NBER WORKING PAPER SERIES

FROM FOG TO SMOG:
THE VALUE OF POLLUTION INFORMATION

Panle Jia Barwick
Shanjun Li
Liguo Lin
Eric Zou

Working Paper 26541
http://www.nber.org/papers/w26541

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2019, Revised January 2020

We thank Antonio Bento, Fiona Burlig, Trudy Cameron, Lucas Davis, Todd Gerarden, Jiming Hao, Guojun He, Joshua Graff Zivin, Matt Khan, Jessica Leight, Cynthia Lin Lowell, Grant McDermott, Francesca Molinari, Ed Rubin, Ivan Rudik, Joe Shapiro, Jeff Shrader, Jörg Stoye, Jeffrey Zabel, Shuang Zhang, and seminar participants at the 2019 NBER Chinese Economy Working Group Meeting, the 2019 NBER EEE Spring Meeting, the 2019 Northeast Workshop on Energy Policy and Environmental Economics, MIT, Resources for the Future, University of Alberta, University of Chicago, Cornell University, GRIPS Japan, Indiana University, University of Kentucky, University of Maryland, University of Oregon, University of Texas at Austin, and Xiamen University for helpful comments. We thank Jing Wu and Ziye Zhang for generous help with data. Luming Chen, Deyu Rao, Binglin Wang, and Tianli Xia provided outstanding research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

From Fog to Smog: the Value of Pollution Information
Panle Jia Barwick, Shanjun Li, Liguo Lin, and Eric Zou
NBER Working Paper No. 26541
December 2019, Revised January 2020
JEL No. D80,I10,Q53,Q58

ABSTRACT

During 2013-2014, China launched a nationwide, 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 implementation 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 household behavioral changes, including increases in online searches for pollution-related topics, adjustments in day-to-day consumption patterns to avoid pollution exposure, and higher willingness to pay for housing in less polluted areas. As a result of both short- and long-term behavioral changes, the program reduced the mortality cost of air pollution by nearly 7%. Conservative estimates suggest annual benefits of RMB 130 billion from the program, a figure at least one order of magnitude larger than the costs of the program and associated avoidance behaviors combined. 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.

Panle Jia Barwick
Department of Economics
Cornell University
462 Uris Hall
Ithaca, NY 14853
and NBER
panle.barwick@cornell.edu

Shanjun Li
Cornell University
405 Warren Hall
Ithaca, NY 14853
and NBER
SL2448@cornell.edu

Liguo Lin
School of Economics
Shanghai University of Finance and Economics
lin.liguo@mail.shufe.edu.cn

Eric Zou
Department of Economics
University of Oregon
529 Prince Lucien Campbell Hall (PLC)
1415 Kincaid Street
Eugene, OR 97405
ericzou@uoregon.edu

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 unavailable or underprovided in real-world settings, especially for information with public good properties (such as forecasts on weather, 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 gathering and providing pollution-related information in developing countries despite the fact that they experience the worst pollution in the world. This is because pollution information is either not collected or deliberately withheld by the government.[1] Consequently, key questions – such as what are the benefits of information provision and how much public support is optimal – remain mostly unanswered. These issues are pressing because 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 above the World Health Organization guideline. Despite the hazardous level of exposure, a comprehensive monitoring network was nonexistent. Dissemination of the scant data that were collected was politically controlled and, in many cases, forbidden. In 2013, amid the social outcry over the lack of transparency and a dramatic shift in government position on air pollution, China launched a nationwide program that undertook real-time monitoring of air quality (henceforth, the information program). The program covers 98% of its population, a watershed moment in the history of China's environmental policies. 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 and provide the first estimate of the value of pollution information through this program. The basis of our identi-

[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.

fication strategy is the staggered introduction of the information program across three waves of cities and completed over the course of two years. The sequence of the staggered roll-out is based on cities' administrative hierarchy (e.g., provincial capitals) and pre-determined designations (e.g., environmental protection priority cities of 2007). The program implementation date is a top-down decision driven by the physical constraints of installing monitoring stations, and it is thus uncorrelated with the day-to-day variation of local pollution levels as shown in our analysis. In addition, there was no other program following a similar roll-out schedule during our sample period. To causally identify how information has altered consumer responses to (potentially endogenous) pollution, we formally derive two orthogonality conditions that allow our estimator to "difference out" the pollution endogeneity before and after the policy's introduction, thus isolating the causal effect of the policy. We show that the validity of the identifying assumptions can be examined with two empirical tests: (1) a pre-trend test analogous to the test for parallel trends assumption in the standard difference-in-differences setting, and (2) a balance test for the absence of *level* shifts in pollution before and after the policy. Our methodology allows us to estimate, using simple OLS, how information causes both a "change in level" of information availability and public awareness, and a "change in slope" between household behavior and outdoor air quality.

We build the most comprehensive database ever compiled on a rich set of outcomes – including social awareness, air quality measures from satellites, short- and long-run economic activities, and health outcomes – that covers the period before and after the information program. 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 the publication of air pollution-related articles in *People's Daily*, the government's official newspaper, rises from less than once per week to almost daily. The number of mobile phone applications ("apps") released each year that stream air pollution data to users surges by 500%, four times faster than the growth of other apps. Immediately after the program is launched, the term "smog" ("wu mai" in Chinese) becomes, for the first time, a buzzword in social media. Within one year of the program implementation in a city, purchases of air purifiers more than double and become highly correlated with the local pollution level.

These changes in information access and public awareness 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 card and debit card transactions in China from 2011 to 2015 to build a measure of outdoor purchase trips for all cities in the country. Linking purchase activities to local ambient air pollution, we show that the information program boosted pollution avoidance by triggering a negative purchase-pollution elasticity of 3%, a pattern that is robust across a host of econometric specifications. As expected, the increase in avoidance concentrates in plausibly "deferrable" consumption categories, such as supermarket shopping, outdoor dining, and entertainment, rather than in "scheduled" trips such as paying bills, business-to-business wholesales, and cancer treatment sessions.

Our long-run analysis focuses on the capitalization of pollution information in the housing market. Leveraging geolocation information from the near-universe of new home sales in Beijing spanning the nine years from 2006 to 2014, we examine the information-induced *changes* in the relationship between housing prices and local pollution levels using two different research designs. First, we employ the pixel-averaging technique ("oversampling") to enhance the original satellite pollution data's spatial resolution from 10-by-10 km to 1-by-1 km (Fioletov et al., 2011; Streets et al., 2013). The high-resolution measure allows us to conduct comparisons within fine geographic units, such as communities that are similar to census block groups in the United States. We estimate a home value-pollution elasticity of -0.6 to -0.8 post disclosure; one standard deviation increase in pollution is associated with a 4-6% decrease in housing prices. 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 addresses of Beijing-area mega polluters: the 10% facilities that account for 90% of total industrial air emissions. Following the literature (Currie et al., 2015), we estimate separate "distance gradient" curves 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 three kilometers of a major polluter depreciate 27% afterward. This corresponds to 42% of the interquartile range of the housing price dispersion. While somewhat larger than the results of Currie et al. (2015), who find that properties within 1km of a toxic plant experience an 11% reduction in value, these estimates are plausible in light of Beijing's three-fold increase in housing prices over the sample period. Thus, the information program facilitates the capitalization of air quality in the housing market, potentially improving social welfare through residential sorting.

These behavior changes could significantly mitigate the devastating consequences of severe air pollution in China. 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 post monitoring that concentrates in cardiorespiratory causes, and is more precisely estimated among the age groups more vulnerable to pollution exposure. The impact is more pronounced in cities that have a larger share of urban population, have more hospitals per capita, consume more electricity, and have a higher mobile phone penetration. These patterns are remarkably similar to heterogeneous short-term avoidance responses uncovered using the card spending data, suggesting a plausible pathway from effective avoidance to beneficial health outcomes.

Combining our findings with existing estimates on the causal effect of pollution on mortality in China (e.g., Ebenstein et al., 2017), we find that access to pollution information has reduced premature deaths attributable to air pollution exposure by nearly 7%, and has generated a health benefit equivalent to a 10 ug/m^3 reduction in PM_{10}. Based on recent estimates in the literature (Ito and Zhang, 2018; Ashenfelter and Greenstone, 2004; Murphy and Topel, 2006), the associated annual benefits vary from RMB 130 billion (using willingness to pay for clean air) and RMB 520 billion (using age-adjusted value of statistical life). By our calculations, such social benefits outweigh the costs from defensive investments (such as air purifier purchases), avoidance behavior (such as foregone consumption), and deploying and maintaining the program by at least one order of magnitude, making the information program among the most successful environmental policies in a developing country.

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

Second, our study shows that information is a crucial 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 commonly made in the existing literature that uses revealed preference to estimate the value of non-marketed environmental goods, which is perhaps more tenable for developed economies. To the extent that access to information is lacking in developing countries, the perfect-information assumption could underestimate consumers' true willingness to pay for environmental goods. Our findings provide a potential explanation for why environmental quality is severely undervalued in developing countries (Greenstone and Jack, 2015), and why

[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).

the dose relationship between pollution and mortality differs 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; Newell and Siikamäki, 2014; Mastromonaco, 2015; 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 over both the short and long terms. Our analysis shows that these behavioral responses could lead to improved health conditions. 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 describes our unified empirical framework. Section 6 presents the effect of the program on short- and long-term behavioral changes and on mortality. Section 7 calculates the value of information. Section 8 concludes.

2 Institutional Background and Data

2.1 Environmental Regulations

The real-time $PM_{2.5}$ monitoring-and-disclosure program started in 2013 was 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 drastically enhanced the public awareness of the health impact of $PM_{2.5}$. To help understand

[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. Chay and Greenstone (2005); Banzhaf and Walsh (2008); Bayer, Keohane and Timmins (2009); Mastromonaco (2015); Chen, Oliva and Zhang (2017); Freeman et al. (2019) analyze long-term housing choices and migration decisions in response to pollution.

this change, we provide a brief history of China's environmental regulations.

Environmental Regulations Prior to 2012 In 1982, China established its first national ambient air quality standards (NAAQS), which set limits for six air pollutants including total suspended particles (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. The 2000 amendment removed NO_x from the list, and relaxed the standards for NO_2 and O_3 in response to noncompliance due to increased automobile usage.

Throughout much of the 1980s to the early 2000s, the primary 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, environmental regulations focused on the control of acid rain and SO_2 emissions (Yi, Hao and Tang, 2007). The prominent regulation was the two-control zone policy (TCZ) implemented in 1998 where prefectures with high PH values of precipitation and 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). These zones imposed a series of measures, such as mandating the installation of flue gas desulfurization in coal-fired power plants and closing down small coal-fired power plants (Tanaka, 2015). These aggressive emissions control and clean energy policies reduced the average SO_2 concentration 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 early 2000, the source of air pollution shifted from coal burning to mixed sources. Particulate matter (PM) rather than SO_2 became the primary pollutant. This shift was driven by the significant decline in emissions from coal-fired power plants, and by skyrocketing emissions from automobiles, manufacturing facilities, and construction linked to the dramatic growth in vehicle ownership, increased industrial activities (after China's World Trade Organization accession in 2001), and rapid urbanization. The regulatory focus 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, combined with weak monitoring and enforcement from the central government. Episodes of extreme air pollution became frequent, especially during winters in many urban centers. The U.S. Embassy in Beijing and U.S. consulates in Guangzhou and Shanghai started to report hourly $PM_{2.5}$ levels in 2008 based on monitoring stations installed on-site.

[4]The fraction of the acid rain zone in China's total terrain decreased from the peak level of about 30% in the mid-1990s to 8.8% in 2015.

The $PM_{2.5}$ readings from these sites were often inconsistent with the official pollution reports, and they thus became sources of diplomatic tension.[5]

Limited Pollution Awareness Prior to 2012 While air pollution has been a long-standing issue, public access to daily pollution measures was almost absent prior to 2013. Although the Ministry of Environmental Protection (MEP) began publishing the daily Air Pollution Index (API) for major cities in 2000, the reported API 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 and onward.[6] In addition, mass media and broadcasts did not report the API index. Furthermore, the API data were gathered and reported by local environmental bureaus whose leaders were appointed by local governments. The MEP did not control the monitoring stations, and it had limited ability to monitor 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).

Though the dominant pollutant had shifted from SO_2 to particulate matter in the 2000s, there was no systematic collection of $PM_{2.5}$ measures. As a result, consumer awareness of $PM_{2.5}$ concentration was extremely limited prior to 2013. Both government agencies and the media often characterized poor visibility due to high levels of $PM_{2.5}$ as *fog* rather than *smog*. For example, on November 27, 2011, newspaper headlines, as well as the China Meteorological Administration, attributed dense fog in Beijing and northern cities as the reason for widespread flight delays and cancellations. In fact, extreme pollution caused the low visibility, as shown in Figure 1 that displays the U.S. National Aeronautics and Space Administration (NASA) satellite view of China and the aerosol optical depth (AOD) reading of 4.5 or higher for many northern cities.[7] A similar pollution event that occurred on December 4-6, in 2011 was again reported to be dense fog by major news media including China Central Television, the predominant state television broadcaster in Mainland China, and popular websites such as sina.com.

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

[5]The then-vice minister of the Ministry of Environmental Protection (MEP), Wu, Xiaoqing, openly requested the 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 does not 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.

[7]The AOD measure is usually between 0.1 to 0.15 in the United States. For China, the average is about 0.5.

in her 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 it was fog back then. That's what we called it.... as a former journalist, I started to blame myself because for all these 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."[8]

The Information Program and Environmental Regulation Post 2012 In 2012, China's MEP revised the NAAQS and for the first time in 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.[9] 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 10 concrete policies such as promoting the role of market-based mechanisms and establishing monitoring and warning systems to cope with severe air pollution events.[10] In addition to this action plan, for the first time in the history of national five-year plans, the 13th Five-Year Plan required cities at prefecture level or higher to reduce $PM_{2.5}$ concentrations by 18% from 2015 to 2020.

The recognition of $PM_{2.5}$ as a primary pollutant, and the aggressive policies to reduce $PM_{2.5}$ concentrations marked a significant shift of the China's long-standing strategy of prioritizing economic growth over environmental concerns. This shift took place against the backdrop of China's 12th and 13th Five-Year Plans, which set pollution reduction as one of the bureaucratic hard targets in the cadre evaluation system (Wang, 2017).[11] To effectively

[8]The documentary has been compared with Al Gore's *An Inconvenient Truth* in terms of its style and impact. The film had been viewed over 150 million times on the popular website tencent.com within three days of its release, and had been viewed another 150 million times the time the government took it offline four days later. Tu et al. (2020) show that household willingness to pay for clean air increased among those who had viewed the documentary; the findings are based on a longitudinal survey conducted in the city of Nanjing, and a regression discontinuity (in time) design.

[9]The cities in the 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.

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

[11]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." He declared a war on 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.

monitor local air pollution levels and to address the pitfalls of the previous reporting system of API (Greenstone et al., 2019), the MEP implemented a nationwide monitoring-and-disclosure program starting in 2013. The focus was to build an efficient, scientific system to monitor air quality, and to publicly disclose publicly data in real time.

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 that 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 they were 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.[12] In the final wave, achieved by November 20, 2014, the program reached the remaining 177 cities. Figure 2 plots the roll-out of the program. 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.

Second, and more importantly, the information program established a dissemination system to report a real-time Air Quality Index (AQI) that is on a single scale of zero to 500. Monitoring results are displayed in real time on MEP's website. In contrast to the previous API reporting system, the MEP directly controls the new monitors, and data are directly transmitted to MEP's information center in real time to avoid manipulation by the local government. Both hourly and daily AQIs, as 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 C.2 provides a screenshot of the website interface. Importantly, the government allows private parties to access and stream data directly from the web. This has spurred a surge in private websites and mobile phone applications that report real-time air quality information. Section 4 provides more detail on the program's effects on information access and public awareness about $PM_{2.5}$ and related health consequences.

[12]The Environmental Improvement Priority Cities were designated in 2007 during the Eleventh-Five-Year period and contain important regional economic centers and cities face significant environmental challenges. The National Environmental Protection Exemplary Cities program was created during the Ninth Five-Year Plan; 68 cites were awarded the title from 1997 to 2007 based on a host of environmental quality criteria. Appendix Figure C.1 tabulates cities by waves and their associated city clusters.

Concurrent Government Policies and Potential Confounders A key feature of the information program's roll-out is that it is based on cities' administrative hierarchies and well-known groupings that were designated long before the initiation of the information program. The pre-determined nature of these groupings indicates that there is little likelihood of selecting cities into different roll-out waves based on unobservable characteristics or future projections of outcome variables. In addition, the program implementation date is a top-down decision driven by the physical constraints of installing monitoring stations, and, thus, uncorrelated with *day-to-day* variation in local pollution, as shown in our balance tests.[13]

We are unaware of any other concurrent government policies that coincide with the geography and the timeline of the information program's roll-out during our study period. Nevertheless, several major environmental and energy programs are worth noting. First, China rolled out pilot CO_2 cap-and-trade programs targeted at heavily polluting industries in two provinces (Guangdong in December 2013 and Hubei in April 2014) and five individual cities (Shenzhen in June 2013, Beijing and Shanghai in November 2013, Tianjin in December 2013, and Chongqing in June 2014).[14] The second policy is the energy reduction plan among over 16,0000 of the largest energy users that collectively accounted for over 60% of total energy consumption in 2010. The policy, started in 2012, aimed to reduce energy intensity and carbon intensity as outlined in the 12th Five-Year Plan. The third policy is fuel switching from coal to natural gas for winter heating in northern China from the winter season of 2013 as part of the *Action Plan on Air Pollution Prevention and Control.*

To the extent that there is no significant overlap in the geography or timeline of implementation of these policies with the monitoring roll-out schedule, we expect the impact of these policies to be picked up by the rich set of spatial and temporal controls we use in our econometric models (e.g., city fixed effects and region-by-week-of-sample fixed effects). Consistent with this view, a series of tests (reported in Appendix Table C.2) shows that the information roll-out is not associated with any significant changes in the level of pollution in the treated city conditional on the fixed effects controls, while the time fixed effects dummies picks up a substantial reduction of overall pollution levels since 2014.

[13] Appendix Table C.1 tabulates economics attributes for cities in each wave. Cities in earlier waves tend to have a larger population, higher GDP per capita, and higher levels of air pollution and industrial emissions. On the other hand, as shown in balance tests in Appendix Table C.2, these economic and environmental conditions do not change systematically before and after the program roll-out. Together, the nature of the program and the evidence suggest that the choice of cities and the timing of the program roll-out are unlikely to be based on city-level unobservables.

[14] As a proof of concept for the national CO_2 cap-and-trade program slated to come online in 2020, these seven regional pilots had limited impact on carbon emission and air quality given the generous allocation and the lack of strong punishment mechanism for noncompliance (Zhang, Wang and Du, 2017).

2.2 Data

We compile multiple data sets to allow for a comprehensive study of the impacts of the information disclosure program on a variety of outcomes. The data sets include data on consumer awareness and online search behavior, air purifier sales, bank card transactions, housing transactions, mortality rates, and air quality measures from satellites. The data on housing transactions are for Beijing, and all the other data sets are national or nationally representative in scope (Appendix Table C.3).

Mass Media, Phone Apps, and Web Searches The high Internet and mobile phone penetration among the Chinese population provides a unique opportunity to investigate pollution awareness using digital sources.[15] We draw on several digital sources to illustrate the evolution of public access to pollution information. First, we look at the publication trends in the *People's Daily*, the government's official newspaper. We pull articles that contain the word "smog" (or "air pollution", or "atmospheric pollution") from the *People's Daily*'s digital archive, and we also identify a list of relevant cities mentioned in each article about "smog".

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."[16] These apps typically display current hourly pollution levels; some also provide health-related guidelines (e.g., avoiding outdoor activities) when pollution levels are high. Appendix Figure C.3 is a screenshot from a typical pollution app. Apps in other major categories such as gaming, reading, and shopping serve 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 given city and on a given 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. The algorithm is based on the underlying search traffic and reflects search intensity.

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

[15] Data from the China Internet Network Information Center show that, by the end of 2012, China had about 0.56 billion internet users (or 40% of its population). 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%). Mobile phone prevalence rose from 73.5 per 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).

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

monthly frequencies for 50 cities from 2012-2015.[17] Among these 50 cities, 34, 11, and 5 are in the first, second, and third waves of the program roll-out, respectively.

Bank Card Transactions Data Households' shopping trips are constructed using the universe of credit card and debit card transactions during the 2011-2015 period 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 as shown in Appendix Figure C.4. For each transaction, we observe the merchant name and location, and transaction amount and time. Appendix Figure C.5 shows the spatial pattern of cards and transactions. Credit and debit cards are widely adopted across the country as the most commonly used transaction method, especially in urban areas (Barwick et al., 2018).[18] To the best of our knowledge, the fine degree of spatial and temporal resolution provided by this data set represents the most comprehensive data available on consumption activities for China.

Two features of the data are worth mentioning. First, our data contain a small fraction of transactions that are made online.[19] We drop these online transactions because tracing these buyers' physical locations is difficult. Second, we do not observe specific items purchased in each transaction. Instead, UnionPay classifies merchants into over 300 categories, such as department stores and supermarkets. We use the category information in our analysis.

Our key outcome variable is the *purchase rate*, defined as the ratio between (1) the total number of transactions occurred in a city-week, and (2) the total number of active cards with positive transactions in a given city-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.

Housing Transactions Data (Beijing) Our housing data contain a total of over 660,000 new home purchases in about 1,300 apartment complexes in Beijing from January 2006 to April 2014, with a nearly universal coverage. Variables recorded include the transaction date and price, housing unit characteristics (floor, size of the unit, etc.), as well as attributes and geolocation of the apartment complex. Among the 1,300 apartment complexes, 64% of the complexes are sold out in three years. The temporal variation in price among apartments within the same complex is used as one key source of variation in one of the empirical strategies.

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

[18] Since 2015, mobile payment apps including Alipay and Wechat have become very popular but cards transactions still account for more than 50% of national consumption.

[19] Online transactions accounted for 5.1% of total transactions by volume and 3.6% by transaction amount between 2013 and 2015. The shares were smaller in earlier years.

Polluter Data (Beijing) The MEP conducts an annual survey of all major industrial polluters, and compiles the Chinese Environmental Statistics database (CES), the most comprehensive coverage of firm 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 the address and operation status of each firm by linking CES data 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 the 2006-2014 period, the sample period of our housing transactions.

Mortality Data The 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, a 5% representative sample of China's population over the 2011-2016 period.[20] The DSP mortality database, drawn from hospital records and surveys of the deceased's household, is one of the highest-quality health databases 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 three groups are closely related to cardiovascular diseases and 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 post the information program, we obtain ambient air quality measures from AOD via the Moderate Resolution Imaging Spectroradiometer (MODIS) algorithm, operated by NASA and installed on the platform of the research satellite Terra. The original data have a geographic resolution of 10 km $\times$ 10 km and a scanning frequency of 30 minutes, which we average to the city $\times$ day level from 2006 to 2015. MODIS records the degree to which sunlight is scattered or absorbed in the entire atmospheric column corresponding to the overpassed area under the cloud-clear condition. As such, AOD captures concentration of particles 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

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

C.6 documents a strong correspondence between AOD and $PM_{2.5}$ after the implementation of 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, the 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., annually) and cannot be disaggregated in a straightforward manner. Second, MODIS AOD data allow us to observe overlapping 10 km $\times$ 10 km grid cells, which is essential for the oversampling exercise in Section 6.2. Finally, while MODIS AOD measure is a common input in predicted ground-level $PM_{2.5}$, there is no consensus on the precise relationship between AOD and $PM_{2.5}$ data 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). Access to 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 a classical model of health demand and production (Grossman, 1972; Deschenes, Greenstone and Shapiro, 2017).

3.1 Model Setup

Individuals derive utility $U(x, h)$ from the consumption of a numeraire good x, whose price is normalized to one, and health stock 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: $I + w \cdot g(h) \geq x + p_a \cdot a$, 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.[21] Individuals allocate their wage and non-wage income between

[21] The effect of health on wage income captures the impact of pollution on labor supply or productivity

consumption and engaging in avoidance behavior a, where p_a is the associated price (e.g., the cost of an air purifier or medication). We use a to include broadly defined (costly) adaptation behavior.[22] Dynamics and savings are assumed away to ease exposition.

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:

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

The health function $h = h(c_0, a)$ in the optimization can be viewed as an ex ante health function upon which consumers rely for decisions before the health outcome is realized. 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) and 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 $a(c_0)$: $w \cdot g[h(c, a(c_0))]$. Let $x(c, c_0)$ denote consumption of the numeraire good. The 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: the first argument is the actual pollution c, and the second argument is the perceived pollution level c_0. To examine the behavioral changes associated with and the welfare impacts of the information program, we make the following assumptions:

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 do not engage in an unreasonable amount of avoidance behavior. Similarly, we assume that hours worked increases in health,

as documented in Graff Zivin and Neidell (2012); Hanna and Oliva (2015); Chang et al. (2019); He, Liu and Salvo (2019).

[22] Examples include reducing outdoor activities (Zivin and Neidell, 2009; Saberian, Heyes and Rivers, 2017), engaging in defensive spending (e.g., buying face masks and air purifiers) (Ito and Zhang, 2018; Zhang and Mu, 2018), and making choices to change residential locations or migrate (Chay and Greenstone, 2005; Banzhaf and Walsh, 2008; Bayer, Keohane and Timmins, 2009; Chen, Oliva and Zhang, 2017).

but at a decreasing rate: $\frac{dg}{dh} \geq 0, \frac{d^2g}{dh^2} \leq 0$. 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.

We focus on interior solutions for the optimal level of avoidance behavior a.[23] The assumption of $\frac{\partial^2 h}{\partial a \partial c} \geq 0$ is crucial in delivering "complementarity" between pollution and avoidance: the higher the pollution, the more intense avoidance is likely to be. At low levels of pollution, the marginal health benefit of avoidance $\frac{\partial h}{\partial a}$ is likely to be limited. As pollution elevates, higher marginal benefits induce individuals to engage in more avoidance to mitigate the health and wage impact of pollution. There are many low-cost defensive mechanisms. 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.

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 before the information program. We assume that $c_0 < c$, that is, the perceived level of pollution is lower than the actual level. Another interpretation of Assumption 3 is that people underestimate the negative health impact of pollution. Pollution concentration c is assumed to be perfectly observed following the implementation of 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:*

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

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

Appendix A provides the proof. The theoretical model predicts that following the implementation of 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.

[23] A necessary condition for an interior solution is $w \cdot \frac{dg}{dh} \cdot \frac{\partial h}{\partial a}|_{a=0} > p_a$.

3.2 Value of Information

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

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

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

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

The indirect utility before the information program is:

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

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 Taylor's series expansion to the indirect utility function $V(c,c)$ at the second argument $c = c_0$: $V(c,c) = V(c,c_0) + \frac{\partial V}{\partial c_0}(c - c_0) + o_p(c - c_0)$, where $o_p(c - c_0)$ denotes higher order terms of $(c - c_0)$. The value of information is therefore:

$$\begin{aligned} VOI &= V(c,c) - V(c,c_0) \\ &= \{U_h \cdot \frac{\partial h}{\partial a} \cdot \frac{\partial a}{\partial c_0} + \lambda \cdot w \cdot g_h \cdot \frac{\partial h}{\partial a} \cdot \frac{\partial a}{\partial c_0} - \lambda \cdot p_a \cdot \frac{\partial a}{\partial c_0}\}(c - c_0) + o_p(c - c_0) \end{aligned} \tag{2}$$

There are three terms in the curly bracket. The first refers to changes in utility as health improves from the avoidance behavior. The second denotes changes in wage income due to pollution impact on effective work hours or productivity. The third term in the curly bracket captures the cost of taking additional avoidance measures such as buying air purifiers or changing outdoor activities. Our empirical analysis below quantifies the magnitude of the terms in the curly bracket.

4 The Sea Change in Information Access and Awareness

4.1 Information Access: News and Mobile Apps

The government's official newspaper, the *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 the *People's Daily* mentions "smog" in any articles. "Smog" is rarely mentioned in the 1990s and 2000s. Almost immediately following the information program's initial roll-out, the frequency of "smog" appearing in the People's Daily jumped from three days per month in 2012 to seventeen days per month in 2013. It remained high for the rest of the sample period. This effect is unlikely to have been driven by the coverage of the information program itself.[24]

One might be concerned that the sharp increase in "smog" mentions post 2013 is confounded by changes in the general environment (e.g., shifts in government policies), instead of driven by the information program that is gradually rolled out across cities. To examine this, we scan each "smog"-containing article in the *People's Daily* to determine the list of cities mentioned. This allows us to construct a city panel data set of "smog" and conduct a test for a sharp change in "smog" mentions after a city begins to monitor pollution, conditional on general within-year seasonality and year-by-year changes in pollution. Assuming that unobserved changes in the overall environment do not correlate with the timing of the monitoring roll-out (an assumption we return to in Section 4.3), the difference between pre- ($t < 0$) and post- ($t \geq 0$) program coefficients identifies the causal impact of the information program.

Figure 3b plots standardized "smog" as a function of time in the wake of the roll-out of the information program in a local city, controlling for month dummies and year dummies. 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 increased by 50% of a standard deviation. We find very similar results when we repeat this analysis using other keywords, including "air pollution" and "atmospheric pollution"s.

We then examine the availability of pollution-related mobile phone apps. Unlike newspapers that provide pollution information at a daily frequency, information from apps is 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 learns about its pollution exposure at the moment. We compare the distribution of the release times for pollution apps with the release times of apps from other popular categories – including gaming, music, video, reading, finance, sports, education, shopping, and navigation – that capture the majority of commonly used apps.

As shown in Figure 4, 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-

[24] There was a modest increase in the frequency of the term "smog" before 2013. A close read of the articles indicate that these phrases were mostly used to describe dense fog, and were rarely associated with pollution.

out. In total, about 82% of pollution apps are released between 2013 and 2015, compared to 18% released between 2009 and 2013. The availability of pollution apps has grown nearly 500% post 2013. This growth is four times greater than the growth of other apps.[25]

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 keyword searches on Baidu. This analysis is analogous to the examination of "smog" news (and news on "air pollution") in Section 4.1. Here we again focus on keyword "smog," though the patterns for keywords "air pollution," "mask," and "air purifier" are very similar.

Figure 5a plots the time-series pattern of the search index for the keyword "smog" at the national level. The index varies between zero and 50 for most of 2011 and 2012, and jumps overnight from 45 in December 2012 to 4,000 in January 2013, the month of the initial roll-out. This is a hundred-fold increase. In addition to remaining at a high level, post-2013 searches exhibit a strong seasonal pattern in which the index is highest in winter seasons; smog is more severe in winter partly due to coal-fueled heating. Not surprisingly, while the search index and AOD are essentially uncorrelated before 2013, they are strongly correlated afterward. The two outliers in Figure 5a, 9,000 in December 2014 and 11,000 in January 2015, correspond to the two worst smog episodes in the country's recent history.

Leveraging the search index at the city $\times$ 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 in a local city. Echoing results in Section 4.1, the index is flat and near zero before the information program, and rises rapidly when monitoring starts. Within one year after the roll-out, smog searches increase by 75% of a standard deviation.

Our second measure of awareness is public and private investments in specific type of defensive equipment: air purifiers. We repeat the same analysis described above using data on monthly air purifier sales for 50 cities. Air purifier sales more than double, rising from 11,000 units per month in 2012 to over 25,000 units per month after 2013 (Figure 6a). Similar to web searches, air purifier sales are invariant to weather and pollution conditions prior to the launch of the information program, but after the launch, they exhibit a strong seasonality with more sales in winter. Finally, analogous to what we see in web search patterns, the increase in sales coincides with the timing of the local roll-out (Figure 6b).

The surge in Internet searches and air purifier sales both nationally and locally and their

[25] A mass of pollution apps released before 2013 initially streamed weather information, and then incorporated air quality contents after the information program began. These apps were categorized as pollution apps by the time we queried Apple's App Store.

strong correlation with air pollution after then start of the information program provide strong evidence that concepts such as "smog" and "air pollution," as well as related adverse health consequences, entered the public domain as a result of easy access to related information provided by the program. As forcefully put by Wainwright (2014) and quoted by Greenstone and Schwarz (2018), by 2014, the "daily talk of the AQI has become a national pastime amongst ex-pats and Chinese locals alike. Air-quality apps are the staple of every smartphone. Chinese micro blogs and parenting forums are monopolized by discussions about the best air filters (sales of the top brands have tripled over the last year alone) and chatter about holidays to 'clean-air destinations.' " (Section 6analyzes changes in households' short- and long-term behaviors as a result of the information program.) In the remaining part of this section, we present a balance test showing that the timing of the information program is unlikely to be correlated with unobservables in day-to-day variation in local air quality.

4.3 Changes in Social and Economic Conditions

As we discuss in greater detail in the next section, our research design relies on the assumption that there are no confounding factors that systematically coincide with the timing of the information roll-out. China experiences rapid social and economic changes during the sample period. While our statistical analyses control for general as well as city-specific time trends using fixed effects, one might be concerned about differential trends or confounding factors. 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 is the case, Appendix Table C.2 presents a series of balance tests on differential shifts in city-level observables before and after the program. We focus on three classes of social and economic conditions: pollution levels using satellite-based AOD (both the weekly average and the maximum pollution reading in a city-week), the political and regulatory environment (the number of local officials ousted during the anti-corruption campaign, demographics of local political leaders, news mentions of regulation policies), and healthcare access (the number of medical facilities). Each cell in Table C.2 is a regression of these observables on a dummy variable indicating whether the information program has begun in a local city, controlling for city fixed effects and various trends (week-of-sample or region-specific dummies interacted with week-of-sample dummies). 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 before and after the program in any of these regressions, suggesting that the role of both observed and unobserved confounding factors is likely limited (Altonji, Elder and Taber, 2005).

The results from these balance tests are not inconsistent with recent evidence that China's air quality has improved in recent years, especially after 2015 (Huang et al., 2018; Greenstone and Schwarz, 2018). Regional and national improvement in air quality is absorbed by various trends in our empirical analysis below. Rather, these balance tests indicate that during the months surrounding the local implementation of the information program, economic and environmental measures do not display noticeable differences.

5 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- and long-run behavioral changes in household activities, including avoidance behavior, housing choice and prices, and in health outcomes, including mortality. Throughout the regression analysis, we use the same empirical framework to examine the change in the relationship between pollution exposure and the outcomes (i.e., the "slope") before and after the program:

$$\mathrm{Outcome}_{ct} = \alpha \times \mathrm{P}_{ct} + \beta \times \mathrm{P}_{ct} \times \mathrm{d}_{ct} + \mathrm{x}'_{ct}\gamma + \varepsilon_{ct}, \tag{3}$$

where c denotes a city and t denotes time (e.g., week or month). P_{ct} is the AOD measure of the ambient air pollution. 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.[26] Note that α represents the outcome-pollution gradient before the information program, and β denotes changes in the gradient after the program.

Equation (3) 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: $E\left(\mathrm{P}_{ct} \times \varepsilon_{ct}\right) \neq 0$. Such endogeneity could be due to omitted variables or errors in the measurement of pollution exposure.[27] Addressing endogeneity in air pollution is challenging

[26] All analyses below include full interactions between P_{ct} and the treatment dummy d_{ct}.

[27] For example, satellite-based AOD captures particulate concentration in the entire air column above a

and has been the subject of recent research on understanding the morbidity and mortality cost of air pollution.[28] Critically, 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 that takes place after the information program (which is β). Second, in our analysis, P_{ct} is intended to be a direct measure of ambient pollution, rather than a measure of population exposure determined by the ambient air quality, avoidance behavior, and population distribution. In fact, in the analysis we directly examine how avoidance and residential sorting respond to ambient air pollution with and 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). 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 could cancel out, leaving the OLS estimate of β consistent. The following two assumptions formalize this intuition:

Assumption B1: $\varepsilon \perp \mathrm{d} \mid \mathrm{x}$. 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 the MEP's promulgation of the national $PM_{2.5}$ standard, which marked a sudden and drastic change in the government's stance regarding the importance of environmental quality. The roll-out schedule of the monitoring stations in three waves was primarily based on the pre-determined city designations (mostly administrative hierarchies, such as provincial capitals, and a list of environmental improvement priority cities designated in 2007), as shown in Figure 2 and Appendix Figure C.1. For a given city, the date of the roll-out is a top-down decision driven by the physical constraints of installing monitor stations, and uncorrelated with the day-to-day variation of local pollution levels as shown in our analysis.

One might be concerned about other contemporaneous regulations at both national and local levels to achieve the pollution-reduction goals set out in the 12th (2011-2015) and 13th (2016-2020) Five-Year Plans. As discussed in Section 2.1, other concurrent policies do not coincide with the information program in the roll-out schedule or the spatial coverage. In

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.

[28] See for example 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).

our regressions, we include a rich set of temporal and spatial fixed effects such as region by week-of-sample fixed effects to control for unobserved policies and other confounders. The estimate of the key parameter of interest β is robust across specifications with different sets of fixed effects, lending additional support to this assumption.

A direct check of Assumption B1 is a pre-trend test in the same spirit of the test for parallel trends in a standard difference-in-differences setting. Patterns that are consistent with Assumption B1 would be stable estimates in event time, followed by a trend break that occurs exactly at the time of roll-out. This is indeed what we find in event studies below, where the coefficient estimates are flat and stable before and after the treatment, which contrasts sharply with a sizeable break that is both economically and statistically significant at the time of the information treatment.

Assumption B2: $\mathrm{d} \perp \mathrm{P} \quad | \quad \mathrm{x}$. This assumption implies that conditioning on x_{ct}, the treatment is independent of the level of pollution.[29]

An intuitive way to conceptualize this assumption is perhaps to imagine a binary context in which P indicates "high" vs. "low" pollution areas. Note that equation 3 reduces to a difference-in-differences style setting that compares outcome in regions with high vs. low pollution, before vs. after policy introduction. The "slope", i.e., the outcome-pollution gradient, in this case is simply the difference in the outcomes experienced in areas with high and low pollution. Assumption B2 ensures that there are no "compositional changes" in regions that experience high or low levels of pollution after the policy introduction. In other words, the nature of the endogeneity in pollution does not change before and after the policy.

As a direct test for this assumption, we report a series of balance checks in Appendix Table C.2 showing that the pollution level does not change after the policy introduction, conditional on x_{ct}. We also extend the balance table to show that there are no statistically significant changes across a rich set of economics, political, and social variables before and after the program.

Proposition 2. *Under Assumptions (B1) and (B2), the OLS estimate of β in equation* (3) *is consistent.*

The proof is provided in Appendix B. There are two sources of inconsistency in the OLS estimate of β: one from the endogeneity of the interaction term $\mathrm{P}_{ct} \times \mathrm{d}_{ct}$, and the other from smearing due to the endogeneity of P_{ct}. Under assumptions (B1) and (B2), the inconsistency from these two sources cancels out, leaving the OLS estimate of the difference (the change

[29]This assumption is stronger than what we need for the consistency of β. As we show in Appendix B, a sufficient condition is $E[\mathrm{d}|\mathrm{M}_x\mathrm{P}] = c$, where M_x is the projection matrix.

in slope that we are interested in) to be consistent. Based on Proposition 2, our subsequent analysis focuses on the OLS estimate of β. In Section 6.3, which examines 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 P_{ct}.

6 Pollution Disclosure, Behavior, and Health

6.1 Pollution Disclosure and Avoidance

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

$$\text{PurchaseRate}_{ct} = \sum_{k=-24}^{15} \beta_k \times \ln \mathrm{P}_{ct} \times \mathbb{1}(t=k) + \sum_{k=-24}^{15} \eta_k \times \mathbb{1}(t=k) + \mathrm{x}'_{ct}\gamma + \varepsilon_{ct} \tag{4}$$

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 (Section 2.2). The pollution measure "$\ln \mathrm{P}_{ct}$" is logged average AOD. The key parameters of interest are the β's, which represent changes in the purchase rate for a 1 percent increase in AOD. 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-treatment 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 a nearly identical number of city-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 prefecture-city fixed effects, week-of-year fixed effects, and year fixed effects. 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 estimates to be representative of the population impact, we weight the regression using the number of active cards in a city and year to account for the cities' significant size differences.

Figure 7 summarizes the estimates of β_k coefficients. We restrict β_k to vary quarterly (three-month) 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 (4) 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 (3). 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 correspond to the specifications 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. Column 3 further adds region-by-year fixed effects, allowing for common trends in transactions and pollution that are specific to each region.[30] Column 4 is our most stringent specification, controlling for city and region-by-week-of-sample fixed effects. We obtain similar results across the board.

Consistent with the evidence in Figure 7, outdoor consumption trips are invariant to pollution before the program. While our transaction-pollution slope estimate prior to the information program need not be causal (as previously discussed), this evidence does suggest that households are less likely to be engaging in any mitigating measures in the absence of information. In contrast, after a city implements the monitoring-and-disclosure program, purchase trips become responsive to pollution levels: doubling 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 our analysis covers *all* consumption categories, which represent 59% of national consumption. Furthermore, the week-level analysis by construction has already incorporated within-week intertemporal substitution. The estimate could reflect to some extent permanently displaced outdoor consumption trips as households seek to mitigate pollution exposure.

As a point of reference, Cutter and Neidell (2009) find that when a "Spare the Air" alert is issued in the San Francisco Bay Area, *daily* traffic falls by 2.5-3.5% with the largest effect during and just after the morning commuting period.[31] Graff Zivin and Neidell (2009) estimate that a *one-day* smog alert issued in Southern California leads to an 8-15% reduction in attendance at two major outdoor facilities (the Los Angeles Zoo and the Griffith Park Observatory) on the same day, though the effect dissipates quickly in consecutive days of such

[30] "Region" is a conventional partition of cities by location: north (36 cities), northeast (38 cities), east (105 cities), central south (81 cities), southwest (54 cities), and northwest (52 cities).

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

alerts. These two studies focus on immediate (daily) behavioral changes after government-issued air quality warnings while our elasticity estimates measure behavioral responses to air quality over the course of a week following the implementation of the information program.

In Appendix C, we report three sets of additional analyses that support our main findings. First, we examine heterogeneity between "deferrable" and "scheduled" consumption (Appendix Tables C.4 and C.5).[32] Deferrable categories include supermarkets, dining, and entertainment. These categories experience a 4 to 9 percentage point increase (in magnitude) in purchase-pollution elasticity with the most stringent set of controls, and they 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 trips for paying bills (e.g., for utilities, insurance, telecommunication, and cable services), paying for government services (e.g., court costs, fines, and taxes), business-to-business wholesale transactions, as well as spendings at cancer treatment centers. There is no statistical evidence that information availability changes the responses of "scheduled" consumptionto air pollution.

Second, we conduct a battery of robustness checks in Appendix Table C.6. To highlight a few examples, we find that the inclusion of flexible weather controls is not consequential to our estimation, that online transactions cannot explain away our findings, and that our conclusion holds for cities without the presence of a U.S. embassy or consulates office. (These offices have independent $PM_{2.5}$ monitoring, and so residents in these cities might have had better information on air quality prior to 2013.) In addition, our results are robust to excluding top 10% of the cities with the highest number of officials who were ousted as a result of China's sweeping anti-corruption campaign, which has also associated with a reduction in luxury consumption (Qian and Wen, 2015). Finally, results are similar when we use maximum pollution readings instead of average pollution readings.

To more comprehensively account for geographically correlated unobservables that might confound our identification, we implement a more saturated research design; for each wave of "treatment" cities, we introduce a group of "control" cities that neighbor the treatment cities, but have not yet experienced monitoring. These "control" cities are assigned the same roll-out time as their treated counterparts. We then estimate and compare *differential* change in the transaction-pollution gradient across the treatment and control cities.[33] Appendix

[32]To facilitate interpretation of the magnitude across different categories, we use the inverse hyperbolic sine function (ArcSinh) of transactions, so that the key coefficient can be roughly interpreted as elasticities. The ArcSinh function is preferable to logs due to a non-trivial fraction of zero transactions in some categories in small cities from the 1% card sample. The patterns are very similar if we use levels instead of ArcSinh.

[33]The logic of this test is similar to a triple difference design that compares purchase behavior before and after the information program, in treatment and control cities, on high and low pollution days. Notice cities in the third wave (the last wave) of the roll-out do not have control cities. We include third-wave cities in

Table C.7 summarizes the results. Two patterns emerge. First, the triple interaction terms suggest a differential change in the transaction-pollution gradient in treatment cities relative to that of their neighbors by an amount similar to our main effect estimate reported in Table 1. Second, the double interaction terms show that there are no detectable changes in the transaction-pollution gradient in the neighboring cities at the time when the information program began in the treated cities. These findings suggest an isolated effect that occurs only in cities that actually experience monitoring.[34]

6.2 Pollution Disclosure and Housing Choices

Our analysis so far suggests that the information disclosure program has resulted in short-term behavioral changes that reflect greater mitigation efforts in response to air pollution. To examine the change in consumer response to pollution in the long run, we turn to the capitalization of air quality in the housing market before and after the program. The housing market has been used as a classical market to study consumer preference for non-marketed environmental goods whereby a housing property can be viewed as a bundle of attributes including the environmental quality (Oates, 1969; Chay and Greenstone, 2005; Banzhaf and Walsh, 2008; Bayer et al., 2016). The price differential among properties that have different environmental qualities but are otherwise similar would help us infer consumer preferences for environmental quality. Many studies have used the housing market response to examine the impact of one of the most prominent U.S. pollution information programs, the Toxic Release Inventory (TRI), which publicizes toxic emissions reports from major emitters (Bui and Mayer, 2003; Oberholzer-Gee and Mitsunari, 2006; Konar and Cohen, 2001; Mastromonaco, 2015).

Our analysis focuses on Beijing for which we have access to nearly the universe of transactions of involving new homes sold from January 2006 to April 2014. Analogous to the previous analysis in Section 6.1, we study the housing price-pollution relationship across neighborhoods with varying degrees of air pollution. Our parameter of interest is the degree to which this relationship shifts before and after the information program is implemented in Beijing in January 2013. The municipality of the Metropolitan Beijing area is divided into 16 districts, which are further divided into 180 communities and 1,200 apartment complexes.

our regression simply as treated units. With reduced power, we can estimate our triple difference design leaving out third-wave cities altogether, and we find the results do not change qualitatively.

[34]Our triple difference effect size and precision drop with the inclusion of region fixed effects as shown in columns 3 and 4 of Appendix Table C.7. We find this is largely due to the lack of independent variations among treated and neighboring areas within a region, especially for western China where cities are large in geographic extent, as shown in Figure 2. In unreported analysis, we show that we restore effect size and precision by simply using a less stringent "region" definition where the northwest and southwest regions are counted as one region.

A community is comparable to a ZIP Code in the United States in terms of geographical coverage, while an apartment complex is similar to a U.S. census block group. Our analysis examines how complex-level housing prices vary with pollution.[35]

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

$$\ln \mathrm{TransactionPrice}_{ict} = \mathrm{w}'_{ict}\gamma + \eta_{cy} + \varepsilon_{ict}, \tag{5}$$

where $\ln \mathrm{TransactionPrice}_{ict}$ 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 represents apartment-complex-by-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-by-year level, $\hat{\eta}_{cy}$, we examine the relationship between housing prices and pollution using a framework similar to equation (3):

$$\hat{\eta}_{cy} = \alpha \cdot \ln \mathrm{P}_{cy} + \beta \cdot \ln \mathrm{P}_{cy} \times \mathbb{1}(\text{after monitoring}) + \mathrm{x}'_{cy}\gamma + \varepsilon_{cy} \tag{6}$$

where β captures the change in the pollution-housing gradient before and after the program. We use two different measures of pollution at the sub-city level: fine-scale ambient air quality (AOD) at the 1km-by-1km $\times$ year resolution and distance to major polluters. We discuss each of these two sets of results in our analysis.

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 reprocesses the original AOD data to increase 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 is the case in our situation, in which we only need annual pollution data), it is possible to average across the overlapping overpasses to

[35] Because we only have 16 months of transactions following the implementation of the treatment, we skip the event study dynamics and only estimate the change in the price-pollution gradient.

enhance the geospatial resolution of the AOD measure.[36] 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-1 km region that contains the apartment-complex c.

We report results for equation 6 in Table 2. Column 1 includes apartment-complex fixed effects, year fixed effects (based on the transaction date of each unit), and year-on-market (when the complex became available) fixed effects. Column 2 is similar but additionally controls for local pollution in the previous year. These specifications exploit the fact that we sometimes observe transactions in the same apartment-complex to occur over different years. Columns 3 and 4 replicate columns 1 and 2 but control for community fixed effects and district-by-year fixed effects, in addition to year-on-market fixed effects. These specifications compare transaction prices within the same district and year, but across apartment-complexes with high and low pollution levels, controlling for time-invariant differences in community-level characteristics. To flexibly account for potential autocorrelation in both housing prices and pollution across time and over space, we two-way cluster standard errors at the community level and the district-year level.

Prior to the information program, a doubling 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 varies from 59 percentage points (Column 1) to 85 percentage points (Column 3) and significant at the 10% confidence level. In other words, housing prices do not respond to variation in pollution levels before the program, while after the program, housing prices capitalize air quality.

Our estimates of Beijing's housing price-pollution elasticity for the post-monitoring period therefore ranges from -0.6 to -0.8. One standard deviation increase in pollution is associated with a 4-6% reduction in housing price. This is slightly larger than those obtained in the U.S. setting, but comparable to those obtained in China's context. For example, by exploiting permanent reductions in total suspended particle (TSP) pollution in the wake of the U.S. Clean Air Act provisions undertaken in 1970 and 1977, Chay and Greenstone (2005) 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. In a hedonic regression exercise using Beijing's housing transactions and land parcel data, Zheng and Kahn (2008) find a price-PM_{10} elasticity of -0.41. Using moving costs and housing value information from the China Population Census micro-level data, Freeman et al. (2019) estimate a price-$PM_{2.5}$ elasticity of -0.71 to -1.10.

[36] Appendix Figure C.7 illustrates the oversampling idea using two consecutive days of MODIS AOD data.

One might be concerned about using the fine-scale AOD information in the hedonic analysis, given there were less than 30 monitoring stations in the metropolitan area of Beijing during our data period. In practice, some residents are likely to have more localized pollution information than that provided by monitoring stations. For example, air purifiers have become a common household appliance since 2013, and they provide real-time $PM_{2.5}$ readings. Nearly a quarter of a million units of air purifiers were sold in Beijing in 2015 alone. In addition, portable $PM_{2.5}$ monitors are popular household products as well. Nevertheless, to address this concern, we next analyze the price-pollution relationship based on coarse pollution information next.

Proximity to Major Polluters and Housing Prices. Following the literature (e.g., Davis, 2011; Currie et al., 2015; Muehlenbachs, Spiller and Timmins, 2015), we use the distance to the nearest major pollution source as our second pollution measure. While large polluters might be visible landmarks in a city, the information program could raise the salience of the health impacts of these large polluters in residents' housing choice decisions.

Our distance-gradient analysis begins with the 41 polluters in Beijing that represent the top 10% of the city's pollution sources, and were in operation from 2007 to 2014. These heavy polluters account for nearly 90% of total emissions according to the 2007 CES emission inventory (Appendix Figure C.8). Using geolocations of all polluters, we construct a time-invariant "distance to top-decile polluter" variable as our second pollution measure while controlling for distance to other polluters. We drop apartment-complex fixed effects that are perfectly colinear with the distance measure, and we control for district-by-year fixed effects, community fixed effects, and year-on-market fixed effects.

Figure 9 presents the estimates graphically. 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 slope 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 three km of the top polluters experienced the largest depreciation of about 27%. The effect fades with distance and becomes insignificant over six km. In comparison, Currie et al. (2015) estimate an 11% reduction for properties located within 1km of a toxic plant in the United States. Gamper-Rabindran and Timmins (2013) show that the cleanup of Superfund sites leads to a 21% increase in the median housing price of the census tracts within 3-mile radius around the sites. Our estimate is larger but not implausible given the unprecedented housing boom in the city, where the average housing price in Beijing grew by 262% during our sample period, and the effect size corresponds to 42% of the interquartile range of the housing price dispersion.

6.3 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 examines whether the same amount of pollution exposure is associated with fewer deaths after pollution information becomes widely available. Our analysis uses mortality data in 131 cities from 2011 to 2015. Similar to the procedures outlined in Section 6.1, we conduct an event study and regress the logged mortality rate in county c and 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:

$$\ln \text{Mortality}_{ct} = \sum_{k=-10}^{6} \beta_k \times \ln \text{P}_{ct} \times \mathbb{1}(t = k) + \sum_{k=-10}^{6} \eta_k \times \mathbb{1}(t = k) + \text{x}'_{ct}\gamma + \varepsilon_{ct} \tag{7}$$

We made several specification choices based on the nature of our data. First, we aggregate the weekly mortality rate to a quarterly rate to better capture the longer-term impact; the qualitative findings remain the same whether we conduct our analysis at the weekly, monthly, or quarterly level, with the β_k estimates smaller (but statistically more precise) using the weekly or monthly data. Second, we allow the β_k coefficients to vary from ten quarters before to six quarters after the information program to ensure a roughly balanced number of underlying counties for each event quarter. The remaining quarters are grouped in a separate dummy variable.

Figure 10 plots the β_k coefficient estimates for the event study in which we control for city, quarter-of-year, and year fixed effects. The mortality-pollution elasticity exhibits a flat trend before the program, followed by a noticeable decline after the program. This decline appears to strengthen after a couple of quarters; however, our sample is too short to examine the dynamic response formally. We repeat the analysis in Table 3, replace the event dummies with interactions between the pollution term and the post-treatment dummy, and experiment with increasingly stringent controls varying from city and quarter fixed effects to city and region-by-quarter-of-sample fixed effects. The coefficient estimate suggests a statistically significant 5 percentage point reduction in the mortality-pollution elasticity after the program. The results are consistent with the graphical evidence in Figure 10. In addition, the estimates are remarkably similar across specifications with different fixed effects, suggesting a limited role of potential confounding factors.[37]

Our heterogeneity analysis provides suggestive evidence of the underlying mechanism behind the mortality effect. Specifically, we repeatedly split the sample into two using the average value of a series of city-level characteristics, including per capita income, share of the urban population, per capita number of hospitals, per capita residential electricity use, and share of mobile phone users. Panel A of Table 4 reports the results that focus on the interaction between the change in the pollution gradient and city-level characteristics. Columns 1 to 5 tabulate the interaction coefficient for each of the five city characteristics. While there is no heterogeneity between cities with above- or below-average per capita income, there exists a larger reduction in mortality damage (between 7 and 8 percentage points) in cities that are more urban, have more hospitals, have a higher rate of residential electricity use, and have a higher mobile phone penetration level. These findings are consistent with the hypothesis that residents in these cities are more likely to benefit from pollution information and, at the same time, engage in defensive activities to reduce the health damage from air pollution exposure.

To better understand the pathway from avoidance to health outcomes, we examine heterogeneity in short-run outdoor consumption trips based on debit card and credit card transactions across the same set of city attributes. Results in Panel B of Table 4 are consistent with the patterns in Panel A: residents in cities that are more urban, have more hospitals, have a higher rate of residential electricity use, and have a higher mobile phone penetration level exhibit a stronger and more precisely estimated behavior response to elevated pollution. While there are many defensive mechanisms, the evidence here suggests one plausible pathway from effective avoidance to improved health outcomes.

In Appendix C, we conduct a series of additional tests to examine the plausibility of the reduction in the mortality-pollution elasticity estimates. First, Appendix Figure C.9a examines age-specific mortality rates. The effect is most precisely estimated among people aged over 40 (including age groups 40-49, 50-59, 60-69, and 70+) who are more vulnerable to pollution exposure than younger age groups. There is no change in the mortality-pollution relationship for infants under one year of age. This could be associated with the low pol-

[37] The event study of Figure 10 suggests the effect size grows over time. In recent work, Goodman-Bacon (2018) concludes that the presence of time-varying treatment effect may bias difference-in-differences (DID) estimates of the true effect. This happens because in a DID setting with staggered treatment, treated units essentially serve as "controls" for other treated units that receive treatment later on, and the time-varying treatment effects put units on a differential trend. Following Goodman-Bacon (2018), we correct for this potential bias using a decomposition method that discards variations coming from comparisons of "switchers" and those "already treated". In the Goodman-Bacon language, we use 66% of the variations coming only from comparing switchers and units that have not yet received treatment. The adjusted effect size is -0.049, similar to our baseline estimate of -0.055.

lution exposure among infants due to the traditional Chinese practice of keeping infants strictly indoors within the first few months of their birth to minimize the outdoor exposure. Somewhat surprisingly, the 10-19 age group experiences a significant reduction in mortality, although this accounts for a small number of total mortality reduction because the baseline mortality rate is the lowest across age groups, as shown in Figure C.9a.[38]

Appendix Figure C.9b illustrates that changes in the mortality-pollution relationship concentrate in cardiorespiratory causes, such as chronic obstructive pulmonary diseases, heart diseases, and cerebrovascular diseases, which are widely considered as the most relevant consequences of pollution exposure. The impact on respiratory infection and digestive diseases is both small and insignificant. For traffic fatalities, the pollution-mortality relationship after the program's implementation appears to flatten, though the change is not statistically significant.[39] Lastly, an exploration of nonlinear specifications finds that the reduction in the mortality-pollution gradient is insignificantly convex in the level of pollution shock (Appendix Figure C.10).

7 The Value of Pollution Information

As outlined in the theoretical model, the value of information (VOI) arises from the power of information in changing decisions which reduces the wedge between the experience utility and the decision utility. Our analyses illustrate that disclosing pollution information has affected a range of behavioral and market outcomes that reflect households' efforts to mitigate the negative health consequences of air pollution. In this section, we quantify the value of the information program by estimating the benefit and cost components outlined in equation (2) in the theory section.

We estimate the benefit of program using two methods based on different assumptions. Our first method assumes that the vast majority of the value of the information program comes from reduced mortality, which we quantified using age-adjusted value of statistical life (VSL). This approach is commonly used among policy makers to evaluate benefits of environmental policies. Due to the lack of reliable population-wide VSL estimates based on Chinese data, we use the benefit transfer approach that infers the VSL for Chinese residents from US-based VSL estimates. This approach is commonly used to provide population average VSL for countries with insufficient or unreliable data based on VSL estimates from US and the income elasticity of VSL (Viscusi and Masterman, 2017). We use the income

[38]Anecdotally, some middle and high schools would limit student outdoor schedules in response to a high pollution period, e.g., <cn.nytimes.com/education/20130525/cc25PE/>.

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

elasticity of VSL (i.e., VSL transfer elasticity) of 1.2 as suggested by Narain and Sall (2016). [40] We then use the age-adjusted VSL from Ashenfelter and Greenstone (2004) and Murphy and Topel (2006), and scale the VSL for the Chinese population by age accordingly. Following these two steps, we obtain VSL estimates of RMB 1.3 million for population ages 0-39, and RMB 0.13 million for ages 85 and above. Based on the mortality impact by age groups shown in Figure C.9a, the benefit from mortality reduction is about RMB 520 billion per year.[41]

The second method uses the concept of consumer willingness to pay (WTP) for clean air. Compared to the VSL approach, value estimates based on consumer WTP better encompass the first two terms in Equation (2) of Secion 3, capturing direct health benefits, productivity benefits, and improvements in the quality of life. To apply the idea of WTP for clean air in our context, note that the analysis on the mortality-pollution relationship in Section 6.3 shows that the program reduced pollution-caused deaths, which could be equivalently achieved through reduction in air pollution. We begin by calculating the fraction of pollution-caused deaths that are avoided by providing information access, *holding pollution exposure constant*:

$$\frac{\text{Mortality Reduction}}{\text{Baseline Mortality}} = \frac{\epsilon_1 - \epsilon_0}{\epsilon_0}, \tag{8}$$

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 (3), and is the interaction coefficient ($\approx$ 0.05) reported in Table 3. To obtain a consistent estimate of ϵ_0, we replicate the regression discontinuity (RD) analysis in Ebenstein et al. (2017) that examines the long-term mortality effects of PM exposure. We favor this study because it is based on a well-established, quasi-experimental method, and uses a similar data source for mortality measurement.[42] Using an 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. We obtain very similar baseline mortality estimates as shown in Appendix Figure C.11 and Appendix Table

[40] Narain and Sall (2016) suggest a transfer elasticity of 1.2 for transferring the VSL from U.S. or a developed country to a developing country with low or middle income. China's per capita income is about an eighth of that in the U.S.

[41] Health benefits may also emerge in the format of reduced health spending. Barwick et al. (2018) estimate that an individual saves RMB 38 in out-of-pocket health spending from a 5 ug/m^3 reduction in $PM_{2.5}$ exposure. This figure aggregates to RMB 52 billion per year nationwide.

[42] In unreported analysis, we use the instrumental variable approach in Barwick et al. (2018) that exploits long-range transport of pollution from upwind cities, and, using this approach, we also obtain similar conclusions on the baseline mortality impact.

C.8.[43] Our analysis therefore suggests a 7% reduction in the mortality-pollution elasticity (0.05 divided by 0.70). Assuming a linear dose-reponse function between pollution and mortality, the information program thus delivers the same benefit as a 7% reduction in pollution concentration. This corresponds to a 10 ug/m^3 reduction in PM_{10} or 5 ug/m^3 reduction in $PM_{2.5}$ in China.[44] To convert this to value, we follow Ito and Zhang (2018) who provide a quasi-experimental estimate of WTP based on air purifier purchases from 2006 to 2014 in 80 cities in China. Their estimates suggest that a 10 ug/m^3 reduction in PM_{10} is associated with consumer surplus gains in the order of around RMB 130 billion per year nationwide.[45]

In theory, the WTP approach captures a more comprehensive range of benefits – such as health benefits as well as quality of life benefits – than the VSL approach that focuses solely on mortality benefits. In contrast, we find a relatively *smaller* benefit using the WTP approach than using the VSL approach. As pointed out by Ito and Zhang (2018), the WTP estimates are likely to be lower bounds because, even after the advent of the information program, consumers in developing countries may not have the extent of information available to consumers in developed countries regarding the health consequences of pollution. Our exercise also relies on more parameters drawn from the quasi-experimental literature in the context of air pollution exposure.

We now turn to the estimate of the total costs which include increased defensive spending, the welfare loss from foregone consumption, and the cost of the program itself. First, total sales of air purifiers and $PM_{2.5}$ masks have increased at a rate of RMB 7 billion and RMB 0.55 billion per year post 2013, respectively. Because cities in waves two and three started the information program toward the end of 2013 and 2014, these numbers are upper bounds on the increased defensive investments. Second, the total costs should also include the welfare loss from foregone consumption as a pollution avoidance strategy. Our estimates from Table 1 suggest that the upper bound of annual foregone consumption from the program would be RMB 457 billion (or 2.3 percent of total debit card and credit card transactions). Assuming

[43] It is perhaps important to point out that the magnitude of our OLS-based α estimate (the coefficient estimate on "Log(Pollution)" in Table 3) 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). For example, Ebenstein et al. (2017) report an OLS regression between logged cardiorespiratory mortality and logged PM_{10} exposure. This yields a correlational elasticity estimate of 0.02, which is similar to our OLS estimate.

[44] We perceive the effect size as plausible for several reasons. First, these magnitudes are moderate compared to the average cross-city variation in $PM_{2.5}$ post 2013, which has a standard derivation of 20.4 ug/m^3 and an interquartile range of 25.2 ug/m^3. Second, several government programs have been shown to shift pollution levels significantly. For example, the winter heating policy implemented to the north of the Huai River has been shown to have increased PM_{10} by about 41.7 ug/m^3 (Ebenstein et al., 2017).

[45] Interestingly, although Ito and Zhang (2018) do not analyze the effect of the information program, they do find that, on average, the WTP for clear air increased after 2013, consistent with stronger consumer awareness of pollution.

a price elasticity of demand of -1 or -2, the implied loss of consumer surplus would be about RMB 5.26 or 2.63 billion, respectively. The price elasticity of most consumer goods and services tends to be between -0.5 and -1.5 (Deaton and Muellbauer, 1980; Blundell, Pashardes and Weber, 1993; Banks, Blundell and Lewbel, 1997). Using the price elasticity of -1 hence provides an upper-bound estimate of the consumer welfare loss from foregone consumption. Third, the one-time cost to set up the monitoring stations and broadcast the pollution information online is estimated to be RMB 2 billion to RMB 5 billion, and the annual operation cost (staff and maintenance) is about RMB 0.5 billion a year. Taken together, the upper bound of the total cost is about RMB 18 billion in its first year (including the capital costs) and 13 billion annually thereafter.

The estimated benefits brought by the program, varying from RMB 130 billion to RMB 520 billion annually, relative to the associated costs of less than 20 billion annually, underscore the cost effectiveness of the information monitoring-and-disclosure program.[46] While this policy is not a substitute for other policies directly aimed at reducing pollution, the increased awareness about pollution among residents could put pressure on polluters, and could complement other environmental regulations in reducing pollution (Konar and Cohen, 1997). Perhaps more importantly, empowering the public with real-time pollution information mobilizes individuals' ability to mitigate the adverse consequences of pollution. This, together with other environmental regulations, can more effectively reduce the societal cost of pollution.

8 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 that resulted in the installation of air pollution monitoring stations, and provision of previously unavailable, real-time pollution information to the public. Based on several rich and unique data sets, our analysis provides strong evidence that the program led to a cascade of changes, including increased access to and awareness of information about pollution and its harm, more pronounced short- and long-term avoidance behaviors, and a reduction in the negative health effects. The findings suggest that the value of the information monitoring-and-disclosure program arising from improved health is at least one order of magnitude larger than its costs – both in terms of implementing and maintaining the program, and in terms

[46]Successful environmental regulations such as the US Acid Rain Program and the Clean Air Act have been shown to achieve benefit an order of magnitude larger than the cost (Krupnick and Morgenstern, 2002; Chestnut and Mills, 2005). Different from our context, these programs substantially reduced pollution levels. We are not aware of other studies that quantify the benefit and cost of pollution information programs.

of the behavior-related changes among the general public.

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 – or nonexistent – in many such countries. As income rises, the demand for environmental quality increases, and households are better able to adapt. Improving information access can be an initial step, providing a form of "low-hanging fruit" in the effort to find cost-effective ways to begin to address environmental challenges faced in these countries. Providing real-time pollution monitoring data, combined with effective dissemination infrastructure such as smartphones and the Internet, both of which are now commonly available among developing countries, could be a powerful tool to help households themselves mitigate health damages from environmental pollution, and to enhance the effectiveness of – and public demand for – other environmental regulations that go further in addressing pollution's health effects.

Finally, while our study is in the context of environmental quality, lessons gleaned from this large-scale information program could offer beneficial information in other settings such as traffic safety, risky health behaviors, as well as food and nutritional guidance. Well-designed information provision could mobilize household efforts and complement government regulations to help address market failures associated with information asymmetry and externalities, especially in developing countries.

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.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113(1): 151–184.

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.

Ashenfelter, Orley, and Michael Greenstone. 2004. "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." *American Economic Review*, 94(2): 454–460.

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

Banks, James, Richard Blundell, and Arthur Lewbel. 1997. "Quadratic Engel Curves and Consumer Demand." *The Review of Economics and Statistics*, 79(4): 527–539.

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.

Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins. 2016. "A Dynamic Model of Demand for Houses and Neighborhoods." *Econometrica*, 84(3): 893–942.

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.

Blundell, Richard, Panos Pashardes, and Guglielmo Weber. 1993. "What do we

Learn About Consumer Demand Patterns from Micro Data?" *The American Economic Review*, 83(3): 570–597.

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

Bui, Linda T. M., and Christopher J. Mayer. 2003. "Regulation and Capitalization of Environmental Amenities: Evidence from the Toxic Release Inventory in Massachusetts." *The Review of Economics and Statistics*, 85(3): 693–708.

Chang, Tom Y., Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2019. "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China." *American Economic Journal: Applied Economics*, 11(1): 151–72.

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.

Chestnut, Lauraine G., and David M. Mills. 2005. "A fresh look at the benefits and costs of the US acid rain program." *Journal of Environmental Management*, 77(3): 252 – 266.

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.

Deaton, Angus, and John Muellbauer. 1980. "An Almost Ideal Demand System." *The*

American Economic Review, 70(3): 312–326.

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.

Gamper-Rabindran, Shanti, and Christopher Timmins. 2013. "Does cleanup of hazardous waste sites raise housing values? Evidence of spatially localized benefits." *Journal of Environmental Economics and Management*, 65(3): 345 – 360.

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

Goodman-Bacon, Andrew. 2018. "Difference-in-differences with variation in treatment timing." National Bureau of Economic Research.

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.

Graff Zivin, Joshua, and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." *American Economic Review*, 102(7): 3652–73.

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.

Hanna, Rema, and Paulina Oliva. 2015. "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City." *Journal of Public Economics*, 122: 68 – 79.

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

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.

Konar, Shameek, and Mark A. Cohen. 1997. "Information As Regulation: The Effect of Community Right to Know Laws on Toxic Emissions." *Journal of Environmental*

Economics and Management, 32(1): 109 – 124.

Konar, Shameek, and Mark A. Cohen. 2001. "Does the Market Value Environmental Performance?" *The Review of Economics and Statistics*, 83(2): 281–289.

Krupnick, Alan, and Richard Morgenstern. 2002. "The Future of Benefit-Cost Analyses of the Clean Air Act." *Annual Review of Public Health*, 23(1): 427–448.

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.

Murphy, Kevin M., and Robert H. Topel. 2006. "The Value of Health and Longevity." *Journal of Political Economy*, 114(5): 871–904.

Narain, Urvashi, and Chris Sall. 2016. *Methodology for Valuing the Health Impacts of Air Pollution: Discussion of Challenges and Proposed Solutions.* The World Bank.

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.

Newell, Richard G., and Juha Siikamäki. 2014. "Nudging Energy Efficiency Behavior: The Role of Information Labels." *Journal of the Association of Environmental and Resource Economists*, 1(4): 555–598.

Oates, Wallace E. 1969. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy*, 77(6): 957–971.

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

Qian, Nancy, and Jaya Wen. 2015. "The Impact of Xi Jinping's Anti-Corruption Campaign on Luxury Imports in China." *Northwestern Working Paper.*

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.

Tu, Meng, Bing Zhang, Jianhua Xu, and Fangwen Lu. 2020. "Mass media, information and demand for environmental quality: Evidence from the "Under the Dome"." *Journal of Development Economics*, 143: 102402.

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).

Viscusi, W. Kip, and Clayton J. Masterman. 2017. "Income Elasticities and Global Values of a Statistical Life." *Journal of Benefit-Cost Analysis*, 8(2): 226–250.

Wainwright, Oliver. 2014. "Inside Beijing's airpocalypse – a city made 'almost uninhabitable' by pollution." *The Guardian*.

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, Junjie, Zhenxuan Wang, and Xinming Du. 2017. "Lessons learned from China's regional carbon market pilots." *Economics of Energy & Environmental Policy*, 0(Number 2).

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.

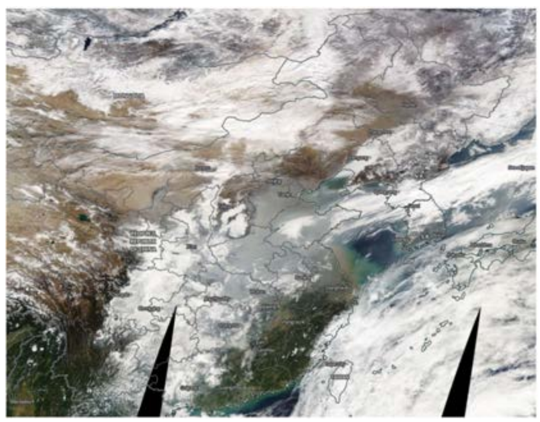
Zheng, Siqi, and Matthew E Kahn. 2008. "Land and residential property markets in a booming economy: New evidence from Beijing." *Journal of Urban Economics*, 63(2): 743–757.

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.

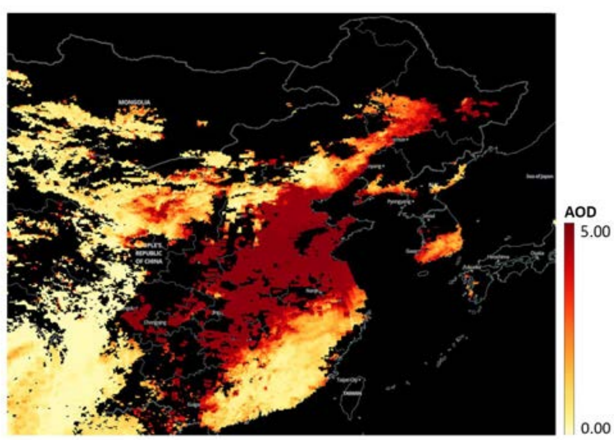
Zivin, Joshua Graff, and Matthew Neidell. 2009. "Days of haze: Environmental information disclosure and intertemporal avoidance behavior." *Journal of Environmental Economics and Management*, 58(2): 119 – 128.

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

(a) News coverage



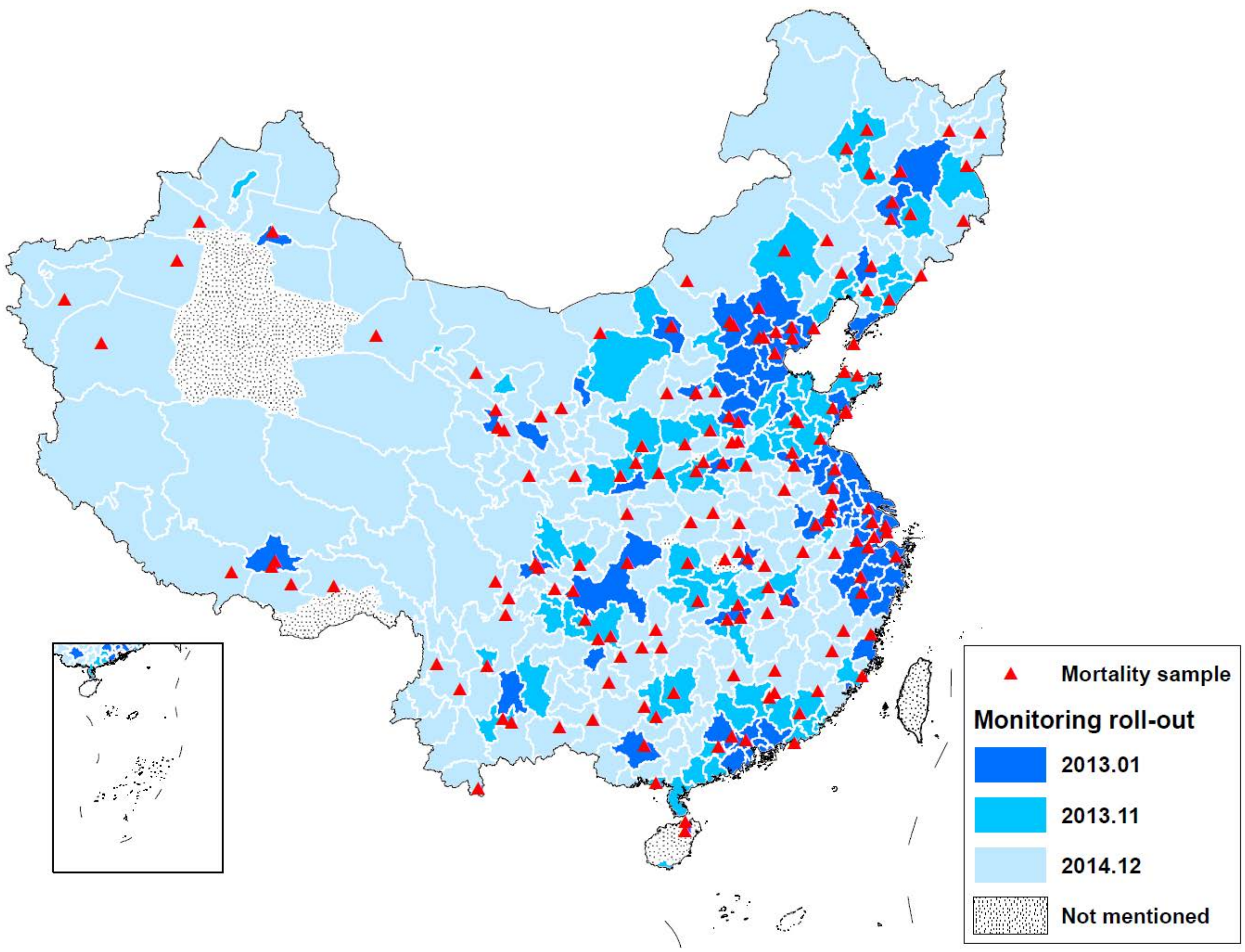
(b) Satellite picture of the event



(c) Satellite-retrieved pollution levels

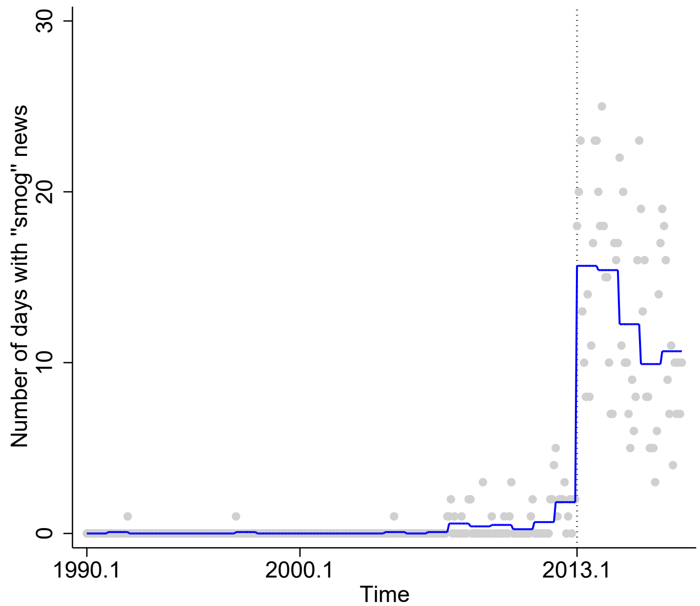
Notes: This figure illustrates a “widespread, dense fog event” on November 27, 2011 which is likely a major pollution event. Panel A, sourced from China Meteorological Administration, shows official news coverage of the event. Panel B, sourced from NASA, shows the satellite view of China on the same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

Figure 2: Information Program Roll-Out and Location of Mortality Data

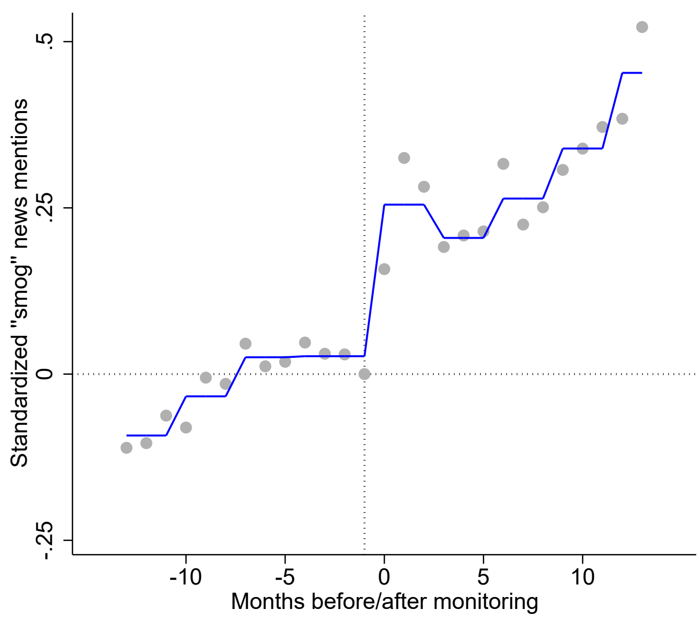


Notes: This map shows prefecture-cities by the completion date of the air pollution monitoring-and-disclosure program. "Not mentioned" are cities where the timing of monitoring is not mentioned in the government's policy notice. "Mortality sample" are centroids of counties included in the DSP mortality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

Figure 3: Changes in Pollution Information Access: News Mentions



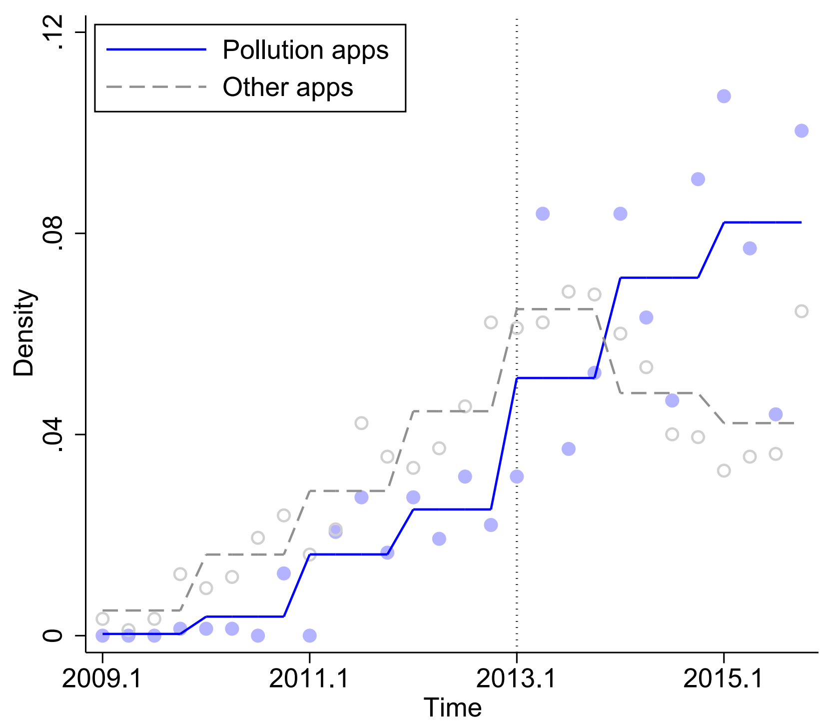
(a) *People's Daily*'s "smog" news mentions



(b) "Smog" mentions before and after monitoring

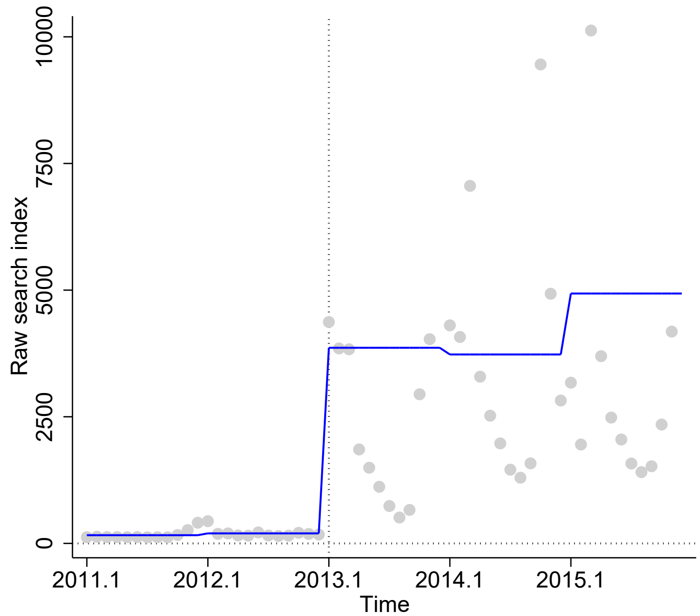
Notes: Panel A plots the number of days in each month when the *People's Daily* (the official newspaper of the Chinese government) published articles containing "smog" in content. Each dot represents a month. Line shows annual averages. Panel B plots standardized city-level "smog" mentions, defined as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

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

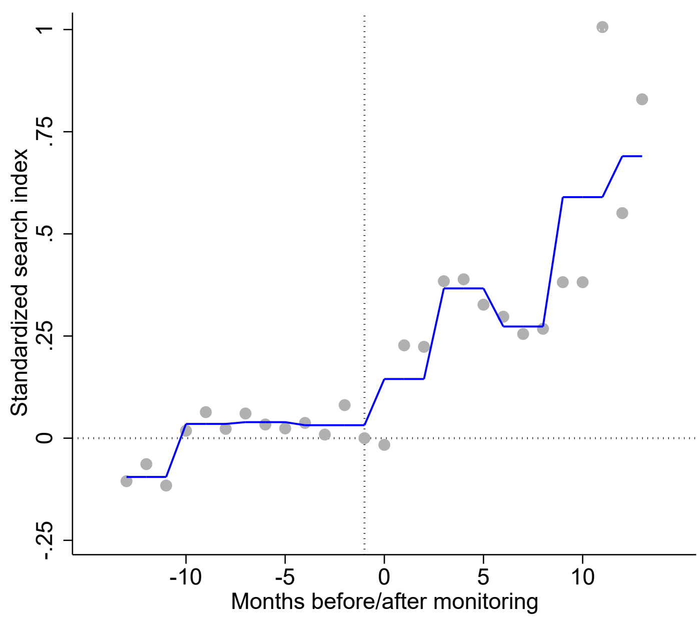


Notes: This chart shows release-date distribution of Apple App Store apps related to pollution (solid dots and line). Release-time distribution for apps in other categories (dashed dots and line) includes game, music, video, reading, finance, sports, education, shopping, and navigation. For each category, 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 incorporate real-time air quality content post 2013. These apps are (re)categorized as pollution apps when we queried for the Appstore data in 2015.

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



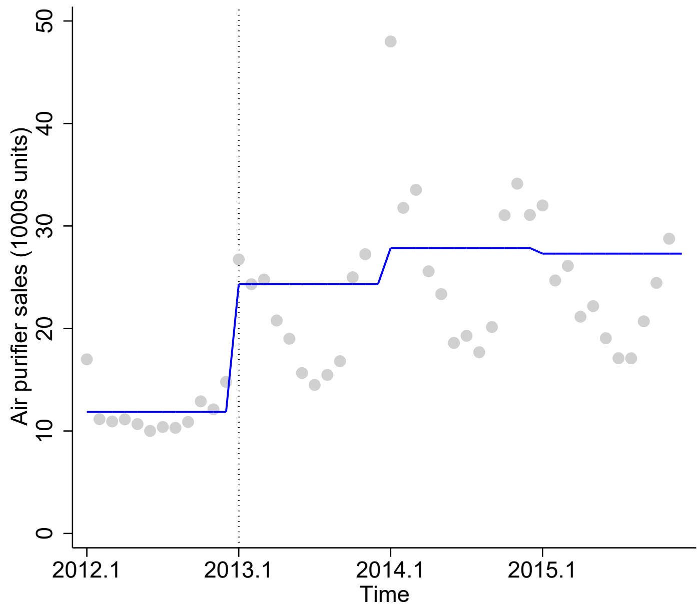
(a) Baidu "smog" search index at the national level



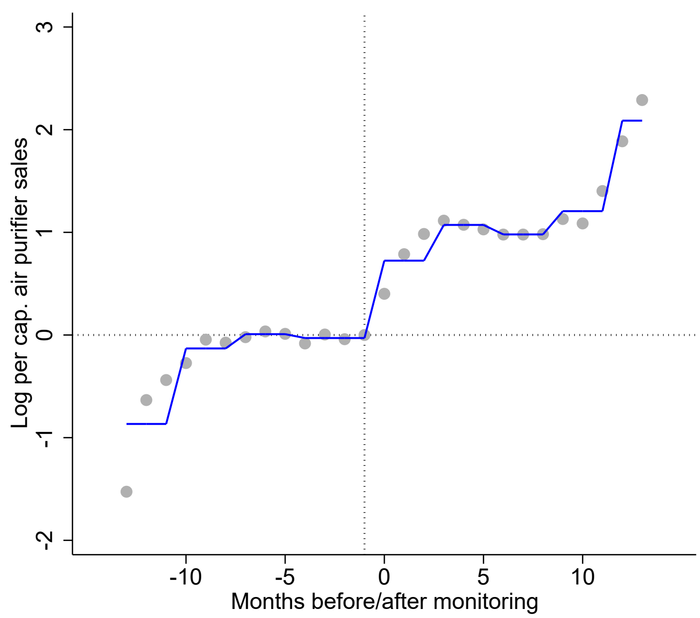
(b) Baidu "Smog" search index before and after a city implements the information program

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 standardized "smog" search index as a function of months since the completion of the information program in a given city. The search index is normalized to zero for event month -1. The underlying regression controls for month-of-year and year indicators. Line shows quarterly averages.

Figure 6: Changes in Pollution Awareness: Air Purifier Sales in 50 Cities



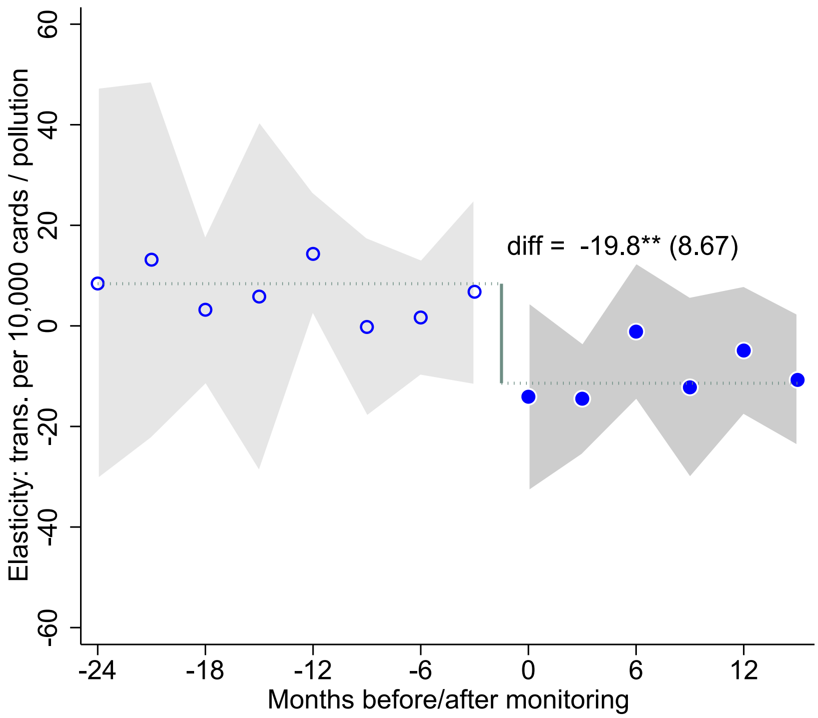
(a) Total air purifier sales



(b) Air purifier sales before and after monitoring

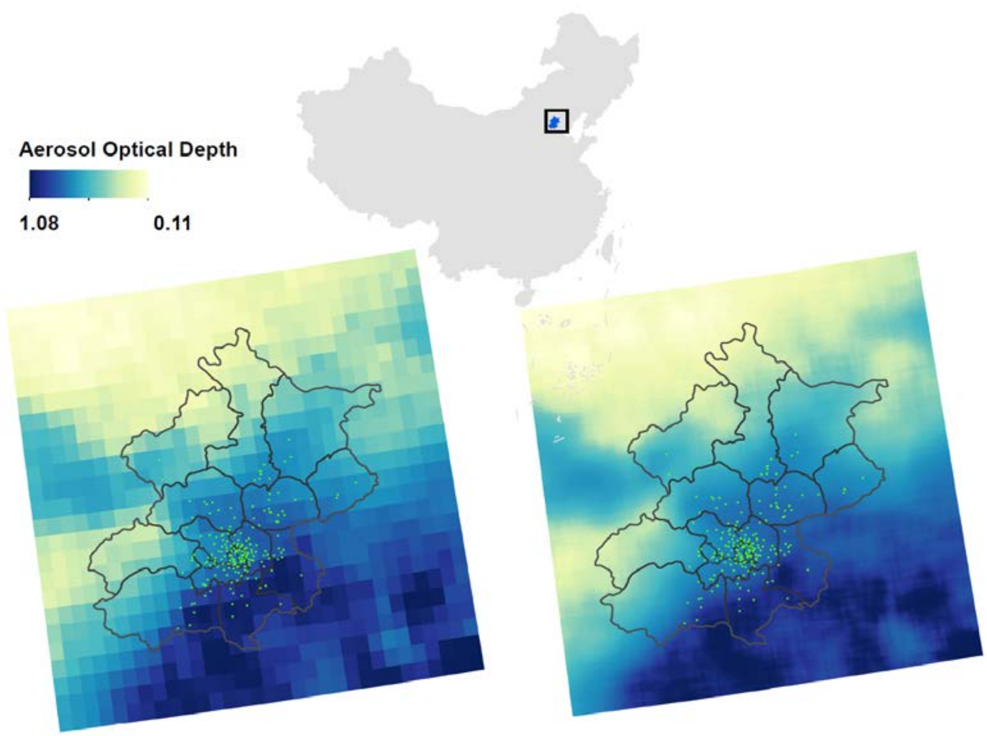
Notes: Panel A plots raw monthly trends in national air purifier sales from offline venues. The graph omits two dots with exceptionally high sales for readability purpose. These dots correspond to December 2013 (sales = 61,605 units) and Decembet 2015 (sales = 74,352 units). Line shows annual averages.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

Figure 7: Changes in Short-Run Avoidance: Weekly Card Transaction-Pollution Gradient



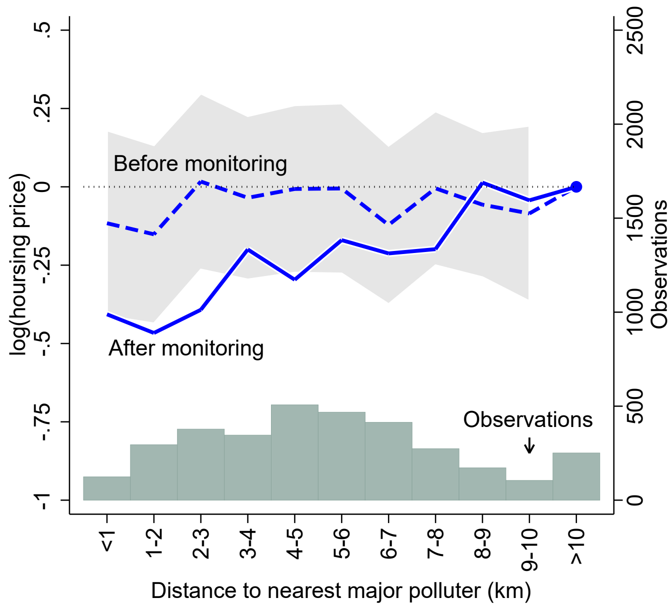
Notes: This graph shows the relationship between weekly bank card transcation rate and log satellite-based pollution as a function of time since the completion of the information program in a given city. Each dot is a regression coefficient. The regression controls for prefecture-city FEs, week-of-year FEs, and year FEs. Regressions are weighted by the number of active cards in a 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

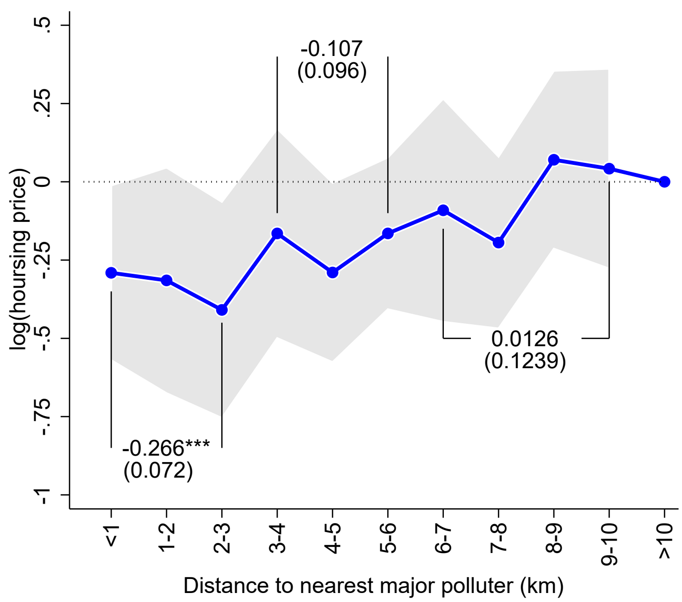


Notes: This map shows the 2006-2014 average aerosol optical depth (AOD) level for the municipality of Beijing. Left panel shows MODIS AOD at the original 10×10km resolution. Right panel shows AOD oversampled to 1×1km resolution. Dots show centroid locations of communities in the housing transaction data.

Figure 9: Changes in Long-run Capitalization: Housing Price against Polluter Distance, Beijing



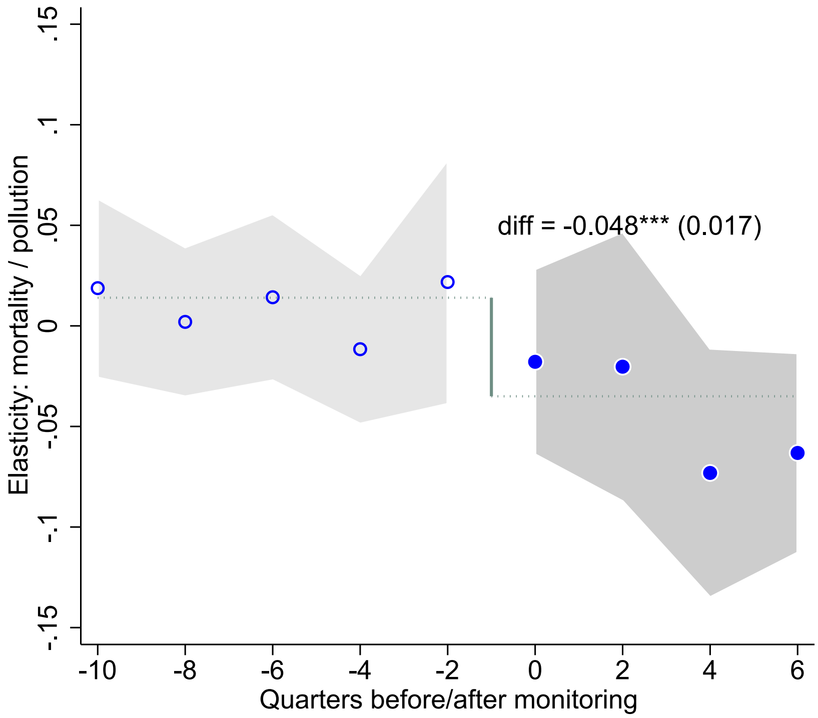
(a) Before vs. after monitoring



(b) Change in price-distance to polluter gradient, by distance bins

Notes: This graph shows coefficients from regressions of attribute-adjusted complex×annual log housing prices on distance in 1-km bins to the nearest major polluter before and after January 2013 when Beijing initiated ambient pollution monitoring. In Panel A, estimations are done separately for periods before (dashed line) and after (solid line) the information program, with prices normalized to zero for the >10-km bin. The histogram (right axis) plots the total number of observations by distance bins. Panel B pools the sample and estimates the difference. All regressions control for district×year FEs, community FEs, and years-on-market FEs. Shaded region shows the 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 Health Outcome: 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 the completion of the information program in a given city. Each dot represents a coefficient estimate. The -10 to 6 quarter event window is chosen so that the underlying sample for each reported coefficient is a balanced panel of cities. The underlying regression controls for prefecture-city FEs, quarter-of-year FEs, and year FEs. Shaded region shows the 95% confidence interval constructed from standard errors clustered at the prefecture-city level. Number of observations = 2,620.

Table 1: Changes in Short-Run Avoidance: Weekly 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 a city×week. Mean of the dependent variable is 869.1. The coefficient for the Log(Pollution) × 1(after monitoring) interaction term is equivalent to a three-percentage point change in the transaction-pollution elasticity. "region" is a conventional partition of cities by location: north (36 cities), northeast (38 cities), east (105 cities), central south (81 cities), southwest (54 cities), and 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 Long-run Capitalization: Beijing's Housing Price-Pollution Gradient

Dep. var.: Log housing price index in a complex×year				
	(1)	(2)	(3)	(4)
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: Beijing is divided into 16 districts, 180 communities, and roughly 1,200 apartment complexes. A community is comparable to a ZIP Code in the United States. in terms of geographical coverage while an apartment complex is similar to a census block-group. This analysis examines how complex-level housing prices vary with pollution. 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 at the oversampled 1km resolution corresponding to the complex's geographic coordinates. Standard errors are two-way clustered by community and district×year. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table 3: Changes in the Health Outcome: 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 a city×quarter. "region" is a conventional partition of cities by location: north (36 cities), northeast (38 cities), east (105 cities), central south (81 cities), southwest (54 cities), and 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 the Health Outcome: Heterogeneity by City Characteristics

	(1)	(2)	(3)	(4)	(5)
City characteristics:	Per cap. income	Frac. urban	Per cap. hospitals	Per cap. residential electricity	Per cap. mobile phones
Panel A. 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
Panel B. 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

Notes: This table reports heterogeneous changes in the mortality-pollution gradient (panel A) and purchase-pollution gradient (panel B) by above and 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. City characteristics are computed as 2011-2015 averages. Cities with missing attributes are omitted from the analysis. "Equality $p-$value" tests for equality between the above/below average coefficients. All regressions control for city, month-of-sample, and region-by-year fixed effects, as well as 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:

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

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

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

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):

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

Applying the implicit function theorem to f, we obtain:

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

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, $C + D \leq 0$.[47] 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

[47] At the optimal a and X, $U_h + \lambda \cdot w \cdot g(h) > 0$ 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.

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. The total derivative of health w.r.t. pollution is:

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

where $0 < dc_0/dc < 1$. 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:

$$\frac{dh}{dc}\,|_{c} = \frac{\partial h}{\partial c} + \frac{\partial h}{\partial a} \cdot \frac{da}{dc} \geq \frac{dh}{dc}\,|_{c_0}$$

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:

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

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

Appendix B: Proof of Proposition 2

To examine the impact of the information program on the pollution-outcome relationship, our analysis uses the framework as outlined in equation 3, in matrix form:

$$Y = \alpha \times P + \beta \times P \cdot \mathrm{d} + X\gamma + \varepsilon, \tag{B.2}$$

where '$\cdot$' is an element-by-element product. Y, P, d, and ε are N by 1 vectors, X is a N by k matrix. α, β are scalors while γ is a k by 1 vector. P measures ambient air quality and could be correlated with ε due to unobservables or measurement error as discussed in the main text. d represents the treatment dummy and is equal to one based on the staggered roll-out schedule. The key parameter of interest is β, the change in the slope of pollution-outcome relationship.

To simplify proof, we first partial out regressors X. Let M_{x} denote the projection matrix: $M_{\mathrm{x}} = I - X(X'X)^{-1}X'$. Multiplying both sides of equation B.2 with M_{x}, we have:

$$M_{\mathrm{x}}Y = \alpha M_{\mathrm{x}}P + \beta M_{\mathrm{x}}P \cdot d + M_{\mathrm{x}}\varepsilon$$

where $M_{\mathrm{x}}P$ is an N by 1 vector, the projection residual of P on X. Collect the two key regressors in $Z = [M_{\mathrm{x}}P, M_{\mathrm{x}}P \cdot d]$. Here we prove Proposition 2 under Assumption B1: $\varepsilon \perp d|X$ and Assumption B2: $\mathrm{d} \perp P|X$.

Proof: Let the OLS estimates of α and β be denoted as $\hat{\alpha}$ and $\hat{\beta}$.

$$\begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} = \begin{pmatrix} (M_{\mathrm{x}}P \cdot d)'(M_{\mathrm{x}}P \cdot d) & -(M_{\mathrm{x}}P)'(M_{\mathrm{x}}P \cdot d) \\ -(M_{\mathrm{x}}P \cdot d)'(M_{\mathrm{x}}P) & (M_{\mathrm{x}}P)'(M_{\mathrm{x}}P) \end{pmatrix} * \begin{pmatrix} (M_{\mathrm{x}}P)'(M_{\mathrm{x}}\epsilon) \\ (M_{\mathrm{x}}P \cdot d)'(M_{\mathrm{x}}\epsilon) \end{pmatrix}$$
$$* \frac{1}{det[(Z'Z)]}$$

The probability limit of $(\hat{\beta} - \beta)$ converges to the following term multiplied by a constant:

$$-E[(M_{\mathrm{x}}P \cdot d)'(M_{\mathrm{x}}P)]E[(M_{\mathrm{x}}P)'(M_{\mathrm{x}}\epsilon)] + E[(M_{\mathrm{x}}P)'(M_{\mathrm{x}}P)]E[(M_{\mathrm{x}}P \cdot d)'(M_{\mathrm{x}}\epsilon)]$$

Assumption B2 implies that $E\left(d|M_{\mathrm{x}}P\right) = c$. Hence:

$$E[(M_{\mathrm{x}}P \cdot d)'(M_{\mathrm{x}}P)] = E[(M_{\mathrm{x}}P)'(M_{\mathrm{x}}P)E(d|M_xP)] = cE[(M_{\mathrm{x}}P)'(M_{\mathrm{x}}P)].$$

Assumptions B1 and B2 imply that:

$$E[(M_{\mathrm{x}}P \cdot d)'(M_{\mathrm{x}}\epsilon)] = cE[(M_{\mathrm{x}}P)'(M_{\mathrm{x}}\epsilon)].$$

Therefore, $p\lim(\hat{\beta} - \beta) = 0$ and the OLS estimate $\hat{\beta}$ is consistent.

Appendix C: Figures and Tables

Figure C.1: List of Cities by Roll-out Waves

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).

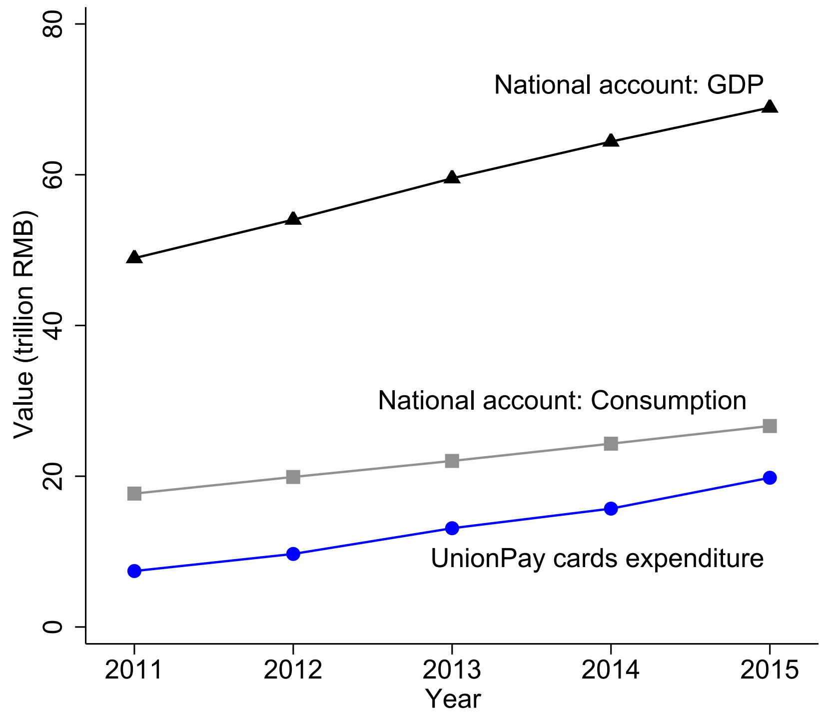
Figure C.2: Screenshot of the Government's Air Quality Disclosure Platform Web Interface

Notes: This figure shows a screenshot of the Ministry of Environmental Protection's real time air quality disclosure platform web interface as of September 25, 2016. The left panel is an interactive map that displays locations of all monitoring stations. The right panel reports real-time measures of six major pollutants for all monitoring stations in the city that is specified (Beijing).

Figure C.3: Screenshot of an Air Quality App

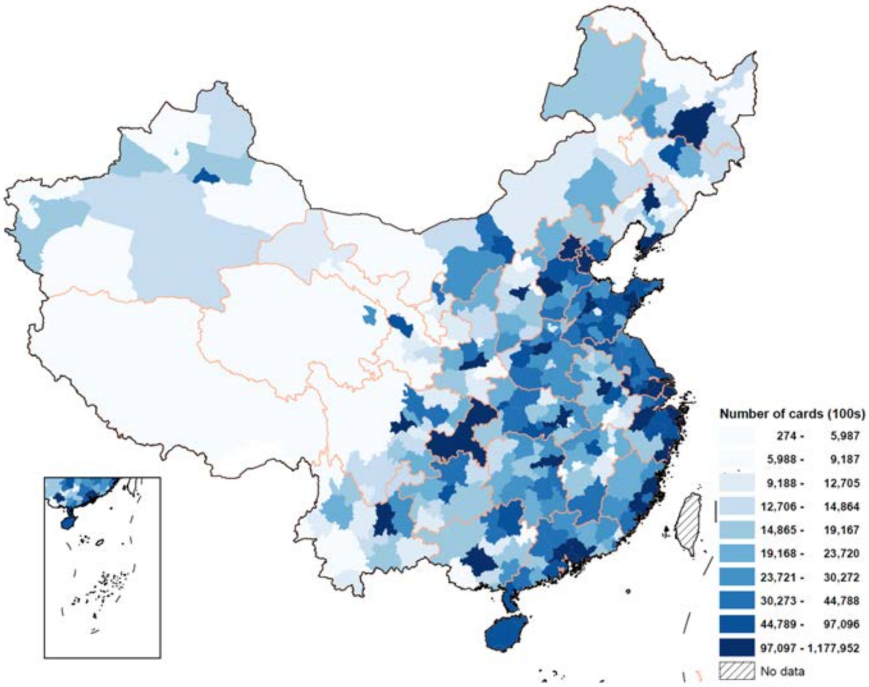
Notes: This figure shows a screenshot of a typical air quality app. The left panel shows the air quality index (AQI) in the city of Shanghai for that hour is 101 and $PM_{2.5}$ is 75 ug/m3. The right panel shows $PM_{2.5}$ and AQI readings at different locations within Shanghai.

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

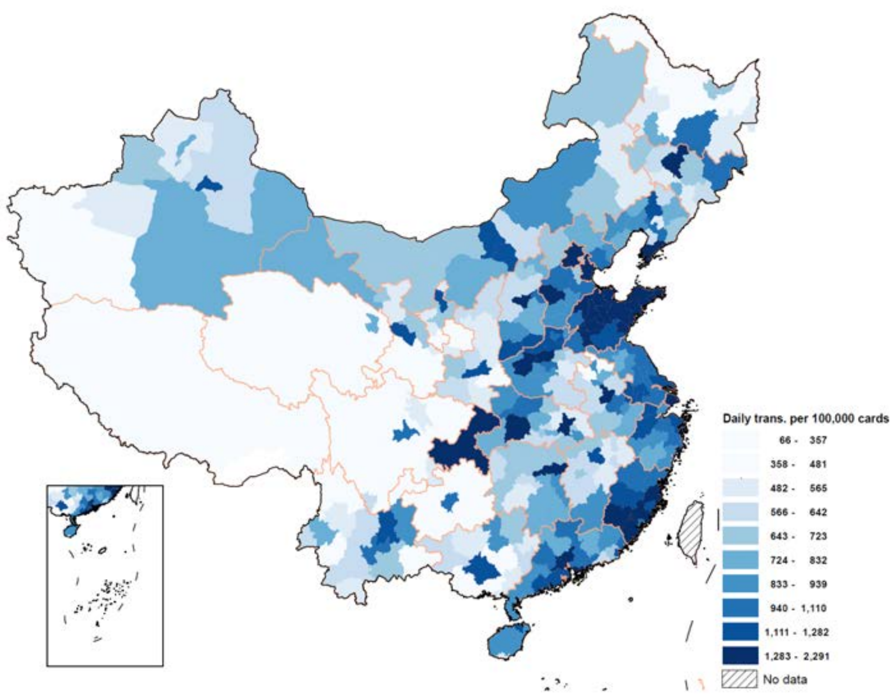


Notes: This figure plots annual GDP (triangles), consumption (squares) reported by the National Bureau of Statistics of China (NBS), and total bank card spendings ×100 (circles) aggregated from the UnionPay 1% bank card data, excluding transactions in the business wholesale categories.

Figure C.5: UnionPay Bank Card Transaction by Prefecture-City, 2011-2015 Average



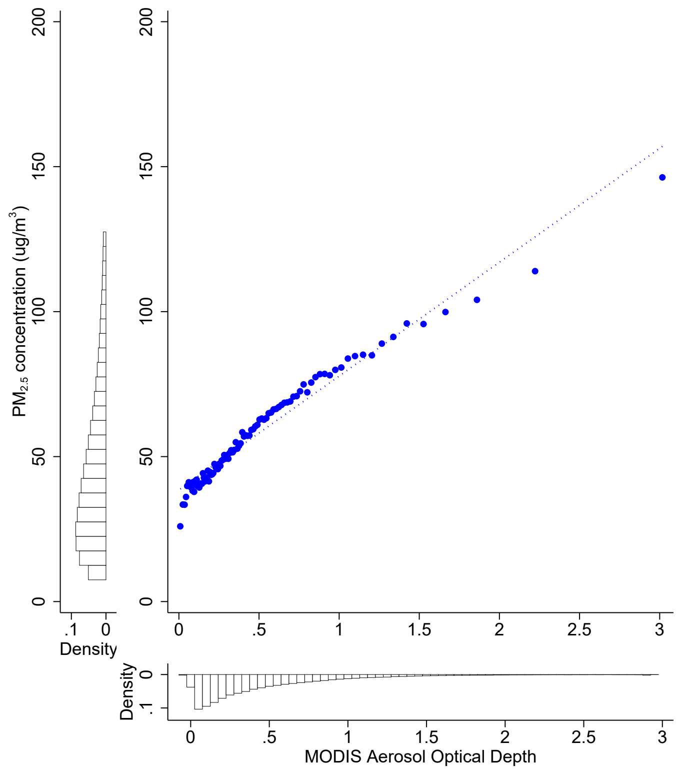
(a) Number of active cards



(b) Number of transactions per 100,000 cards

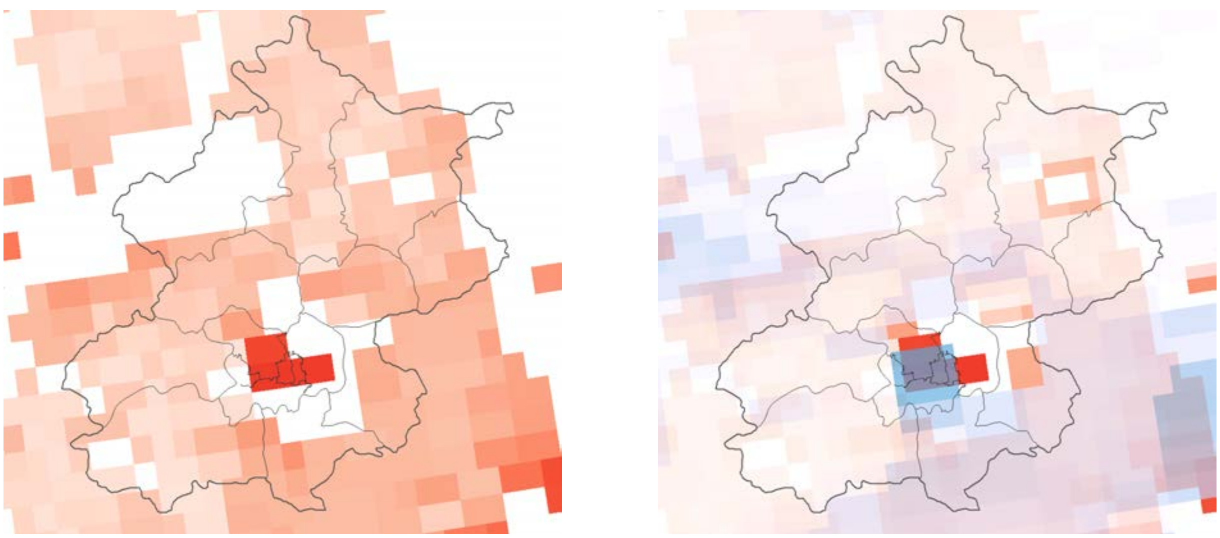
Notes: The maps show 2011-2015 average number of active UnionPay bank cards (panel A) and transactions per 100,000 cards (panel B) at the prefecture-city level. Orange lines show inter-provincial borders.

Figure C.6: Correlation between $PM_{2.5}$ and AOD



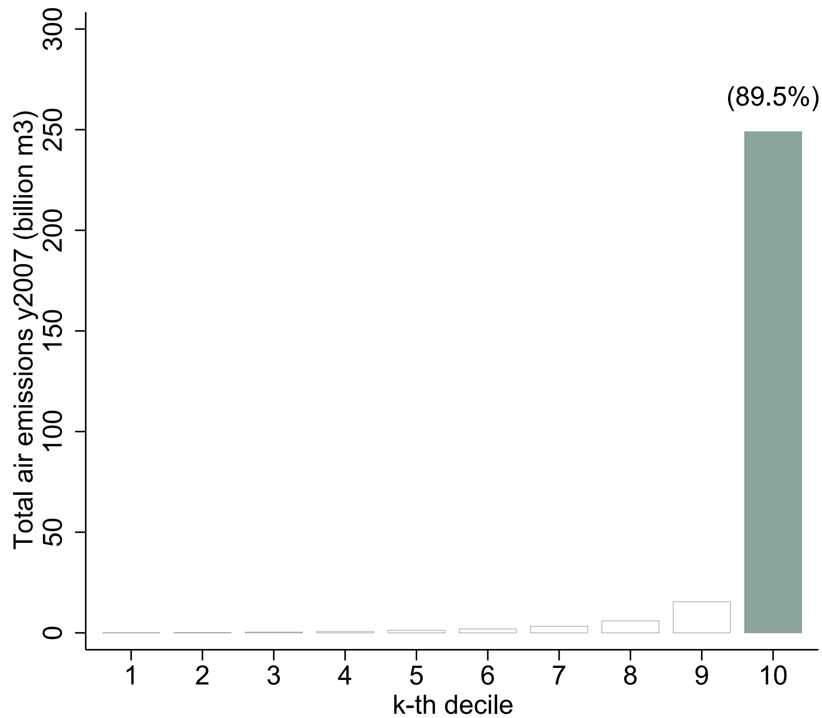
Notes: This graph shows city×day level average $PM_{2.5}$ concentratoin (y-axis) by 100 equal bins of AOD (x-axis), for periods after the information program. There is reliable information on $PM_{2.5}$ before the program. Histograms show the distribution of the two variables.

Figure C.7: Illustration of Satellite AOD Oversampling



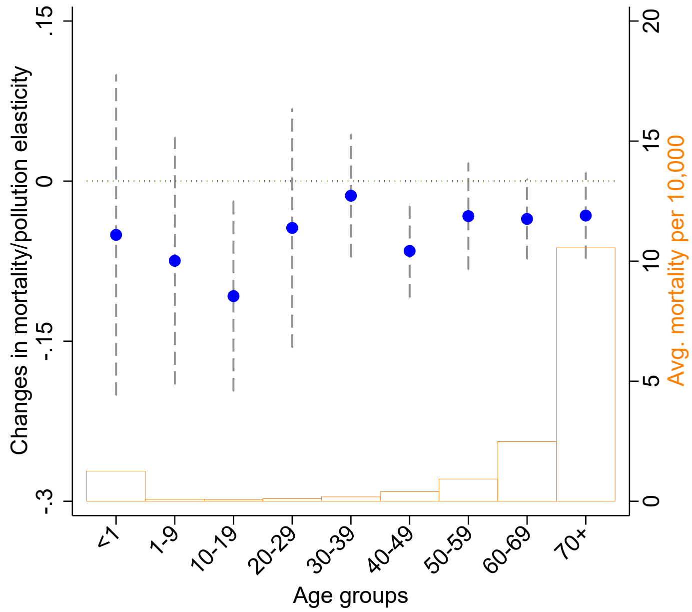
Notes: Left panel shows original MODIS AOD (10×10km) around Beijing on August 30, 2008. Right panel shows an overlay with data on August 31, 2008. In both panels, darker colors indicate higher pollution levels.

Figure C.8: Total Air Emissions by Emission Deciles, Beijing Polluter Census 2007

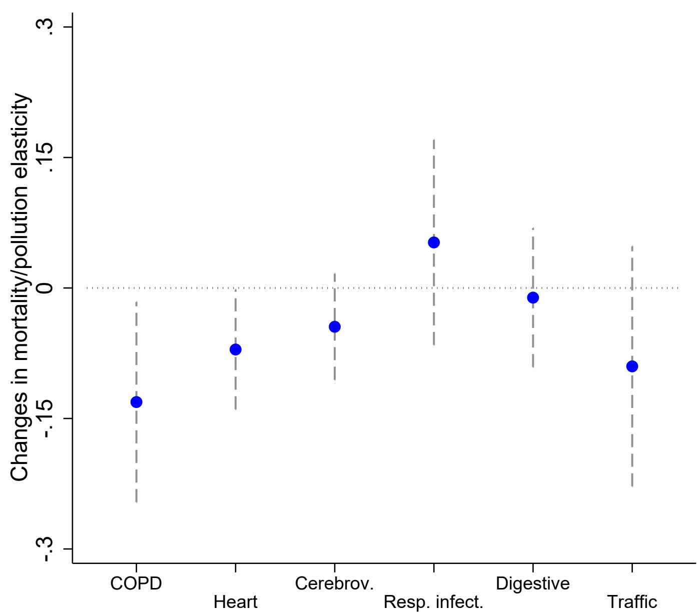


Notes: This graph shows Beijing polluters' total air emissions in billion m^3 for each decile of the annual emission distribution according to the Polluter Census 2007. The sample includes about 440 polluters. Firms in the top decile account for 89.5% of total emissions.

Figure C.9: Heterogeneous Changes in Quarterly Mortality-Pollution Gradient



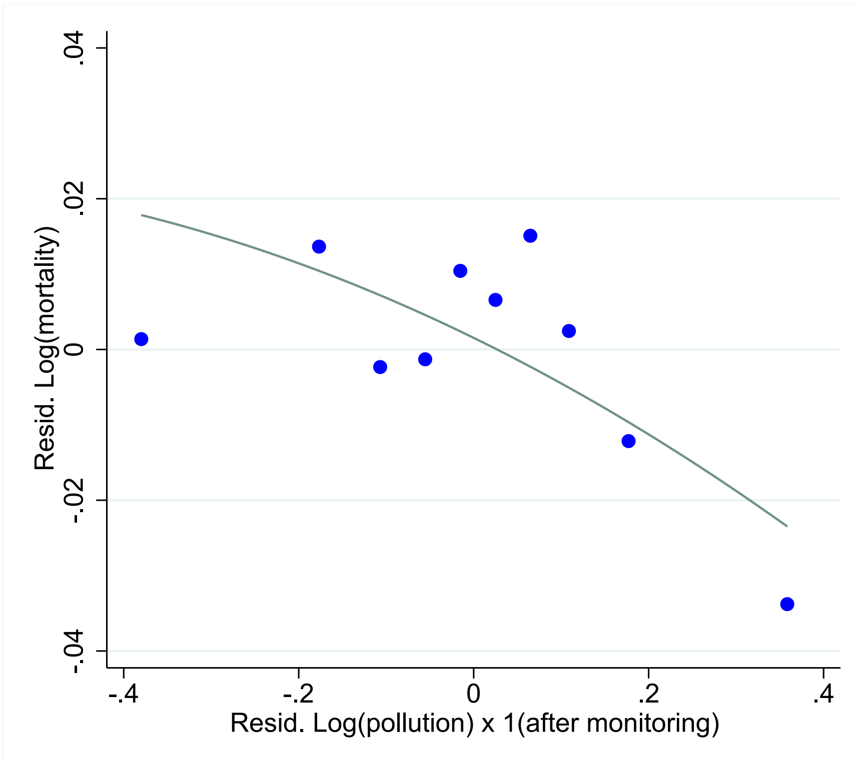
(a) Heterogeneity by age groups



(b) Heterogeneity by causes-of-death

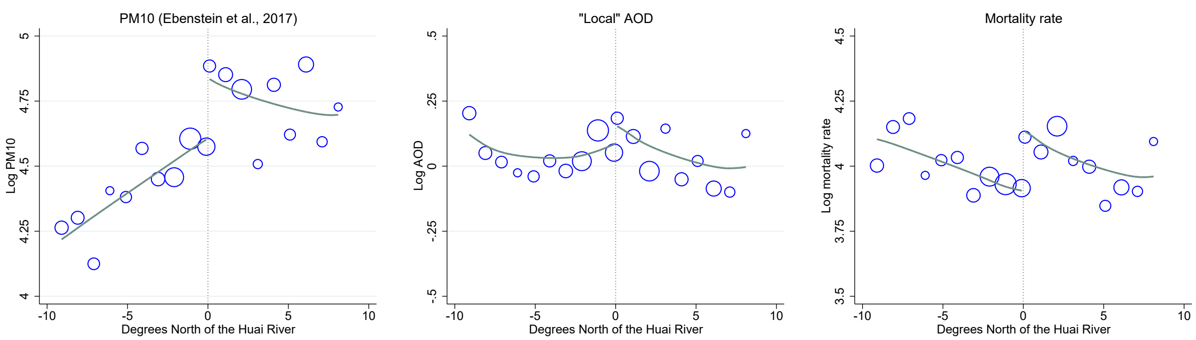
Notes: the figures illustrate heterogeneity in changes in the mortality-pollution elasticity (i.e., coef of Log(Pollution)$\times\mathbb{1}$(after monitoring)) across age groups (panel a) and by causes of death (panel b). Each dot is from a separate regression using sub-group log mortality rate as the outcome variable. All regressions control for prefecture-city FEs and quarter-of-sample FEs. Range bars show 95% confidence interval constructed from standard errors clustered at the prefecture-city level.

Figure C.10: Changes in Quarterly Mortality-Pollution Gradient: Nonlinear Specification



Notes: This graph shows residualized plot between logged mortality rate by 10 equal bins of residualized Log(Pollution) × $\mathbb{1}$(after monitoring). There is limited evidence of nonlinearity. All regressions control for prefecture-city FEs, quarter-of-year FEs, and year FEs.

Figure C.11: 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 = AOD residualized of inverse-distance weighted $PM_{2.5}$ from cities within 1,000 km radius, following Ebenstein et al. (2017).

Table C.1: Characteristics of Cities by Monitoring Roll-out 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: all characteristics are measured by the 2011-2015 average, except for $PM_{2.5}$ (average over the post-monitoring periods) and industrial SO_2 emissions (year 2006). The table report average characteristics for cities in different waves. Standard deviations are in parentheses.

Table C.2: Changes in the Economic and Regulatory Environment Before and 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 a city×week. "Anti-corruption cases" are the number of local officials downfalls during the anti-corruption campaign, "doc. mayor" indicates whether the current city mayor has a doctoral degree, "pollution regulation news" is 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), central south (81 cities), southwest (54 cities), and 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 C.3: Primary Data Sources Coverage

Name	Year	Coverage
Baidu web searches	2011-2015	all cites; 97% of Internet users
Air purifier sales	2012-2015	50 cities; 28% of population
UnionPay card	2011-2015	all cites; 59% of national consumption
Housing transaction	2006-2014	Beijing; 4.8 million households
Mortality	2011-2015	161 counties; 5% nationally representative
Satellite data	2011-2015	all cities

Notes: A 2013 survey by the China Internet Network Information Center on more than 2,800 phone respondents shows more than 99% of Internet users have heard of the Baidu search engine, and 98% have used it in the past six months.

Table C.4: Changes in Weekly Bank Card Transaction-Pollution Gradient: "Deferrable" Consumption

Dep. var.: arcsinh(Number of transactions per 10,000 active cards in a city×week)				
	(1)	(2)	(3)	(4)
Panel A. Merchant type = supermarkets				
Log(Pollution)	0.037***	0.033**	0.043***	0.052***
	(0.013)	(0.015)	(0.009)	(0.010)
Log(Pollution) × 1(after monitoring)	-0.061***	-0.062***	-0.072***	-0.088***
	(0.015)	(0.019)	(0.012)	(0.014)
Panel B. Merchant type = dining				
Log(Pollution)	0.042***	0.048***	0.029**	0.037***
	(0.013)	(0.014)	(0.011)	(0.013)
Log(Pollution) × 1(after monitoring)	-0.078***	-0.089***	-0.054***	-0.063***
	(0.015)	(0.017)	(0.015)	(0.018)
Panel C. Merchant type = entertainment				
Log(Pollution)	0.029*	0.042**	0.025*	0.031*
	(0.017)	(0.019)	(0.014)	(0.017)
Log(Pollution) × 1(after monitoring)	-0.043***	-0.065***	-0.032*	-0.042*
	(0.022)	(0.025)	(0.019)	(0.022)
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 a city×week. To facilitate interpretation of the magnitude across different categories, we use the inverse hyperbolic sine function (ArcSinh) of transactions, so that the key coefficient can be roughly interpreted as elasticities. The ArcSinh function is preferable to logs due to a non-trivial fraction of zero transactions in small city-categories. The patterns are similar if we use levels instead of ArcSinh. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Central South (81 cities), Southwest (54 cities), and Northwest (52 cities). Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table C.5: Changes in Weekly Bank Card Transaction-Pollution Gradient: "Scheduled" Consumption (Placebo Tests)

Dep. var.: arcsinh(Number of transactions per 10,000 active cards in a city×week)				
	(1)	(2)	(3)	(4)
Panel A. Merchant type = billings				
Log(Pollution)	0.004	0.015	0.034	0.048
	(0.037)	(0.040)	(0.028)	(0.030)
Log(Pollution) × 1(after monitoring)	0.008	-0.014	-0.048	-0.055
	(0.056)	(0.060)	(0.044)	(0.049)
Panel B. Merchant type = government services				
Log(Pollution)	0.028	0.022	0.011	0.023
	(0.027)	(0.030)	(0.031)	(0.035)
Log(Pollution) × 1(after monitoring)	-0.034	-0.036	-0.026	-0.026
	(0.040)	(0.043)	(0.042)	(0.049)
Panel C. Merchant type = business-to-business wholesales				
Log(Pollution)	0.009	0.016	0.006	0.010
	(0.018)	(0.020)	(0.014)	(0.017)
Log(Pollution) × 1(after monitoring)	-0.008	-0.027	-0.010	-0.013
	(0.028)	(0.030)	(0.021)	(0.025)
Panel D. Merchant type = cancer treatment centers				
Log(Pollution)	0.001	0.001	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Log(Pollution) × 1(after monitoring)	-0.001	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
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 a 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. To facilitate interpretation of the magnitude across different categories, we use the inverse hyperbolic sine function (ArcSinh) of transactions, so that the key coefficient can be roughly interpreted as elasticities. The ArcSinh function is preferable to logs due to a non-trivial fraction of zero transactions in small city-categories. The patterns are similar if we use levels instead of ArcSinh. "region" is a conventional partition of cities by location: North (36 cities), Northeast (38 cities), East (105 cities), Central South (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 C.6: 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 cities with top 10% anti-corruption cases	-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 variables	-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 examines the robustness of the changes in the transaction - pollution gradient. 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 C.7: 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 a city×week. "1(Treated)" equals 1 for cities actually in the roll-out wave, 0 for neighboring cities not yet experiencing the roll-out. "region" is a conventional partition of cities by location: north (36 cities), northeast (38 cities), east (105 cities), central south (81 cities), southwest (54 cities), and northwest (52 cities). Standard errors are clustered at the prefecture-city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table C.8: 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) $\sim$ $\hat{\text{Log}}$(Pollution)			
$\hat{\text{Log}}$("Local" AOD)	0.660*	0.591**	0.875
	(0.344)	(0.299)	(0.650)
$\hat{\text{Log}}$(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 the coefficient estimate for a dummy variable indicating north of the Huai River in a separate regression (observations = 161). "Local" AOD = AOD residualized of inverse-distance weighted $PM_{2.5}$ from cities within 1,000 km radius, following Ebenstein et al. (2017). PM_{10} data are directly 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. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.