

Online Appendix for

From Fog to Smog: the Value of Pollution Information

Panle Jia Barwick, Shanjun Li, Liguo Lin, Eric Zou

This appendix provides supplementary material to the paper “From Fog to Smog: the Value of Pollution Information.” Section [A](#) describes the driving factors that influenced China’s dramatic shift in environmental regulations surrounding 2012. Section [B](#) presents a theoretical model on the role of the monitoring program that reduces the wedge between individuals’ perceived and experienced air pollution. Section [C](#) proves that the key parameter can be consistently estimated via OLS under assumptions 1 and 2 as stated in Section [4.2](#). Section [D](#) provides more details on several empirical analyses, including program rollout, a data-driven method to identify outdoor trips, private sources for pollution information, IV and synthetic analyses of mortality outcomes, and robustness results using annual data (instead of weekly data).

Appendix A: Policy Shift Timeline and Driving Factors

Appendix Table [E.1](#) provides a chronology of important events related to the policy shift on PM_{2.5} regulation discussed in Section [2.1](#) and the implementation of the monitoring and disclosure program. The shift of the government’s stance on regulating PM_{2.5} is reflected by the change from the first to the second draft of the National Ambient Air Quality Standards amendment published for public comments on November 2010 and November 2011, respectively. While the first draft deemed China not yet ready to implement national standards for PM_{2.5}, the second draft a year later added the national standards for PM_{2.5}.

As household income rose rapidly and a large middle class emerged in China during the past two decades, the demand for better environmental quality and quality of life in general increased ([Dasgupta et al., 2002](#); [Kahn and Zheng, 2016](#)). Together with the important changes in economic and social conditions, the following three factors likely played a key role in propelling the dramatic shift in the stance of the Chinese government on PM_{2.5} monitoring and disclosure, and on air pollution regulations in general.

First, the environmental science community has been working behind the scene to push for more stringent environmental regulations and transparency in China. As a large body of evidence emerged regarding the harmful health impacts of PM_{2.5} in the past several decades ([Pope and Dockery, 2012](#); [Landrigan et al., 2018](#)) and recognized by policymakers around the world, many countries have amended their ambient air quality standards to include PM_{2.5}

as a primary pollutant (e.g., U.S. in 1997, EU in 1999, and Japan in 2009). Environmental scientists in China have been pushing for the regulation for $PM_{2.5}$ through academic conferences and policy dialogues, according to our conversation with Jiming Hao, Professor and a former Dean of the School of Environment at Tsinghua University and one of the pioneers in China’s environmental regulations and pollution prevention.⁴⁴ In the early 2000s, environmental scientists organized international conferences on air pollution and $PM_{2.5}$ in China in order to draw attention from and educate policymakers. Nevertheless, policymakers were predominantly concerned about economic growth and did not fully appreciate the benefit of educating the general public when the government was not yet ready to bring the pollution problem under control. This concern perhaps underlies the views stated in the first draft of the NAAQS amendment released in November 2010. While the draft recognized that $PM_{2.5}$ had become a major pollutant in many areas in China, it considered standards set by the WHO and developed countries too aggressive for China and deemed that China was not ready to set national $PM_{2.5}$ standards.

Second, the U.S. Embassy in Beijing is likely another important catalyst behind the shift. In early 2008, the embassy installed a rooftop air quality monitor and reported hourly $PM_{2.5}$ readings through the Twitter account @beijingair in order to advise U.S. citizens who travel in China before the upcoming summer Olympics. The $PM_{2.5}$ readings were often in conflict with the Chinese government’s official report on air quality, especially during the extreme pollution episodes in winter. For example, on November 19, 2010, the Twitter account @beijingair reported $PM_{2.5}$ over 500 and described the air pollution as “crazily bad”, while China’s official assessment of air quality based on API was “mildly polluted” ($PM_{2.5}$ was not incorporated in API). The discrepancy prompted environmental activists in China to call for more transparency on pollution information. The ensuing social media storm likely contributed to the decision by the MEP to provide their own data that the general public would trust, rather than to dispute the U.S. Embassy’s data.

Third, environmental NGOs and social influencers have played an important role in raising the awareness of the general public and exerting pressure on policymakers. Their role is facilitated and amplified by the emergence of social media (notably through Weibo in China) and the wide adoption of information technologies. Motivated by severe pollution episodes and the lack of transparency and reporting of air pollution levels, the campaign “Measuring Air Quality for our Motherland” in the Fall of 2011 was a concerted effort mobilized through social media Weibo by environmental NGOs and influential activists. The campaign called

⁴⁴In 2015, Professor Hao was awarded the Hagen-Smit Clean Air Award by the California Air Resources Board for his contribution to air pollution control and understanding PM pollution in major cities in China. When he visited Cornell University in May 2019, our research team met with him to understand the context of the policy shift in our study and more broadly the history of environmental regulations in China.

for setting national PM_{2.5} standards and implementing accurate monitoring and reporting of air quality by the government. In order to achieve their goals more efficiently, environmental NGOs in China often position themselves as educators and partners of government agencies, taking into account the cultural and social landscape. The campaign appealed to the government to sustain its vigilance towards safeguarding people’s welfare due to the detrimental health impacts of air pollution especially PM_{2.5}. This campaign was considered a milestone and a great success in the discourse of civic participation in China’s environmental governance (Fedorenko and Sun, 2016), contributing to the release of the second draft of the NAAQS amendment in November 2011 where PM_{2.5} was added as a primary pollutant to be regulated.

Appendix B: 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 monitoring 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; Harrington and Portney, 1987)..

B.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.⁴⁵ Individuals allocate their wage and non-wage income between consumption and engaging in avoidance behavior a , where p_a is the associated price (e.g.,

⁴⁵The effect of health on wage income captures the impact of pollution on labor supply or productivity as documented in Graff Zivin and Neidell (2012); Hanna and Oliva (2015); Chang et al. (2019); He, Liu and Salvo (2019).

the cost of an air purifier or medication). We use a to include broadly defined (costly) adaptation behavior.⁴⁶ 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 monitoring 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 monitoring 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, 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

⁴⁶Examples 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](#)).

much higher when pollution is severe than when it is modest.

We focus on interior solutions for the optimal level of avoidance behavior a .⁴⁷ 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 monitoring program. We assume that $c_0 < c$, that is, the perceived level of pollution is lower than the actual level.⁴⁸ 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 monitoring program is predicted to result in the following impacts:*

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

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

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

Appendix B.3 provides the proof. The theoretical model predicts that following the implementation of the monitoring 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 monitoring program.

⁴⁷A necessary condition for an interior solution is $w \cdot \frac{dg}{dh} \cdot \frac{\partial h}{\partial a} \Big|_{a=0} > p_a$.

⁴⁸An alternative assumption to $c_0 < c$, is that the monitoring program reduces the price of avoidance p_a . This also delivers Proposition 1.

B.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:

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

The indirect utility before the monitoring 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 (B.1) rather than Equation (B.3). 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:

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

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 quantifies the magnitude of the terms in the curly bracket.

B.3 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 B.1. The Lagrangian equation is:

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

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

$$\begin{aligned}\frac{\partial L}{\partial x} &= 0 \Rightarrow U_x - \lambda = 0 \\ \frac{\partial L}{\partial a} &= 0 \Rightarrow (U_h + \lambda \cdot w \cdot g_h) \frac{\partial h(c_0, a)}{\partial a} - \lambda p_a = 0 \\ \frac{\partial L}{\partial \lambda} &= 0 \Rightarrow I + w \cdot g(h) - x - p_a \cdot a = 0\end{aligned}\tag{B.3}$$

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 B.3):

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

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, $C + D \leq 0$.⁴⁹ Similarly, $A + B \geq 0$. Hence, avoidance increases weakly in (perceived) pollution. The key assumption for this result is $\partial h^2 / \partial a \partial c \geq 0$. When pollution deteriorates, avoidance restores health more effectively (that is, the marginal benefit of avoidance is large with bad pollution). After the monitoring 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

⁴⁹At 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 avoidance.

B.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 monitoring program, perceived pollution c_0 is unlikely to fully respond to day-to-day changes in actual pollution. The total derivative of health w.r.t. pollution is:

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

where $0 < dc_0/dc < 1$. Post the monitoring 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:

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

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 (and accurate) information, we have:

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

Appendix C: Identification of the Change in the Outcome-Pollution Gradient

In this section, we restate the two key identification assumptions that are outlined in the paper and prove Proposition 2 (that the OLS estimate of β in Equation (1) is consistent). Rewriting Equation (1):

$$y_{it} = \alpha p_{it} + \beta p_{it} d_{it} + \mathbf{x}_{it} \boldsymbol{\gamma} + \varepsilon_{it}, \tag{C.4}$$

where p_{it} measures ambient air quality and could be correlated with ε_{it} due to unobservables or measurement error as discussed in the main text. d_{it} represents the treatment dummy and is equal to one after treatment based on the staggered rollout schedule. \mathbf{x}_{ct} includes city attributes and other controls such as city and time fixed effects, and city-specific time trends. The key parameter of interest is β , the change in the slope of pollution-outcome

relationship. To facilitate proof, we write Equation C.4 above in matrix form:

$$\mathbf{y} = \alpha \mathbf{p} + \beta \mathbf{p} \circ \mathbf{d} + \mathbf{X} \boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \quad (\text{C.5})$$

where ‘ \circ ’ is an element-by-element product. $\mathbf{y}, \mathbf{p}, \mathbf{d}$, and $\boldsymbol{\varepsilon}$ are N by 1 vectors, \mathbf{X} is a N by k matrix.

Assumption B1: $\varepsilon_{it} \perp d_{it} \mid \mathbf{x}_{it}$. This assumption implies that conditioning on city attributes and other controls \mathbf{x}_{ct} , the treatment d_{ct} is exogenous.

Assumption B2: $d_{it} \perp p_{it} \mid \mathbf{x}_{it}$. This assumption implies that the treatment status is independent of pollution levels conditioning on \mathbf{x}_{ct} (i.e., variations in pollution unexplained by \mathbf{x}_{ct}).⁵⁰

Assumption B1 is the standard conditional exogeneity assumption: the program rollout is as good as random conditional on the control variables. This assumption can be assessed via both institutional and econometric evidence (Section 4.3). We show that the assignment of cities to rollout wave largely follows city hierarchies and designated status that are determined long before the monitoring place came into place. We discuss other major environmental policies and argue that they have limited scope to confound our analysis. We present event study estimates, which allows us to assess whether there are any preexisting trends in the outcomes that could be indicative of endogeneity issues.

Assumption B2 ensures the nature of endogeneity is the same before versus after the monitoring program. One way to conceptualize it is to imagine a binary context in which *Pollution* indicates “high” vs. “low” pollution areas. Note that Equation (1) reduces to a difference-in-differences style setting that compares outcome in regions with high vs. low pollution, before vs. after policy introduction. 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. Assumption B2 can be tested empirically as well (Section 4.3). In particular, we present balancing test in Appendix Table E.5 which shows that, conditional on fixed effects controls, the monitoring policy does not lead to significant changes in pollution levels.

⁵⁰While we present Assumption B2 for ease of interpretation, the assumption is stronger than what we need to prove the consistency of β . As we show in the proof that follows, a sufficient condition is $\mathbb{E}[d_{it} \mid \mathbf{M}_x \mathbf{p}] = c$, where \mathbf{M}_x is the projection matrix; $\mathbf{M}_x \mathbf{p}$ is the residual from OLS regression of \mathbf{p} on \mathbf{X} ; and c is a constant.

We now demonstrate theoretically that the β coefficient in Equation 1 is consistent under these two identification assumptions.

Proposition 2. *Under Assumptions B1 and B2, the OLS estimate of β in Equation (C.4) is consistent.*

Proof of Proposition 2 Let \mathbf{M}_x denote the projection matrix: $\mathbf{M}_x = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Multiplying both sides of equation C.5 with \mathbf{M}_x , we have:

$$\mathbf{M}_x\mathbf{Y} = \alpha\mathbf{M}_x\mathbf{P} + \beta\mathbf{M}_x\mathbf{p} \circ \mathbf{d} + \mathbf{M}_x\boldsymbol{\varepsilon},$$

where $\mathbf{M}_x\mathbf{P}$ is an N by 1 vector, the projection residual of \mathbf{p} on \mathbf{X} . Collect the two key regressors in $\mathbf{Z} = [\mathbf{M}_x\mathbf{p}, \mathbf{M}_x\mathbf{p} \circ \mathbf{d}]$. Let the OLS estimates of α and β be denoted as $\hat{\alpha}$ and $\hat{\beta}$.

$$\begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} = \begin{pmatrix} (\mathbf{M}_x\mathbf{p} \circ \mathbf{d})'(\mathbf{M}_x\mathbf{p} \circ \mathbf{d}) & -(\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\mathbf{p} \circ \mathbf{d}) \\ -(\mathbf{M}_x\mathbf{p} \circ \mathbf{d})'(\mathbf{M}_x\mathbf{p}) & (\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\mathbf{p}) \end{pmatrix} \begin{pmatrix} (\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\boldsymbol{\varepsilon}) \\ (\mathbf{M}_x\mathbf{p} \circ \mathbf{d})'(\mathbf{M}_x\boldsymbol{\varepsilon}) \end{pmatrix} \\ * \frac{1}{\det[(\mathbf{Z}'\mathbf{Z})]}.$$

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

$$- \mathbb{E}[(\mathbf{M}_x\mathbf{p} \circ \mathbf{d})'(\mathbf{M}_x\mathbf{p})] \mathbb{E}[(\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\boldsymbol{\varepsilon})] + \mathbb{E}[(\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\mathbf{p})] \mathbb{E}[(\mathbf{M}_x\mathbf{p} \circ \mathbf{d})'(\mathbf{M}_x\boldsymbol{\varepsilon})].$$

Assumption B2 implies that $E(\mathbf{d}|\mathbf{M}_x\mathbf{p}) = c$. Hence:

$$\mathbb{E}[(\mathbf{M}_x\mathbf{p} \circ \mathbf{d})'(\mathbf{M}_x\mathbf{p})] = \mathbb{E}[(\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\mathbf{p}) \mathbb{E}(\mathbf{d}|\mathbf{M}_x\mathbf{p})] = c \mathbb{E}[(\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\mathbf{p})].$$

Assumptions B1 and B2 imply that:

$$\mathbb{E}[(\mathbf{M}_x \circ \mathbf{d})'(\mathbf{M}_x\boldsymbol{\varepsilon})] = c \mathbb{E}[(\mathbf{M}_x\mathbf{p})'(\mathbf{M}_x\boldsymbol{\varepsilon})].$$

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

Appendix D: Additional Analysis

This section provides more details on several analyses that we do not have space for in the main paper. Section D.1 discusses how different cities are designated into different rollout waves. Section D.2 uses sensitivity to precipitation to define consumption categories that

are more ‘outdoor’ in nature. Sections D.3 and D.4 examine two private sources of pollution information and evaluate whether households engage in effective pollution avoidance prior to the monitoring program. Sections D.5 and D.6 conduct the IV and synthetic analyses of mortality effects. Section D.7 uses annual data to evaluate the robustness of avoidance findings.

D.1 Determinants of Information Rollout Assignment

Appendix Table E.3 reports an analysis on the determinants of the official rollout assignment. Columns (1)-(3) show a representation of Appendix Figure E.9 using three separate linear probability models. The models regress an indicator of a city being in wave k in the rollout on three indicators for administrative hierarchies and pre-designated status that are denoted by the deep blue, light blue, or white group.⁵¹ Consistent with our discussion above, these simple models produce a very high R^2 ranging from 0.65 to 0.84.

We now examine the potential role of observed covariates in determining the official rollout timing. To make it simple, in column (4), we first estimate a benchmark model where the outcome variable is city’s rollout wave as a continuous variable (take values 1, 2 or 3, and hence negative coefficients would mean earlier waves). Regressors include a constant term and indicators for deep blue and light blue cities; the indicator for white cities is the omitted category. This model is therefore a parsimonious summary of columns (1)-(3). Again, this simple model has a high R^2 of 0.75.

Column (5) bases off the specification of column (4), but asks whether observed covariates in the pre-policy period can provide additional explanatory power. Specifically, we examine pollution, income, and their trends. Pollution and pollution trends are measured by a city’s 2006-2012 average AOD and its 2006-2012 annual change in AOD. Income and trends are measured by a city’s 2006-2012 average per capita disposable income and its 2006-2012 annual changes in per capita income.

All covariates are normalized by dividing the mean, so the interpretation of the coefficients is “how much change in the rollout wave is associated with a doubling of the characteristics”, with a negative coefficient meaning that an increase in the characteristics is associated with the city being in an earlier rollout.

Results in column (5) suggest cities with higher levels of pollution and higher levels of per capita income are significantly more likely to be in earlier waves. This is expected, as cities in earlier rollout waves tend to be more developed such as provincial capitals.⁵² We note that

⁵¹Note in column (3), the coefficient on light blue indicator is missing because there were no light blue cities in the wave-3 rollout. Also see Appendix Figure E.9.

⁵²We have shown similar quantitative patterns in Appendix Table E.4.

our analysis includes city fixed effects, and therefore any permanent differences in the levels of characteristics do not pose a threat to the identification. On the other hand, trends in pollution and income prior to the monitoring program do not predict wave assignment. This suggests the rollout timing is not determined by changes in local conditions. Comparing R^2 in columns (4) vs. (5), the additional covariates provide little improvement in explanatory power ($R^2=0.747$ in column (4) and $R^2=0.769$ in column (5)).

Finally, column (6) presents a model with only the pollution and income covariates. Similar qualitative conclusions emerge: cities with high pollution and high income are in earlier waves but trends are, once again, not predictive. The model’s goodness-of-fit is lower ($R^2=0.427$).

Overall, the regression analysis reinforces the view that the policy rollout is a top-down decision that reflects pre-existing administrative hierarchies and city characteristics rather than changes in local economic or pollution conditions.

D.2 Outdoor Consumption Trips

Our transaction data do not contain flags on the indoor/outdoor nature of the underlying purchases. To make progress, we consider a data-driven exercise that exploits heterogeneity in how precipitation affects transaction activities across 273 narrowly defined merchant category codes (MCCs). The underlying logic of this exercise is that merchant categories whose transaction volumes are more sensitive to precipitation events are more likely to be those that are either deferrable in nature or those that tend to occur in outdoor settings (and thus involve more pollution exposure). These are the types of transaction activities that we expect to exhibit more pronounced pollution avoidance behavior.

We begin by estimating the following regression separately for each of the 273 merchant categories. We use daily-level observations to detect responses to precipitation better.

$$\text{Log}(\text{Transactions}_{mct}) = \beta_m \cdot \text{Precipitation}_{ct} + \alpha_{mc} + \alpha_{mt} + \varepsilon_{mct}$$

where Transaction_{mct} is the number of transactions in merchant category m in city c on day t . $\text{Precipitation}_{ct}$ is a dummy variable for whether there was any precipitation (snow or rain) in the city-day. All regressions control for city fixed effects (α_{mc}) and week fixed effects (α_{mt}). We cluster standard errors at the city level. Panel (a) of Appendix Figure E.10 plots the estimates β_m estimates. Each bar represents a merchant category m , and the range of each bar corresponds to the 95% confidence interval of the estimate. We rank merchant categories by their estimated sensitivity to precipitation, starting from those with the most negative point estimate of β_m . Blue bars highlight merchant categories whose β_m estimates

are statistically significant at the 5% level.

Two findings emerge from this exercise. First, the vast majority of precisely estimated effects have negative signs (i.e., precipitation reduces transactions). This is reassuring because virtually all transactions in our database occur in physical stores and necessarily involve outdoor trips, so we expect an overall negative association with precipitation. Second, at the bottom of the same chart, we use red bars to mark merchant categories in supermarkets, dining, and entertainment sectors that we defined as *deferrable consumption* trips in Table 2, Panel B of the paper. Our new analysis shows that these categories are disproportionately deferrable and/or outdoor in nature from a precipitation-sensitivity perspective. In Appendix Figure E.11, we tabulate the top 20 most precipitation-sensitive merchant categories according to our estimates. The most sensitive categories include snow car rentals and dealers, vacation houses, sports stadiums, garment stores, amusement parks, and optical exams and products, among others that are arguably deferrable in nature or occurring in open-air settings.

Next, we construct a new purchase frequency variable which equals the number of card transactions per 10,000 active cards in these top 20 merchant categories. Its mean is 56.3 transactions per week per 10,000 cards (for reference, the all-category transaction frequency mean is 869 transactions). We then repeat the shopping-pollution gradient analysis using this deferrable/outdoor purchase frequency as the dependent variable. Panel (b) of Appendix Figure E.10 reports the event study figure. The pollution gradient after the monitoring program exhibited a sharp drop of -3.12 transactions per 10,000 cards (SE = 1.27, p -value = 0.015), which represents a 5.5 percent reduction relative to the mean transaction frequency (56.3 per 10,000 cards). Recall from Table 2 that the effect size for all-category transactions is a 1.4 percent reduction. As expected, the effect size for the deferral categories is much more pronounced. This new analysis provides additional evidence that consumption trips that are more deferrable/outdoor in nature responded more strongly to pollution variations after the information became available.

D.3 Visibility as Alternative Pollution Information

We assess the possibility that people used visibility as a proxy for pollution levels even before the pollution information program. We obtain weather-station-based daily visibility data covering all cities from China’s National Meteorological Administration for 2011-2016. Panel (a) of Appendix Figure E.12 plots the distribution of the raw visibility measure. For the vast majority of stations, visibility is top coded at 30 km (18.6 miles).

How accurate is visibility as a proxy for air pollution? Panel (b) of Appendix Figure

E.12 shows the median, inter-quartile range, and 10-90 percentile range of satellite-based AOD by ventile bins of visibility. Lower visibility bins (the left part of the x-axis) mean poor visibility conditions. We find an overall downward-sloping relationship between pollution and visibility, which suggests that visibility provides useful information on air quality, especially when visibility is high. However, the relationship features substantial uncertainty when visibility is poor. For example, for the bottom 20% days with the lowest visibility, the 10-90 quantile range of AOD can vary from 0.35 (decent air quality, compared to a national average of 0.55) to 1.55 (poor air quality). These patterns underline the challenges that people face in trying to avoid pollution without actual pollution monitoring data: both heavy fog and heavy smog could cause poor visibility outdoors. Therefore, the scope for using visibility as a proxy for pollution is limited.

It is nevertheless interesting to empirically test whether residents used visibility as a proxy for pollution and exercised avoidance, and how that relationship interacts with actual pollution variation. To operationalize the test, consider the following two sets of estimation equations, where we use the “*pre, post*” superscripts to note that we estimate coefficients of interest for both before and after the monitoring program:

$$Transactions_{ct} = \Gamma^{pre,post} \cdot Visibility_{ct} + \alpha_c + \alpha_t + \varepsilon_{ct};$$

$$Transactions_{ct} = \gamma^{pre,post} \cdot Visibility_{ct} + \beta^{pre,post} \cdot Pollution_{ct} + \alpha_c + \alpha_t + e_{ct}.$$

We leverage these two sets of estimation equations to answer the following four questions:

Question 1. Did purchase trips respond to outdoor visibility before pollution information was available (Γ^{pre})?

Question 2. Did the purchase-visibility relationship *change* after pollution monitoring began (Γ^{pre} vs. Γ^{post})?

Question 3. If the answer to **Question 2** is yes, then how much of that change was explained by people’s response to “real” pollution ($\Gamma^{pre,post}$ vs. $\gamma^{pre,post}$)?

Question 4. Once we take into account that people may use visibility as a proxy for pollution, does “real” pollution variation still matter for purchase trips ($\beta^{pre,post}$)?

Panels (c) and (d) of Appendix Figure E.12 summarize the answers to Questions 1, 2, and 3. The blue lines show the estimates for $\Gamma^{pre,post}$, and the orange lines show the estimates for $\gamma^{pre,post}$.

Before the monitoring program, we observed fewer card transactions during weeks when visibility is in the lowest quintile. However, the shopping-visibility relationship is non-monotone, economically modest, and imprecisely estimated. The reduction in card transactions during weeks with the lowest visibility may reflect a genuine concern for poor visibility per se (e.g., road safety concerns) instead of pollution. Overall, the data provide little support that people used visibility as a proxy for pollution and engaged in avoidance before pollution monitoring.

The blue line in Panel (d) of Appendix Figure E.12 shows that the negative relationship between visibility and card transactions became stronger after monitoring, which is consistent with our main findings based on the pollution measure.

More importantly, adding pollution controls barely changes the transactions-visibility relationship before the monitoring program (the orange line is very close to the blue line). If people were using visibility to predict pollution levels, the transactions-visibility relationship should change after we add pollution controls. This is indeed the pattern we find post the monitoring program: controlling for pollution attenuates the transactions-visibility relationship, especially at low visibility ranges.

To numerically summarize our findings, Appendix Table E.9 reports specifications where the key regressors are log-inverse visibility “Log(1/Visibility)” (higher values stand for lower visibility), pollution “Log(Pollution)”, and their interactions with the post-monitoring dummy.

Columns (1) through (4) examine the relationship between card transactions and poor visibility and how the relationship changes before and after the monitoring policy. There is no evidence that people responded to visibility before the monitoring program (the coefficient for poor visibility, Log(1/Visibility), is mostly imprecisely estimated and positive). The interaction coefficient is negative, suggesting that people were more responsive to poor visibility after the monitoring program. However, unlike the case for pollution (Table 2 in the paper), the interaction coefficients are not precisely estimated.

Columns (5) through (8) show that, once pollution and its interaction with the post-information dummy are included in the specifications, the gradient change in (1/Visibility) is largely explained away (and of wrong sign in Columns (7) and (8)). On the other hand, the visibility terms cannot explain away the pollution findings. The “Log(Pollution) \times 1(after monitoring)” coefficients from columns (5) through (8) ($\beta^{\text{pre,post}}$ in Question 4 above) are similar in magnitude to those from our preferred specifications where we only looked at the effect of pollution (Table 2 of the paper), though more noisily estimated when both the visibility and pollution terms are included as regressors.

Overall, the data do not support the conjecture that consumers engaged in effective pollution avoidance through visibility cues likely due to: (1) visibility is a noise indicator of

pollution when visibility is low (e.g., it could be due to fog), and (2) the general public had limited awareness of air pollution before the program.

D.4 U.S. Embassy PM_{2.5} Data in Beijing as Alternative Pollution Information

We consider the possibility that independent pollution monitoring conducted by the U.S. Embassy in China might have been used as an alternative pollution measure. Five cities in China with U.S. Embassy and Consulates (Beijing, Chengdu, Guangzhou, Shanghai, and Shenyang) operate independent PM_{2.5} monitoring. Our analysis focuses on Beijing, the only city with PM_{2.5} data going back to January 2011 that allows our regression analysis to include seasonal controls before and after the monitoring program. Appendix Figure E.13 summarizes the results. We examine web searches for smog in panel (a) and card transactions in panel (b). For each outcome, we show a decile bin scatterplot of the relationship between the outcome and Embassy PM_{2.5} and do so separately for the pre-policy (i.e., 2011 and 2012) and post-policy (i.e., 2013-2016) periods. The difference in the slope of the two fitted lines (gray for pre-policy, blue for post-policy) corresponds to what we call “change in pollution gradient” – that is, the change in how the outcome variable responds to pollution, before vs. after information availability. With weekly time series data from one city we cannot have detailed control variables. For both the pre- and post-policy periods, we residualize each outcome with 52 week-of-year indicators to parse out seasonality. The analysis contains 312 observations.

Panel (a) in Appendix Figure E.13 shows that the correlation between web searches and pollution readings was close to zero before 2013. It is a tight ‘zero’ relationship as the dots lie close to the linear fitted line. This suggests that the U.S. Embassy PM_{2.5} monitoring data, which were reported on Twitter since 2008, were not a significant information source for the average resident in Beijing. If they were, Beijing residents would have responded to such information before 2013. After 2013, the search-pollution gradient became much more positively correlated, with searches going up substantially on weeks when PM_{2.5} concentration exceeds 100 $\mu\text{g}/\text{m}^3$.

Panel (b) shows that the correlation between U.S. Embassy pollution readings and weekly card transactions was also flat before 2013. After 2013 when pollution information became available, the correlation became significantly negative, and doubling the pollution level is associated with 195 fewer transactions per 10,000 active cards per week. The magnitude of this estimate is much larger than our full-sample estimate (about 12 fewer transactions per 10,000 active cards per week). This is partly because Beijing has a higher baseline

transaction rate (1,232 weekly transactions per 10,000 cards compared to a national average of 869) and because the levels of pollution in Beijing tend to be higher (daily mean = 79 $\mu\text{g}/\text{m}^3$, IQR = 29 to 109 $\mu\text{g}/\text{m}^3$ in 2015, compared to the national average daily mean = 39 $\mu\text{g}/\text{m}^3$, IQR = 24 to 63 $\mu\text{g}/\text{m}^3$).

D.5 Instrumental Variables Approach

Our empirical analysis shows that the monitoring program has reduced the mortality-pollution gradient, i.e., the mortality impact of air pollution was mitigated after the program relative to that before the program likely due to behavioral changes as documented in our analysis. While the impact of pollution on mortality itself is not the focus of our paper (we evaluate its changes), it might be of interest to policymakers. In this section, we recover the causal estimate of the pollution’s mortality effect using IVs to address the endogeneity in pollution.

We implement a wind-transport instrumental variable (IV) approach in the spirit of [Bayer, Keohane and Timmins \(2009\)](#), [Deryugina et al. \(2019\)](#), [Anderson \(2020\)](#) and especially [Barwick et al. \(2020\)](#). The goal is to tease out (exogenous) variation in a city’s local air pollution attributable to transported pollutants from upwind cities. Our IV is essentially a function of wind directions, wind speed, weather conditions, and distance of origin and destination cities. To see the IV construction, we use the city of Beijing as a narrative example. We begin with a daily panel dataset of $\text{PM}_{2.5}$ in a set of cities whose pollution level may affect air quality in Beijing. Let C denote the set of contributing cities. For each city c and day t , we calculate the radian angle ϕ_{ct} between city c ’s local wind direction and the vector pointing from city c to Beijing (e.g., $\phi_{ct} = 0$ if city c is exactly upwind from Beijing on day t). The IV is a time-series variable constructed using the following formula:

$$\text{IV}_t = \sum_{c \in C} \max\{0, \cos(\phi_{ct})\} \cdot \text{Pollution}_{ct} \cdot \left(\frac{1/\text{distance}_c}{\sum_{i \in C} 1/\text{distance}_i} \right), \quad (\text{D.6})$$

where the term $\max\{0, \cos(\phi_{ct})\} \cdot \text{Pollution}_{ct}$ – which we call “upwind pollution” – is the vector component of air pollution in city c on day t that is expected to move toward Beijing. We assume upwind pollution is zero if ϕ_{ct} is an obtuse angle, i.e., winds in city c on day t are blowing away from the direction toward Beijing. On any date t , the IV is the average of individual cities’ upwind pollution terms, inversely weighted by city c ’s distance to Beijing (distance_c).

The choice of contributing cities (denoted as C) for our analysis involves a trade-off between bias and variance. Using cities that are very far from Beijing helps the exclusion restriction assumption of the IV (that is, the assumption that transported pollution from

distant cities does not affect mortality in Beijing, except through its impact on local air quality). However, focusing on pollution from cities too far away hurts the first stage assumption because the local impact of upwind pollution from faraway cities is weak. We take the following steps to address such bias-variance trade-off. First, we restrict contributing cities to those that are at least 300 km away from Beijing.⁵³ Second, among the remaining cities, we employ a data-driven method that selects the most predictive upwind cities in a “zero-stage” Lasso regression. Specifically, before constructing the IV variable, we estimate the following equation with linear Lasso:

$$\text{Pollution}_{\text{Beijing},t} = \lambda_0 + \sum_{c \in \tilde{C}} \lambda_c \cdot \max\{0, \cos(\phi_{ct})\} \cdot \text{Pollution}_{ct} + \epsilon_t,$$

which selects a subset of 73 upwind cities from a total of 330 cities that are at least 300 km away from Beijing. Appendix Figure E.16 maps out the location of these upwind cities and their corresponding λ_c coefficients in the post-Lasso regression. We then construct the IV variable for Beijing using these selected cities as outlined in Equation (D.6). We repeat this procedure for all cities in our sample. The first stage of our IV is strong, with the Kleibergen-Paap F-statistic exceeding 600.

Having constructed the IV, we use 2SLS to estimate the mortality impact of pollution separately for the pre and post periods following Equation 1. We include city, week-of-year, and year FEs, as well as city-specific trends, as in the main analysis. Standard errors are clustered at the city level. Appendix Figure E.17 reports the parameter estimates capturing the mortality impact of pollution before and after the monitoring program, separately by the cause of deaths. Panel (a) shows the OLS results while Panel (b) shows the IV results.

Several patterns emerge from this analysis. First, our IV estimates are in general larger in magnitude than OLS estimates, consistent with the literature (Deschenes and Greenstone, 2011; Ebenstein et al., 2017; Deryugina et al., 2019) that OLS estimates of the health impact of air pollution are small and often insignificant. For example, before the monitoring program, doubling pollution increased the all-cause mortality rate (the first blue solid dot in both figures) by 0.64% according to the OLS regression (Panel (a)) and 2.35% according to the IV regression (Panel (b)). The magnitude of the IV estimates is intuitive: the mortality-pollution elasticity is highest for the respiratory category at 6.74% (as $\text{PM}_{2.5}$ directly affects the human respiratory system). The pollution elasticity is close to zero for injuries, which is also sensible.

Second, for both IV and OLS estimates, the monitoring program is associated with a

⁵³Similar strategies have been used in the prior literature. For example, Williams and Phaneuf (2019) uses a buffer zone of 60-120 km radius; Barwick et al. (2020) uses a buffer zone of 150 km radius.

reduction in the pollution elasticity, i.e., a decrease in the pollution gradient. In the figure, the hollow dots (which denote post-monitoring estimates) shift downward from the solid dots (which denote pre-monitoring estimates) for all of the cause categories except for injuries. That is, post the monitoring program, the pollution’s mortality damage is partially mitigated, consistent with our main estimates on the change-in-gradient shown in Table 3.

Third, for all-cause mortality, the change in coefficients is virtually the same whether we use OLS or IV regressions (at 0.02 log points). The orange dashed lines in both figures provide a visual aid to this result, with the two lines having the same length and slope. This finding provides support to the identification assumption in our change-in-gradient analysis: the magnitude of bias appears similar before vs. after the monitoring program and hence cancels out when we focus on the change in the gradient itself.

D.6 Synthetic Control Analysis

The key step of this exercise is that, for each city, we need to identify a set of “control” cities that can serve as reasonable counterfactuals in the absence of the monitoring program. To identify these control cities, we use the synthetic control method following [Abadie, Diamond and Hainmueller \(2010\)](#) and [Arkhangelsky et al. \(2021\)](#). For each city \tilde{c} , we create a group g which contains the city itself and all of its control cities: $g = \{\tilde{c}, c_1, \dots, c_k\}$ with associated synthetic weights $w = \{1, w_1, \dots, w_k\}$ such that $\sum_{i=1}^k w_i = 1$. We will henceforth call \tilde{c} the “treated” city and $\{c_1, \dots, c_k\}$ the “control” cities.

We make two decisions regarding the choice of the control cities and the synthetic weights. First, each treated city is matched to control cities that have not yet been treated by the monitoring program. That is, for each city in wave 1, we choose control cities from waves 2 and 3; for each city in wave 2, we choose controls from wave 3 cities. We drop wave 3 cities because they are in the last wave of the information rollout and there are no untreated cities to serve as their controls.

Second, we select control cities based on outcomes prior to the monitoring policy, but compute synthetic weights using data up to one year before the information program is actually implemented. This allows us to use the year prior to the pre-treatment period as a validation sample to examine the synthetic control method’s performance.

We estimate the following difference-in-differences estimation equation:

$$\begin{aligned} \text{Log}(\text{Mortality})_{gct} &= \beta \cdot \text{Treated}_{gc} \times \text{Post}_{gt} + \gamma \cdot \text{Treated}_{gc} \\ &\quad + \theta \cdot \text{Post}_{gt} + \mathbf{x}'_{ct} \boldsymbol{\theta} + \varepsilon_{ct}, \end{aligned} \tag{D.7}$$

where g denotes a group that contains a treated city and its control cities. A unit in this

panel-data regression is a city c in a group g (thus the subscript gc). $Treated_{gc}$ is a dummy variable for the treated city in each group. $Post_{gt}$ indicates post-treatment periods for group g based on the treated city’s treatment time. The regression is weighted by group-by-city level synthetic weights and includes city-by-group, week-of-year, and year fixed effects, as well as city-by-group-specific trends. Standard errors are clustered at the city level.

Panel (a) in Appendix Figure E.18 shows the event study results. The blue line at the bottom of the chart represents the difference between treated cities’ mortality rates and their synthetically weighted control cities’ mortality rates as a function of time relative to the information rollout. All periods prior to eight months before the monitoring policy are pooled into one group and used as the reference group. The regression includes city, week-of-year, and year fixed effects. The orange line on the top shows the average pollution level in the corresponding event time. The shaded areas highlight the quarters when a pollution peak occurred.

Several patterns emerge from the graph. First, the blue stays close to zero during the matching period. This means that, as expected, treated and control cities’ mortality rates track each other closely during the period used to compute synthetic weights. Second, during the validation period (i.e., one year prior to the information rollout), mortality rates in the treated and control cities continue to be close to each other. Our estimate suggests that, relatively to the the matching period, mortality rate decreases by 0.25 percent (SE = 1.16 percent) during the validation period. Third, we estimate that the mortality rate decreases by 1.64 percent on average after the monitoring policy. This estimate is imprecise with SE=1.70. However, comparing the blue (mortality) and the orange (pollution) lines, we notice much more pronounced gaps in mortality rates between the treated city and the control group during heavily polluted periods post the information program, a visual pattern that was not apparent before the monitoring rollout. In contrast, differences in mortality levels during low pollution periods are modest before vs. after the program. These qualitative patterns suggest that the imprecise average effect of the information program may be masked by important heterogeneity.

To formally analyze this possibility, we test if the reduction in mortality rate after the monitoring program concentrates on periods when the city experiences high pollution levels. We group each city’s weekly observations into quintiles based on its weekly pollution concentration. Then we estimate an augmented version of Equation (D.8), allowing the β coefficient to vary by the pollution quintiles.

Panel (b) in Appendix Figure E.18 reports the heterogeneous β coefficients. Each coefficient represents a synthetic difference-in-differences estimate regarding the monitoring program’s impact on mortality within a quintile of a city’s pollution levels. The monitoring

program has no discernible impact on the mortality rate during periods of low pollution levels, but reduced mortality significantly (both statistically and economically) during high pollution periods. For example, the mortality rate decreases by more than four percent in weeks with the 20 percent highest pollution levels after the monitoring program. Importantly, this result is “isomorphic” to our *gradient* analysis in the main text of the paper, where we document a negative coefficient for the pollution and post-treatment interaction. In theory, this negative relationship could be driven by an increase in mortality during low pollution periods, a decrease in mortality during high pollution periods, or some combination of both. The new insight from the quintile analysis above shows that our main finding is mostly driven by a net reduction in mortality during high pollution periods post the monitoring program. The findings provide empirical justification of our focus on the gradient analysis which leverages the fact that the mortality impact of the program is critical through its interaction with pollution. That is, rather than a (uniform) level shift in the mortality rate across different pollution levels, the program resulted in a change in the slope of the mortality-pollution relationship.

D.7 Annual Analysis

We conclude the empirical exercise with a robustness analysis that uses annual pollution measures. Specifically, we use the annual PM_{2.5} modeling data from [Van Donkelaar et al. \(2016\)](#), which combines satellite-based AOD and chemical transport modeling tools to provide estimates of ground-level PM_{2.5} concentration at the 10km-grid-by-annual frequency. We use this data to calculate two versions of city-year level PM_{2.5}. The first version uses the 10-by-10km grids that correspond to the ground locations of the pollution monitors and averages over these readings to get a city average. This is closest to the level of air quality that would have been captured by the monitors. The second version averages the pollution information from all grids within the city border, which allows us to measure overall pollution conditions in the city. The annual-level estimation model mirrors the primary specification Equation (1), with a few minor differences.

Given the annual frequency, we assign years 2013-2016 as the post period for wave one cities, 2014-2016 for wave two cities, and 2015-2016 for wave three cities. The regression includes city and year fixed effects (η_t) and clusters standard errors at the city level.

Appendix Table [E.10](#) displays the results. Each panel-column represents a regression. The panels differ by the type of pollution measures used: Panel A uses modeling PM_{2.5} near pollution monitors and Panel B uses modeling PM_{2.5} citywide. For reference, panel C uses the satellite-based AOD as in the paper but aggregates to the annual level. Column (1)

reports results for card transactions. The outcome variable is annual rather than weekly transaction rates, and therefore the magnitudes of the coefficients are much larger than the weekly coefficients in Table 2 of the paper. Consistent with our original analysis, the coefficient of “Log(Pollution) \times 1(after monitoring)” shows a negative sign regardless of which pollution measure we use. The magnitude is smaller than if one were to linearly scale up the weekly coefficient from Table 2 of the paper to the annual level (-12 transactions per week \times 52 weeks). This is consistent with intertemporal substitution of consumption. For example, a consumer might postpone a shopping trip due to high pollution, but the trip would eventually occur at some future point.

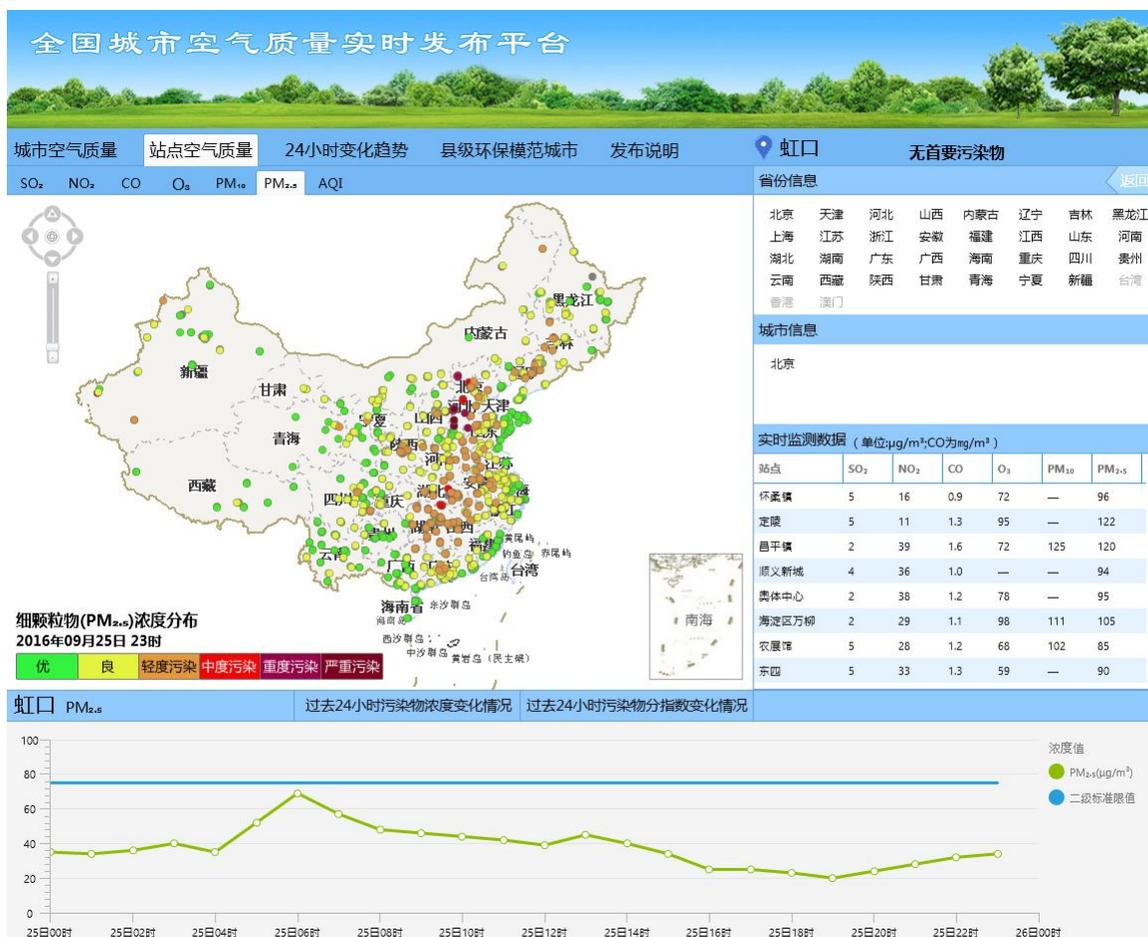
Column (2) reports mortality results. Once again, the coefficients are consistently negative, which echoes our findings in the paper using weekly data. In terms of the magnitude, we find relatively larger effect estimates at the annual level than those from the weekly analysis. This reflects the fact that the mortality effect of pollution in a given week takes more than one week to fully manifest, which is consistent with the findings in the literature (e.g., (Deryugina et al., 2019)). With the much smaller sample size, however, the annual estimates are imprecise in all three panels.

We take away from these exercises that the annual level analysis using alternative pollution measures gives us a broadly consistent conclusion, but we are underpowered by a 50-fold reduction in sample size. The annual level analysis discards rich, short-term pollution variations that are useful to identify the impact of the monitoring program on individuals’ avoidance and health responses. In addition, there is evidence that the modeling PM_{2.5} measure might not provide accurate estimates of ground-level pollution in countries where pollution levels are high (Greenstone et al., 2021; Fowlie, Rubin and Walker, 2019).⁵⁴

⁵⁴Greenstone et al. (2021) calculate trends in PM_{2.5} pollution in China since 2013 and find that the values base on modeling data are lower than the levels recorded by the government monitors. A similar pattern is documented in Fowlie, Rubin and Walker (2019) using the U.S. portion of the modeling PM_{2.5} data. They find that the modeling data understate PM_{2.5} at levels above 15 $\mu\text{g}/\text{m}^3$.

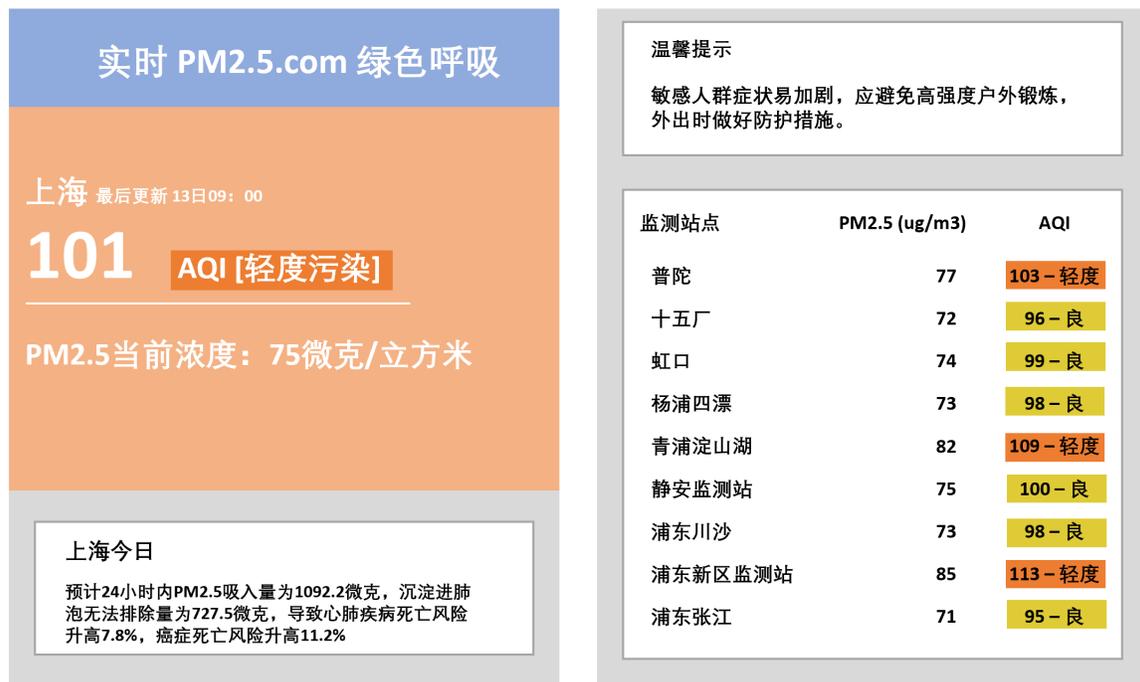
Appendix E: Figures and Tables

Figure E.1: Screenshot of China's Air Quality Disclosure Platform Web Interface



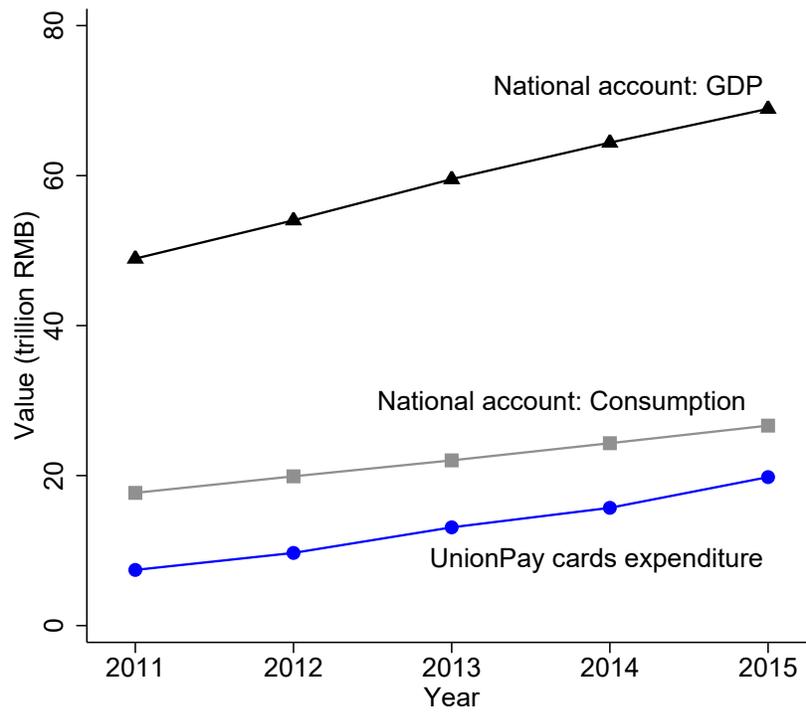
Notes: This figure shows a screenshot of the Ministry of Environmental Protection's real-time air quality disclosure platform web interface as of September 25, 2016. The left panel is an interactive map that displays the locations of all monitoring stations. The right panel reports real-time measures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

Figure E.2: Smartphone Air Quality App



Notes: This figure shows a sketch of what a typical smartphone air quality app looks like. 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 monitoring stations in Shanghai.

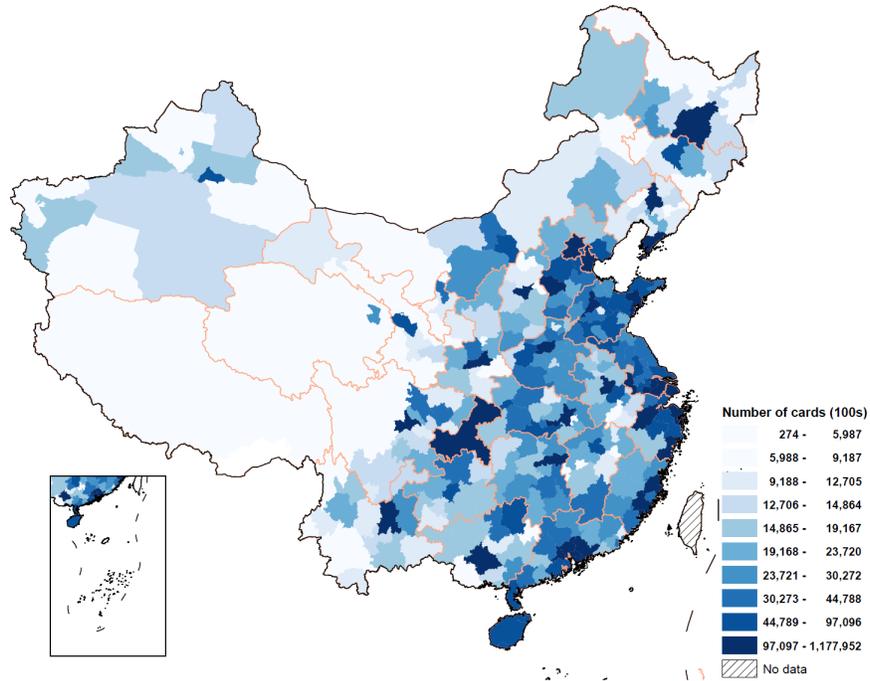
Figure E.3: Consumption Trends: UnionPay vs. National Accounts



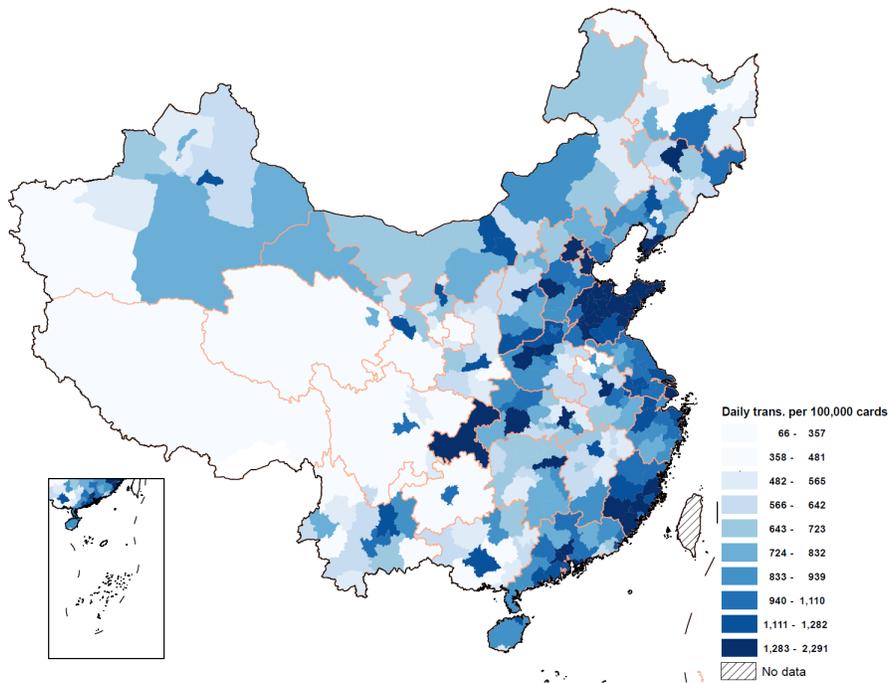
Notes: This figure plots annual GDP (triangles) and aggregate consumption (squares) reported by the National Bureau of Statistics of China (NBS), and total bank card spendings $\times 100$ (circles) aggregated from the UnionPay 1% bank card data, excluding business to business transactions (the wholesale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

Figure E.4: UnionPay Bank Card Penetration by City, 2011-2015 Average

(a) Number of active cards

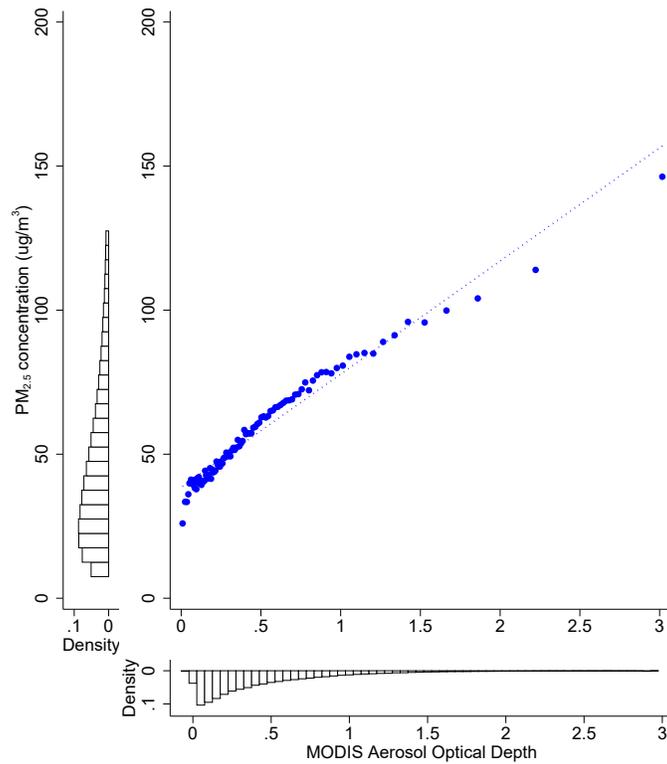


(b) Number of daily transactions per 100,000 cards



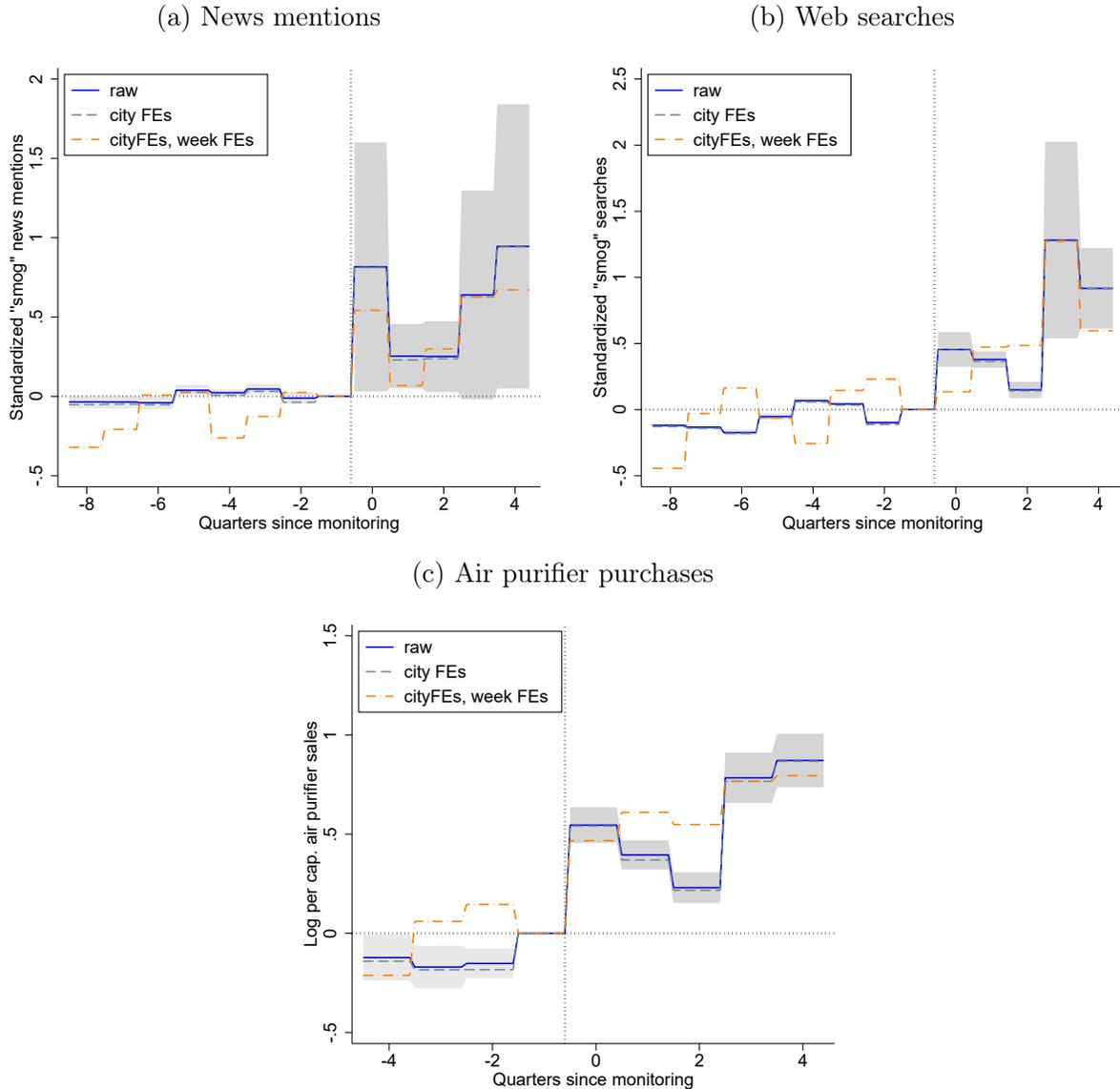
Notes: The maps show the 2011-2015 average number of active UnionPay bank cards (panel (a)) and daily transactions per 100,000 cards (panel (b)) at the city level. Orange lines show inter-provincial borders.

Figure E.5: Correlation between $PM_{2.5}$ and AOD



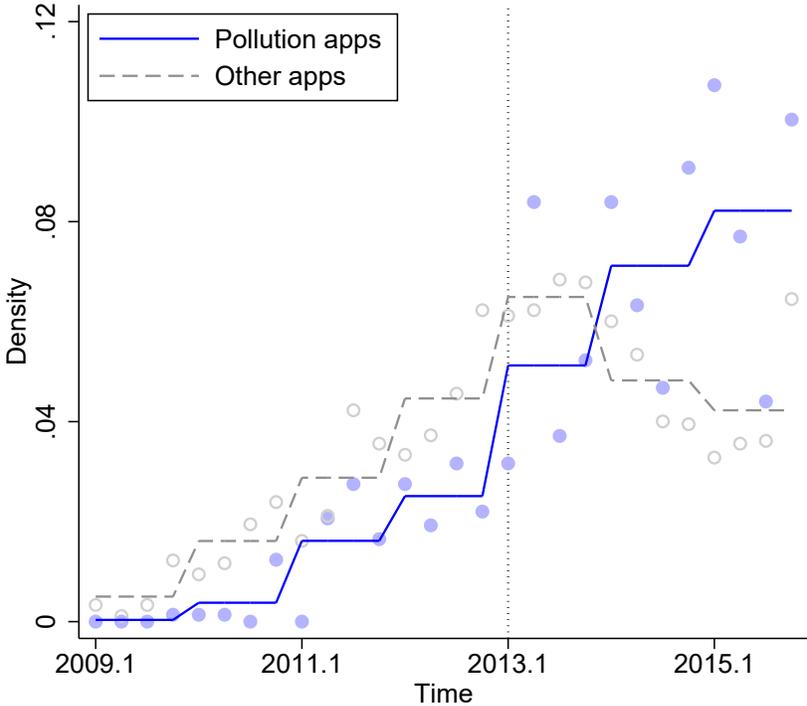
Notes: This graph shows city \times day level average $PM_{2.5}$ concentration (y-axis) by 100 equal bins of AOD (x-axis), for periods after the monitoring program. There is no reliable information on $PM_{2.5}$ before the program. Histograms show the distribution of the two variables.

Figure E.6: Event Time Plots with Different Controls



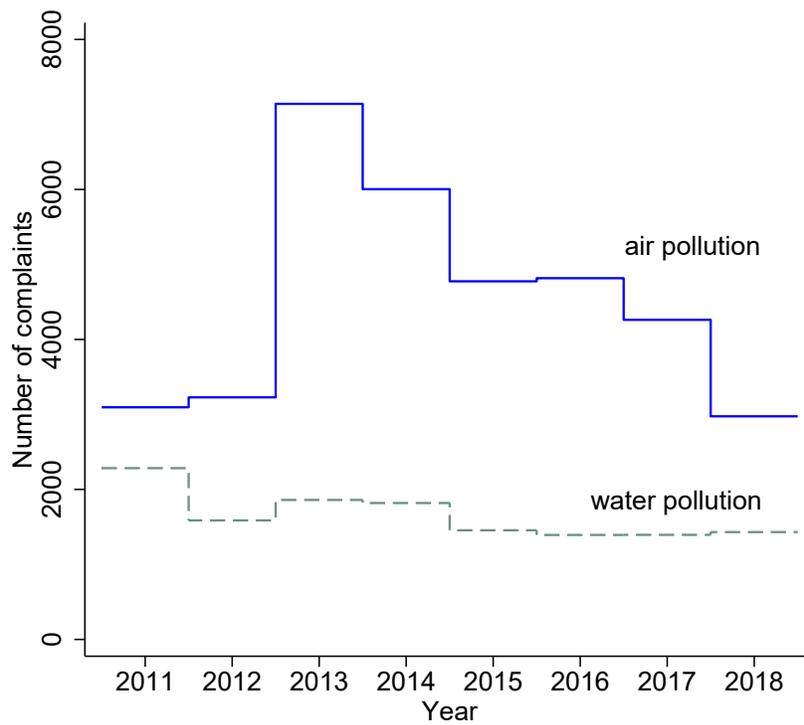
Notes: Figures plot standardized news mentions for smog (panel (a)), standardized smog-related web searches (panel (b)), and log per-capita air purifier purchases in large 50 cities (panel (c)) as a function of quarters since the completion of the monitoring program in a given city. Event quarter -1 is normalized to 0. Solid blue lines show raw data patterns, green dashed lines control for city fixed effects, and orange dashed lines control for city and week fixed effects. Shaded regions show the 95% confidence intervals for the raw data pattern (the blue lines) and are constructed from standard errors clustered at the city level.

Figure E.7: Changes in Pollution Information Access – Mobile Phone Apps



Notes: This chart shows the release-date distribution of Apple App Store apps separately for pollution apps (solid dots and line) and other apps (hollow dots and dashed line). Apps in other categories include games, 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 queried from Apple App Store on December 27, 2015. Pollution apps released before 2013 typically stream weather information and incorporate real-time air quality content post 2013.

Figure E.8: Citizen Complaints on Air and Water Pollution Issues



Notes: This figure plots the total number of complaints posted by citizens that are related to air pollution issues (solid line) and water pollution issues (dashed line) by year. Data are sourced from the message board for government leaders (Agarwal et al. 2020).

Figure E.9: List of Cities by Rollout Waves

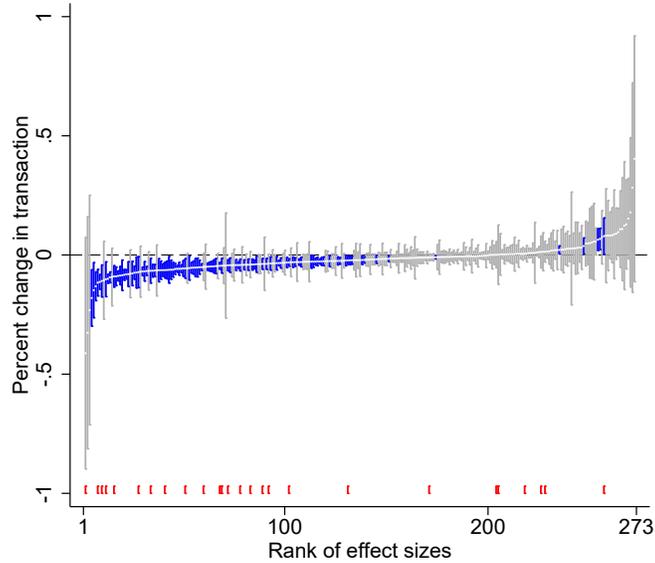
Wave 1 cities			Wave 2 cities					Wave 3 cities							
Beijing	Xining	Taizhou	Wuhu	Jinzhou	Jimo	Wujiang	Yingkou	Tongling	Jixi	Nanping	Ezhou	Guangyuan	Chuxiong	Dingxi	Shihezi
Tianjin	Hefei	Lanzhou	Maanshan	Zhuzhou	Pingdu	Changshu	Panjin	Anqing	Hegang	Longyan	Jingmen	Suining	Honghe	Longnan	Wujiayu
Shijiazhuang	Fuzhou	Hangzhou	Datong	Xiangtan	Laixi	Zhangjiagang	Huludao	Chuzhou	Shuangyashan	Ningde	Xiaogan	Neijiang	Wenshan	Linxia	
Tangshan	Yinchuan	Ningbo	Yangquan	Yueyang	Zibo	Kunshan	Zigong	Chizhou	Yichun	Jingdezhen	Huanggang	Leshan	Xishuangbanna	Gannan	
Qinhuangdao	Wulumuqi	Xi'an	Changzhi	Changde	Zaozhuang	Taicang	Zhuji	Xuancheng	jiamusi	Pingxiang	Xianning	Meishan	Dali	Haidong	
Handan	Jinan	Jiaying	Linfen	Zhangjiajie	Dongying	Haimen	Jiayuguan	Lüliang	Qitaihe	Xinyu	Suizhou	Guangan	Dehong	Haibei	
Xingtai	Nantong	Huzhou	Baotou	Shaoguan	Yantai	Jurong	Deyang	Wuhai	Heihe	Yingtang	Enshi	Dazhou	Nujiang	Huangnan	
Baoding	Zhengzhou	Shaoxing	Chifeng	Shantou	Laizhou	Fuyang	Laiwu	Tongliao	Suihua	Ganzhou	Hengyang	Yaan	Diqing	Hainan	
Zhangjiakou	Wuhan	Jinhua	Anshan	Zhanjiang	Penglai	Lin'an	Dezhou	Hulunbeier	Daxinganling	Ji'an	Shaoyang	Bazhong	Changdou	Guolu	
Chengde	Changsha	Lasa	Fushun	Pingdingshan	Zhaoyuan	Jiaozhou	Binzhou	Bayannaer	Bengbu	Yichun	Yiyang	Ziyang	Shannan	Yushu	
Cangzhou	Guangzhou	Zhoushan	Benxi	Anyang	Weifang	Yiwu	Heze	Wulanchabu	Huainan	Fuzhou	Chenzhou	Aba	Rikaze	Haixi	
Langfang	Shenzhen	Taizhou	Yan'an	Jiaozuo	Shouguang	Jiujiang	Sanmenxia	Xingan	Huaipei	Yongrao	Yongzhou	Ganzi	Naqu	Wuzhong	
Hengshui	Zhuhai	Kunming	Jingzhou	Jinchang	Jining	Quanzhou	Weinan	Xilinguole	Jincheng	Hebi	Huaihua	Liangshan	Ali	Guyuan	
Taiyuan	Foshan	Xiamen	Yichang	Shizuishan	Taian	Eerduosi	Zhangqiu	Alashan	Shuozhou	Xinxiang	Loudi	Liupanshui	Linzhi	Zhongwei	
Huhehaote	Jiangmen	Nanchang	Baoji	Kelamayi	Weihai	Wafangdian	Nanchong	Fuxin	Huangshan	Puyang	Xiangxi	Anshun	Hanzhong	Tulufan	
Shenyang	Zhaoqing	Wenzhou	Xianyang	Kuerle	Wendeng	Maoming	Yuxi	Liaoyang	Jinzhong	Xuchang	Wuzhou	Bijie	Yulin	Hami	
Yangzhou	Huizhou	Qingdao	Jilin	Kaifeng	Rongcheng	Meizhou		Tieling	Fuyang	Luohe	Fangchenggang	Tongren	Ankang	Changji	
Changchun	Dongwan	Dalian	Qiqihaer	Luoyang	Rushan	Shanwei		Chaoyang	Suzhou	Nanyang	Qinzhou	Qianxinan	Shangluo	Boertala	
Haerbin	Zhongshan	Lianyungang	Daqing	Liuzhou	Rizhao	Heyuan		Siping	Liuan	Shangqiu	Guigang	Qiandongnan	Baiyin	Akesu	
Shanghai	Nanning	Huai'an	Mudanjiang	Guilin	Zunyi	Yangjiang		Liaoyuan	Haozhou	Xinyang	Yulin	Qiannan	Tianshui	Kezilesu	
Nanjing	Haikou	Xuzhou	Jiaonan	Beihai	Linyi	Qingyuan		Tonghua	Yuncheng	Zhoukou	Baise	Baoshan	Wuwei	Kashi	
Wuxi	Chongqing	Quzhou	Jiangyin	Sanya	Qujing	Chaozhou		Baishan	Xinzhou	Zhumadian	Hezhou	Shaotong	Zhangye	Hetian	
Yancheng	Chengdu	Suqian	Yixing	Tongchuan	Liaocheng	Jieyang		Songyuan	Putian	Huangshi	Hechi	Lijiang	Pingliang	Yili	
Changzhou	Guiyang	Lishui	Liyang	Panzhuhua	Mianyang	Yunfu		Baicheng	Sanming	Shiyan	Laibing	Puer	Jiuquan	Tacheng	
Suzhou	Zhenjiang		Jintan	Luzhou	Yibin	Dandong		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

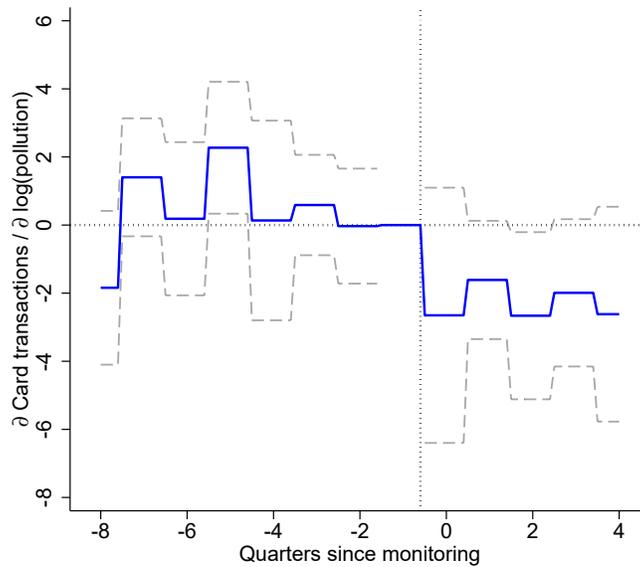
Notes: The three panels tabulate cities included in the official rollout waves of the monitoring program. Color coding indicates administrative hierarchies and pre-designated status. Deep blues are cities in the Jing-Jin-Ji Metropolitan Region, the Yangtze River Delta Economic Zone, the Pearl River Delta Metropolitan Region, direct administered municipalities, and provincial capitals. Light blues are 2007 Environmental Improvement Priority Cities and 1997-2007 National Environmental Protection Exemplary Cities. White denotes remaining cities.

Figure E.10: Pollution Avoidance Among Categories Most Sensitive to Precipitation

(a) % changes in transactions on days with precipitation by category



(b) Avoidance in most precipitation-sensitive categories



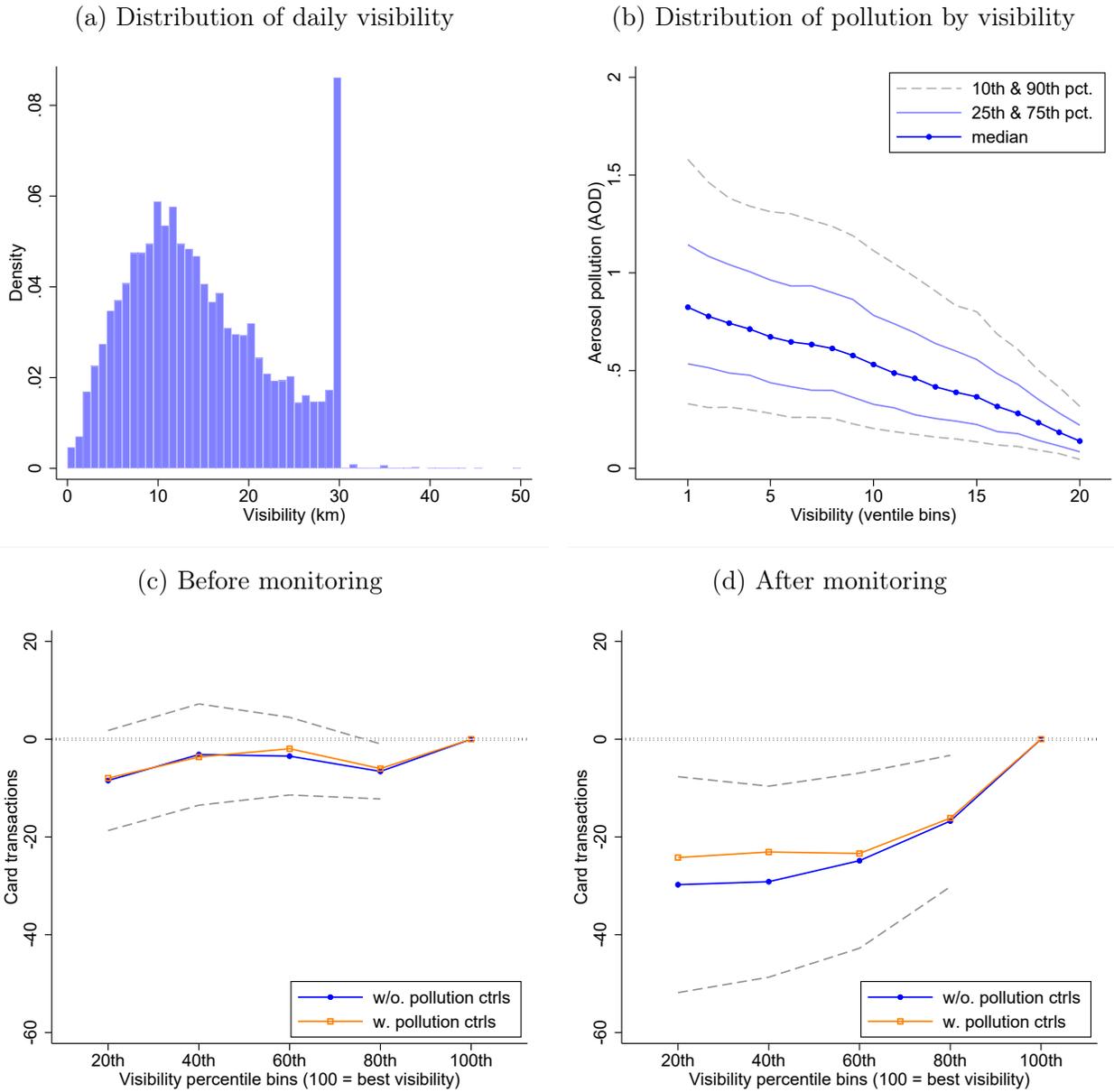
Notes: Panel (a) shows the impact of precipitation on spending for 273 merchant categories. To facilitate comparison across categories, for each category, we regress log city-day transactions on a dummy variable of rain or snow, controlling for city and week-of-year fixed effects. Each bar represents a separate merchant category. A bar's middle point denotes the percentage change in transactions on rainy/snowy days and its range corresponds to the 95% confidence interval. Blue (grey) color represents precise (imprecise) estimates. Red bars at the bottom mark merchant categories in supermarkets, dining, and entertainment sectors that are defined as deferrable in Table 2. Panel (b) conducts an event study using city-weekly transactions and limiting to the top 20 categories whose transactions are most sensitive to rain/snow. The underlying regression controls for city, week-of-year, and year fixed effects and city-specific time trends. Shaded region shows the 95% confidence interval constructed from standard errors clustered at the city level.

Figure E.11: List of Top 20 Most Precipitation-Sensitive Merchant Categories

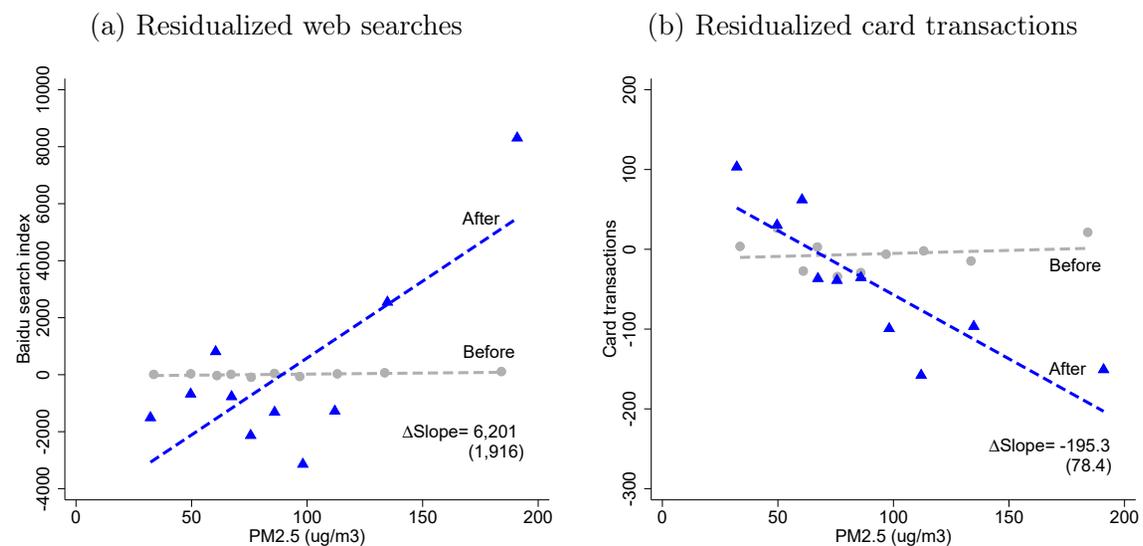
MCC Code	MCC Name	Effect Size	<i>t</i> -stat
5598	snow car dealers	-0.181	-2.99
7012	timeshare vacation houses	-0.148	-2.51
7375	information retrieval services	-0.131	-4.30
7941	sports stadiums, clubs, playgrounds	-0.116	-4.09
5691	garments store	-0.113	-7.15
7996	amusement parks, circuses, carnivals	-0.109	-3.20
4733	large resort ticket	-0.102	-2.74
8043	opticians, optical products	-0.100	-10.46
5941	sporting goods	-0.094	-8.33
7991	tourism, exhibition	-0.090	-3.93
5661	shoe store	-0.090	-4.59
5641	children and baby goods	-0.088	-8.30
5271	trailer dealer	-0.086	-4.41
5947	gifts, souvenir	-0.085	-2.63
5977	cosmetics store	-0.083	-7.14
7538	auto service shops (non-dealer)	-0.081	-8.41
5621	women's fashion store	-0.080	-3.04
5681	fur store	-0.077	-2.98
5655	sportswear	-0.074	-5.33
9399	other government services	-0.074	-2.22
5300	membership bulk retail	-0.072	-2.18

Notes: this table lists the top 20 most precipitation-sensitive merchant categories that are identified in panel (a) of Figure E.10. The event study in panel (b) of Figure E.10 uses these categories. “Effect Size” is the percentage changes in daily transaction volume in response to any precipitation, rain or snow, during the day. *t*-statistics are computed using standard errors clustered at the city level.

Figure E.12: Visibility as Private Information of Air Pollution

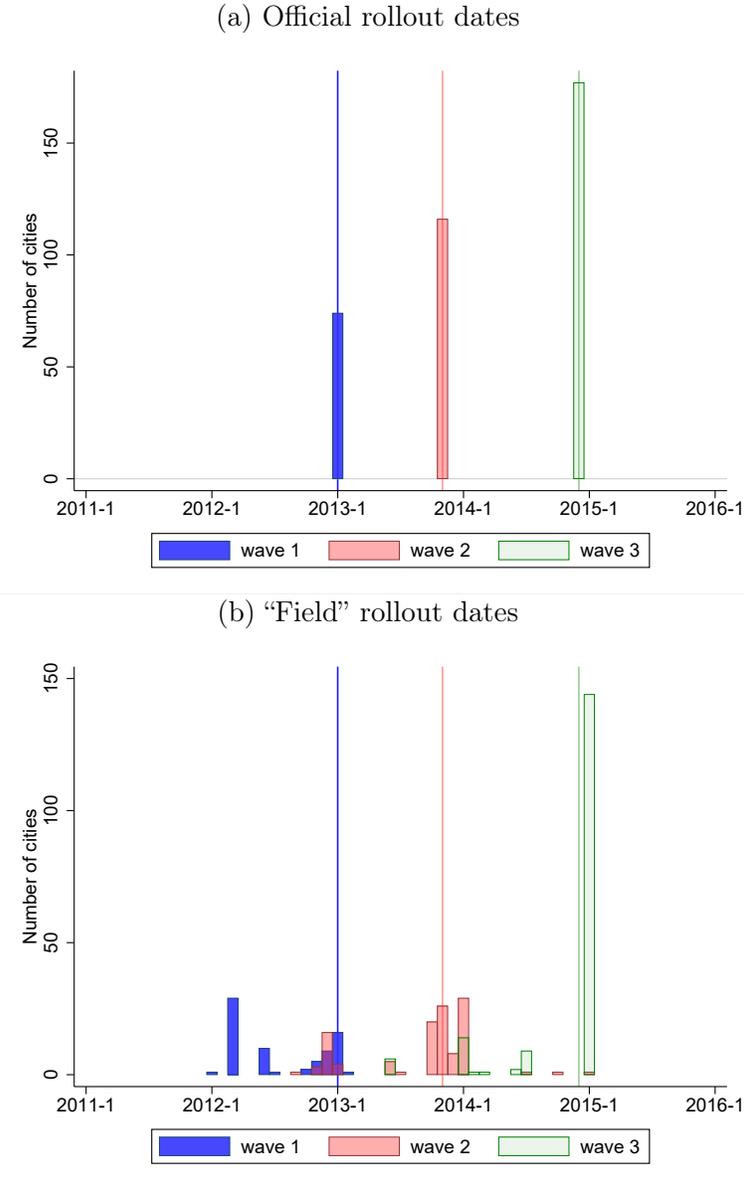


Notes: Panel (a) shows the distribution of raw visibility measure at the station-day level. For the vast majority of stations, visibility is top coded at 30 km (18.6 miles). Panel (b) shows the median, inter-quartile range, and 10-90 percentile range of satellite-based AOD by ventile bins of visibility. Panels (c) and (d) plot card transactions by visibility bins, separately for before and after the program. The underlying regressions control for city, week-of-year, and year fixed effects and city-specific linear time trends. The blue lines with solid circles show the effect of visibility without controlling for the pollution (AOD) levels. The orange lines show the effect of visibility controlling for pollution levels. Dashed lines show 95% confidence intervals for the no-pollution-controls specifications, calculated using standard errors clustered at the city level.

Figure E.13: Awareness and Avoidance vs. U.S. Embassy PM_{2.5} Readings

Notes: Each panel shows decile bin scatterplots of an outcome against U.S. Embassy's PM_{2.5} readings in Beijing, separately for before vs. after the monitoring program. Outcome variables are residualized Baidu searches for smog (panel a) and card transactions per 10,000 active cards (panel b) in Beijing after partialling out week-of-year fixed effects. “ ΔSlope ” reports the change in the outcome-pollution gradient after the monitoring policy. Robust standard errors are reported in parentheses.

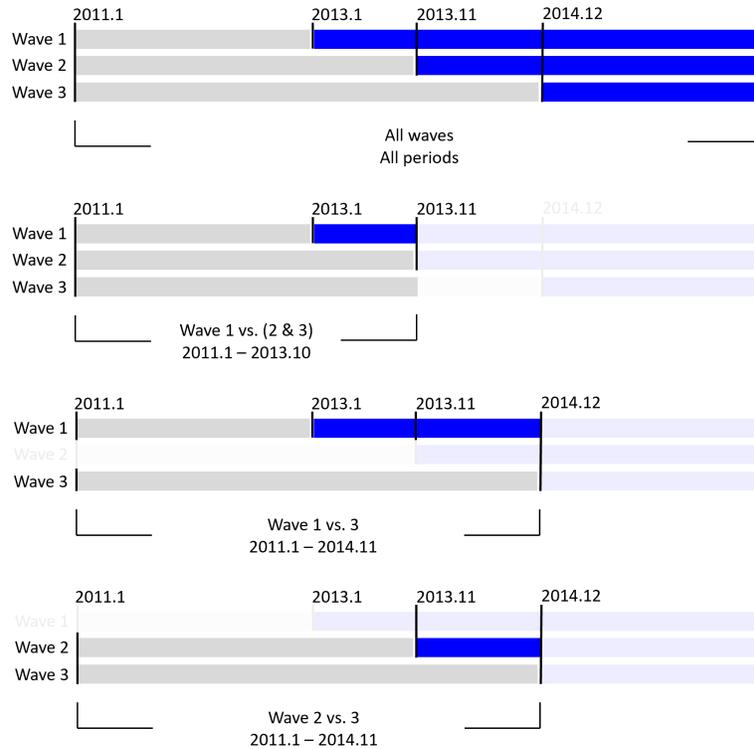
Figure E.14: Distribution of Cities by Rollout Dates



Notes: Panel (a) shows the number of cities for each official rollout dates. Panel (b) reports the number of cities for each "field" rollout date, the earliest date when a city's pollution monitoring data became available. We collected field rollout dates for over 92% of cities in our sample by searching news media and city government yearbooks. Overall, the field dates are close to the official dates. Using the field rollout dates produces similar results to those reported in the main text.

Figure E.15: Staggered Rollout Design: Treated vs. Not-yet-Treated Comparison

(a) Illustration of comparison groups

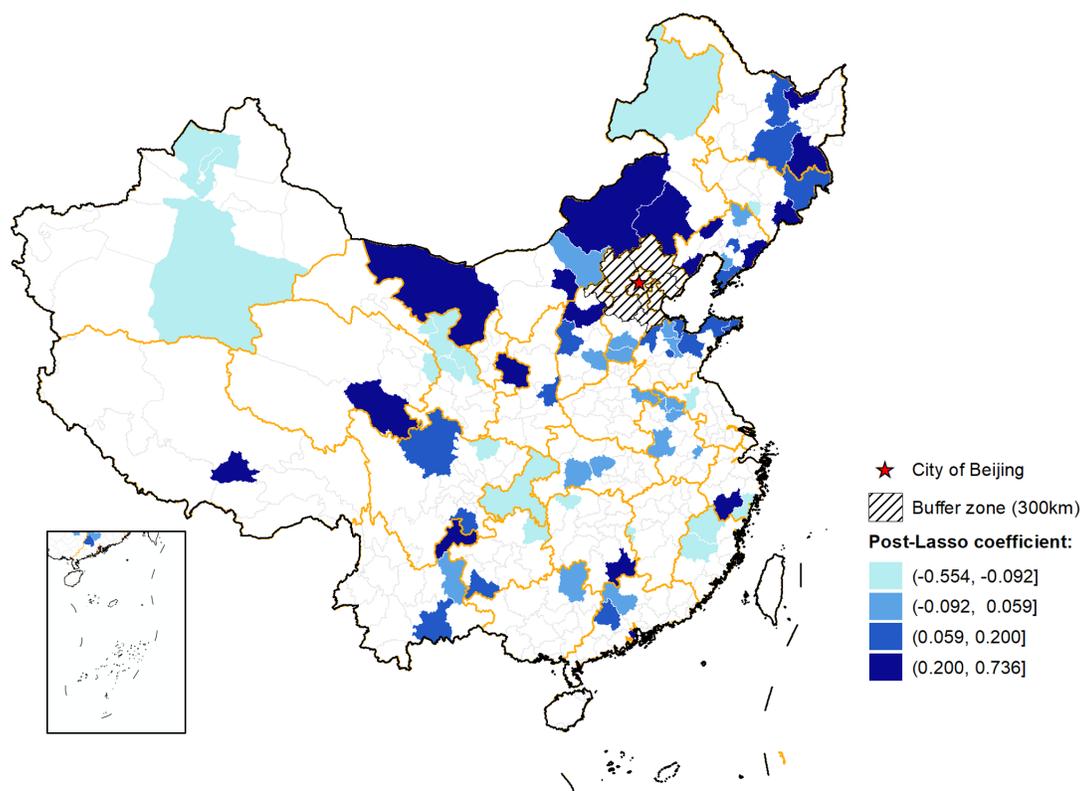


(b) Estimation results

	(1) All waves All periods	(2) Wave 1vs(2&3) ~2013.10	(3) Wave 1vs3 ~2014.11	(4) Wave 2vs3 ~2014.11
Panel A. Card transactions				
Log(Pollution) × 1(after monitoring)	-10.50 (4.45)	-15.07 (6.49)	-13.30 (4.49)	-5.86 (3.96)
Observations	83,122	44,224	45,433	47,623
Panel B. Mortality				
Log(Pollution) × 1(after monitoring)	-0.023 (0.007)	-0.026 (0.016)	-0.030 (0.013)	-0.053 (0.016)
Observations	36,369	17,249	16,693	16,901

Notes: Panel (a) illustrates the three comparison groups. Panel (b) reports estimation results using different comparison groups. Each panel-column is a separate regression using different cities and sample periods, as indicated by column head. “Log(Pollution)” is logged AOD in a city × week. In panel A, dependent variable is city × weekly bank card transactions per 10,000 active cards. Mean transaction rates are 869.1 (column 1), 821.1 (column 2), 850.1 (column 3), and 581.0 (column 4). In panel B, dependent variable is city × weekly log mortality rate. The coefficient of Log(pollution) × 1(after monitoring) reports changes in mortality-pollution elasticity after monitoring. All regressions control for city, year, week-of-year fixed effects and city-specific time trends. Standard errors are clustered at the city level.

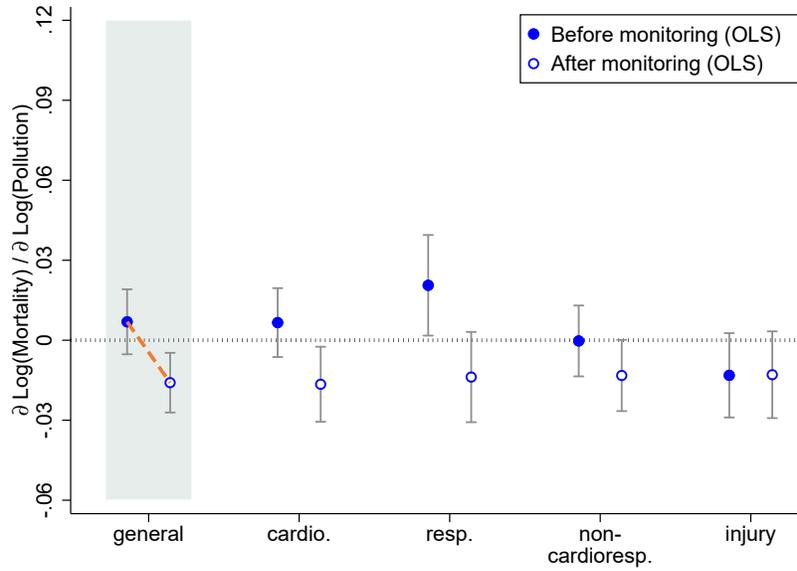
Figure E.16: Upwind Cities Selected by the “Zero-Stage” Lasso Regression: Beijing



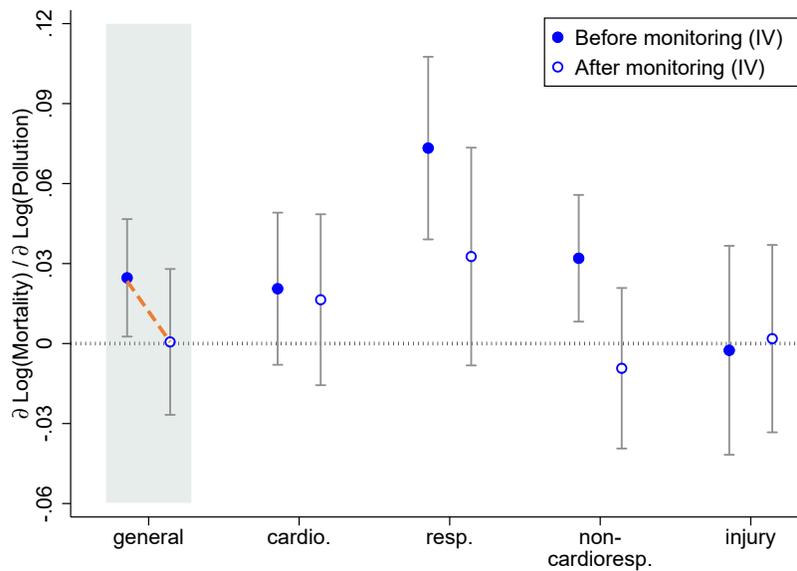
Notes: This map highlights 73 cities selected by a “zero-stage” Lasso regression of Beijing’s daily AOD on all other 330 cities’ upwind component vector AOD.

Figure E.17: OLS and IV Estimates of the Mortality-Pollution Gradient

(a) OLS Estimates



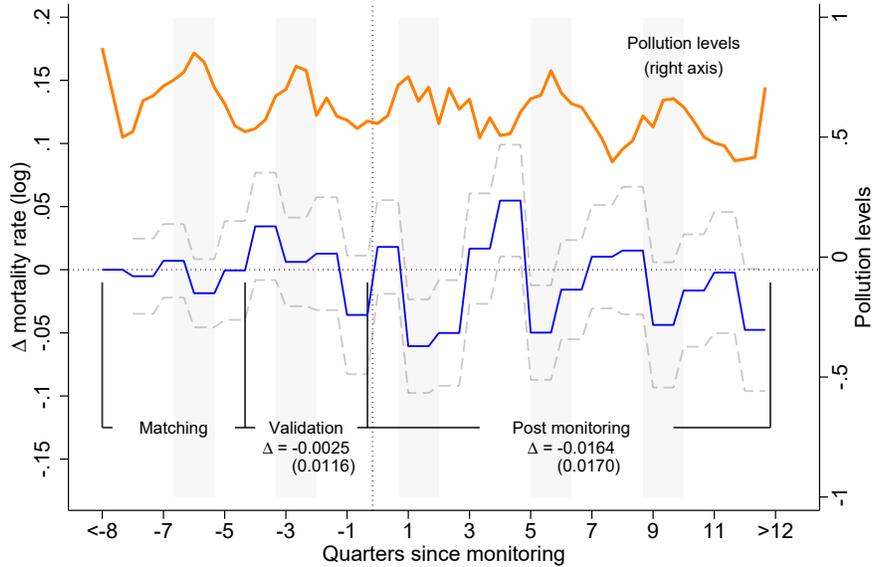
(b) IV Estimates



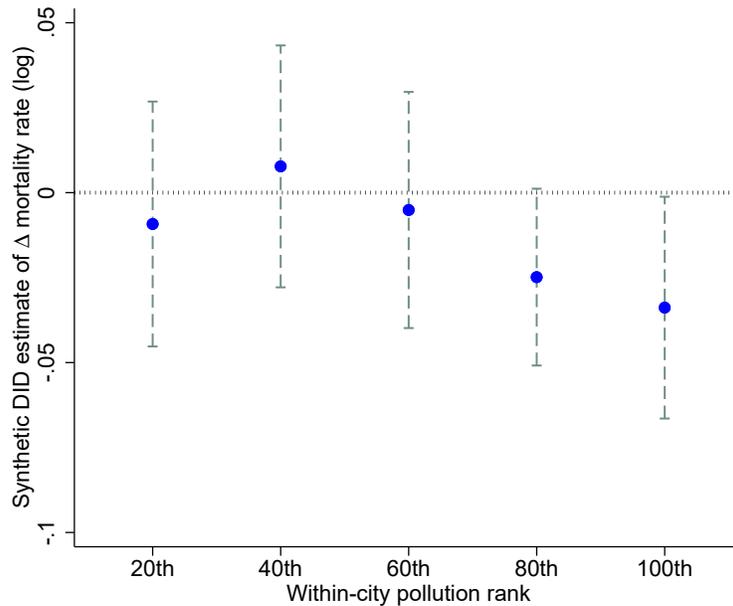
Notes: Panel (a) reports the OLS estimates of the mortality-pollution gradients by cause of death before the program (solid circles) and after the program (open circles). Panel (b) shows the IV estimates. We run separate regressions for each mortality cause. Both OLS and IV results suggest a reduction of the mortality-pollution gradient after the program. For all-cause mortality, the OLS and IV estimates exhibit virtually the same reduction in mortality-pollution gradient. The orange dashed lines, which have the same slope and length, serve as visual aid. This confirms the assumptions in Section 4.2: the magnitude of bias appears similar before vs. after monitoring and cancels out when we focus on the change in gradient.

Figure E.18: The Monitoring Program's Mortality Impact – Synthetic Control Method

(a) Event study of the mortality impact



(b) Mortality impact by within-city pollution quintile



Notes: In panel (a), the blue line (and left axis) shows the difference between treated cities' mortality rates and their synthetically weighted control cities' quarterly mortality rates as a function of time relative to the information rollout, conditioning on city and quarter-of-year fixed effects. The orange line (and right axis) shows the treated cities' average monthly pollution level in the corresponding event time. Shaded areas highlight the quarters when a pollution peak occurred. Panel (b) shows the synthetic DID estimates of the monitoring program's mortality impact by within-city pollution quintiles. The dependent variable is city-week log mortality rate and the sample consists of city-weeks in different pollution quintiles. Regressions control for city-by-group, week-of-year fixed, and year fixed effects, as well as group-specific linear time trends. Range bars display the 95% confidence intervals constructed using standard errors clustered at the city level.

Table E.1: Chronology of Key Events Related to the Monitoring Program

Time	Key Events
Jun 2004	The first known use of the word “Smog” in news media is found in the weather column of Beijing Daily, a popular local newspaper in Beijing
Apr 2008	US Embassy in Beijing installs a rooftop air quality monitor and reports hourly PM _{2.5} via its twitter account @beijingair
Apr 2009	A report by the National People’s Congress recognizes PM _{2.5} as a major pollutant, a first in official public documents
Jan 2010	China Meteorological Agency issues national standards on the observation and forecast of smog based on visibility levels
Nov 2010	First draft for NAAQS amendment opens for public comments, deeming the choice of national PM _{2.5} standards requires public input
Nov 2011	Environmental NGOs start campaign “Measuring Air Quality for our Motherland” and call for systematic PM _{2.5} monitoring by the government
Nov 2011	Second draft for NAAQS amendment opens for public comments, proposing national standards for PM _{2.5} for the first time
Feb 2012	NAAQS amended to set national standards for PM _{2.5} for the first time
Jun 2012	Vice Minster of China’s Ministry of Environmental Protection objects the release of PM _{2.5} data by US Embassy during a press conference
Jan 2013	The first wave of Air Quality Monitoring and Disclosure program officially rolls out in 74 cities, with two other waves to follow
Sep 2013	China State Council issues “Air Pollution Prevention and Control Action Plan” with PM _{2.5} -reduction targets in key areas
Mar 2014	Chinese Premier Li Keqiang first declared “war on smog”

Notes: NAAQS stands for the National Ambient Air Quality Standards, first established in 1982.

Table E.2: Changes in Web Searches-Pollution Gradient

Dep. var.: Standardized web searches for “smog”				
	(1)	(2)	(3)	(4)
Log(Pollution)	0.412 (0.286)	0.200 (0.212)	0.098 (0.105)	0.073 (0.089)
Log(Pollution) \times 1(after monitoring)	0.205 (0.105)	0.339 (0.108)	0.438 (0.154)	0.371 (0.136)
FEs: city, city-linear time trends	✓	✓	✓	✓
FEs: month-of-year	✓			
FEs: year	✓			
FEs: month-of-sample		✓	✓	
FEs: region \times year			✓	
FEs: region \times month-of-sample				✓

Notes: Number of observations is 83,122. “Log(Pollution)” is logged AOD in a city \times week. “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 city level.

Table E.3: Determinants of Monitoring Roll-out Waves

	(1) 1(wave-1)	(2) 1(wave-2)	(3) 1(wave-3)	(4) wave	(5) wave	(6) wave
1(■)	0.925 (0.032)	0.015 (0.015)	0.060 (0.029)	-1.711 (0.066)	-1.467 (0.108)	
1(□)	0.119 (0.040)	0.881 (0.040)	- -	-0.964 (0.049)	-0.856 (0.052)	
1(□)	0.015 (0.009)	0.125 (0.023)	0.860 (0.025)	- -	- -	
Pollution level					-0.100 (0.046)	-0.254 (0.079)
Pollution trends					-0.004 (0.004)	0.00022 (0.0066)
Per capita income level					-0.654 (0.231)	-1.876 (0.224)
Per capita income trends					0.143 (0.156)	-0.039 (0.229)
Constant				2.845 (0.028)	3.386 (0.111)	4.470 (0.118)
R ²	0.800	0.648	0.842	0.747	0.769	0.427

Notes: In columns 1-3, the dependent variable 1(wave-k) is an indicator variable for whether a city is among the wave-k cities in the actual roll-out. In columns 4-6, the dependent variable is the roll-out wave as a “continuous” variable whose value equals 1, 2, or 3. The indicator variable 1(deep blue) denotes cities in the Jing-Jin-Ji Metropolitan Region, the Yangtze River Delta Economic Zone, the Pearl River Delta Metropolitan Region, direct administered municipalities, and provincial capitals. The indicator variable 1(light blue) denotes the 2007 designated Environmental Improvement Priority Cities and the 1997-2007 National Environmental Protection Exemplary Cities. The indicator variable 1(white) denotes the remaining cities. Variables ‘Pollution level’, ‘Pollution trends’, ‘Per capita income level’, and ‘Per capita income trends’ are divided by their mean (so the coefficients indicate percent changes).

Table E.4: 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 (4.85)	3.90 (2.10)	2.90 (1.95)
GDP per capita (yuan)	69,836 (27,627)	42,881 (23,110)	27,400 (13143)
AOD level	0.665 (0.239)	0.600 (0.242)	0.456 (0.237)
PM _{2.5} level (ug/m ³)	61.3 (22.1)	57.9 (20.2)	46.0 (17.4)
Industrial SO ₂ emissions (ton)	37,569 (40,186)	29,609 (24,695)	18,214 (17,550)
Average temperature (F)	59.7 (8.52)	58.0 (9.59)	55.3 (10.6)
Total precipitation (inches)	47.0 (21.9)	42.2 (23.2)	40.3 (24.4)
Average wind speed (m/s)	1.94 (0.63)	1.71 (0.62)	1.47 (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 SO₂ emissions (year 2006). The table reports the average characteristics of cities in different waves. Standard deviations are in parentheses.

Table E.5: Changes in the Economic and Regulatory Environment After Monitoring

Indep. var.: 1(after monitoring)	(1)	(2)	(3)	(4)
Panel A. Pollution levels				
Log(Pollution)	0.0054 (0.0095)	0.0044 (0.0094)	0.0025 (0.0096)	-0.0042 (0.0088)
Log(max Pollution)	-0.0001 (0.0127)	-0.0077 (0.0107)	-0.0100 (0.0109)	-0.0140 (0.0096)
Panel B. Political/regulatory environment				
^a N(anti-corruption cases)	-0.004 (0.044)	-0.028 (0.047)	0.010 (0.026)	0.005 (0.027)
^b Age(mayor)	0.335 (0.158)	0.313 (0.167)	0.350 (0.163)	0.354 (0.165)
^c Mayor having a Ph.D. degree	-0.017 (0.023)	-0.016 (0.025)	-0.026 (0.025)	-0.026 (0.025)
^d N(“pollution regulation” news mention)	-0.0062 (0.0064)	-0.0092 (0.0071)	-0.0084 (0.0074)	-0.0084 (0.0075)
Panel C. Healthcare access				
^e Log N(hospitals per 100,000 people)	0.085 (0.080)	0.089 (0.083)	0.094 (0.085)	0.095 (0.087)
FEs: city, city-linear time trends	✓	✓	✓	✓
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 Mayor having a Ph.D. degree	mean = 0.234,		sd = 0.423	
^d N(“pollution regulation” news)	mean = 0.052,		sd = 0.45	
^e N(hospitals per 100,000 people), annual frequency	mean = 4.56,		sd = 3.16	

Notes: Each cell is a regression. Estimation data are at the city × weekly level, except for Panel C which uses city × annual hospital counts. Row names show the dependent variable. “Log(Pollution)” is logged AOD in a city×week. “Anti-corruption cases” is the number of local officials ousted during the anti-corruption campaign, “Mayor having a Ph.D. degree” 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). Standard errors are clustered at the city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table E.6: Changes in Air Purifier Purchase-Pollution Gradient

Dep. var.: Number of air purifier purchases per 1,000 people				
	(1)	(2)	(3)	(4)
Log(Pollution)	1.34 (1.33)	-1.76 (1.28)	-2.12 (1.25)	-1.79 (1.61)
Log(Pollution) \times 1(after monitoring)	1.86 (2.09)	6.21 (2.44)	6.69 (2.70)	6.50 (3.26)
FEs: city, city-linear time trends	✓	✓	✓	✓
FEs: month-of-year	✓			
FEs: year	✓			
FEs: month-of-sample		✓	✓	
FEs: region \times year			✓	
FEs: region \times month-of-sample				✓

Notes: The analysis contains 50 cities for which we have air purifier sales data between 2012-2016. Number of observations is 9,871. Mean of dependent variable is 64.1 purchases per month per 1,000 city residents. “Log(Pollution)” is logged AOD in a city \times month. These 50 cities account for 28% of the national population; among them, 34, 11, and 5 cities are in the first, second, and third wave of the information program rollout, respectively. All regressions include lower-order interaction and main effect terms. Standard errors are clustered at the city level.

Table E.7: Changes in Bank Card Transaction-Pollution Gradient – Robustness

Coef. of interest: $\text{Log}(\text{Pollution}) \times 1(\text{after monitoring})$				
	(1)	(2)	(3)	(4)
Drop cities w/ top 10% anti-corruption cases	-10.47 (4.50)	-13.29 (5.26)	-8.65 (4.79)	-12.62 (5.53)
Control for online shopping shares	-10.58 (4.54)	-13.21 (4.83)	-7.86 (4.41)	-12.02 (4.95)
Control for weather variables	-10.36 (5.14)	-12.76 (5.54)	-8.63 (5.03)	-12.94 (5.68)
Weekly max pollution	-12.12 (5.97)	-12.68 (5.59)	-7.42 (4.37)	-10.59 (4.87)
Manually collected “field” roll-out date	-10.10 (3.72)	-12.34 (3.85)	-7.23 (3.48)	-8.08 (3.75)
Quarterly aggregation	-180.15 (99.35)	-248.60 (149.89)	-204.77 (150.74)	-205.40 (135.80)
FEs: city, city-linear time trends	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region \times year			✓	
FEs: region \times 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 $\text{Log}(\text{Pollution})$ term is not reported in the interest of space. Weather controls include linear terms of weekly temperature, precipitation, wind speed, barometric pressure, and their full interactions. Standard errors are clustered at the city level. Numbers of observations are (from top to bottom rows): 74903, 82703, 71857, 83122, 83122, 6674.

Table E.8: Changes in Transaction-Pollution Gradient - Spatial Spillovers

Dep. var.: Number of transactions per 10,000 active cards in a city \times week				
	(1)	(2)	(3)	(4)
Log(Pollution) \times 1(after monitoring)	0.15 (0.71)	-0.29 (0.83)	-0.36 (0.54)	-0.28 (0.63)
Log(Pollution) \times 1(after monitoring) \times 1(Treated)	-9.05 (5.71)	-10.03 (7.34)	-9.00 (6.85)	-10.66 (8.00)
FEs: city-group, city-linear time trends	✓	✓	✓	✓
FEs: week-of-year	✓			
FEs: year	✓			
FEs: week-of-sample		✓	✓	
FEs: region \times year			✓	
FEs: region \times week-of-sample				✓

Notes: Number of observations is 150,289. A treated city and its neighboring cities not yet experiencing the roll-out constitute a city-group. “Log(Pollution)” is logged AOD in a city \times week. “1(Treated)” equals 1 for cities in the roll-out wave and 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). All regressions include lower-order interaction and main effect terms. All fixed effects are interacted with 1(Treated) dummy. Standard errors are clustered at the city level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table E.9: Changes in Card Transaction-Pollution Gradient: Visibility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(1/Visibility)	10.12 (6.59)	3.08 (6.18)	3.96 (5.96)	15.46 (7.01)	11.46 (10.57)	0.23 (9.57)	1.51 (7.86)	10.20 (8.88)
Log(1/Visibility) \times 1(after monitoring)	-11.98 (12.92)	-13.49 (13.28)	-7.54 (9.92)	-10.86 (10.64)	-6.05 (16.40)	-1.52 (18.72)	2.95 (13.54)	1.60 (14.71)
Log(Pollution)					3.25 (5.19)	3.47 (5.63)	2.09 (4.20)	3.21 (4.64)
Log(Pollution) \times 1(after monitoring)					-9.38 (4.86)	-12.88 (8.01)	-9.06 (6.42)	-12.87 (7.54)
FEs: city, city-linear time trends	✓	✓	✓	✓	✓	✓	✓	✓
FEs: week-of-year	✓				✓			
FEs: year	✓				✓			
FEs: week-of-sample		✓	✓			✓	✓	
FEs: region \times year			✓				✓	
FEs: region \times week-of-sample				✓				✓

Notes: Number of observations is 83,122. “Log(1/Visibility)” is logged inverse visibility in a city \times week that is based on weather-station data sourced from China’s National Meteorological Administration. “Log(Pollution)” is logged AOD in a city \times week. Dependent variable is city \times weekly bank card transactions per 10,000 active cards. “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 city level.

Table E.10: Annual Analysis using Alternative Pollution Measurements

	(1) Card	(2) Mortality
<hr/> Panel A. Pollution = Modeling PM _{2.5} near monitoring site <hr/>		
Log(Pollution)	-656.6 (349.2)	0.039 (0.049)
Log(Pollution) × 1(after monitoring)	-167.7 (173.3)	-0.032 (0.087)
<hr/> Panel B. Pollution = Modeling PM _{2.5} citywide <hr/>		
Log(Pollution)	-490.3 (280.2)	0.032 (0.049)
Log(Pollution) × 1(after monitoring)	-282.9 (119.4)	-0.054 (0.092)
<hr/> Panel C. Pollution = AOD <hr/>		
Log(Pollution)	343.6 (181.3)	0.119 (0.058)
Log(Pollution) × 1(after monitoring)	-329.6 (147.5)	-0.111 (0.076)
Observations	1,998	774
FEs: city	✓	✓
FEs: year	✓	✓

Notes: Log(Pollution) in Panels A and B is based on the modeling PM_{2.5} measure from [Van Donkelaar et al. \(2016\)](#) at the 10km-grid-by-annual frequency. For each city, Panel A averages over the annual pollution readings from grids that are closest to the city's monitoring stations, while Panel B averages over all grids within the city. Log(Pollution) in Panel C is the city-annual AOD measure. Dependent variables are the number of annual transactions per 10,000 active cards in each city (column 1) and the log mortality rate in the city year (column 2). Standard errors are clustered at the city level.

Table E.11: 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. Card transactions					
Log(Pollution) \times 1(after monitoring) \times 1(below average)	-4.99 (3.34)	-7.26 (4.82)	-6.40 (5.57)	-7.84 (3.02)	-11.00 (3.30)
Log(Pollution) \times 1(after monitoring) \times 1(above average)	-9.22 (6.79)	-10.80 (7.35)	-13.17 (7.70)	-10.14 (7.04)	-11.81 (6.84)
Panel B. Mortality					
Log(Pollution) \times 1(after monitoring) \times 1(below average)	-0.019 (0.008)	-0.018 (0.009)	-0.014 (0.008)	-0.020 (0.010)	-0.014 (0.009)
Log(Pollution) \times 1(after monitoring) \times 1(above average)	-0.035 (0.011)	-0.033 (0.008)	-0.038 (0.013)	-0.046 (0.011)	-0.041 (0.010)

Notes: This table reports heterogeneous changes in the purchase-pollution gradient (panel A) and mortality-pollution gradient (panel B) by above and below average city characteristics. Each column corresponds to a separate city characteristic: column 1 = per capita personal disposable income; column 2 = share of urban population; column 3 = per capita number of hospitals; column 4 = per capita residential electricity usage; column 5 = share of mobile phone users. City characteristics are computed as the 2011-2015 average. Cities with missing attributes are omitted from the analysis. All regressions control for city, week-of-year, and year fixed effects, full sets of lower-order interaction terms, and city-specific time trends. Standard errors are clustered at the city level.