Our first pollution measure in the housing analysis is therefore the oversampled AOD level in year y in the 1-by-1km region that contains the apartment-complex c.[28]

We use two sources of variation in our regression analysis. The first source of variation comes from the fact that we often observe transactions in the same apartment-complex for a streak of years before all units are sold out. We can therefore use a standard panel fixed effects regression strategy to compare transaction prices within the same complex, but across different years with high versus low pollution levels. In this specification, we include apartment-complex fixed effects, year fixed effects, and "year-on-market" fixed effects (9 indicators, each indicates if year y is the apartment-complex's r^{th} year on market).

The second source of variation comes from our ability to observe fine-grained, cross-sectional variations in air pollution even within small geographic area. We observe about 1,200 apartment-complexes scattered in 180 communities across 16 districts in Beijing. We compare transaction prices within the same district $\times$ year, but across apartment-complexes in areas with high versus low pollution levels, controlling for time-invariant differences in community-level characteristics. Hence in the second type of specification, we include district $\times$ year fixed effects, community fixed effects, and year-on-market fixed effects.

The two specifications exploit rather different sources of pollution variation, with the former focusing more on year-to-year variation within the same location, the latter focusing more on cross-sectional variation at a given point in time. To flexibly account for potential autocorrelation in both housing price and pollution across time and over space, we two-way cluster standard errors at the community level and the district $\times$ year level. We report results for equation 5 in Table 2. Column 1 shows that prior to the information program, a doubling

[27] Appendix Figure D.9 illustrates the oversampling idea using two consecutive days of MODIS AOD data.

[28] Air purifiers have become common household appliance since 2013 and provide real-time $PM_{2.5}$ readings.

of annual pollution corresponds to an insignificant 9% increase in housing prices. After the program, the price elasticity becomes negative and the change in elasticity is 59 percentage points and significant at the 10% confidence level. The results suggest that housing prices do not respond to variation in pollution levels before the program, while after the program, air quality is capitalized in housing prices.

In column 2, we examine the effect of lagged pollution in addition to current year's pollution exposure. We obtain similar results: a marginally significant 73 percentage-point change in elasticity on current pollution, but a noisy effect from lagged pollution. Columns 3 and 4 correspond to our cross-sectional specification estimates. These specification yields a similar reduction-in-elasticity estimates of 85 percentage points.

Our estimates of housing price-pollution elasticity for the post-monitoring period therefore ranges from -0.6 to -0.8. This is somewhat larger than those obtained in the U.S. setting. For example, Chay and Greenstone (2005) exploits permanent reduction in Total Suspended Particle pollution (TSP) due to the 1970s U.S. Clean Air Act. They estimate a price-pollution elasticity of -0.25. Taking into account moving costs and variation in air quality across U.S. metro areas, Bayer, Keohane and Timmins (2009) show a price-pollution elasticity of roughly -0.34 to -0.42. Our estimates are similar to those obtained in China settings. In a hedonic regression exercise using Beijing's housing transactions and land parcel data, Zheng and Kahn (2008) find a price-PM elasticity of -0.41. In a recent residential-sorting exercise, Freeman et al. (2019) use moving costs and housing value information from China Population Census micro-level data to estimate a price-PM2.5 elasticity of -0.71 to -1.10.

Proximity to Major Polluters and Housing Prices. Our second pollution measure is distance to the nearest major pollution source, following the literature (e.g., Davis, 2011; Currie et al., 2015; Muehlenbachs, Spiller and Timmins, 2015). We examine how the distance gradient shifts before versus after the information program. Large polluters tend to be visible and well known landmarks in a city. The information program could raise the salience of the potential health impacts of these large polluters in residents' housing choice decisions.

As described in section 2.2, our distance-gradient analysis begins with 41 top 10% polluters in Beijing that were in operation from 2007 – 2018 and account for nearly 90% of total emissions in 2007 (Appendix Figure D.10). Using geo-locations of all four hundred plus major polluters, we construct a time-invariant "distance to top-decile polluter" variable as our second pollution measure while controling for distance to non-top polluters. We control for district by year fixed effects, community fixed effects, and year-on-market fixed effects. We drop apartment-complex fixed effects which are perfectly colinear with the time-invariant distance to the nearest major pollutor.

Figure 9 presents the results. Figure 9a shows the estimated distance gradients separately for before and after the information program. We detect no statistically significant distance gradient curve before the program. The shape of the curve shifted substantially after the program, where a near-monotonic price-distance relationship emerges. Figure 9b plots the difference in the distance gradient. Houses within 3 km of the top polluters experienced the largest depreciation of about 27%. The effect fades with distance and becomes insignificant over 6 km. The magnitude is large but not implausible given the unprecedented housing boom in the city: the average housing price in Beijing grew by 262% during our sample period and the effect size corresponds to 42% of the inter-quartile range of the housing price dispersion.

5.4 Pollution Disclosure and Health Benefit

Our previous analyses have documented a range of behavioral responses to the information program. To quantify the value of pollution information, our endpoint analysis is to examine whether the same amount of pollution exposure is associated with fewer deaths after information becomes widely available using county-level mortality data from 2011 to 2016. Similar to Section 5.2, we conduct an event study and regress logged mortality rate in county $c \times$ quarter t on the corresponding logged pollution level, allowing the coefficient to vary by event quarter k, i.e., the k^{th} quarter since pollution monitoring:

$$\ln \text{Mortality}_{ct} = \sum_{k=-10}^{6} \beta_k \times \ln \text{Pollution}_{ct} \times \mathbb{1}(t=k) + \sum_{k=-10}^{6} \eta_k \times \mathbb{1}(t=k) + \text{x}'_{ct}\gamma + \varepsilon_{ct} \quad (6)$$

We made several specification choices based on the nature of our data. First, we aggregate weekly mortality rate to quarterly to average out noises. However, the qualitative findings are the same whether we conduct our analysis at weekly, monthly, or quarterly level, with the β_k estimates being slightly smaller using the weekly and monthly data. Second, we allow the β_k coefficients to vary from 10 quarters before to 6 quarters after the information program to ensure a roughly balanced number of underlying counties for each event quarter. We have also included a separate dummy variable that groups the remaining quarters.

Figure 10 plots the β_k coefficient estimates. The mortality-pollution elasticity exhibits a roughly flat trend before the program, followed by a noticeable decline after the program. In the event study, we control for city, quarter-of-year, and year fixed effects. We repeat the analysis as in Table 3 and replace the event dummies with full interactions between the pollution term and the post-treatment dummy. We experiment with increasingly stringent fixed effects controls by including quarter-of-sample and region $\times$ quarter-of-sample fixed effects dummies (Tables 1). The coefficient estimate on the interaction term between pollution and the program dummy suggests a statistically significant 5 percentage point reduction in the mortality-pollution elasticity after the program. The results are similar across specifications and consistent with the graphical evidence from Figure 10.

Our heterogeneity analysis provides suggestive evidence of underlying mechanism behind the mortality effect. Specifically, we split the sample into above vs. below average values of a series of city-level characteristics, including per capita income, shares of urban population, per capita number of hospitals, per capita residential electricity use, and shares of mobile phone users. Table 4 reports the results where we focus on the interaction between the change-in-gradient "Log(Pollution) $\times$ $\mathbb{1}$(after monitoring)" coefficient and city-level characteristics. Column 1 shows there is no heterogeneity by city's average per capita income. Interesting patterns emerge when we examine more specific dimensions of heterogeneity. Columns 2 through 5 suggest large reduction (about -8 percentage points) in mortality damage in cities that are more urban, having more hospitals, having higher rate of residential electricity use, and with higher mobile phone penetration. These findings are broadly consistent with the fact that residents in these cities are more likely to benefit from pollution information and engage in defensive activity that counteracts health damages of air pollution exposure.

In Appendix D, we conduct a series of additional tests to examine the plausibility of the reduction in the mortality-pollution elasticity estimates. First, Appendix Figure D.11a examines age-specific mortality rates. The effect is most precisely estimated among people aged over 40 who are more vulnerable to pollution exposure than younger age groups. Appendix Figure D.11b illustrates that changes in the mortality-pollution relationship concentrate in cardio-respiratory causes, such as COPD, heart diseases, and cerebrovascular diseases, which are widely considered as the most relevant consequences of pollution exposure. The impact on mortality-pollution relationship from respiratory infection and digestive diseases is both small and insignificant. For traffic fatalities, the relationship post disclosure appears to become flatter though the change is not statistically significant.[29] Third, we have explored non-linear specification and found that the reduction in the mortality-pollution gradient is insignificantly convex in the level of pollution shock (Appendix Figure D.12). Finally, we repeat the same randomized inference exercise as discussed in section 5.2 and compare the true effect size to a distribution of placebo effect size obtained from 500 repetitions of random city roll-out assignment. Appendix Figure D.8b shows our effect estimate is significant

[29] Air pollution could affect visibility as well as cognitive function (Zhang, Chen and Zhang, 2018), both of which could result in increased risk from traffic accidents.

at the 5% level.

6 The Value of Pollution Information

The value of information (VOI) arises from the power of information in changing decisions. Our analyses illustrate that disclosing pollution information has affected a range of behavioral and market outcomes that reflect households' effort to mitigate the negative health consequences of air pollution. We measure VOI as the fraction of pollution-caused deaths that are avoided by providing information access, holding pollution exposure constant:

$$\text{VOI} = \frac{\epsilon_1 - \epsilon_0}{\epsilon_0} \tag{7}$$

where the ratio is between the *change* in the mortality-pollution elasticity due to the program ($\epsilon_1 - \epsilon_0$) and the *level* of the mortality-pollution elasticity prior to the program (ϵ_0). The numerator corresponds to β in equation (2) and is the interaction coefficient reported in Table 3. The denominator corresponds to α in equation (2) and is the coefficient estimate on "Ln(Pollution)" in Table 3. The counter-intuitive magnitude is similar to the OLS estimates in the literature using the correlation between PM exposure and mortality in China (e.g., Yin et al., 2017; Ebenstein et al., 2017). Studies based on quasi-experimental methods have yielded much larger effect sizes in the right direction (e.g., Chen et al., 2013; He, Fan and Zhou, 2016; Ebenstein et al., 2017).

To get at the true level of ϵ_0, we use the main finding of a recent paper by Ebenstein et al. (2017) that examines the long-term mortality effects of PM exposure. We favor this study because it is based on well-established quasi-experimental method, uses a similar data source for mortality measurement, and is based on 2004-2012 before the information program was implemented. Using a regression discontinuity (RD) design that leverages a free coal-based heating policy available only to cities to the north of the Huai River, the authors find a mortality-PM_{10} elasticity of 0.70. Assuming a linear dosage-response function, our estimate of a 5 percentage point reduction in the mortality-pollution elasticity therefore indicates a roughly 7% reduction in deaths attributable to the information program for the same amount of pollution exposure.

With a slightly different sample period, we replicate the RD analysis in Ebenstein et al. (2017) and obtain very similar baseline mortality estimates as shown in Appendix Figure D.13 and Appendix Table D.7. The authors also report an OLS regression between logged cardio-respiratory mortality and logged PM_{10} exposure and yields an correlational elasticity estimate of 0.02, which is similar to our OLS estimate. In unreported analysis, we use the instrumental variable approach introduced in Barwick et al. (2018) that exploits long-range transport of pollution from upwind cities and obtain similar estimates on the baseline mortality impact.

To conceptualize the effect size, we note that under the assumption of linear mortality-pollution dose response function, the benefit of a 7% reduction in mortality-pollution elasticity is roughly the same with the benefit of a 7% reduction in pollution concentration. This corresponds to roughly 10 ug/m^3 reduction in PM_{10} or 5 ug/m^3 reduction in $PM_{2.5}$ in China. We perceive the effect size as plausible for several reasons. First, the effect size is moderate compared to the average cross-city variation in $PM_{2.5}$ after the program (SD = 20.4 ug/m^3, IQR = 25.2 ug/m^3). Second, several recent government programs have been shown to shift pollution levels significantly. For example, the winter heating policy implemented to the north of the Huai River is shown to increase PM_{10} by about 41.7 ug/m^3 (Ebenstein et al., 2017). Large-scale inspection and cleanup efforts across China since 2013 are associated with over 50 ug/m^3 reduction in $PM_{2.5}$ for some northern cities (Greenstone and Schwarz, 2018).

The information program brings sizeable economic and health benefits to the society. Using Ito and Zhang (2018)'s WTP estimate based on air purifier purchases in China, a 10 ug/m^3 reduction in PM_{10} is about RMB 90 ($13.4) per year, which aggregates to RMB 122 billion per year nationwide. In Barwick et al. (2018) , an individual saves RMB 38 ($5.7) in out-of-pocket health spending from a 5 ug/m^3 reduction in $PM_{2.5}$ exposure, aggregating to RMB 52 billion per year nationwide.

We then compare the benefits to financial costs associated with increased defensive and avoidance actions after the program. First, we observe that total sales of air purifiers ($PM_{2.5}$ masks) increased at a rate of RMB 7 (0.55) billion per year post 2013. Because many cities started the information program after 2013, we consider these numbers to be upper bounds on the costs of increased defensive investments due to the information program. Second, we consider forgone consumption due to pollution avoidance. From our bank-card analysis in section 5.2, we expect 1.34 million fewer transactions per year post the information program. The average transaction in our data carries a value of about RMB 3,568. We therefore estimate that the value of forgone transaction is about RMB 4.75 billion per year. This is likely an upper bound on foregone consumption as some of these transactions are probably deferred rather than permanently foregone. As shown in section 5.2, the effect on bank-card use appears to concentrate on "deferrable" categories, many of which are temporally substitutable in nature (such as supermarkets trips). Summing up the costs numbers of defensive goods and forgone consumption, we conclude that the cost of the information program is around RMB 13 billion, which is an order of magnitude smaller than the health benefits. Finally, the estimated expenses to set up the 1,300 monitoring stations and broadcast the pollution information online is estimated to be RMB 2-5 billion, a rather trivial number relative to the discounted future benefit in saved lives.

7 Conclusion

This paper examines the role of pollution information in shaping how ambient air pollution affects household behavior and health outcomes. The focus is on a watershed policy change in China whereby air pollution monitoring stations are installed and real-time pollution information is made public by the government. Based on several rich and unique data sets, our analysis provides consistent evidence that the pollution monitoring and disclosure program led to a cascade of changes such as increased pollution access and awareness, more pronounced short- and long-term avoidance behavior, as well as muted pollution-health relationship. The findings suggest that the value of the program arising from improved health is an order of magnitude larger than its cost.

China's experience offers an important lesson for other developing countries that are experiencing severe environmental challenges. The infrastructure for monitoring environmental quality and disclosing information is often inadequate in those countries. As income rises, the demand for environmental quality increases in these countries and households are better able to adapt to the changing environment. Providing real-time pollution monitoring data, combined with effective dissemination infrastructure such as smartphones and internet that are now commonly available among developing countries, could prove to be a powerful tool to help households mitigate health damages from environmental pollution in other developing countries. In addition, while our study is in the context of environmental quality, the large-scale information program could offer guidance in better leveraging information provision to address issues related to food and nutrition, traffic safety, as well as risky health behaviors.

References

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

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

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2015. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *Economic Journal*, 126: 257–280.

Bai, Jie. 2018. "Melons as Lemons: Asymmetric Information, Consumer Learning and Quality Provision." Working Paper.

Banzhaf, H. Spencer, and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout." *American Economic Review*, 98: 843–863.

Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur. 2018. "The Morbidity Cost of Air Pollution: Evidence from the World's Largest Payment Network." NBER Working Paper.

Bayer, Patrick, Nate Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management*, 58: 1–14.

Bernheim, B. Douglas, and Antonio Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *The Quarterly Journal of Economics*, 124(1): 51–104.

Bernheim, B. Douglas, Andrey Fradkin, and Igor Popov. 2015. "The Welfare Economics of Default Options in 401(k) Plans." *American Economic Review*, 105(9): 2798–2837.

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen. 2011. "Calorie posting in chain restaurants." *American Economic Journal: Economic Policy*, 3(1): 91–128.

Chay, Kenneth Y., and Michael Greenstone. 2005. "Does air quality matter? Evidence from the housing market." *Journal of political Economy*, 113(2): 376–424.

Chen, Shuai, Paulina Oliva, and Peng Zhang. 2017. "The Effect of Air Pollution on Migration: Evidence from China." NBER Working Paper No. 24036.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. "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*, 110: 12936–12941.

Chen, Yuyu, Ginger Jin, Naresh Kumar, and Guang Shi. 2012. "Gaming in Air Pollution Data? Lessons from China." *B.E. Journal of Economic Analysis Policy*, 3(12): e313

– e323.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review*, 99(4): 1145–77.

Craft, Erik D. 1998. "The Value of Weather Information Services for Nineteenth-Century Great Lakes Shipping." *American Economic Review*, 88(5): 1059–76.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker. 2015. "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings." *American Economic Review*, 105(2): 678–709.

Cutter, W. Bowman, and Matthew Neidell. 2009. "Voluntary information programs and environmental regulation: Evidence from 'Spare the Air'." *Journal of Environmental Economics and Management*, 58(3): 253 – 265.

Davis, Lucas W. 2011. "The effect of power plants on local housing values and rents." *Review of Economics and Statistics*, 93(4): 1391–1402.

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

Dranove, David, and Ginger Zhe Jin. 2010. "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature*, 48(4): 935–63.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences*, 114: 10384–10389.

Fioletov, V. E., C. A. McLinden, N. Krotkov, M. D. Moran, and K. Yang. 2011. "Estimation of SO2 emissions using OMI retrievals." *Geophysical Research Letters*, 38(21).

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

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

Graff Zivin, Joshua, and Matthew Neidell. 2009. "Days of haze: Environmental information disclosure and intertemporal avoidance behavior." *Journal of Environmental Economics and Management*, 58(2): 119–128.

Greenstone, Michael, and B. Kelsey Jack. 2015. "Envirodevonomics: A Research Agenda for an Emerging Field." *Journal of Economic Literature*, 53(1): 5–42.

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

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

Grossman, Michael. 1972. "On the Concept of Health Capital and Demand for Health." *The Journal of Political Economy*, 80: 223–255.

Grossman, Sanford J., and Joseph E. Stiglitz. 1976. "Information and Competitive Price Systems." *The American Economic Review*, 66(2): 246–253.

Hao, Jiming, and Litao Wang. 2005. "Improving Urban Air Quality in China: Beijing Case Study." *Journal of the Air & Waste Management Association*, 55(9): 1298–1305.

Hastings, Justine S., and Jeffrey M. Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments*." *The Quarterly Journal of Economics*, 123(4): 1373–1414.

He, Guojun, Maoyong Fan, and Maigeng Zhou. 2016. "The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games." *Journal of Environmental Economics and Management*, 79: 18–39.

Hilton, Ronald W. 1981. "The Determinants of Information Value: Synthesizing Some General Results." *Management Science*, 27(1): 57–64.

Hirshleifer, Jack. 1971. "The Private and Social Value of Information and the Reward to Inventive Activity." *The American Economic Review*, 61(4): 561–574.

Huang, Jing, Xiaochuan Pan, Xinbiao Guo, and Guoxing Li. 2018. "Health impact of China's Air Pollution Prevention and Control Action Plan: an analysis of national air quality monitoring and mortality data." *The Lancet Planetary Health*, 2(7): e313 – e323.

Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review*, 104(2): 537–63.

Ito, Koichiro, and Shuang Zhang. 2018. "Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China." *Journal of Political Economy*. forthcoming.

Jagnani, Maulik, Christopher Barrett, Yanyan Liu, and Liangzhi You. 2018. "Within-Season Producer Response to Warmer Temperatures: Defensive Investments by Kenyan Farmers." Working Paper.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review*, 104(4): 1417–38.

Jin, Ginger Zhe, and Phillip Leslie. 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *The Quarterly Journal of Economics*, 118(2): 409–451.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel. 2012. " Comparison Friction: Experimental Evidence from Medicare Drug Plans." *The Quarterly Journal of Economics*, 127(1): 199–235.

Landrigan, Philip, Richard Fuller, Nereus J R Acosta, Olusoji Adeyi, Maureen Cropper, Alan Krupnick, Michael Greenstone, and et al. 2018. "The Lancet Commission on pollution and health." *The Lancet*, 391(10119): 462–512.

Lave, Lester B. 1963. "The Value of Better Weather Information to the Raisin Industry." *Econometrica*, 31(1/2): 151–164.

Liu, Mengdi, Ronald Shadbegian, and Bing Zhang. 2017. "Does environmental regulation affect labor demand in China? Evidence from the textile printing and dyeing industry." *Journal of Environmental Economics and Management*, 86: 277–294.

Mastromonaco, Ralph. 2015. "Do environmental right-to-know laws affect markets? Capitalization of information in the toxic release inventory." *Journal of Environmental Economics and Management*, 71: 54 – 70.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The Housing Market Impacts of Shale Gas Development." *American Economic Review*, 105(12): 3633–59.

Nelson, Richard R., and Jr. Winter, Sidney G. 1964. "A Case Study in the Economics of Information and Coordination the Weather Forecasting System." *The Quarterly Journal of Economics*, 78(3): 420–441.

Oberholzer-Gee, Felix, and Miki Mitsunari. 2006. "Information regulation: Do the victims of externalities pay attention?" *Journal of Regulatory Economics*, 30(2): 141–158.

Saberian, Soodeh, Anthony Heyes, and Nicholas Rivers. 2017. "Alerts work! Air quality warnings and cycling." *Resource and Energy Economics*, 49: 165 – 185.

Shin, Jeong-Shik. 1985. "Perception of Price When Price Information Is Costly: Evidence from Residential Electricity Demand." *The Review of Economics and Statistics*, 67(4): 591–598.

Shrader, Jeffrey. 2018. "Expectations and adaptation to environmental risks." Working Paper.

Smith, V. Kerry, and F. Reed Johnson. 1988. "How do Risk Perceptions Respond to Information? The Case of Radon." *The Review of Economics and Statistics*, 70(1): 1–8.

Stigler, George J. 1961. "The economics of information." *Journal of Political Economy*, 69(3): 213–225.

Stigler, George J. 1962. "Information in the Labor Market." *Journal of Political Economy*, 70(5): 94–105.

Streets, David G., Timothy Canty, Gregory R. Carmichael, Benjamin de Foy, Russell R. Dickerson, Bryan N. Duncan, David P. Edwards, John A. Haynes, Daven K. Henze, Marc R. Houyoux, Daniel J. Jacob, Nickolay A. Krotkov, Lok N. Lamsal, Yang Liu, Zifeng Lu, Randall V. Martin, Gabriele G. Pfister, Robert W. Pinder, Ross J. Salawitch, and Kevin J. Wecht. 2013. "Emissions estimation from satellite retrievals: A review of current capability." *Atmospheric Environment*, 77: 1011 – 1042.

Sun, Cong, Matthew E. Kahn, and Siqi Zheng. 2017. "Self-protection investment exacerbates air pollution exposure inequality in urban China." *Ecological Economics*, 131: 468 – 474.

Tanaka, Shinsuke. 2015. "Environmental regulations on air pollution in China and their impact on infant mortality." *Journal of Health Economics*, 42: 90 – 103.

Van Donkelaar, Aaron, Randall V Martin, and Rokjin J Park. 2006. "Estimating ground-level PM2. 5 using aerosol optical depth determined from satellite remote sensing." *Journal of Geophysical Research: Atmospheres*, 111(D21).

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

Wichman, Casey J. 2017. "Information provision and consumer behavior: A natural experiment in billing frequency." *Journal of Public Economics*, 152: 13 – 33.

Yi, Honghong, Jiming Hao, and Xiaolong Tang. 2007. "Atmospheric environmental protection in China: Current status, developmental trend and research emphasis." *Energy Policy*, 35(2): 907 – 915.

Yin, Peng, Guojun He, Maoyong Fan, Kowk Yan Chiu, Maorong Fan, Chang Liu, An Xue, Tong Liu, Yuhang Pan, Quan Mu, et al. 2017. "Particulate air pollution and mortality in 38 of China's largest cities: time series analysis." *bmj*, 356: j667.

Zhang, Bing, Xiaolan Chen, and Huanxiu Guo. 2018. "Does central supervision enhance local environmental enforcement? Quasi-experimental evidence from China." *Journal of Public Economics*, 164: 70–90.

Zhang, Junjie, and Quan Mu. 2018. "Air pollution and defensive expenditures: Evidence from particulate-filtering facemasks." *Journal of Environmental Economics and Management*, 92: 517 – 536.

Zhang, Xin, Xi Chen, and Xiaobo Zhang. 2018. "The impact of exposure to air pollution on cognitive performance." *Proceedings of the National Academy of Sciences*, 115(37): 9193–9197.

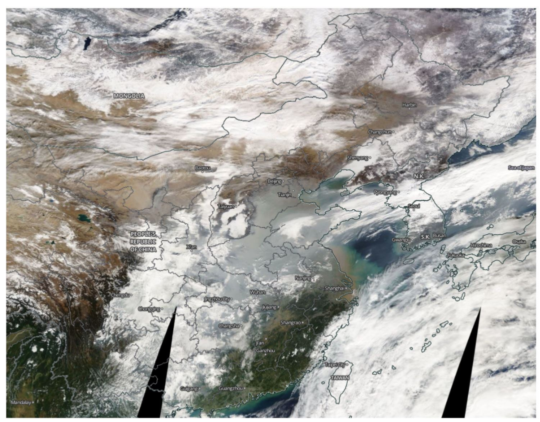
Zheng, Siqi, and Matthew E Kahn. 2008. "Land and residential property markets in a booming economy: New evidence from Beijing." *Journal of Urban Economics*, 63(2): 743–757.

Zhou, Maigeng, Haidong Wang, Jun Zhu, Wanqing Chen, Linhong Wang, Shiwei Liu, Yichong Li, Lijun Wang, Yunning Liu, Peng Yin, et al. 2016. "Cause-specific mortality for 240 causes in China during 1990–2013: a systematic subnational analysis for the Global Burden of Disease Study 2013." *The Lancet*, 387(10015): 251–272.

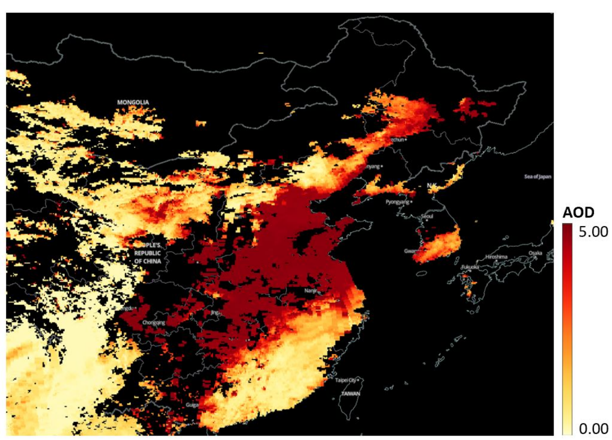
Zivin, Joshua Graff, and Matthew Neidell. 2009. "Days of haze: Environmental information disclosure and intertemporal avoidance behavior." *Journal of Environmental Economics and Management*, 58(2): 119 – 128.

Figure 1: November 2011 “Widespread, Dense Fog Event”

(a) News coverage



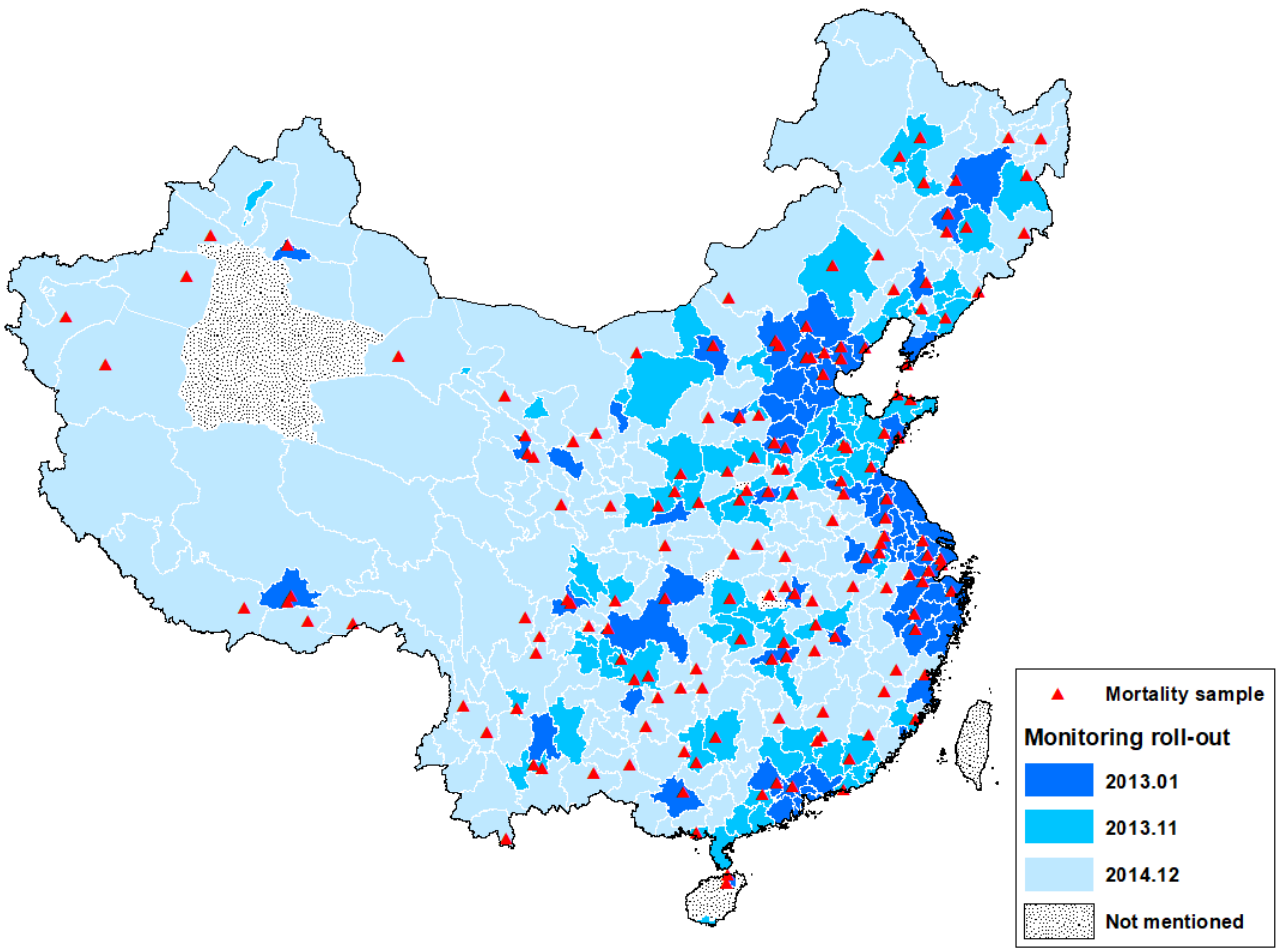
(b) Satellite picture of the event



(c) Satellite-retrieved pollution levels

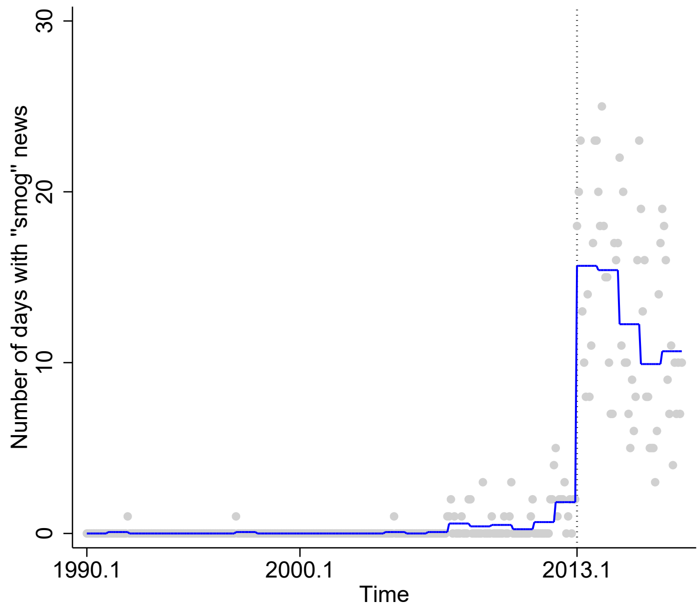
Notes: This figure illustrates a “widespread, dense fog event” around November 27, 2011 which is likely a major pollution event. Panel A, sourced from China Meteorological Administration, shows official news coverage of the event. Panel B, sourced from NASA, shows satellite views of China on the same day. Panel C, sourced from NASA MODIS algorithm, shows a measurement of satellite-based particulates pollution (aerosol optical depth) levels.

Figure 2: Air Pollution Monitoring Roll-out

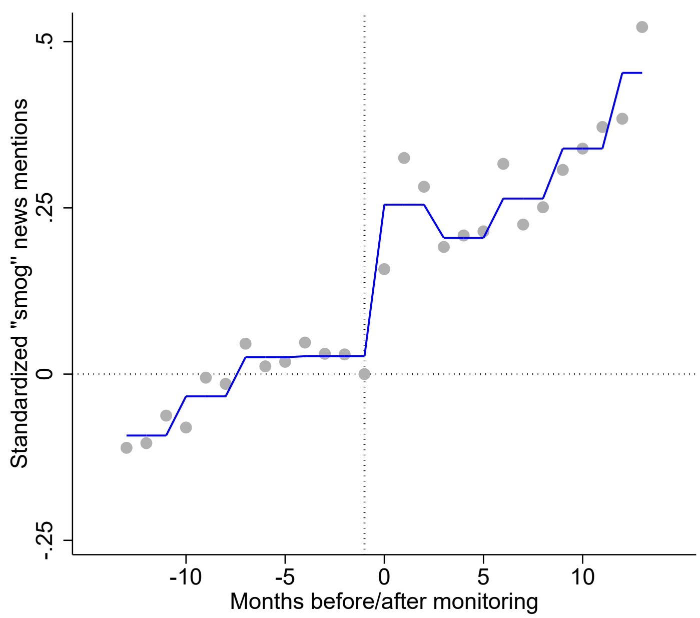


Notes: This map shows prefecture-city by the initiation date of real-time air pollution monitoring. "Not mentioned" are cities where the timing of monitoring is not mentioned in the government's policy notice. "Mortality sample" are centroids of counties included in the DSP mortality data.

Figure 3: Changes in Pollution Information Exposure: News



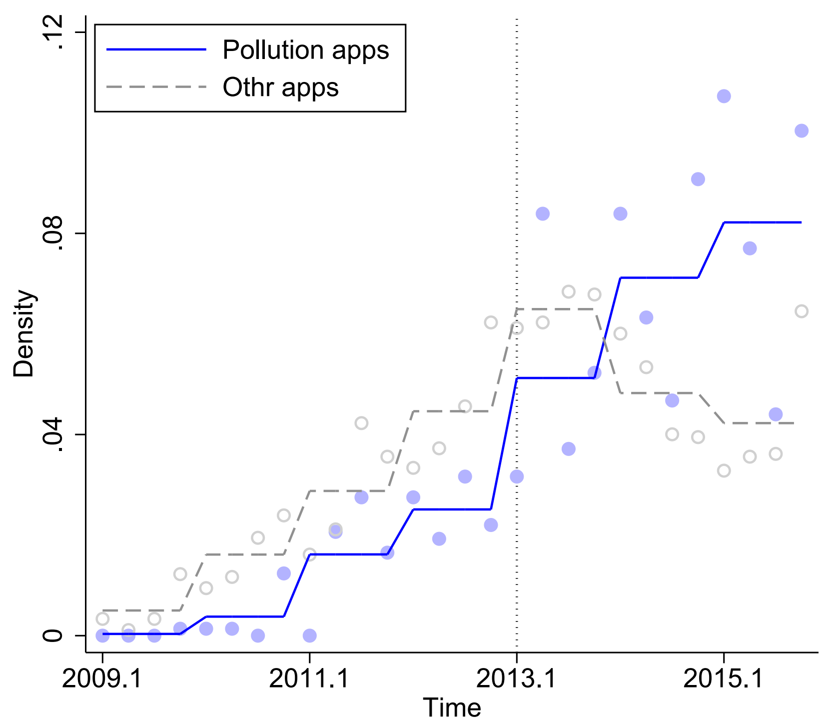
(a) *People's Daily* news "smog" mention



(b) "Smog" mentions before and after monitoring

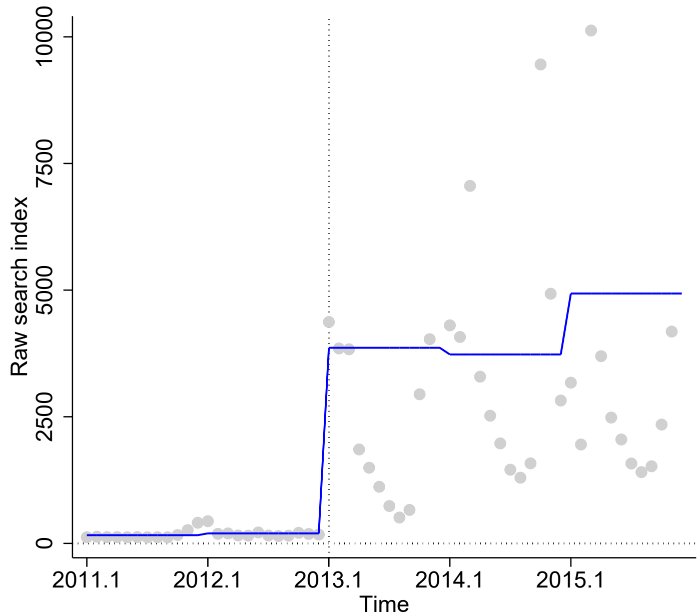
Notes: Panel A plots the number of days in each month when the *People's Daily* (official newspaper of the Chinese government) published articles containing "smog" in content. Each dot represents a month. Line shows annual averages. Panel B mean standardized "smog" mentions associated with the city, defined as news that mention both "smog" and the city name, as a function of month since monitoring initiation. Event month -1 is normalized to 0. The underlying regression controls for month-of-year and year indicators. Line shows quarterly averages.

Figure 4: Changes in Pollution Information Exposure: Mobile Phone Apps

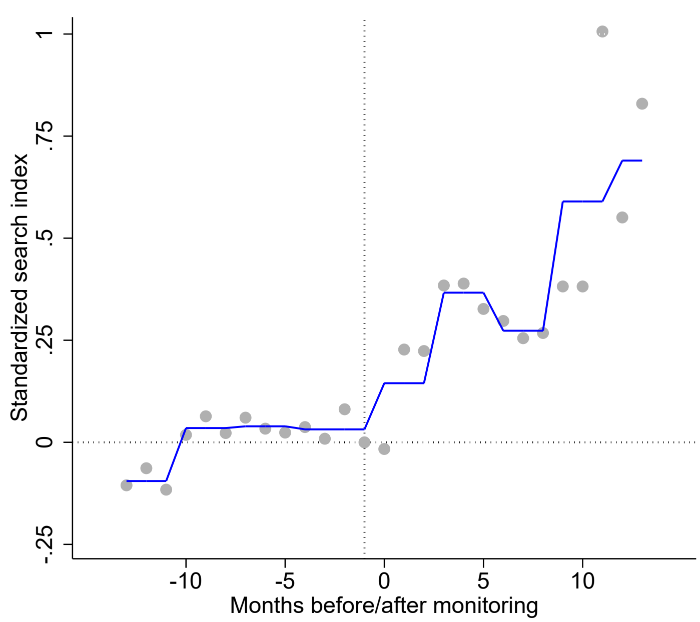


Notes: This chart shows release-date distribution of Apple App Store apps related to pollution (solid dots and line). Averaged release-time distribution for apps in other categories (dashed dots and line) includes game, music, video, reading, finance, sports, education, shopping, and navigation. For each category, sample is restricted to the first 200 apps returned by the Apple API given the search key. Data are accessed on December 27, 2015. Pollution apps released before 2013 typically stream weather information and later incorporated real-time air quality contents post 2013. These apps are therefore categorized as pollution apps by the time we queried for the Appstore data.

Figure 5: Changes in Pollution Awareness: Baidu Smog Search Index



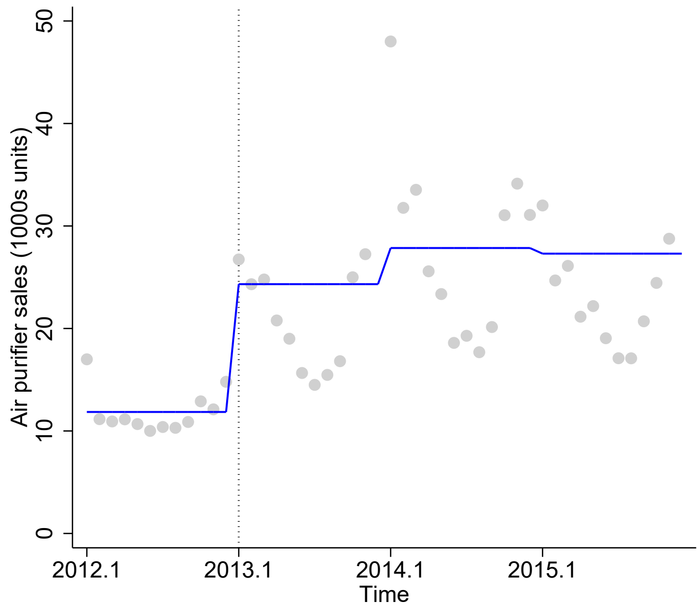
(a) Baidu “smog” search index at the national level



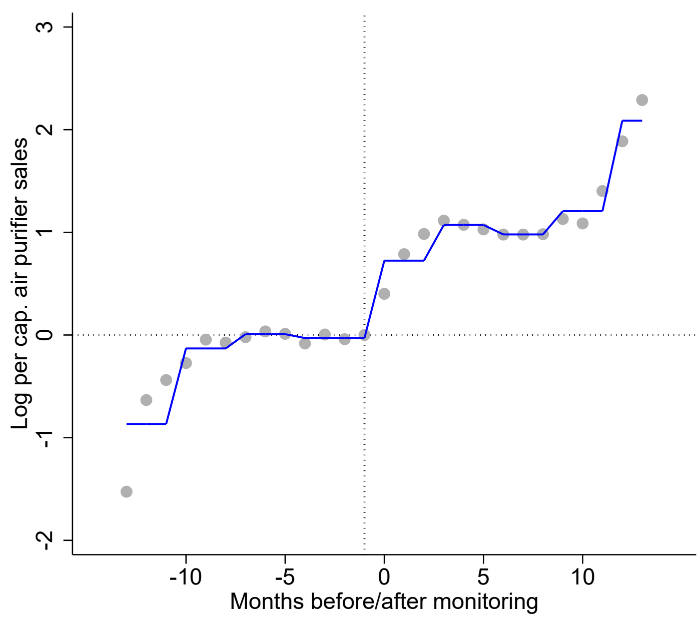
(b) Baidu “Smog” search index before and after a city implements the information program

Notes: Panel A plots raw monthly trends in Baidu Search Index for the word “smog”. The graph omits two dots with exceptionally high search index for readability purpose. These dots correspond to December 2013 (index = 20,942) and Decembet 2015 (index = 24,679). Line shows annual averages. Panel B plots mean standardized “smog” search index as a function of months since monitoring initiation. Event month -1 is normalized to 0. The underlying regression controls for month-of-year and year indicators. Line shows quarterly averages.

Figure 6: Changes in Air Purifier Sales (50 Cities)



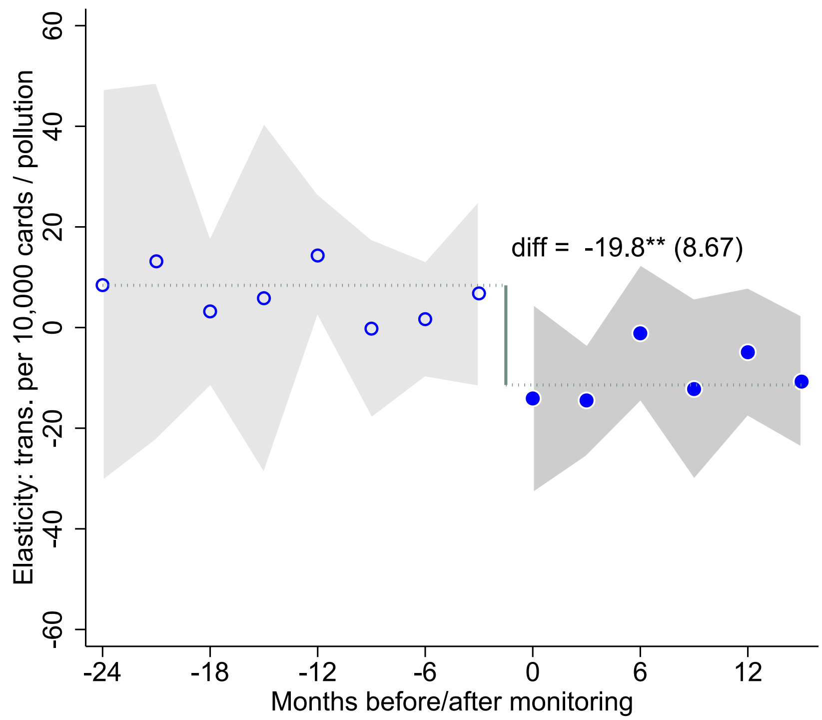
(a) Total air purifier sales



(b) Air purifier sales before and after monitoring

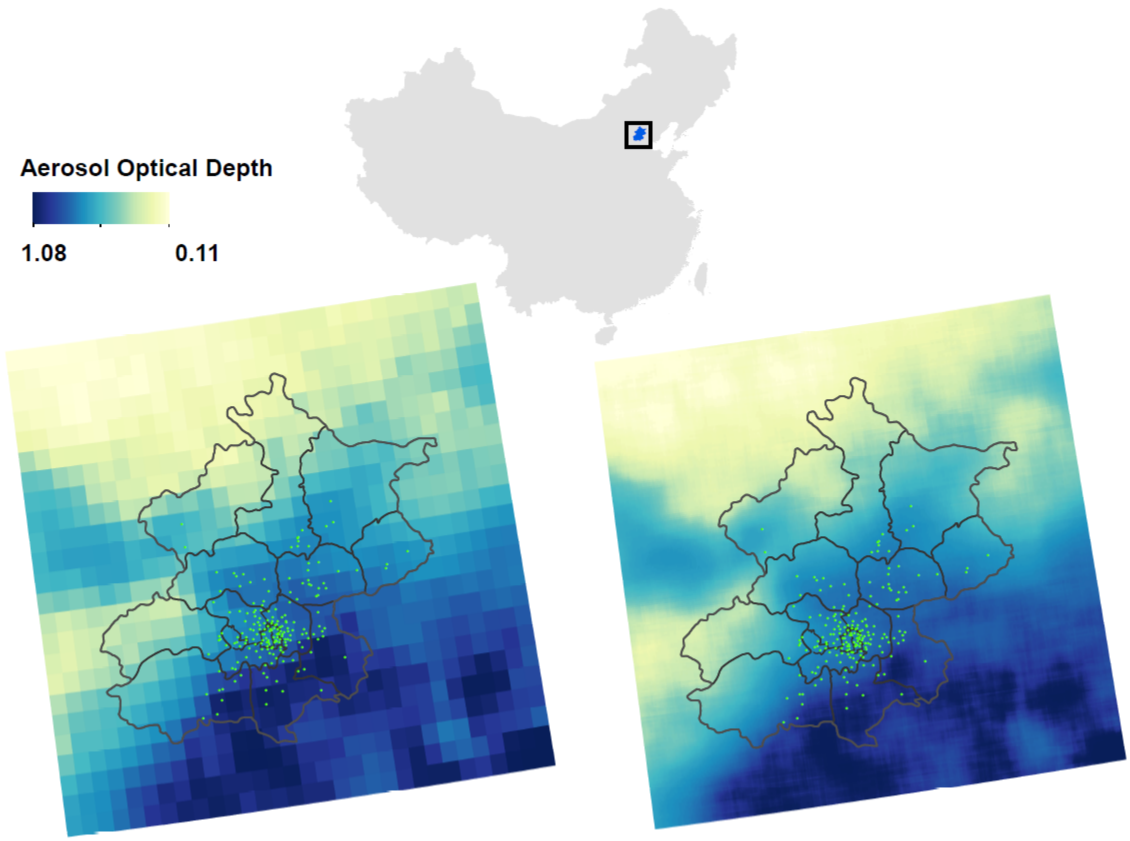
Notes: Panel A plots raw monthly trends in total air purifier sales from offline venues. The graph omits two dots with exceptionally high sales for readability purpose. These dots correspond to December 2013 (sales = 61,605 units) and Decembet 2015 (sales = 74,352 units). Line shows annual averages. Panel B plots log per capita air purifier sales as a function of months since monitoring initiation. Event month -1 is normalized to 0. The underlying regression controls for month-of-year and year indicators. Line shows quarterly averages.

Figure 7: Changes in Weekly Bank Card Transaction-Pollution Gradient



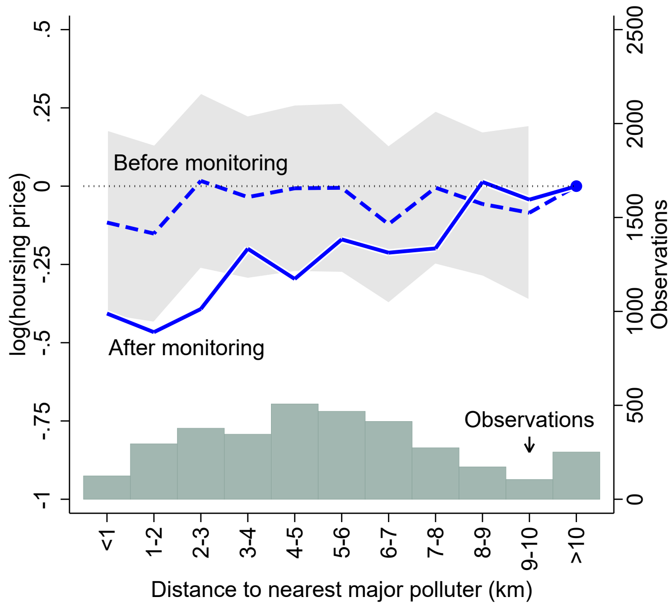
Notes: This graph shows the relationship between weekly bank card transcation rate and log satellite-based pollution as a function of time since monitoring initiation. The regression controls for prefectiry-city FEs, week-of-year FEs, and year FEs. Regressions are weighted by the number of active cards in the city. Shaded region shows 95% confidence interval constructed from standard errors clustered at the prefecture-city level. Number of observations = 83,122.

Figure 8: Original (10km) vs. Oversampled (1km) AOD, Beijing 2006-2014 Average

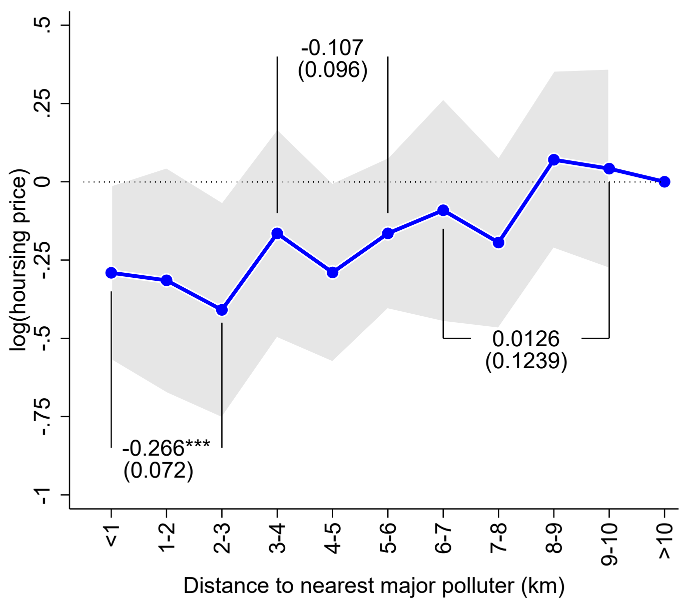


Notes: This map shows 2006-2014 average aerosol optical depth (AOD) level for the municipality of Beijing. Left panel shows MODIS AOD at the original 10×10km resolution. Right panel shows AOD oversampled to 1×1km resolution. Dots show centroid locations of communities (i.e., "jiedao") in the housing transaction data.

Figure 9: Changes in Annual Housing Prices-Distance to Polluter Gradient, Beijing



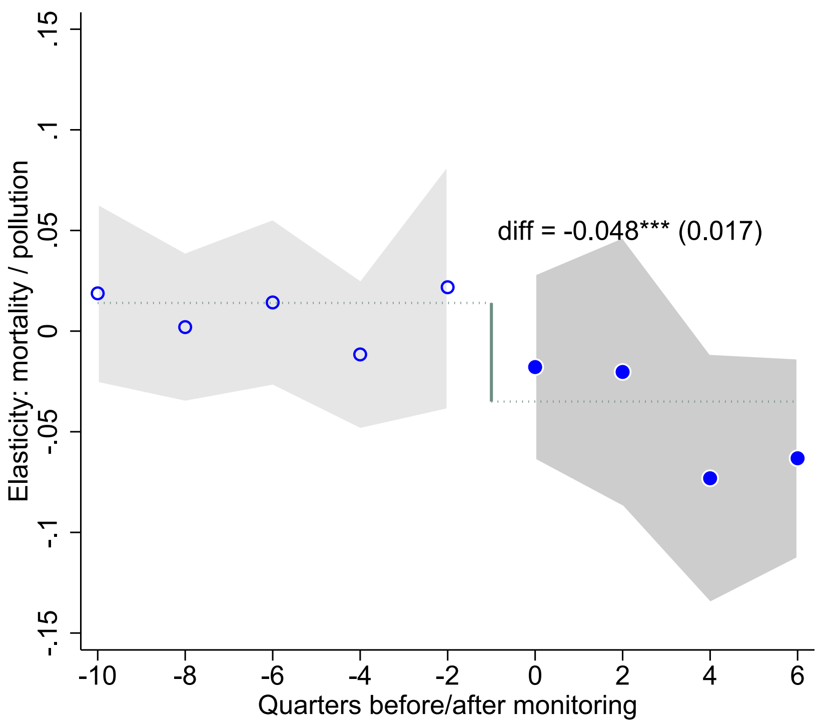
(a) Before vs. after monitoring



(b) Difference estimates

Notes: This graph shows coefficients from a regression of complex×annual log housing prices on distance (in 1-km bins) to nearest major polluter before and after January 2013 when Beijing initiated ambient pollution monitoring. In panel A, estimations are done separately for time before (dashed line) and after (solid line) monitoring began, with prices normalized to 0 for the >10-km bin. The histogram (right axis) plots total number of observations by distance bins. In panel B, the difference estimation pools before/after samples. All regressions control for district×year FEs, community FEs, and years-on-market FEs. Shaded region shows 95% confidence interval constructed from standard errors two-way clustered at the community level and the district×year level. Number of observations = 3,827.

Figure 10: Changes in Quarterly Mortality-Pollution Gradient



Notes: This graph shows coefficients from a regression of log mortality rate on log satellite-based pollution as a function of quarters since monitoring initiation. The (-10 to 6) month event window is chosen so that the underlying sample is a balanced panel of cities. Coefficients are obtained from a single regression, controlling for prefectiry-city FEs, quarter-of-year FEs, and year FEs. Shaded region shows 95% confidence interval constructed from standard errors clustered at the prefecture-city level. Number of observations = 2,620.

Table 1: Changes in Weekly Bank Card Transaction-Pollution Gradient

Dep. var.: Number of transactions per 10,000 active cards in a city×week				
	(1)	(2)	(3)	(4)
Log(Pollution)	8.39	6.07	7.96	10.3
	(8.19)	(8.78)	(5.75)	(7.20)
Log(Pollution) × 1(after monitoring)	-19.8**	-22.8**	-19.4**	-25.1**
	(8.67)	(10.8)	(7.77)	(10.1)
FEs: city	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region×year			✓	
FEs: region×week-of-sample				✓
N	83,122	83,122	83,122	83,122

Notes: "Log(Pollution)" is logged AOD in the city×week. Mean of dependent variable is 869.1 transactions per week per 10,000 cards. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Centralsouth (81 cities), Southwest (54 cities), Northwest (52 cities). Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table 2: Changes in Beijing's Housing Price-Pollution Gradient

Dep. var.: Log housing price index in a complex×year				
	(1)	(2)	(3)	(4)
Identifying variation:	Within complex across years		Within district×year across communities	
Log(pollution)	0.090	0.063	0.009	-0.103
	(0.104)	(0.121)	(0.239)	(0.244)
Log(lagged pollution)		0.034		0.335
		(0.124)		(0.216)
Log(pollution)×1(after 2013)	-0.591*	-0.730*	-0.850*	-0.753*
	(0.299)	(0.434)	(0.436)	(0.432)
Log(lagged pollution)×1(after 2013)		-0.377		-0.216
		(0.490)		(0.754)
FEs: complex	✓	✓		
FEs: year	✓	✓		
FEs: years on-market	✓	✓	✓	✓
FEs: community			✓	✓
FEs: district×year			✓	✓
N	3,372	2,715	3,827	3,266
N (complex)	988	801	1,224	1,129
N (community)	179	167	180	172
N (district)	16	16	16	16

Notes: A complex is a real estate project site that often contains multiple buildings. The dependent variable is logged nominal housing price adjusted for quadratic floor size, floor indicators, and sale month-of-year indicators. "Log(pollution)" is logged AOD level at the (oversampled) 1km resolution corresponding to the complex's geographic coordinates. Standard errors are two-way clustered at the community (i.e., "jiedao") level and the district×year level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table 3: Changes in Quarterly Mortality-Pollution Gradient

Dep. var.: Log mortality rate in a city×quarter				
	(1)	(2)	(3)	(4)
Log(Pollution)	0.014	0.034	0.039*	0.041*
	(0.019)	(0.021)	(0.020)	(0.023)
Log(Pollution) × 1(after monitoring)	-0.048***	-0.055***	-0.055***	-0.046**
	(0.017)	(0.020)	(0.021)	(0.021)
FEs: city	✓	✓	✓	✓
FEs: quarter-of-year	✓			
FEs: year	✓			
FEs: quarter-of-sample		✓	✓	
FEs: region×year			✓	
FEs: region×quarter-of-sample				✓
N	2,620	2,620	2,620	2,620

Notes: "Log(Pollution)" is logged AOD in the city×quarter. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Centralsouth (81 cities), Southwest (54 cities), Northwest (52 cities). Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table 4: Changes in Mortality-Pollution Gradient: Heterogeneity by City Characteristics

City characteristics:	(1) Per cap. income	(2) Frac. urban	(3) Per cap. hospitals	(4) Per cap. residential electricity	(5) Per cap. mobile phones
Panel A. Dep. var. = Number of transactions per 10,000 active cards in a city×week					
Log(Pollution) × 1(after monitoring)	-13.0*	-14.2	-15.1*	-20.2**	-0.191
× 1(below average)	(7.48)	(9.10)	(9.11)	(8.20)	(7.05)
Log(Pollution) × 1(after monitoring)	-25.5**	-25.8**	-33.1***	-22.6**	-35.1***
× 1(above average)	(12.0)	(10.3)	(10.6)	(10.6)	(11.6)
Equality $p-$value	0.354	0.340	0.175	0.859	0.006
N	66,854	66,854	67,046	64,540	67,046
67,046					
Panel B. Dep. var. = Log mortality rate in a city×quarter					
Log(Pollution) × 1(after monitoring)	-0.052**	-0.032	-0.036	-0.015	-0.025
× 1(below average)	(0.023)	(0.020)	(0.026)	(0.025)	(0.020)
Log(Pollution) × 1(after monitoring)	-0.046	-0.081***	-0.066***	-0.073*	-0.080**
× 1(above average)	(0.029)	(0.030)	(0.021)	(0.044)	(0.035)
Equality $p-$value	0.888	0.139	0.348	0.246	0.145
N	2,560	2,220	2,220	2,120	2,220

Notes: This table reports heterogeneous purchase-pollution (panel A) and mortality-pollution (panel B) gradient changes by above/below average city characteristics. Each column corresponds to a separate regression: column 1 = per capita personal dispensable income; column 2 = share of urban population; column 3 = per capita number of hospitals; column 4 = per capita residential electricity use; column 5 = share of mobile phone users. All city characteristics are computed as 2011-2015 averages. "Equality $p-$value" tests for equality across the above/below average coefficients. All regressions control for city, month-of-sample, and region-by-year fixed effects. All regressions include full sets of lower-order interaction terms which are not reported in the table in the interest of space. Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Appendices. For Online Publication Only

Appendix A: Proof of Proposition 1

Individuals choose optimal consumption x and defensive investment a to maximize utility under the perceived pollution level c_0 as described in Section 3.1. The Lagrangian equation is:

$$L = U(x, h(c_0, a)) + \lambda \left[I + w * g(h(c_0, a)) - x - p_a * a\right]$$

where λ is the Lagrange multiplier and denotes the marginal utility per dollar. The first order conditions are:

$$\begin{aligned}
&\frac{\partial L}{\partial x} = 0 \Rightarrow U_x - \lambda = 0 \\
&\frac{\partial L}{\partial a} = 0 \Rightarrow (U_h + \lambda * w * g_h)\,\frac{\partial h(c_0, a)}{\partial a} - \lambda p_a = 0 \\
&\frac{\partial L}{\partial \lambda} = 0 \Rightarrow I + w * g(h) - x - p_a * a = 0
\end{aligned} \tag{A.1}$$

where U_x, U_h, and g_h denote partial derivatives. We first show that under Assumptions 1-3, optimal avoidance (weakly) increases in perceived pollution:

$$\frac{da}{dc} \geq 0.$$

Let f denote the first order condition w.r.t avoidance (equation A.1):

$$f = (U_h + \lambda * w * g_h)\,\frac{\partial h}{\partial a} - \lambda p_a = 0$$

Applying the implicit function theorem to f, we obtain:

$$\begin{aligned}\frac{da}{dc} &= -\frac{\partial f/\partial c}{\partial f/\partial a} = -\frac{[U_{hh} + \lambda * w * g_{hh}] * \frac{\partial h}{\partial c} * \frac{\partial h}{\partial a} + (U_h + \lambda * w * g_h) * \frac{\partial^2 h}{\partial a \partial c}}{(U_{hh} + \lambda * w * g_{hh}) * \left(\frac{\partial h}{\partial a}\right)^2 + (U_h + \lambda * w * g_h) * \frac{\partial^2 h}{\partial a^2}} \\ &= -\frac{A+B}{C+D}\end{aligned}$$

where U_{hx}, U_{hh}, g_{hh} are second order derivatives. Under the assumption of diminishing marginal utility, decreasing marginal labor product of health, and decreasing health benefit of avoidance, $C + D \leq 0$.[30] Similarly, $A + B \geq 0$. Hence, avoidance increases weakly in (perceived) pollution. The key assumption for this result is $dh^2/dadc > 0$. When pollution deteriorates, avoidance restores health more effectively (that is, the marginal benefit of avoidance is large with bad pollution). After the information program, individuals observe the actual pollution c, which is higher than previously perceived level: c_0. The above analysis indicates that individuals would increase the level of avoidance post the policy intervention:

$$a(c) \geq a(c_0).$$

As the marginal health benefit of avoidance is positive from Assumption (A1) in Section 3.1, the health condition improves with avoidance:

$$h(c, a(c)) \geq h(c, a(c_0)).$$

Due to the lack of real-time information on pollution prior to the information program, perceived pollution c_0 is unlikely to respond to day-to-day changes in the actual pollution. Hence, the total derivative of health w.r.t. pollution is:

$$\frac{dh}{dc}|_{c_0} = \frac{\partial h}{\partial c} + \frac{\partial h}{\partial a} * \frac{da}{dc_0} * \frac{dc_0}{dc} = \frac{\partial h}{\partial c}$$

[30] At the optimal a and X, $U_h + \lambda * w * g(h) > 0$ by construction. In addition, $U_{hh}, g_{hh}, \partial^2 h/\partial a^2 < 0$. Another way to show $C + D \leq 0$ is that this is the second order condition for the optimal a.

where the second equation follows from the fact that $dc_0/dc = 0$. Post the information program, the perceived pollution is equal to the actual pollution and individuals can engage in effective avoidance to moderate the negative impact of pollution. The total derivative of health w.r.t. pollution is:

$$\frac{dh}{dc}|_c = \frac{\partial h}{\partial c} + \frac{\partial h}{\partial a} * \frac{da}{dc} \geq \frac{\partial h}{\partial c}$$

The second line follows from the fact that avoidance increases in (perceived) pollution and improves the health stock.

Lastly, let $V(c, c)$ denote the indirect utility when individuals accurately perceive pollution $c_0 = c$. In that case, the experience utility and decision utility coincides. $V(c, c_0)$ is the utility achieved by maximizing the decision utility under perceived pollution of c_0. Since utility is maximized under full information, we have:

$$V(c, c) \geq V(c, c_0).$$

Putting these together, we derive the following predictions of the information program:

- Avoidance behavior increases after the program: $a(c) > a(c_0)$
- Health improves and the (downward slopping) health-pollution response curve flattens:

$$h(c, a(c)) > h(c, a(c_0)), \frac{dh}{dc}|_{c_0=c} \geq \frac{dh}{dc}_{c_0<c}$$

- Individual utility increases: $V(c, c) > V(c, c_0)$

Appendix B: A Simple Example

Suppose the health production function is:

$$h(c,a) = \max\{0, h_0 - \frac{c}{a}\}, \quad h_c > 0, c > 0, a > 0$$

where h_0 is the initial health stock with zero pollution, c is pollution and a is avoidance. Assume that the labor production function is an identify function: $g(h) = h$. The budget constraint is:

$$I + w * h(c,a) \geq x + p_a * a$$

Suppose the utility function is:

$$U(x,h) = x$$

These functions satisfy Assumptions 1 and 2. In this simple example, maximizing utility subject to the budget constraint is equivalent to choosing a to maximize the amount of numeraire that can be afforded by the budget constraint:

$$\max_a I + w * h(c,a) - p_a * a$$

The optimal avoidance a satisfies:

$$a^* = \min\{\sqrt{\frac{w * c}{p_a}}, \frac{w * h_0}{p_a}\}$$

which (weakly) increases in c. Optimal health stock $h = \max\{0, h_0 - \sqrt{\frac{c * p_a}{w}}\}$ decreases in c and reaches the minimum of zero when pollution exceeds $\frac{w * h_0^2}{p_a}$. Consumption of the numeraire good and utility also decrease with c. When pollution exceeds $\frac{w * h_0^2}{p_a}$, the health stock is at the minimum, avoidance stops increasing in c and is kept at $\frac{w * h_0}{p_a}$. Consumption of the numeraire good also reaches the minimum level of $I - \sqrt{w * p_a}$. Finally, to ensure an interior solution, we need $I \geq \sqrt{w * p_a}$.

Appendix C: Identifying the Slope Change

To examine the impact of the information program om the pollution-outcome relationship, our analysis uses the following framework (simplifying notations in equation 2):

$$\mathrm{y}_{ct} = \alpha \times \mathrm{p}_{ct} + \beta \times \mathrm{p}_{ct} \times \mathrm{d}_{ct} + \mathrm{x}'_{ct}\gamma + \varepsilon_{ct}, \tag{C.2}$$

where c denotes a city and t denotes time (e.g., day or week). y_{ct} is the outcome variable. p_{ct} measures ambient air quality and could be correlated with the error term due to unobservables or measurement error as discussed in the main text. d_{ct} represents the treatment dummy and it turns to one in three waves across cities based on the staggered roll-out schedule. x_{ct} includes a set of controls including weather conditions and rich spatial and temporal fixed effects such as city fixed effects and time fixed effects. ε_{ct} is the error term. The key parameter of interest is β, the change in the slope of pollution-outcome relationship.

Although p_{ct} could be endogenous due to unobservables as discussed in the main text, the OLS estimate of β is consistent under the follow two assumptions.

Assumption (B1): $E\left(\varepsilon_{ct}|\mathrm{d}_{ct}, \mathrm{x}_{ct}\right) = 0.$

Assumption (B2): Pollution $\mathrm{p}_{ct} = \mathrm{x}'_{ct}\theta + \nu_{ct}$, where $E\left(\upsilon_{ct}|\mathrm{x}_{ct}\right) = 0$, and $E\left(\varepsilon_{ct}\nu_{ct}\right) = \sigma_{\varepsilon\nu}$.

Proof: We denote the set of regressors to be w'_{ct} as three blocks: $\left(\mathrm{x}'_{ct}, \mathrm{p}_{ct}, \mathrm{p}_{ct}\mathrm{d}_{ct}\right)$, where x_{ct} is a k by 1 vector of controls and w_{ct} is a (k+2) by 1 vector. Denote the vector of all parameters as η. For the OLS estimate $\hat{\eta}$: $\mathrm{plim}\hat{\eta} = \eta + [E(\mathrm{w}_{ct}\mathrm{w}'_{ct}]^{-1}[E(\mathrm{w}_{ct}\varepsilon_{ct})$.

The focus of our empirical analysis is β and in order to show that $\hat{\beta}$ is consistent, it suffices to show that the element in $[E(\mathrm{w}_{ct}\mathrm{w}'_{ct}]^{-1}[E(\mathrm{w}_{ct}\varepsilon_{ct})$ corresponding to the coefficient for $d_i p_i$ is zero.

$$E(\mathrm{w}_{ct}\varepsilon_{ct}) = \begin{pmatrix} 0 \\ E\left(\mathrm{p}_{ct}\varepsilon_{ct}\right) \\ E\left(\mathrm{p}_{ct}\mathrm{d}_{ct}\varepsilon_{ct}\right) \end{pmatrix}, E(\mathrm{w}_{ct}\mathrm{w}_{ct}') = \begin{pmatrix} E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\right) & E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\right) & E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\mathrm{d}_{ct}\right) \\ E\left(\mathrm{p}_{ct}\mathrm{x}_{ct}'\right) & E\left(\mathrm{p}_{ct}^2\right) & E\left(\mathrm{p}_{ct}^2\mathrm{d}_{ct}\right) \\ E\left(\mathrm{p}_{ct}d_i\mathrm{x}_{ct}'\right) & E\left(\mathrm{p}_{ct}^2\mathrm{d}_{ct}\right) & E\left(\mathrm{p}_{ct}^2\mathrm{d}_{ct}^2\right) \end{pmatrix}.$$

$$E(\mathrm{w}_{ct}\mathrm{w}_{ct}')^{-1} = \frac{adj\left(E(\mathrm{w}_{ct}\mathrm{w}_{ct}')\right)}{\det(E(\mathrm{w}_{ct}\mathrm{w}_{ct}'))},$$

where $adj\left(E(\mathrm{w}_{ct}\mathrm{w}_{ct}')\right)$ is the transpose matrix of cofactors of $E(\mathrm{w}_{ct}\mathrm{w}_{ct}')$. To consider the OLS estimate of β, we only need to examine the cofactors corresponding to third column of the $E(\mathrm{w}_{ct}\mathrm{w}_{ct}')$. In particular,

$$\mathrm{plim}\beta - \beta = \frac{c_{23} * E\left(\mathrm{p}_{ct}\varepsilon_{ct}\right) + c_{33} * E\left(\mathrm{p}_{ct}\mathrm{d}_{ct}\varepsilon_{ct}\right)}{\det(E(\mathrm{w}_{ct}\mathrm{w}_{ct}'))}. \tag{C.3}$$

We now examine each component in the numerator of equation (C.3),

$$\begin{aligned} c_{23} &= -\det\begin{bmatrix} E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\right)_{k\times k} & E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\mathrm{d}_{ct}\right)_{k\times 1} \\ E\left(\mathrm{p}_{ct}\mathrm{x}_{ct}'\right)_{1\times k} & E\left(\mathrm{p}_{ct}^2\mathrm{d}_{ct}\right)_{1\times 1} \end{bmatrix} \\ &= -\det\left(E\left(\mathrm{p}_{ct}^2\mathrm{d}_{ct}\right) - E\left(\mathrm{p}_{ct}\mathrm{x}_{ct}'\right)E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\right)^{-1}E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\mathrm{d}_{ct}\right)\right)\det\left(E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\right)\right) \\ c_{33} &= \det\begin{bmatrix} E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\right)_{k\times k} & E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\right)_{k\times 1} \\ E\left(\mathrm{p}_{ct}\mathrm{x}_{ct}'\right)_{1\times k} & E\left(\mathrm{p}_{ct}^2\right)_{1\times 1} \end{bmatrix} \\ &= \det\left(E\left(\mathrm{p}_{ct}^2\right) - E\left(\mathrm{p}_{ct}\mathrm{x}_{ct}'\right)E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\right)^{-1}E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\right)\right)\det\left(E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\right)\right) \end{aligned}$$

$$
\begin{aligned}
E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\mathrm{d}_{ct}\right) &= E\left(\mathrm{x}_{ct}E\left(\mathrm{p}_{ct}\mathrm{d}_{ct}|\mathrm{x}_{ct}\right)\right) = E\left(\mathrm{x}_{ct}E\left(\mathrm{p}_{ct}|\mathrm{x}_{ct}\right)E\left(\mathrm{d}_{ct}|\mathrm{x}_{ct}\right)\right) = E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'E\left(\mathrm{d}_{ct}|\mathrm{x}_{ct}\right)\right)\theta \\
E\left(\mathrm{p}_{ct}\mathrm{x}_{ct}'\right) &= E\left(E\left(\mathrm{p}_{ct}|\mathrm{x}_{ct}\right)\mathrm{x}_{ct}'\right) = E\left(\theta'\mathrm{x}_{ct}\mathrm{x}_{ct}'\right) \\
E\left(\mathrm{x}_{ct}\mathrm{p}_{ct}\right) &= E\left(\mathrm{x}_{ct}E\left(\mathrm{p}_{ct}|\mathrm{x}_{ct}\right)\right) = E\left(\mathrm{x}_{ct}\mathrm{x}_{ct}'\theta\right) \\
E\left(\mathrm{p}_{ct}^2\mathrm{d}_{ct}\right) &= E\left(\left(\mathrm{x}_{ct}'\theta+\varepsilon_{ct}\right)^2\mathrm{d}_{ct}\right) = E\left(\mathrm{x}_{ct}'\theta\theta'\mathrm{x}_{ct}E\left(\mathrm{d}_{ct}|\mathrm{x}_{ct}\right)\right)+\sigma_\varepsilon^2E\left(\mathrm{d}_{ct}\right) \\
E\left(\mathrm{p}_{ct}^2\right) &= E\left(\left(\mathrm{x}_{ct}'\theta+\varepsilon_{ct}\right)^2\right) = E\left(\mathrm{x}_{ct}'\theta\theta'\mathrm{x}_{ct}\right)+\sigma_\varepsilon^2.
\end{aligned}
$$

Dropping the subscript for simplicity,

$$
\begin{aligned}
&E\left(\mathrm{p}^2\mathrm{d}\right)-E\left(\mathrm{px}'\right)E\left(\mathrm{xx}'\right)^{-1}E\left(\mathrm{xpd}\right) = E\left(\mathrm{x}'\theta\theta'\mathrm{x}E\left(\mathrm{d}|\mathrm{x}\right)\right)+\sigma_\varepsilon^2E\left(\mathrm{d}\right) \\
&-\theta'E\left(\mathrm{xx}'\right)E\left(\mathrm{xx}'\right)^{-1}E\left(\mathrm{xx}'E\left(\mathrm{d}|\mathrm{x}\right)\right)\theta = \sigma_\varepsilon^2E\left(\mathrm{d}\right).
\end{aligned}
$$

The last equality follows from the fact that $\mathrm{x}'\theta\theta'x$ is a scalar and equal to $\theta'\mathrm{xx}'\theta$.

$$
\begin{aligned}
E\left(\mathrm{p}^2\right)-E\left(\mathrm{px}'\right)E\left(\mathrm{xx}'\right)^{-1}E\left(\mathrm{xp}\right) &= E\left(\mathrm{x}'\theta\theta'\mathrm{x}\right)+\sigma_\varepsilon^2-\theta'E\left(\mathrm{xx}'\right)E\left(\mathrm{xx}'\right)^{-1}E\left(\mathrm{xx}'\right)\theta \\
&= \sigma_\varepsilon^2.
\end{aligned}
$$

Therefore, $c_{23}=-E\left(\mathrm{d}\right)\sigma_\varepsilon^2\det\left(E\left(\mathrm{xx}'\right)\right)$, and $c_{33}=\sigma_\varepsilon^2\det\left(E\left(\mathrm{xx}'\right)\right)$.

$$
\begin{aligned}
E\left(\mathrm{p}\varepsilon\right) &= E\left(\left(\mathrm{x}'\theta+v\right)\varepsilon\right) = \sigma_{ve} \\
E\left(\mathrm{pd}\varepsilon\right) &= E\left(\left(\mathrm{x}'\theta+v\right)\mathrm{d}\varepsilon\right) = E\left(\mathrm{x}'\mathrm{d}\varepsilon\right)\theta+E\left(v\mathrm{d}\varepsilon\right) \\
&= E\left[\mathrm{x}'\mathrm{d}E\left(\varepsilon|\mathrm{x},\mathrm{d}\right)\right]\theta+E\left[\mathrm{d}E\left(v\varepsilon|\mathrm{d}\right)\right] = \sigma_{ve}E\left(\mathrm{d}\right).
\end{aligned}
$$

Collecting terms, the consistency of OLS estimate of β follows:

$$
\operatorname{plim}\hat{\beta}-\beta = c_{23}E\left(\mathrm{p}\varepsilon\right)+c_{33}E\left(\mathrm{pd}\varepsilon\right) = \sigma_{ve}\left(c_{23}+c_{33}E\left(\mathrm{d}\right)\right) = 0.
$$

Appendix D: Figures and Tables

Figure D.1: List of Cities by Roll-out Waves and by Associated City Clusters

Wave 1 cities		
Beijing	Xining	Taizhou
Tianjin	Hefei	Lanzhou
Shijiazhuang	Fuzhou	Hangzhou
Tangshan	Yinchuan	Ningbo
Qinhuangdao	Wulumuqi	Xi'an
Handan	Jinan	Jiaxing
Xingtai	Nantong	Huzhou
Baoding	Zhengzhou	Shaoxing
Zhangjiakou	Wuhan	Jinhua
Chengde	Changsha	Lasa
Cangzhou	Guangzhou	Zhoushan
Langfang	Shenzhen	Taizhou
Hengshui	Zhuhai	Kunming
Taiyuan	Foshan	Xiamen
Huhehaote	Jiangmen	Nanchang
Shenyang	Zhaoqing	Wenzhou
Yangzhou	Huizhou	Qingdao
Changchun	Dongwan	Dalian
Haerbin	Zhongshan	Lianyungang
Shanghai	Nanning	Huaian
Nanjing	Haikou	Xuzhou
Wuxi	Chongqing	Quzhou
Yancheng	Chengdu	Suqian
Changzhou	Guiyang	Lishui
Suzhou	Zhenjiang	

Wave 2 cities				
Wuhu	Jinzhou	Jimo	Wujiang	Yingkou
Maanshan	Zhuzhou	Pingdu	Changshu	Panjin
Datong	Xiangtan	Laixi	Zhangjiagang	Huludao
Yangquan	Yueyang	Zibo	Kunshan	Zigong
Changzhi	Changde	Zaozhuang	Taicang	Zhuji
Linfen	Zhangjiajie	Dongying	Haimen	Jiayuguan
Baotou	Shaoguan	Yantai	Jurong	Deyang
Chifeng	Shantou	Laizhou	Fuyang	Laiwu
Anshan	Zhanjiang	Penglai	Lin'an	Dezhou
Fushun	Pingdingshan	Zhaoyuan	Jiaozhou	Binzhou
Benxi	Anyang	Weifang	Yiwu	Heze
Yan'an	Jiaozuo	Shouguang	Jiujiang	Sanmenxia
Jinzhou	Jinchang	Jining	Quanzhou	Weinan
Yichang	Shizuishan	Taian	Eerduosi	Zhangqiu
Baoji	Kelamayi	Weihai	Wafangdian	Nanchong
Xianyang	Kuerle	Wendeng	Maoming	Yuxi
Jilin	Kaifeng	Rongcheng	Meizhou	
Qiqihaer	Luoyang	Rushan	Shanwei	
Daqinf	Liuzhou	Rizhao	Heyuan	
Mudanjiang	Guilin	Zunyi	Yangjiang	
Jiaonan	Beihai	Linyi	Qingyuan	
Jiangyin	Sanya	Qujing	Chaozhou	
Yixing	Tongchuan	Liaocheng	Jieyang	
Liyang	Panzhihua	Mianyang	Yunfu	
Jintan	Luzhou	Yibin	Dandong	

Wave 3 cities							
Tongling	Jixi	Nanping	Ezhou	Guangyuan	Chuxiong	Dingxi	Shihezi
Anqing	Hegang	Longyan	Jinmen	Suining	Honghe	Longnan	Wujiaqu
Chuzhou	Shuangyashan	Ningde	Xiaogan	Neijiang	Wenshan	Linxia	
Chizhou	Yichun	Jingdezhe	Huanggan	Leshan	Xishuangbanna	Gannan	
Xuancheng	jiamusi	Pingxiang	Xianing	Meishan	Dali	Haidong	
Lüliang	Qitaihe	Xinyu	Suizhou	Guangan	Dehong	Haibei	
Wuhai	Heihe	Yingtan	Enshi	Dazhou	Nujiang	Huangnan	
Tongliao	Suihua	Ganzhou	Hengyang	Yaan	Dixing	Hainan	
Hulunbeier	Daxinganling	Ji'an	Shaoyang	Bazhong	Changdou	Guoluo	
Bayannaoer	Bengbu	Yichun	Yiyang	Ziyang	Shannan	Yushu	
Wulanchabu	Huainan	Fuzhou	Chenzhou	Aba	Rikaze	Haixi	
Xingan	Huaibei	Shangrao	Yongzhou	Ganzi	Neiqu	Wuzhong	
Xilinguole	Jincheng	Hebi	Huaihua	Liangshan	Ali	Guyuan	
Alashan	Shuozhou	Xinxiang	Loudi	Liupanshui	Linzhi	Zhongwei	
Fuxin	Huangshan	Puyang	Xiangxi	Anshun	Hanzhong	Tulufan	
Liaoyang	Jinzhong	Xuchang	Wuzhou	Bijie	Yulin	Hami	
Tieling	Fuyang	Luohe	Fangcheng	Tongren	Ankang	Changji	
Chaoyang	Suzhou	Nanyang	Qinzhou	Qianxinan	Shangluo	Boertala	
Siping	Liuan	Shangqiu	Guigang	Qiandongna	Baiyin	Akesu	
Liaoyuan	Haozhou	Xinyang	Yulin	Qiannan	Tianshui	Kezilesu	
Tonghua	Yuncheng	Zhoukou	Baise	Baoshan	Wuwei	Kashi	
Baishan	Xinzhou	Zhumadia	Hezhou	Shaotong	Zhangye	Hetian	
Songyuan	Putian	Huangshi	Hechi	Lijiang	Pingliang	Yili	
Baicheng	Sanming	Shiyan	Laibing	Puer	Jiuquan	Tacheng	
Yanbian	Zhangzhou	Xiangyang	Chongzuo	Lincang	Qingyang	Aletai	

Legend:

Jing-Jin-Ji Metropolitan Region, Yangtze River Delta Economic Zone, Pearl River Delta Metropolitan Region, Direct-administered municipalities, Provincial Capitals

Environmental Improvement Priority Cities (designated 2007), National Environmental Protection Exemplary Cities (awarded between 1997-2012)

Other prefecture-level cities

Notes: The three panels show cities included in each roll-out wave of the information program. Color coding indicates how cities are logistically divided into roll-out waves, according to the 2012 government notice (GB3095-2012).

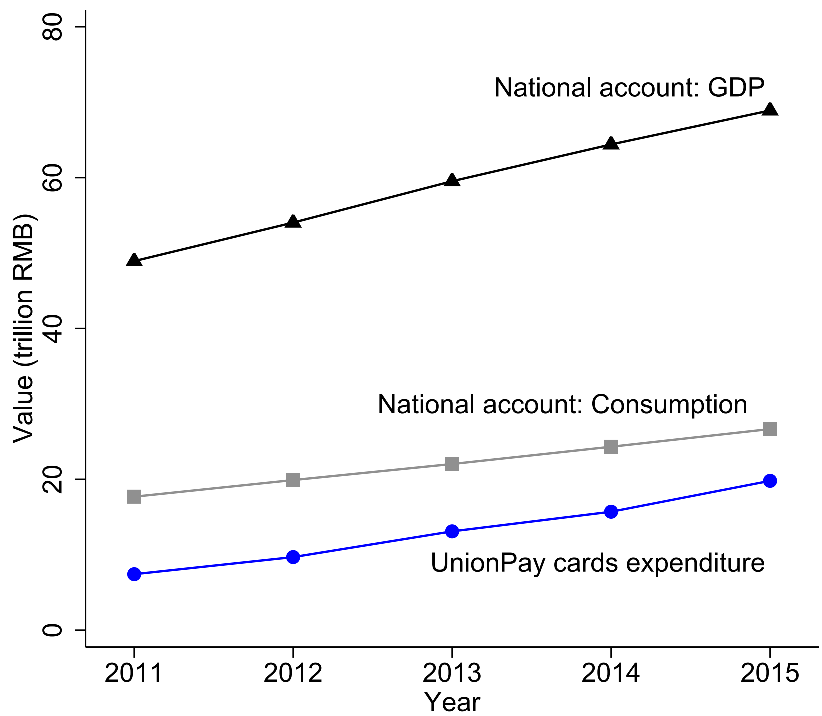
Figure D.2: Screenshot of the Government's Air Quality Disclosure Platform

Notes: This figure shows a screenshot of the Ministry of Environmental Protection's real time air quality disclosure platform as of September 25, 2016.

Figure D.3: Screenshot of an Air Quality App

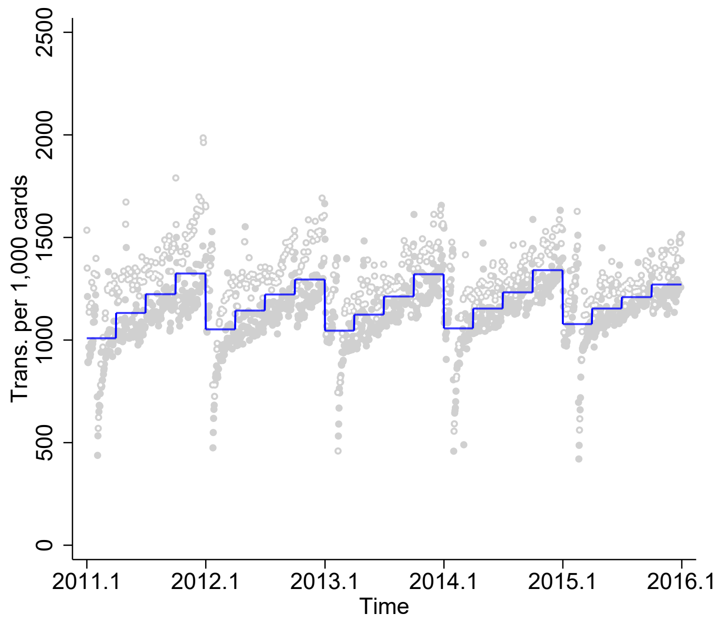
Notes: This figure shows a screenshot of a typical air quality app. The left panel shows the AQI in the city of Shanghai for that hour is 101 ($PM_{2.5}$=75 ug/m3). The right panel shows air quality readings at different locations within Shanghai.

Figure D.4: Consumption Trends: UnionPay vs. National Accounts

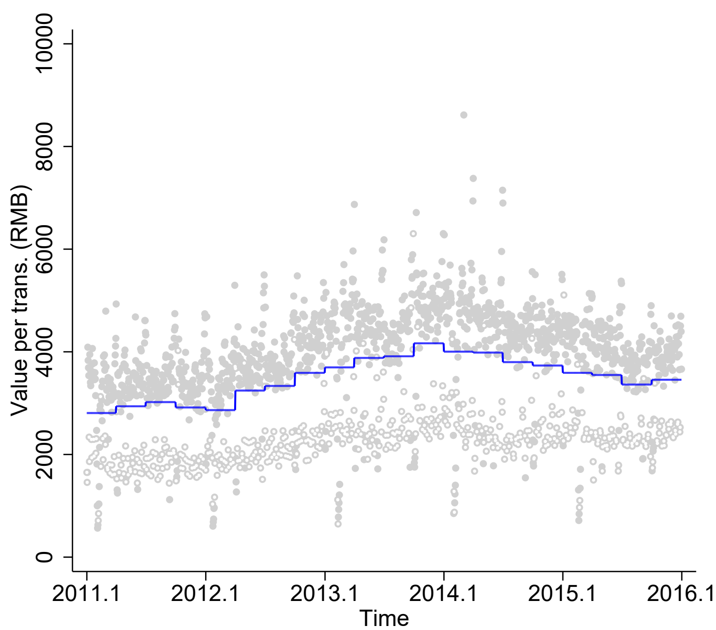


Notes: This figure plots annual GDP (triangles), consumption (squares) reported by the National Bureau of Statistics of China (NBS), and total bank card spendings ×100 (circles) aggregated from the UnionPay 1% bank card data. UnionPay data excludes transactions in the business wholesale categories.

Figure D.5: UnionPay Bank Card Transaction Trends



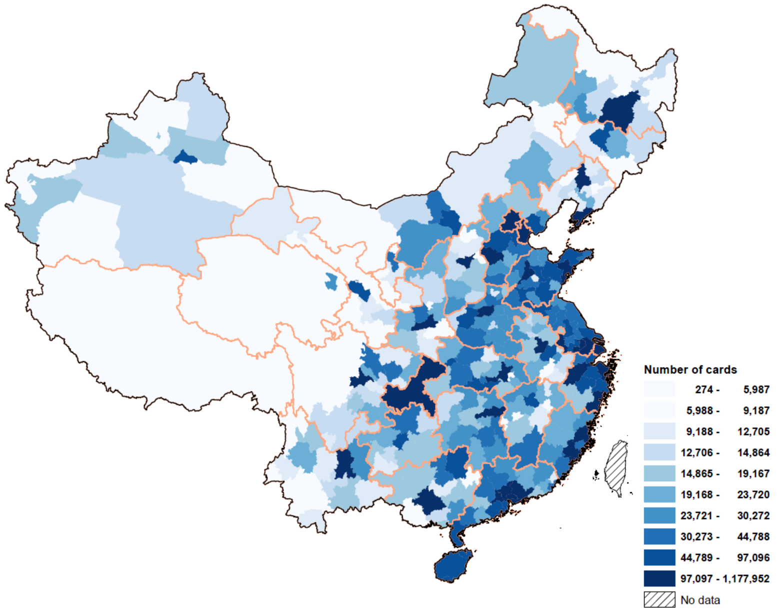
(a) Number of transactions per 100,000 cards



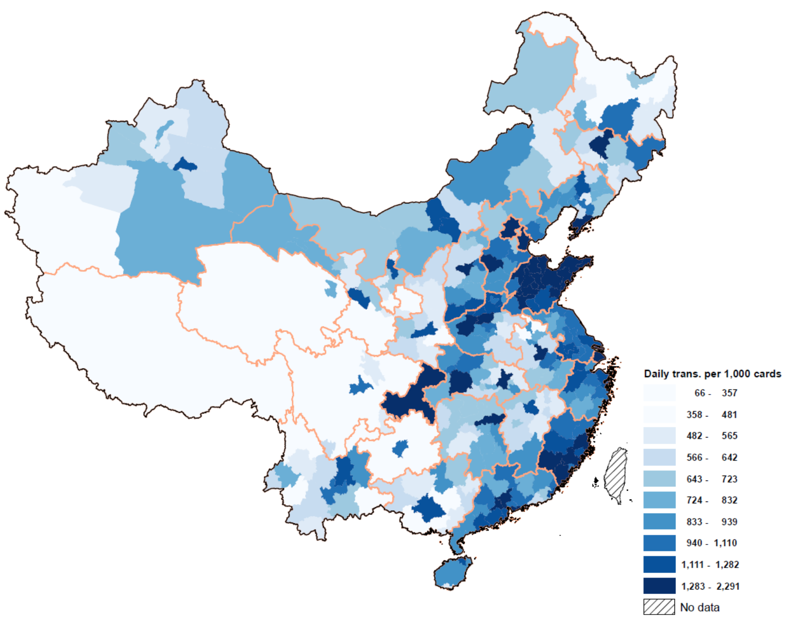
(b) Spending per transaction

Notes: Each dot represents transaction rate (panel A) and spending per transaction (panel B) on a day. Solid dots show weekdays and hollow dots show weekends. Lines show quarterly averages.

Figure D.6: UnionPay Bank Card Transaction by Prefecture-City, 2011-2015 Average



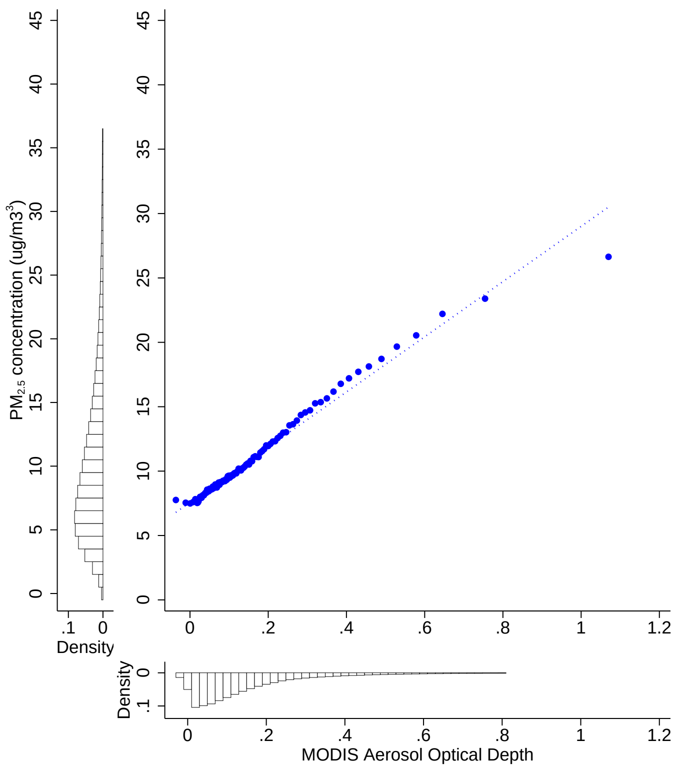
(a) Number of active cards



(b) Number of transactions per 100,000 cards

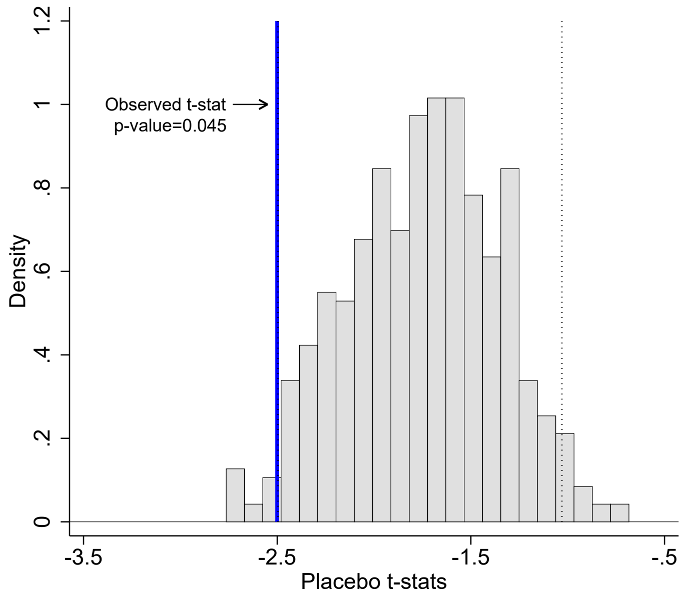
Notes: The maps show 2011-2015 average number of active UnionPay bank cards (panel A) and transactions per 1,000 cards (panel B) at the prefecture-city level. Orange lines show inter-provincial borders.

Figure D.7: Correlation between $PM_{2.5}$ and AOD

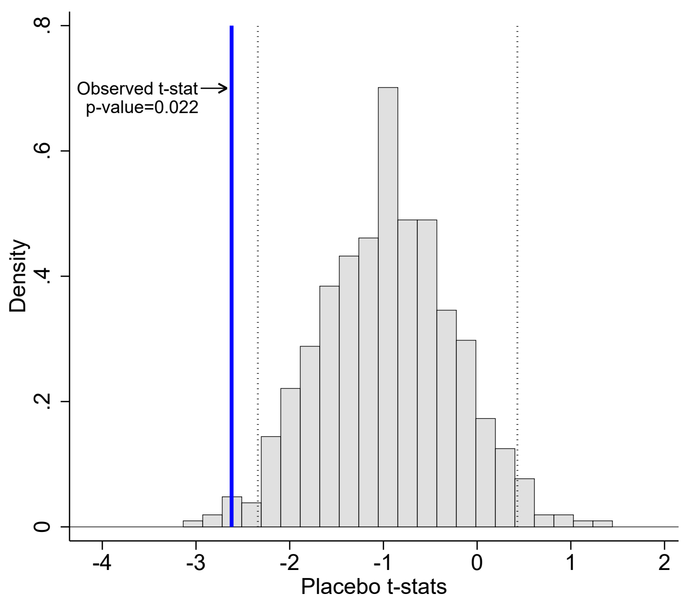


Notes: This graph shows city×day level average $PM_{2.5}$ concentratoin (y-axis) by 100 equal bins of AOD (x-axis), for time periods after monitoring began. Histograms show distribution of the two variables.

Figure D.8: Permutation Tests of the Effect of Monitoring



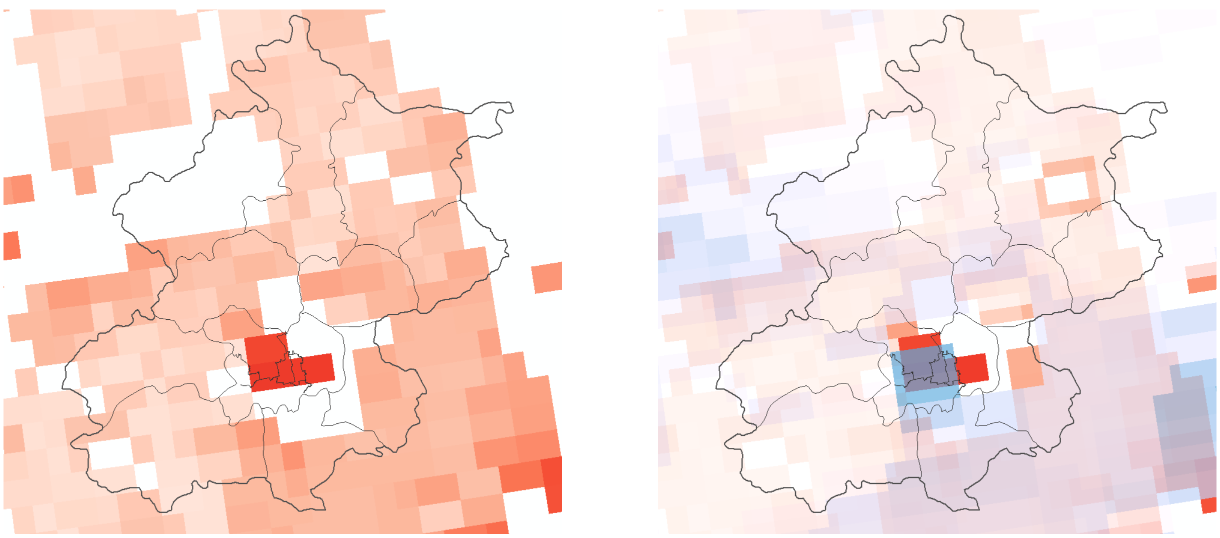
(a) Changes in bank card transaction-pollution gradient



(b) Changes in mortality-pollution gradient

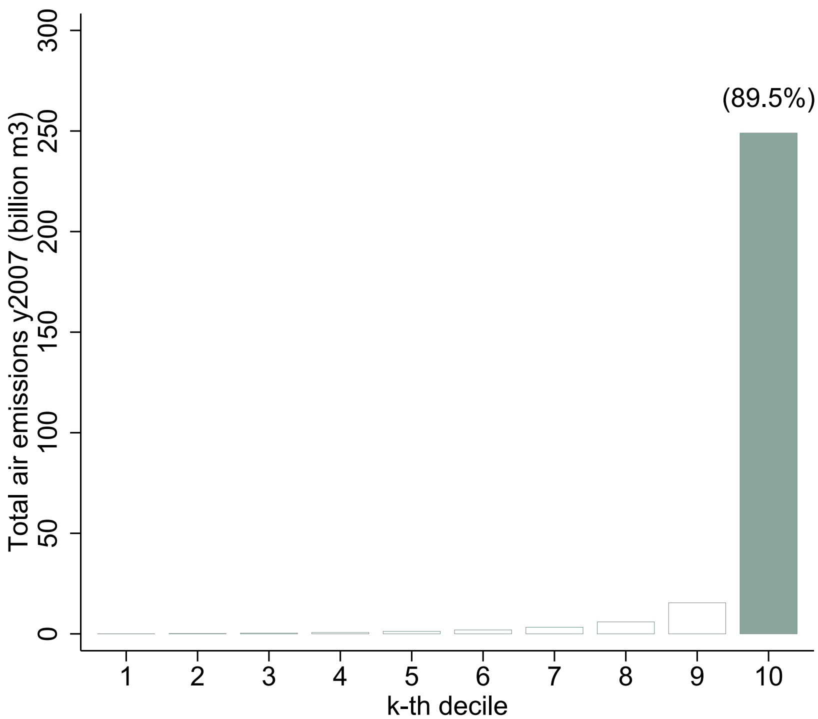
Notes: This plot shows the distribution of t-stats on the "Log(Pollution)$\times\mathbb{1}$(after monitoring)" term across 500 repetitions of random assignment of cities into information roll-out waves. Dashed vertical lines show 95% critical values of the distribution. Solid vertical line shows the observed t-stat from the true city assignment. The regression includes prefecture-city FEs, week(quarter)-of-sample FEs, and region$\times$year FEs. Standard errors are clustered at the city level.

Figure D.9: Illustration: Satellite AOD Oversampling



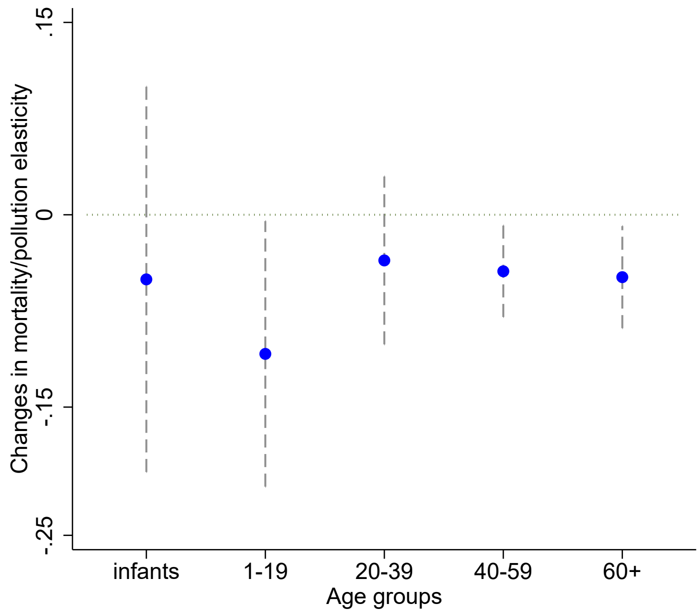
Notes: Left panel shows original MODIS AOD (10×10km) around Beijing on y2008 d243 (i.e., August 30, 2008). Right panel shows an overlay with data on y2008 d244. In both panels, darker colors indicate higher pollution levels.

Figure D.10: Total Air Emissions by Emission Deciles, Beijing Polluter Census 2007

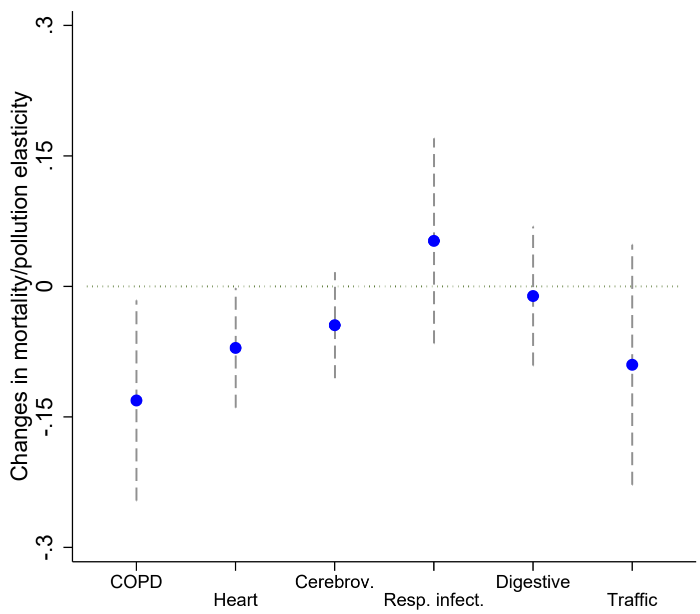


Notes: This graph shows total air emissions (in billion m^3) by Beijing polluters in the k-th decile of annual emission distribution according to the Polluter Census 2007. The sample includes about 440 polluters.

Figure D.11: Heterogeneous Changes in Quarterly Mortality-Pollution Gradient



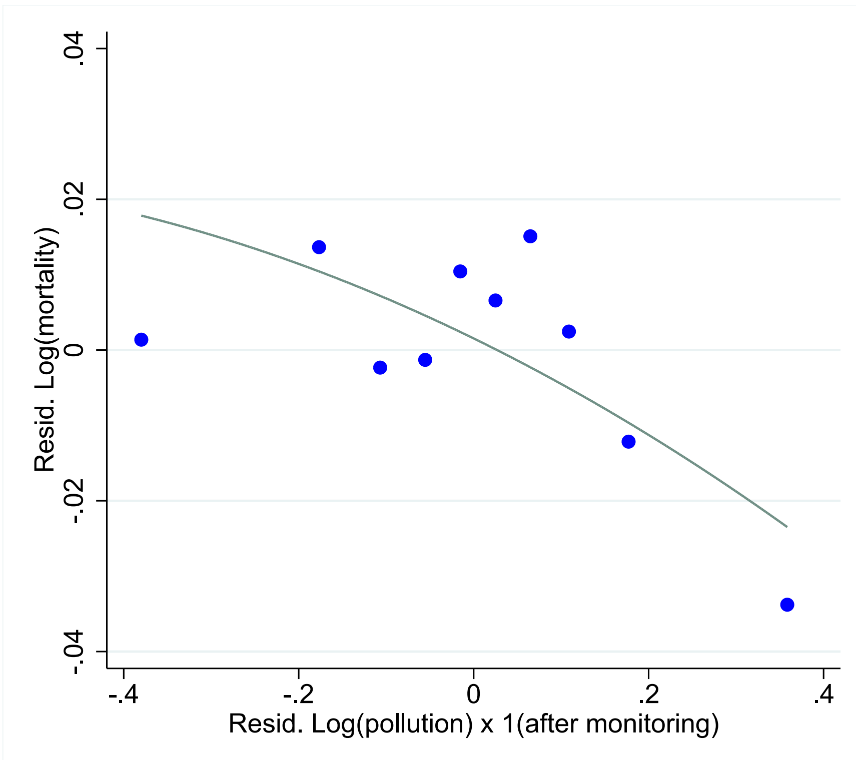
(a) Heterogeneity by age groups



(b) Heterogeneity by causes-of-death

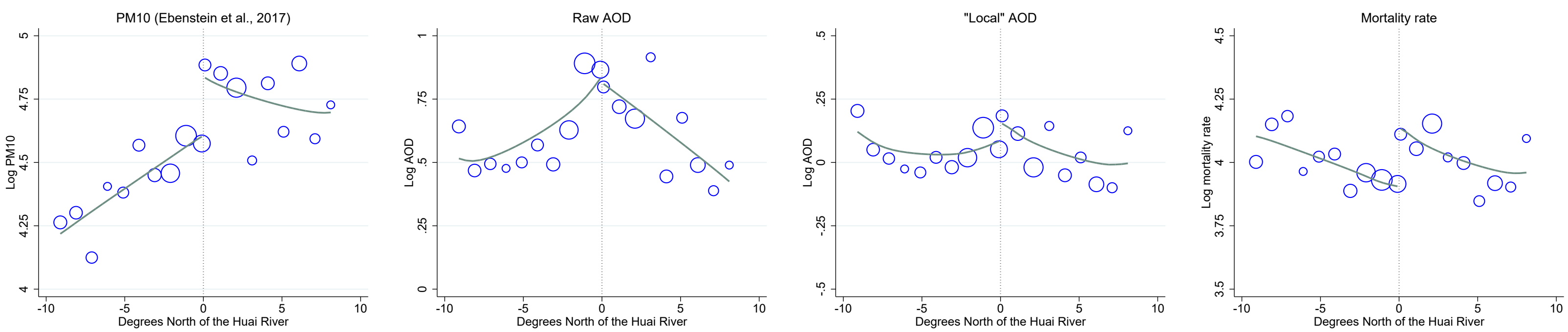
Notes: Each range plot item shows the mortality-pollution elasticity change coefficient (i.e., Log(Pollution)×𝟙(after monitoring)) from a separate regression using sub-group log mortality rate as the outcome variable. All regressions control for prefectiry-city FEs and quarter-of-sample FEs. Range bars show 95% confidence interval constructed from standard errors clustered at the prefecture-city level.

Figure D.12: Changes in Quarterly Mortality-Pollution Gradient: Nonlinear Specification



Notes: This graph shows residualized plot between logged mortality rate by ten equal bins of residualized Log(Pollution) $\times$ $\mathbb{1}$(after monitoring). Conceptually, the slope across these dots corresponds to the interactive coefficient (Log(Pollution) $\times$ $\mathbb{1}$(after monitoring)) in Table 3. All regressions control for prefectiry-city FEs, quarter-of-year FEs, and year FEs.

Figure D.13: Regression Discontinuity at the Huai River (2011-2012 Sample)



Notes: Scatter plot in each panel shows the local means of the corresponding outcome variable with a bin size of 1 degree (Observations = 161). The horizontal axis is the distance (in degree) to the north of the Huai River, following Ebenstein et al. (2017). Solid lines are from local linear regressions estimated separately on each side of the river. Size of circles corresponds to total population in the distance bin. "Local" AOD = raw AOD residualized of inverse-distance weighted $PM_{2.5}$ from cities within 1,000 km radius.

Table D.1: Characteristics of Cities by Monitoring Rollout Waves

	(1)	(2)	(3)
	Wave 1	Wave 2	Wave 3
Number of cities	74	116	177
Population (million)	7.05	3.90	2.90
	(4.85)	(2.10)	(1.95)
GDP per capita (yuan)	69,836	42,881	27,400
	(27,627)	(23,110)	(13143)
AOD level	0.665	0.600	0.456
	(0.239)	(0.242)	(0.237)
$PM_{2.5}$ level (ug/m^3)	61.3	57.9	46.0
	(22.1)	(20.2)	(17.4)
Industrial SO_2 emissions (ton)	37,569	29,609	18,214
	(40,186)	(24,695)	(17,550)
Average temperature (F)	59.7	58.0	55.3
	(8.52)	(9.59)	(10.6)
Total precipitation (inches)	47.0	42.2	40.3
	(21.9)	(23.2)	(24.4)
Average wind speed (m/s)	1.94	1.71	1.47
	(0.63)	(0.62)	(0.68)

Notes: The underlying observations are at the city level. Standard deviations are in parentheses. All characteristics are measured by 2011-2015 average, except for $PM_{2.5}$ level (average of post-monitoring period) and Industrial SO_2 emissions (year 2006).

Table D.2: Changes in Environment After Monitoring

Indep. var.: 1(after monitoring)	(1)	(2)	(3)	(4)
Panel A. Pollution levels				
Log(Pollution)	0.0015	0.0003	-0.0011	-0.0062
	(0.0106)	(0.0097)	(0.0093)	(0.0093)
Log(max Pollution)	-0.0045	-0.0121	-0.0132	-0.0155
	(0.0148)	(0.0118)	(0.0107)	(0.0103)
Panel B. Political/regulatory environment				
[a]N(anti-corruption cases)	-0.037	-0.069	-0.032	-0.034
	(0.052)	(0.056)	(0.028)	(0.029)
[b]Age(mayor)	0.226	0.203	0.240	0.247
	(0.184)	(0.195)	(0.191)	(0.195)
[c]Likelihood(doc. mayor)	-0.013	-0.011	-0.018	-0.018
	(0.026)	(0.027)	(0.027)	(0.028)
[d]N("pollution regulation" news mention)	-0.0048	-0.0074	-0.0067	-0.0071
	(0.0064)	(0.0070)	(0.0072)	(0.0073)
Panel C. Healthcare access				
[e]Log N(hospitals per 1,000 people)	-0.044	-0.047	-0.042	-0.042
	(0.028)	(0.029)	(0.032)	(0.032)
FEs: city	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region×year			✓	
FEs: region×week-of-sample				✓
[a]N(anti-corruption cases)	mean = 0.24,		sd = 0.75	
[b]Age(mayor)	mean = 50.8,		sd = 3.63	
[c]Likelihood(doc. mayor)	mean = 0.234,		sd = 0.423	
[d]N("pollution regulation" news)	mean = 0.052,		sd = 0.45	
[e]N(hospitals per 1,000 people), annual frequency	mean = 1.61,		sd = 2.28	

Notes: Row names show the dependent variable. "Log(Pollution)" is logged AOD in the city×week. "anti-corruption cases" are the number of downfall local officials during the anti-corruption campaign. "doc. mayor" indicates whether the current mayor of the city has a doctoral degree. "pollution regulation news" are the number of People's Daily news articles that mention both smog and the city name. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Centralsouth (81 cities), Southwest (54 cities), Northwest (52 cities). Estimation data are at the city × weekly level, except for Panel C which uses city × annual observations of hospital counts. Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table D.3: Changes in Weekly Bank Card Transaction-Pollution Gradient: "Deferrable" Consumptions

Dep. var.: Number of transactions per 10,000 active cards in a city×week				
	(1)	(2)	(3)	(4)
Panel A. Merchant type = supermarkets (mean = 257.9)				
Log(Pollution)	4.67	3.66	7.49***	8.19***
	(3.74)	(4.07)	(2.24)	(2.76)
Log(Pollution) × 1(after monitoring)	-11.3***	-11.3**	-14.4***	-17.5***
	(3.82)	(4.71)	(3.06)	(3.80)
Panel B. Merchant type = dining (mean = 46.7)				
Log(Pollution)	1.34*	1.62*	1.30**	1.59**
	(0.784)	(0.884)	(0.510)	(0.631)
Log(Pollution) × 1(after monitoring)	-2.84***	-3.35***	-2.22***	-2.54***
	(0.526)	(0.615)	(0.634)	(0.757)
Panel C. Merchant type = entertainment (mean = 9.70)				
Log(Pollution)	0.449	0.711*	0.409	0.498*
	(0.318)	(0.365)	(0.258)	(0.299)
Log(Pollution) × 1(after monitoring)	-0.667	-1.10**	-0.535	-0.686*
	(0.422)	(0.489)	(0.342)	(0.405)
FEs: city	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region×year			✓	
FEs: region×week-of-sample				✓
N	83,122	83,122	83,122	83,122

Notes: "Log(Pollution)" is logged AOD in the city×week. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Centralsouth (81 cities), Southwest (54 cities), Northwest (52 cities). Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table D.4: Changes in Weekly Bank Card Transaction-Pollution Gradient: "Scheduled" Consumptions (Placebo Tests)

Dep. var.: Number of transactions per 10,000 active cards in a city×week				
	(1)	(2)	(3)	(4)
Panel A. Merchant type = billings (mean = 59.4)				
Log(Pollution)	0.252	0.725	2.64	3.51
	(2.32)	(2.70)	(1.80)	(2.16)
Log(Pollution) × 1(after monitoring)	1.04	-0.383	-3.74	-3.85
	(3.89)	(4.59)	(2.98)	(3.18)
Panel B. Merchant type = government services (mean = 12.4)				
Log(Pollution)	0.367	0.329	0.206	0.554
	(0.674)	(0.724)	(0.728)	(0.849)
Log(Pollution) × 1(after monitoring)	-0.565	-0.694	-0.541	-0.583
	(0.992)	(1.06)	(1.03)	(1.25)
Panel C. Merchant type = business-to-business wholesales (mean = 4.79)				
Log(Pollution)	-0.041	0.065	-0.050	-0.009
	(0.385)	(0.409)	(0.338)	(0.401)
Log(Pollution) × 1(after monitoring)	0.180	-0.119	0.071	0.068
	(0.571)	(0.600)	(0.475)	(0.559)
Panel D. Merchant type = cancer treatment centers (mean = 0.320)				
Log(Pollution)	0.009	0.011	0.016	0.014
	(0.012)	(0.013)	(0.011)	(0.013)
Log(Pollution) × 1(after monitoring)	-0.012	-0.017	-0.022	-0.022
	(0.016)	(0.018)	(0.016)	(0.018)
FEs: city	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region×year			✓	
FEs: region×week-of-sample				✓
N	83,122	83,122	83,122	83,122

Notes: "Log(Pollution)" is logged AOD in the city×week. "billings" include transactions in utilities, insurance contribution, telecommunications and cable services. "government services" include transactions in political organizations, court costs, fines, taxes, and consulate charges. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Centralsouth (81 cities), Southwest (54 cities), Northwest (52 cities). Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table D.5: Changes in Weekly Bank Card Transaction-Pollution Gradient: Robustness Checks

Coef. of interest: Log(Pollution)×1(after monitoring)	(1)	(2)	(3)	(4)
Drop U.S. embassy/consulate cities	-14.1*	-16.6**	-16.9**	-21.0**
	(7.18)	(7.93)	(8.29)	(10.5)
Drop top 10% anti-corruption case cities	-16.3*	-18.8*	-18.0**	-23.4**
	(8.62)	(10.9)	(8.14)	(10.6)
Control for online shopping shares	-20.7**	-23.4**	-19.9***	-25.8***
	(8.43)	(10.4)	(7.63)	(9.91)
Control for weather elements	-22.3**	-25.8**	-24.3***	-30.6***
	(9.17)	(11.4)	(8.23)	(10.9)
Use weekly max pollution level	-28.2***	-29.6***	-16.5**	-21.0**
	(9.76)	(10.4)	(7.33)	(9.08)
FEs: city	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region×year			✓	
FEs: region×week-of-sample				✓

Notes: This table reports the changes in band card transactions - pollution gradient after monitoring. Each cell represents a separate regression. The main effect Log(Pollution) term is not reported in the interest of space. Embassy cities include Beijing, Chengdu, Guangzhou, and Shanghai where $PM_{2.5}$ monitoring data were available before 2013. Weather controls include linear terms of weekly temperature, precipitation, wind speed, barometric pressure, and their full interactions. Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table D.6: Changes in Weekly Bank Card Transaction-Pollution Gradient: Triple Difference

Dep. var.: Number of transactions per 10,000 active cards in a city×week				
	(1)	(2)	(3)	(4)
Log(Pollution) × 1(after monitoring)	3.27	1.02	2.35	4.31
	(7.78)	(8.56)	(7.35)	(8.36)
Log(Pollution) × 1(after monitoring) × 1(Treated)	-27.5**	-27.2**	-24.6**	-14.5
	(12.2)	(12.8)	(12.2)	(15.9)
FEs: city-pair	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region×year			✓	
FEs: region×week-of-sample				✓
N	193,563	193,563	193,563	193,563

Notes: "Log(Pollution)" is logged AOD in the city×week. Mean of dependent variable is 1,111.3 transactions per week per 10,000 cards. "1(Treated)" equals 1 for cities actually in the roll-out wave, 0 for neighboring cities not yet experienced roll-out. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Centralsouth (81 cities), Southwest (54 cities), Northwest (52 cities). Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table D.7: Regression Discontinuity at the Huai River (2011-2012 Sample)

Run. var.: Degrees north of the Huai River			
	(1)	(2)	(3)
Local polynomial:	Linear	Quadratic	Cubic
Panel A. RD estimates: Log(Outcome) $\sim$ 1(North)			
Log(Raw AOD)	-0.059	-0.059	0.348*
	(0.074)	(0.093)	(0.184)
Log("Local" AOD)	0.326***	0.351***	0.247***
	(0.157)	(0.064)	(0.096)
Log(PM_{10})	0.347***	0.440**	0.474**
	(0.130)	(0.219)	(0.239)
Log(Mortality rate)	0.219***	0.240**	0.083
	(0.072)	(0.101)	(0.173)
Panel B. IV estimates: Log(Mortality rate) $\sim$ $\hat{\text{L}}$og(Pollution)			
$\hat{\text{L}}$og("Local" AOD)	0.660*	0.591**	0.875
	(0.344)	(0.299)	(0.650)
$\hat{\text{L}}$og(PM_{10})	0.538	0.420	0.463
	(0.348)	(0.369)	(0.427)

Notes: In panel A, each row corresponds to an outcome variable, and each cell reports coefficient for a dummy indicating DSPs north of the Huai River in a separate regression (Observations = 161). "Local" AOD = raw AOD residualized of inverse-distance weighted $PM_{2.5}$ from cities within 1,000 km radius. PM_{10} data are from Ebenstein et al. (2017). Panel B reports fuzzy RD estimates of the effect of Log(Pollution) on Log(Mortality). Columns 1-3 show RD with locally linear, quadratic, and cubic control function for the running variable. All regressions use triangular kernel and Imbens and Kalyanaraman (2012) bandwidth selection. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.