Air Pollution, Health Spending and Willingness to Pay for Clean Air in China

Panle Jia Barwick Shanjun Li Deyu Rao Nahim Bin Zahur*

November 2017

Abstract

Understanding the health impact of air pollution and consumer willingness to pay (WTP) for clean air is critical for understanding the benefit of environmental regulations. Based on the universe of credit and debit card transactions in China from 2013 to 2015, this paper quantifies the health impacts of $PM_{2.5}$ and estimates consumer WTP for improved air quality. To our knowledge, this is the first comprehensive analysis of how air pollution affects health spending for the entire population of a developing country. To address potential endogeneity in air quality, we construct an instrumental variable by modeling the spatial spillovers of $PM_{2.5}$ due to long-range transport. Our analysis shows that $PM_{2.5}$ has significant impacts on health spending in both the short and long term and that consumers exhibit avoidance behavior in spending. The annual reduction in health spending from complying with the World Health Organization's annual standard of 10 $\mu g/m^3$ would amount to over \$50 billion, nearly 9% of China's total health spending in 2015. Our estimates suggest a lower bound of annual household WTP of \$10 for a 10 $\mu g/m^3$ reduction in $PM_{2.5}$.

Keywords: Air Pollution, Health Impact, Avoidance Behavior, Willingness to Pay

*Barwick: Department of Economics, Cornell University and NBER (email: panle.barwick@cornell.edu); Li: Dyson School of Applied Economics and Management, Cornell University and NBER (email: sl2448@cornell.edu); Rao: Dyson School of Applied Economics and Management, Cornell University (email: dr526@cornell.edu); Zahur: Department of Economics, Cornell University (email: nz96@cornell.edu). We thank Solomon Hsiang, Matthew E. Kahn, Lucija Muehlenbachs, Matthew Turner and seminar participants at Beijing University, Beihang University, Clemson University, Jinan University, MIT, Penn State, University of Arizona, University of Chicago, 2017 China India Insights Conference, and the 10th Annual Conference on China's Economic Development at George Washington University for helpful comments.

1 Introduction

The health impact of air pollution and consumer willingness to pay (WTP) for clean air are important components of the overall benefit of environmental regulations. A rich literature from epidemiology and more recently economics has consistently shown a positive association between exposure to air pollution such as particulate matter and carbon monoxide and mortality. These findings have provided guidance on air quality regulations such as setting up or tightening ambient air quality standards. For example, research on the health impacts of particulate matter has led the U.S. Environmental Protection Agency (EPA) to establish a standard for PM_{10} in 1987 and for $PM_{2.5}$ in 1997 (Dockery, 2009).

There is a growing literature in economics that tries to quantify the causal impact of air pollution on health by using quasi-experimental methods to mimic random assignment of pollution exposure. The literature have shown significant impacts of air pollution on mortality (Chay and Greenstone, 2003; Currie and Neidell, 2005; Currie and Walker, 2011; Knittel et al., 2015) and contemporaneous health (Neidell, 2004; Moretti and Neidell, 2011; Schlenker and Walker, 2015). This literature has mainly focused on mortality risk, in particular for infants, in the U.S. and Europe.

Due to increased pressure from economic development and lax environmental regulations, developing countries and especially emerging economies such as China and India are experiencing the worst air pollution in the world. This is especially concerning given the size of population and the lack of access to adequate health care in these countries. While policy makers are increasingly aware of the negative impacts of air pollution on human health and the quality of life, there is a lack of reliable data and rigorous studies on the benefit of pollution reduction in these countries. As a result, the dose-response relationship (between pollution exposure and health outcomes) estimated using data from developed countries are often used as inputs for evaluating environmental regulations in developing countries, raising the question of external validity of this approach (Arceo et al., 2015).

This study fills this gap in the literature by examining the impact of $PM_{2.5}$ on health spending in China. To do so, we combine hourly air pollution readings from all monitoring stations from 2013 to 2015 with the universe of credit/debit card transactions in China during the same period. This is to our knowledge the first comprehensive analysis of how air pollution affects health expenditures from all medical conditions for the entire population of a developing country.[1] The causal impact of air pollution on out-of-pocket health spending also provides a lower bound estimate of consumer WTP for improved air quality. The reliance on health spending to directly infer WTP is in contrast with the revealed preference literature that relies on the implicit trade-off between risk factors and

[1] A growing literature uses health insurance claim data to examine the impact of air pollution on health spending in the U.S. (Deschenes et al., forthcoming; Williams and Phaneuf, 2016; Deryugina et al., 2017). However, health insurance tends to be inadequately provided in developing countries.

prices in product choices.

The key empirical challenge in identifying the causal effect of air pollution on health spending is the potential endogeneity in the key regressor $PM_{2.5}$ used to capture pollution exposure. The endogeneity can arise from multiple sources, including unobservables that affect both the pollution level and consumer spending (e.g., economic conditions) and avoidance behavior in response to air pollution (e.g., reduced outdoor activities). In addition, there could be measurement errors in proxying pollution exposure using air quality monitoring data. The pollution level could vary across locations within a city and the pollution exposure of the residents in a city ideally should be measured by the population weighted average of the pollution level in different parts of the city. However, monitoring stations are located sparsely and this prevents us from constructing population weighted averages. To the extent the measurement errors are classical, they would attenuate the estimates toward zero.

To deal with these challenges, we construct an instrumental variable by modeling the spatial spillovers of $PM_{2.5}$ due to the property of long-range transport of fine particles that is affected by wind direction and speed. In particular, the level of $PM_{2.5}$ in a given city is first predicted based on past $PM_{2.5}$ concentration of surrounding cities, taking into account contemporaneous wind direction and speed. Our approach is in the same spirit of the source-receptor matrix in the atmospheric science literature that is used to predict air quality from various pollution sources.[2] The predicted $PM_{2.5}$ of a given city from non-local sources is used as the instrument for the observed $PM_{2.5}$ in that city. The instrumental variable can be considered as a weighted sum of lagged $PM_{2.5}$ levels in other cities where the weights are a function of distance between the two cities, wind direction and speed in other cities. To address the concern of spatial correlation of economic activities, we create a buffer zone of 150 km and only use pollution sources outside the buffer zone in generating $PM_{2.5}$ predictions. Our results are robust to reasonable choices of the buffer zone.

Based on daily health spending by city, the OLS analysis with a rich set of temporal and location fixed effects suggests that a temporary increase of 10 $\mu g/m^3$ in $PM_{2.5}$ concentration that lasts for a week is associated with an increase of 0.17% in the total number of hospital and pharmacy transactions. A permanent elevation of 10 $\mu g/m^3$ in $PM_{2.5}$ concentration would raise the number of health transactions by 0.86%. The results from IV analysis indicate much stronger impacts: a temporary increase of 10 $\mu g/m^3$ in $PM_{2.5}$ would lead to a 0.32% increase in health transactions, while a permanent increase of the same magnitude would lead to a 2.20% increase in the number of transactions in healthcare. The impact of $PM_{2.5}$ differs across health facilities: spending in Children's hospitals is four times as responsive as that in other types of health facilities. For non-health spending, we find a negative impact of $PM_{2.5}$ in the short-term but no significant impact

[2] In our model, we do not specify specific pollution sources (e.g., power plants) but instead use the pollution levels in other cities as the influencing factors for the pollution level of a given city.

beyond two weeks. In addition, a projected worsening of air quality in the next day increases today's spending on both health and non-health categories. Take together, these results provide evidence of avoidance behavior whereby consumers reduce outdoor activities (such as shopping) to avoid pollution exposure.

The estimates of health impacts of $PM_{2.5}$ survive a variety of robustness checks including various parametric specifications of the long-term impact, different buffer zones in constructing the IV, and the inclusion of other pollutants such as CO and SO_2. In monetary terms, a permanent reduction of 10 $\mu g/m^3$ in daily $PM_{2.5}$ would lead to total annual savings of at least 70 billion *yuan* ($11 billion) in health spending in China. It would require dramatic and systematic measures to bring down China's $PM_{2.5}$ to the WHO's annual standard of 10 $\mu g/m^3$. Our analysis suggest that this level of improvement in air quality could lead to a huge saving in health spending of about $50 billion, nearly 9% of total health spending or 0.4% of GDP in China in 2015.

Our analysis on health spending can inform consumer WTP for improved air quality, a key policy parameter in the cost-benefit analysis of environmental regulations. Through a framework of consumer utility maximization, we show that consumer WTP for clean air can be bounded from below by the estimated impacts of air pollution on health spending. The results suggest a lower bound of $10 for the annual household WTP for a 10 $\mu g/m^3$ reduction in $PM_{2.5}$, which is similar to the WTP estimates for PM_{10} reduction among Chinese households from Ito and Zhang (2016).

This study makes four contributions to the literature. First, to our knowledge, this is the first comprehensive study that analyzes the effect of pollution on health spending by the entire population of a developing country. Our analysis is made possible by a novel data set that is composed of the universe of credit card and debit card transactions in China from January 2011 to present. There are 2.7 billion credit and debit cards that contribute to 1.5 trillion *yuan* of economic activities. Besides covering fifty percent of all health spending in China, this data set also includes spending in three hundred non-health sectors.

Second, a common practice in evaluating the health impact of air pollution in developing countries is to take the dose-response function estimated in developed countries to interpolate the mortality or morbidity impact from reduced air pollution in developing countries (e.g. Lelieveld et al. (2015) and World Bank (2007)). This benefit-transfer approach may lead to large inaccuracies given the differences in the levels of air pollution, baseline health conditions, and access to health care between these two groups of countries.

In contrast, our paper directly estimates the health impact of air pollution in a developing country, adding to the nascent literature using the same approach (Arceo et al., 2015; Chen et al., 2013; Greenstone and Hanna, 2014; He et al., 2016). Different from other studies in this literature, the high-frequency nature of our data allows us to identify the short- and medium-term impacts on health spending, instead of focusing on infant mortality as in most other papers. In addition, quan-

tifying the benefits of reduced pollution by translating the reduced mortality or morbidity into monetary terms requires adopting concepts such as the *value of a statistical life* (VSL).[3] We benefit from directly observing health spending, which gets around such interpolations.

Third, traditionally, Consumer WTP for improved air quality is estimated using the revealed preference approach that infers WTP based on the implicit trade-offs between risk levels and prices in consumer choice decisions such as housing and consumer products (Chay and Greenstone, 2005; Bayer et al., 2009; Ito and Zhang, 2016). Resorting to the utility maximization framework, this approach typically invokes behavioral assumptions such as perfect information on the health impact of air pollution to infer consumer WTP. If consumers systematically underestimate the health impacts for example due to the lack of awareness, downward bias would arise. Different from the revealed preference literature, this study uses realized health spending data and contributes to the growing literature on estimating health impacts and WTP for improved air quality with medical expenditures (Deschenes et al., forthcoming; Williams and Phaneuf, 2016; Deryugina et al., 2017). The approach of using health spending data to estimate WTP does not rely on the informational assumption: the estimates are derived from the fact that elevated pollution leads to illnesses which are then treated through healthcare spending. Whether or not consumers know that the underlying cause for their illnesses is pollution is immaterial for our estimates. The disadvantage of this approach is that the WTP estimates do not capture impacts through other channels such as on mortality and labor productivity, which the reveal preference approach should capture in theory under the assumption of perfect information.

Fourth, the rich spatial and temporal variations in our data allow us to examine both the short- and long-term impacts of air pollution on health spending. The aforementioned studies using health insurance data all focus on the contemporaneous impact by using daily or quarterly data. To deal with high serial correlation in daily air pollution measures which would render a model with many lags unstable, we apply a polynomial distributed lag model known as the Almon technique that uses finite order polynomials to capture the effects of long lags (Almon, 1965). We adapt this technique to the instrumental variable framework to address the potential endogeneity in contemporaneous and lagged air pollution measures. To our knowledge, this is the first analysis in this literature to use high frequency data and the Almon technique to estimate the long-term impacts.

The rest of the paper is organized as follows. Section 2 describes the data and the air pollution challenge in China. Section 3 provides a stylized model to illustrate that the estimated impact on health spending can be used to infer the lower bound for consumer WTP for clean air. Section 4 discusses our empirical framework and the identification strategy. Section 5 presents empirical results and Section 6 discusses our findings in relation to the literature. Section 7 concludes.

[3] Although there is a rich literature on estimating VSL in the US, there are very limited studies on VSL in developing countries (Viscusi and Aldy, 2003).

2 Data

Our analysis is based on three comprehensive micro-level datasets of air pollution, consumer spending by category, and meteorology conditions that have a national coverage. Collectively, they form a daily city-level panel for more than 300 major Chinese cities from 2013 to 2015. This enables us to evaluate the impact of air pollution on spending in both the short- and long-term, as well as heterogeneous impacts across spending categories.

2.1 Air Pollution

For nearly four decades, China has maintained its GDP growth at an annual rate of nearly 10%. The economy has transformed from an agricultural economy to a manufacturing-dominated economy. China became the world's largest exporter of goods in 2009 and the largest trading nation in 2013. This unprecedented economic growth is largely propelled by fossil fuels, with coal accounting for about two-thirds of aggregate energy consumption and oil nearly twenty percent. China is by far the largest energy consumer, accounting for roughly a quarter of world's total energy consumption and half of world's coal consumption.

Fast economic growth and rising energy consumption have put an enormous pressure on the environment, with air, water, and soil pollution becoming the most serious challenges in China today that adversely affect human health, ecosystems, and the quality of life.[4] Improving air quality has become an important policy goal for the central government, which revised extensively the Environmental Protection Law in 2014 and implicitly defined goals of pollution abatement in both the 12th (2011 - 2015) and 13th (2016 - 2020) five-year plans.

Fine-scale air quality data at the monitoring station level became publicly available in 2013. The Ministry of Environmental Protection (MEP) publishes hourly measures of $PM_{2.5}$, CO, SO_2, NO_2, and O_3. The number of monitoring stations and cities covered increases steadily from 1003 stations in 159 cities in 2013 to 1582 stations in 367 cities in 2015. Daily concentration of $PM_{2.5}$ and other pollutants at the city level is calculated by averaging data across monitoring stations within the city.

Figure 1 plots the three-year average level of $PM_{2.5}$ for each city from 2013 to 2015. The nationwide average during this period is around 56 $\mu g/m^3$ (with a standard error of 46 $\mu g/m^3$), which is much higher than the annual average standard of 12 $\mu g/m^3$ that is set by the U.S. Environmental Protection Agency or 35 $\mu g/m^3$ by the China MEP.[5] There is considerable regional disparity. Cities

[4] Lelieveld et al. (2015) estimate that air pollution led to 1.3 million premature deaths in China in 2010, accounting for 40% of the world's total premature deaths in the same year. World Bank (2007) puts the health cost of air pollution at 1.2-3.8% of GDP in 2003.

[5] The EPA's daily standard is 35 $\mu g/m^3$ and annual standard is 12 $\mu g/m^3$. China's MEP sets limits on $PM_{2.5}$ for the first time in 2012 to take effect in 2016: the daily standard is 75 $\mu g/m^3$ and annual standard is 35 $\mu g/m^3$.

in the northern and central China with a high concentration of manufacturing industries suffer from the most severe pollution, with many of them experiencing a $PM_{2.5}$ concentration of 90 $\mu g/m^3$ or higher. The less-developed regions in the west and wealthy regions in the south have better air quality. The latter, especially regions along the coast, has seen improvement in air quality as a result of shutting down or relocating polluting industries and reorienting the industry structure toward high tech and service industries.

One advantage of our empirical analysis is the rich variation in pollution measures both across cities and over time. To illustrate the time-series variation, we present in figure 2 the daily $PM_{2.5}$ concentration for the nation (the top panel) and each of the four broad regions (the bottom panel).[6] In all regions of the country, the daily $PM_{2.5}$ concentration is higher than 35 $\mu g/m^3$, the official MEP standard, for most days. The northern regions have much more pronounced peaks in the winter than the southern region, largely because of the coal-fired central heating systems north of Huai River (Chen et al., 2013) The pollution level is trending downwards in all regions, driven by tightened government regulations, private and public investment in waste treatment, and changes in China's overall industry structure.

2.2 Consumer Spending

We have access to the universe of credit and debit cards (or 'bank card') transactions in China settled through the UnionPay network. The Unionpay network is the only inter-bank payment network in China and is state owned. The network is the largest in the world in terms of both the number and value of transactions, ahead of Visa and Mastercard. There are in total 2.7 billion cards from 2013 to 2015 with transactions covering over 300 merchant categories.[7] The database includes eight trillion *yuan* of annual economic activities. We observe the location, time, merchant name, and amount for each transaction and we aggregate the data to daily spending by category by city from 2013 to 2015 for our analysis. To our knowledge, this is the most comprehensive and fine-scale data in temporal and spatial dimensions on consumer spending in China and we are the first to utilize them for academic research.

Health care in China is financed by out-of-pocket spending, health insurance, and government programs similar to the U.S. medicare. Medical expenses that are covered by the Chinese 'medicare' programs are often directly billed on medicare cards, most of which are settled through

[6]The Northeastern region includes Heilongjiang, Jilin, Liaoning, and the northeastern part of Inner Mongolia. The Northern region includes Beijing, Tianjin, Hebei, Shanxi, Shandong, Henan, and the rest of Inner Mongolia. The Northwestern region includes Shanxi, Gansu, Qinghai, Ningxia, and Xinjiang. The Southern region includes Guangdong, Guangxi, Hainan, Guizhou, Yunnan, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Hubei, Hunan, Chongqing, and Sichuan. Tibet is excluded in the regional plots due to the sparse coverage.

[7]There are six major categories and 300 subcategories. The major categories are: food, lodging, and entertainment; department stores; utility; transportation; real estate; and wholesale. Merchants are classified by these categories.

the UnionPay network and enter the database as regular transactions. Commercial health insurance companies usually require patients to pay for medical expenses first and get reimbursed later through filing claims. If consumers pay for these expenses via their bank cards, then these transactions will be included in our database.[8]

Our data account for 31% of aggregate private healthcare spending in 2013, and as card penetration grew, the coverage rises to 51% in 2015, similar to the share of bank card transactions in other sectors. The high penetration of bank cards in retail spending in China is remarkable given its short history (the first credit card was issued in 1998 and it was not until late 2000s when consumers begin to adopt bank cards. The official statistics from Central Bank of China (2015) show that that bank card transactions accounted for 48% of overall spending on retail sales of consumer goods in the third quarter of 2015, increasing from only 17% in 2006. In the U.S., spending from credit and debit cards accounts for 55% of all consumer spending (Bagnall et al., 2014).

Health spending includes transactions at hospitals, pharmacies, and other healthcare facilities (e.g. small health clinics). In 2015, hospitals account for 83.5% of health spending in our data, and 56.8% of transactions. Different from pharmacies in the U.S. such as CVS or Rite-Aid, most pharmacies in China only carry medicine and rarely sell daily necessities. Pharmacies account for 6.0% of total healthcare spending, and 31.0% of transactions in 2015. We separate hospitals and pharmacies from other healthcare facilities.

Within hospitals, we distinguish Renmin (which means *people*) hospitals and Children's hospitals from other hospitals. Renmin hospitals are state-owned general hospitals and tend to be the largest health care facilities in a city. Each city has at least one Renmin hospital. Children's hospitals accept mostly children patients. Birth centers and infant health centers are grouped into Children's hospitals. Renmin and Children's hospitals account for 24.1% and 4.2% of total health spending respectively, and 26.2% and 9.0% of total number of transactions in 2015.[9]

In addition to health spending, we also analyze spending in non-health categories, such as daily necessities. United Nations' *Classification of Individual Consumption According to Purpose* (COICOP) defines necessity goods as 1) food and non-alcoholic beverages, 2) alcoholic beverages, tobacco and narcotics, 3) clothing and footwear, 4) recreation and culture, and 5) restaurants and hotels. We follow this definition closely. Relative to the health spending, total non-health spending is seven times as large and transact six times as frequent. Spending on daily necessities is three times as large and transact three times as frequent. A unique feature of Chinese consumers' shop-

[8]The healthcare system and the insurance market in China have been improving with significant government support. In 2009, China's central government revealed plans to overhaul its healthcare system by providing 850 billion Yuan to develop the healthcare system between 2009 and 2011 and to increase the basic health insurance coverage from 65 percent to 90 percent by 2011. By 2011, the insurance coverage through three major government supported insurance programs reached nearly 95% and the reimbursement rate was about 44 to 68 percent depending on the insurance program (Yu, 2015).

[9]We use hospital names and keyword matching to identify Renmin hospitals and Children's hospitals.

ping behavior is their frequent trips (often on a daily basis) for groceries at the supermarkets. We therefore use supermarket spending as another proxy for daily consumption, in addition to spending on necessities.[10] Spending in supermarkets is over four times as large as health spending in value and five times as large in terms of the number of transactions in 2015.

To illustrate the inter-temporal patterns, Figure 3 plots weekly healthcare spending and the number of transactions at the national level from 2013 to 2015. There is a significant drop in both the spending amount and the transaction frequency during holidays. In addition, both variables have more than tripled during our sample period due to the diffusion of bank cards. We control for these two salient features in our regression analysis through holiday fixed effects and city-specific time trends.

2.3 Meteorology Data and Summary Statistics

Besides pollution, weather conditions could also directly affect health outcomes (Deschenes et al., 2009). We obtain meteorology data from the Integrated Surface Database (ISD) that is hosted by National Oceanic and Atmospheric Administration (NOAA). The ISD dataset includes hourly measures of temperature, precipitation, wind speed and wind direction for 407 monitoring stations in China.[11] We match cities with the nearest weather station according to their geographic coordinates and compute daily temperature and wind speed from a simple average of the hourly data.

ISD's hourly measure of precipitation suffers from noticeable measurement errors, so we use daily precipitation from NOAA's *Global Surface Summary of the Day* database (GSOD) instead.[12] Wind direction of the day is calculated by adding up twenty-four hourly vectors of wind directions, where the length of each vector is the hourly wind speed.

Table 1 reports the summary statistics for all variables used in our study at the city-day level. The daily $PM_{2.5}$ concentration is on average 56 μg/m^3 between 2013 and 2015, where the interquartile range is from 27 to 69 and the maximum is 985. Sixty-seven percent of these city-day observations record a concentration level that is above the U.S. daily standard of 35 μg/m^3. In terms of health spending, the average daily number of transactions is 7,229 and the average daily spending is 6.7 million *yuan*.

[10] Since supermarkets sell a large variety of goods other than necessities, we exclude supermarkets in necessity spending.

[11] These stations cover most major Chinese cities from dates as early as 1940s to now.

[12] GSOD reports daily precipitation using Greenwich Mean Time, which is the cumulative rainfall from 8 a.m. Beijing time to 8 a.m. the next day. We create a dummy variable for rainfall based on this daily measure.

3 Theoretical Model

Air pollution affects human health mainly through its impact on respiratory and cardiovascular systems. Several decades of study in epidemiology and more recently economics has associated exposure to air pollution with increases in mortality and morbidity risks (Brunekreef and Holgate, 2002; Pope and Dockery, 2012). Fine particles ($PM_{2.5}$) are especially detrimental to health as they can penetrate deep into lungs and carry toxins to other organs. High levels of $PM_{2.5}$ irritate respiratory and cardiovascular systems and can lead to aggravated asthma, lung disease, heart attacks, and stroke.

In this section we provide a theoretical model to illustrate the relationship between the estimated impact of $PM_{2.5}$ on health spending and consumer WTP for improved air quality. The seminal paper by Grossman (1972) first proposed the utility maximization framework of health production where consumers choose optimal health care spending to alleviate the negative impact of air pollution exposure. Following this tradition, Deschenes et al. (forthcoming) and Williams and Phaneuf (2016) show that the marginal effect of air pollution exposure on total health spending provides a lower bound of consumers' WTP for improved air quality. The marginal effect provides a conservative estimate of WTP because consumers can engage in defensive spending such as purchases of air purifiers or face masks or avoidance behavior such as staying indoors. Air purifier spending or lost utility from staying indoors should constitute part of consumers' WTP for clean air but is not included in the marginal response of health spending. While the literature has largely neglected the role of avoidance and loss in the quality of life, here we present a static model to account for both.

There is a continuum of consumers of measure 1. Each consumer i chooses health spending (m_i), non-health offline spending (c_i) and online spending (s_i). The consumer is exposed to air pollution whenever he goes outdoors, and we assume that pollution exposure $e(a, m_i + c_i)$ is an increasing and convex function of the exogenous air pollution level denoted by a and spending activities $m_i + c_i$.[13]

Consumer i has an endowed health stock h_0, which evolves according to an exogenous level of air pollution and his own health spending that mitigates the negative consequences of pollution. Individuals differ in how sick they become when they are exposed. This is captured by a function $g_i(e_i)$, which is individual-specific and represents how much the individual's health stock decreases upon being exposed to air pollution. $g_i(e_i)$ can be thought of as the severity of the illness and is

[13] We combine all non-health spending (except online spending) in c and assume each \$1 of spending results in the same amount of pollution exposure independent of purpose. Convexity implies that on more polluted days, the marginal impact of spending activities on pollution exposure is larger.

non-decreasing in exposure to air pollution e_i.[14] Thus the health stock equation can be written as:

$$h_i = h_0 + m_i - g_i(e_i) \tag{1}$$

Consumers have health insurance, with π denoting the premium and p the proportion of health spending that needs to be paid out-of-pocket.[15] Thus, if the consumer undergoes hospital treatment that costs a total of m_i, the consumer's out-of-pocket spending is equal to pm_i, where $p < 1$. Income $y(h_i)$ is composed of non-wage income y_0, which is exogenous and does not depend on health, and wage income $w(h_i)$, which is affected by health. Wage income is lower with diminished health, for example due to productivity loss or sick days. The budget constraint is:

$$y(h_i) \equiv y_0 + w(h_i) = \pi + pm_i + c_i + s_i \tag{2}$$

Consumer i chooses health spending, non-health related offline spending, and online spending to maximize utility, subject to the budget constraint and the rule of health stock evolution:

$$\begin{aligned} &\max_{\{m_i, c_i, s_i\}} U[h_i, c_i, s_i, e(a, m_i + c_i)], \\ &\quad \text{s.t. } y(h_i) \equiv y_0 + w(h_i) = \pi + pm_i + c_i + s_i, \\ &\quad \text{and } h_i = h_0 + m_i - g_i(e(a, m_i + c_i)), \end{aligned} \tag{3}$$

Consumer utility depends on health stock (h_i), offline consumption (c_i), online spending (s_i), and pollution exposure (e_i). Note that we allow consumer utility to be both directly affected by pollution exposure, and indirectly affected through health stock. The direct mechanism arises because consumer values the quality of life, which is directly affected by air pollution. Heavy haze and smoky air reduce consumer's utility even if his health stock is fully recovered through health spending. For example, Levinson (2012) finds that people report lower levels of happiness on days with worse local air pollution.

Our specification of pollution exposure $e(a, m_i + c_i)$ makes it explicit that all spending (with the only exception of online spending), whether health-related or not, affects pollution exposure because it involves time spent outdoors.[16] This is the key difference between our model and those in Deschenes et al. (forthcoming) and Williams and Phaneuf (2016). In addition, our model incor-

[14] An example is when $g_i(e_i) = \alpha_i e_i$, where $\alpha_i \sim U[0,1]$. Individuals with $\alpha_i = 0$ remain healthy even after being exposed to air pollution; individuals with $\alpha_i = 1$ get very sick upon being exposed to air pollution, and experience a significant decline in their health stock.

[15] We assume that every consumer has health insurance. In 2011, nearly 95% of China's population was covered by one of the three major public health insurance programs (Yu, 2015).

[16] In the short-term, consumer could reduce pollution exposure by delaying hospital visits or reducing time spent outdoors. In the long term, both m_i and c_i will respond to changes in pollution.

porates utility from health (for example, through mortality) and allows income to depend on health, both of which are absent in Williams and Phaneuf (2016)'s model.

The Lagrangian can be written as:

$$L_i = U[h_i, c_i, s_i, e(a, m_i + c_i)] + \lambda_i[y(h_i) - \pi - pm_i - c_i - s_i],$$

The first-order conditions are:

$$\begin{aligned}
\frac{\partial L_i^*}{\partial m_i} &= U_h(1 - g_i'(e_i)e_m) + U_e e_m + \lambda_i(y_h(1 - g_i'(e_i)e_m) - p) = 0, \\
\frac{\partial L_i^*}{\partial c_i} &= -U_h g_i'(e_i)e_c + U_c + U_e e_c - \lambda_i(y_h g_i'(e_i)e_c + 1) = 0, \\
\frac{\partial L_i^*}{\partial s_i} &= U_s - \lambda_i = 0, \\
\frac{\partial L_i^*}{\partial \lambda_i} &= y(h_i^*) - \pi - pm_i^* - c_i^* - s_i^* = 0.
\end{aligned} \tag{4}$$

where U_h, U_e, U_c, U_s are partial derivatives of the utility function with respect to health stock, pollution, consumption, and online spending, respectively. We assume $U_h > 0, U_c > 0, U_s > 0, U_e < 0$, since health and consumption (including future consumption) are desirable but pollution is not. $e_m = e_c$ is the marginal impact of spending activities on pollution exposure e. The net impact of medical spending on health, $\frac{dh_i}{dm_i} = 1 - e_m$, is assumed to be positive, since the health benefit of medical treatment should be much larger than the incremental risk due to additional pollution exposure.[17] Exposure increases with pollution ($e_a > 0$). In addition, we allow $e_m > 0$ and $e_c > 0$ (instead of making them 0), as consumers are exposed to air pollution whether buying food or seeing a doctor. Finally, y_h is the effect of health on income, and is assumed to be positive.

Intuitively we would expect that when air quality worsens, medical spending should increase $\frac{\partial m_i^*}{\partial a} > 0$ and non-health spending should decrease $\frac{\partial c_i^*}{\partial a} < 0$. In Appendix A, we discuss conditions under which these are true. $\frac{\partial c_i^*}{\partial a} < 0$ holds under fairly generic conditions on the utility function. $\frac{\partial m_i^*}{\partial a} > 0$ holds true as long as the health benefit from medical spending is much larger than any disutility from additional pollution exposure, which is likely to be true in practice.

Denote $V_i(a, h_0, y_0)$ as the indirect utility function and $L_i^*(a, h_0, y_0)$ as the optimal value of the Lagrangian. The marginal WTP for reduction in air pollution can be obtained as:

$$MWTP_i = -\frac{\frac{\partial V_i}{\partial a}}{\frac{\partial V_i}{\partial y_0}} = -\frac{\frac{\partial L_i^*}{\partial a}}{\frac{\partial L_i^*}{\partial y_0}} \tag{5}$$

As we show in Appendix A, assuming that insurance premiums do not adjust in response to

[17] The optimal health spending is 0 if $1 - e_m < 0$.

pollution, the marginal WTP for individual i can be expressed as:

$$MWTP_i = p\frac{\partial m_i^*}{\partial a} + y_h(-\frac{dh_i^*}{da}) + \frac{U_h}{\lambda_i}(-\frac{dh_i^*}{da}) + (-\frac{U_e}{\lambda_i})\frac{de_i^*}{da} + \frac{U_c - U_s}{\lambda_i}(-\frac{\partial c_i^*}{\partial a}) \tag{6}$$

Equation (6) illustrates the relationship between the impact of air pollution on health spending, given by $\frac{\partial m_i^*}{\partial a}$, and WTP for improved air quality. Changes in the individual's out-of-pocket health spending provide a lower bound of the individual's WTP. The difference between the two quantities is determined by the last four terms in the equation. The first term, $y_h(-\frac{dh_i^*}{da})$ measures reduction in income due to lower productivity as a result of pollution ($\frac{dh_i^*}{da} < 0$). The second term $\frac{U_h}{\lambda_i}(-\frac{dh_i^*}{da})$ denotes the disutility from reduced health stock. The third term $(-\frac{U_e}{\lambda_i})\frac{de_i^*}{da}$ captures the (monetized) utility loss in the quality of life due to increased pollution exposure. Note that $\frac{de_i^*}{da}$ is the total derivative of exposure when pollution deteriorates. $\frac{de_i^*}{da} = e_a + e_m\frac{\partial m_i^*}{\partial a} + e_c\frac{\partial c_i^*}{\partial a}$, where $\frac{\partial m_i^*}{\partial a} > 0$ and $\frac{\partial c_i^*}{\partial a} < 0$. We posit that non-health spending is relatively inelastic to pollution, $|\frac{\partial m_i^*}{\partial a}| > |\frac{\partial c_i^*}{\partial a}|$, and hence $\frac{de_i^*}{da} > 0$. The last term $\frac{U_c - U_s}{\lambda_i}(-\frac{\partial c_i^*}{\partial a})$ denotes reduction in monetized utility due to the sub-optimal level of consumption distorted by pollution exposure. We assume $U_c - U_s > 0$, since consumers would always choose to purchase online and select $c = 0$ otherwise.

Our model encompasses the case of Deschenes et al. (forthcoming) which abstracts away from pollution exposure associated with consumption ($e_c = 0$), as well as the utility loss of reduced quality of life ($U_e = 0$).[18] When $e_o = U_e = 0$, the FOCs indicates $U_c = U_s = \lambda$, and

$$MWTP_i = p\frac{\partial m_i^*}{\partial a} - y_h\frac{dh_i^*}{da} - \frac{U_h}{\lambda}\frac{dh_i^*}{da}. \tag{7}$$

In addition, if h_i is preset (i.e. kept at a subsistence level with $\frac{\partial h_i}{\partial a} = 0$) and income y is exogenous, as suggested by Williams and Phaneuf (2016), then the expression for marginal willingness-to-pay is similar to theirs:[19]

$$MWTP_i = p\frac{\partial m_i^*}{\partial a}. \tag{8}$$

We now describe how to obtain a lower bound measure for society's willingness to pay for pollution reduction. Society's willingness-to-pay has two components. Firstly, every individual i

[18] In Deschenes et al. (forthcoming), $MWTP = w\frac{ds}{dc} + p_a\frac{\partial a}{dc} - \frac{U_s}{\lambda}\frac{ds}{dc}$, where w is wage rate (equivalent to y_h in our framework), s denotes number of sick days (equivalent to a negative change in health stock), a is defensive behavior, p_a is the price of taking defensive measures, and c is the concentration of pollutants (same as level of air pollution a in our framework.)

[19] In Williams and Phaneuf (2016), $MWTP = p\frac{\partial m^*}{\partial a} + \frac{\partial \pi}{\partial a}$. They consider the case of a competitive insurance provider, and argue that in equilibrium insurance premiums will adjust in response to expected pollution. By contrast, in our setting, given that insurance reimbursement rates for China's public insurance programs are rarely adjusted year-to-year and are the same across cities despite large variance in pollution across cities, we find it more reasonable to assume that $\frac{\partial \pi}{\partial a} = 0$.

is willing to pay at least $MWTP_i$ for a marginal reduction in pollution. Secondly, the net cost of providing insurance increases with the level of air pollution: for a marginal increase in pollution, the change in the net cost equals the increase in insurance pay-outs to individuals, or $(1-p)\int \frac{\partial m_i^*}{\partial a} dF_i$. The health insurance programs are subsidized by the government so the second component reflects the cost to the government in supporting the health insurance market. A lower bound measure for society's WTP for pollution reductions is then given by the sum of individuals' WTP plus the reduction in society's spending on insurance programs:

$$\begin{aligned} MWTP_{society} >= & \int p \frac{\partial m_i^*}{\partial a} dF_i + (1-p) \int \frac{\partial m_i^*}{\partial a} dF_i \\ = & \int \frac{\partial m_i^*}{\partial a} dF_i \end{aligned} \tag{9}$$

Thus the total change in medical spending in response to air pollution is a lower bound for society's WTP for reducing air pollution.

To summarize, the utility maximization framework illustrates that the impact of air pollution on health spending, the focus of our empirical analysis, provides a lower bound estimate for society's WTP for clean air, while the impact on out-of-pocket health spending provides a lower bound estimate for each individual's WTP for clean air. The difference can arise from four additional factors: reduced income from the loss of productivity, the disutility of reduced health stock (e.g., mortality risk), the loss in the quality of life from increased exposure, and the loss in utility due to consumption distortion (avoidance behavior). In the empirical analysis, we estimate the impact of air pollution on non-health spending ($\frac{\partial c_i^*}{\partial a}$) to assess the importance of avoidance behavior.

4 Empirical Framework

In this section, we first present a parsimonious model that allows us to estimate the short- and long-term impacts of air pollution on health spending. We then discuss our estimation strategy and the construction of the instrumental variable.

4.1 Distributed Lag Model

Air pollution has both short-term and long-term consequences on health spending. We posit that this relation can be approximated by the following distributed lag model (DL) where today's health spending is affected by both contemporaneous and past air pollution exposure:

$$y_{ct} = \sum_{i=0}^{k} \beta_i p_{c,t-i} + \mathbf{x}_{ct}\alpha + \kappa_c t + \xi_c + \eta_w + \varepsilon_{ct}, \tag{10}$$

where y_{ct} is daily health spending in a city, and $p_{c,t-i}$ is either contemporaneous ($i = 0$) or lagged pollution exposure ($i \geq 1$). $\mathbf{x}_{ct}$ includes a rich set of controls such as weather conditions, holiday fixed effects, day-of-week fixed effects, seasonality, etc. $\kappa_c t$ is city-specific linear time trend, ξ_c is city fixed effect, and η_w is week fixed effect. The key parameters of interest are β's, which capture the short- and long-term causal impacts of pollution exposure on health spending.

Let us assume for a moment that there is no measurement error in pollution exposure $p_{c,t-i}$ and that there is no avoidance behavior or omitted variable (three important issues that we will return to in the next section), then the DL model can be estimated using OLS. But a straightforward linear estimation is undesirable because with a large number of lags (as many as a few hundreds in our setting), collinearity among $p_{c,t-i}$ leads to imprecise OLS estimates. To reduce the number of parameters that need to be estimated while allowing for flexible long-term impact, the seminal paper by Almon (1965) proposes to specify β_i's as polynomial functions of time with order q, where q is a constant chosen by econometricians. The intuition is that, by Weierstrass's Approximation Theorem, any smooth function (here β_i can be treated as a function of time) defined on a closed interval $[a,b]$ can be uniformly approximated, arbitrarily closely, by a polynomial function of finite degree. Take $q = 3$ as an example:

$$\beta_i = F(i) = \gamma_0 + \gamma_1 i + \gamma_2 i^2 + \gamma_3 i^3. \tag{11}$$

where the contemporaneous effect of pollution on spending is captured by γ_0, the effect of yesterday's pollution is $\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3$, while the effect of pollution i days' in the past is $\gamma_0 + \gamma_1 i + \gamma_2 i^2 + \gamma_3 i^3$.

Plug (11) back into (10) and rearrange terms, we have:

$$\begin{aligned}
y_{ct} &= \sum_{i=0}^{k} \beta_i p_{c,t-i} + \mathbf{x}_{ct}\alpha + \kappa_c t + \xi_c + \eta_w + \varepsilon_{ct} \\
&= \gamma_0 p_{ct} + (\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3) p_{c,t-1} + \ldots + (\gamma_0 + \gamma_1 i + \gamma_2 i^2 + \gamma_3 i^3) p_{c,t-i} + \ldots \\
&\quad + (\gamma_0 + \gamma_1 k + \gamma_2 k^2 + \gamma_3 k^3) p_{c,t-k} + \mathbf{x}_{ct}\alpha + \kappa_c t + \xi_c + \eta_w + \varepsilon_{ct} \\
&= \gamma_0 (p_{ct} + p_{c,t-1} + p_{c,t-2} + \ldots + p_{c,t-k}) \\
&\quad + \gamma_1 (1 \times p_{c,t-1} + 2 p_{c,t-2} + \ldots + k p_{c,t-k}) \\
&\quad + \gamma_2 (1^2 \times p_{c,t-1} + 2^2 p_{c,t-2} + \ldots + k^2 p_{c,t-k}) \\
&\quad + \gamma_3 (1^3 \times p_{c,t-1} + 2^3 p_{c,t-2} + \ldots + k^3 p_{c,t-k}) + \mathbf{x}_{ct}\alpha + \kappa_c t + \xi_c + \eta_w + \varepsilon_{ct}.
\end{aligned} \tag{12}$$

With this reformulation, we only need to estimate four coefficients γ's rather than $k+1$ (the

number of lags plus current day) coefficients. The four key regressors are:

$$\begin{aligned} v_{1t} &= p_{ct} + p_{c,t-1} + p_{c,t-2} + ... + p_{c,t-k}, \\ v_{2t} &= \quad p_{c,t-1} + 2p_{c,t-2} + ... + kp_{c,t-k}, \\ v_{3t} &= \quad p_{c,t-1} + 4p_{c,t-2} + ... + k^2 p_{c,t-k}, \\ v_{4t} &= \quad p_{c,t-1} + 8p_{c,t-2} + ... + k^3 p_{c,t-k}. \end{aligned} \tag{13}$$

where the first term is the sum of past pollution exposure, and the others are weighted sum of past exposure with the weights being polynomial terms of time.

This approach has several advantages over competing distributed lag models, the most popular one being the geometric decay model. One advantage of this approach is that these new regressors as defined in equation (13) exhibit much less multicollinearity than lags of $p_{c,t-i}$ themselves. Second, this model allows for much more flexible decaying patterns than those in geometric decay models. Third, it is straightforward to impose additional restrictions that either is generated by economic theories or reflect a prior knowledge of the data generating process. For example, if tomorrow's pollution exposure (forward one period) should not affect current health spending, then $\gamma_0 - \gamma_1 + \gamma_2 - \gamma_3 = 0$. If pollution exposure beyond k lags should not affect current health spending, then $\gamma_0 + (k+\tau)\gamma_1 + (k+\tau)^2\gamma_2 + (k+\tau)^3\gamma_3 = 0, \forall \tau \in \mathbb{N}$ and $\tau > 0$. These assumptions can be imposed individually or jointly as constraints in the estimation or they can be tested as linear restrictions. Fourth, this specification does not require instruments for the lagged dependent variable as in the geometric decay model, which is often challenging. Finally, we allow for arbitrary correlation between the contemporaneous error term and all of the past error terms, which is difficult in geometric decay models.

Once we choose the number of lags k, the order of polynomials q, and additional conditions on γ's, the estimation can be carried out in (constrained) OLS and β's can then be calculated based on parameter estimates from OLS.[20]

4.2 Identification

A key empirical challenge in estimating the impact of air pollution on health spending is potential endogeneity in the pollution measure for several reasons. First, our measure of pollution exposure likely suffers from measurement errors as is common in the literature on estimating health impact of air pollution. We aggregate pollution data from monitoring stations to the city level. The pollution level could vary across locations in a city. For example, among the 9 monitoring stations in the

[20]To improve efficiency, Almon (1965) initially proposes a weighting matrix from Weierstrass's Approximation and later Giles developed a simpler but numerically equivalent method, which we adopt in our estimation. See http://davegiles.blogspot.com/2017/01/explaining-almon-distributed-lag-model.html.

urban core of Beijing (within the 5th ring road), the average difference between the maximum and minimum pollution level in a day is about 35 μg/m^3 in 2014 while the daily average at the city level is 87. As population are not evenly distributed within a city and the spatial location of monitoring stations does not align with residential areas, the arithmetic mean across all stations within a city may not accurately reflect the city population's exposure to pollution. An ideal measure should be population-weighted average of station-level air quality, but this is impractical due to the lack of air pollution data at the finer spatial level (e.g., city block or zip code). In addition, our daily pollution is the simple average over hourly recordings and abstracts away from the temporal variation of air pollution. To the extent that the measurement errors are classical, attenuation bias would arise in OLS.[21]

Second, pollution exposure is potentially endogenous due to the avoidance behavior in both the short- and long-term. Chinese consumers have increased awareness of air quality and its impact on health. $PM_{2.5}$ readings are easily accessible through cell phone apps or from government websites in recent years.[22] In the short term, during days of severe air pollution, consumers may reduce their outdoor activities, shift the timing of consumption (e.g. postpone visits to hospitals for non-acute conditions), or undertake defensive measures such as wearing face masks and using air purifiers indoors (Mu and Zhang, 2016; Ito and Zhang, 2016; Sun et al., 2017). These types of behavior, in response to contemporaneous air quality variations, could reduce health spending and render the pollution measure endogenous. Long-term air pollution trends could affect migration across cities as documented in the U.S. (Banzhaf and Walsh, 2008). Consumers who are more vulnerable to air pollution or have a high valuation of clean air would choose to move away from more polluting cities. As a result, air pollution could be correlated with the error term (such as the stock of health status of the residents).

In a short or medium time frame such as the one used in our analysis, location fixed effects could control for migration and other long-run avoidance behavior. However, the short-run avoidance behavior as responses to contemporaneous air pollution is more challenging and cannot be absorbed by location fixed effects. In addition, it is not obvious that endogeneity arising from avoidance behavior could be addressed by the instrumental variable strategy since avoidance is the direct result of air pollution (and hence could be correlated with shocks that are correlated with air pollution). We use spending on daily necessities and at supermarkets to quantify avoidance behavior. Our results show that avoidance behavior has an impact on spending in the short term (i.e., within a few days to up to two weeks) through intertemporal substitution but there is no significant aggregate

[21] Satellite data on Aerosol Optical Depth (AOD) offer an alternative measure of the ground level pollution and they have finer spatial resolution (e.g., 3 by 3km from Terra satellite and 10 by 10km from Aqua). But at the daily level, there are a lot of missing values at the grid level in addition to the noise introduced from inferring $PM_{2.5}$ based on the AOD data.

[22] Hourly air pollution data in major Chinese cities are published on the website of the Ministry of Environment Protection and other non-governmental websites since 2013.

impact over a longer period (such as a month or longer).

The third source of endogeneity in pollution measure is unobservables. Despite our rich set of controls for weather and local conditions (e.g., city specific time trend and seasonality), there is various temporal variation that can not be adequately controlled. For example, permanent local shocks to health spending, such as income shocks, could be correlated with economic activities and hence with air quality. Temporary local shocks, such as major sport and political events, could affect both the air pollution level and health spending (and consumer activities in general). These unobservables that are not absorbed by our location and trend/seasonality interactions render the air quality variable endogenous.

To address these issues, we use the instrumental variable strategy. Our IVs exploit the spatial spillovers of $PM_{2.5}$ due to its long-range transportability. $PM_{2.5}$ particles are light, can travel at the speed of 10 mph, and often reside in the atmosphere for 3-4 days. Their region of influence is determined by the wind speed and direction. Based on the atmospheric modeling, Zhang et al. (2015) document significant regional pollution transport in China. For example, nearly half of the pollution in Beijing originates from sources outside of the municipality. These results suggest that $PM_{2.5}$ from surrounding cities can potentially serve as instruments for pollution for a given city.

The approach of constructing instruments exploiting $PM_{2.5}$'s region of influence is in spirit similar to the source-receptor matrix constructed by the US EPA for air pollution prediction. We take each city as both a pollution source and a receptor, and develop a model to predict the air pollution level of a given city using wind speed, wind direction, and $PM_{2.5}$ in other cities.[23] Once we estimate such a model, we use $PM_{2.5}$ from cities farther away than 150 km to predict the *exogenous* part of pollution in city i that is induced by pollutants elsewhere. The purpose of such a buffer zone is to minimize the correlation in unobserved regional economic shocks. Our assumption is that economic shocks in regions outside the buffer zone have little correlation with local economic shocks. We test the robustness of our results to this choice of radius in section 5.

Denoting the pollution level of city i in time t as p_{it}. We model p_{it} as a function of past pollution and pollution from other cities:

$$p_{it} = \theta_1 p_{i,t-1} + \sum_{j \neq i} p^{+}_{j \to i,\, t} + \mu_{it}, \tag{14}$$

where θ_1 governs the amount of pollution that is carried over from the previous day, $p^{+}_{j \to i,\, t}$ denotes the contribution from city j to city i at time t, and μ_{it} is the error term.

To compute $p^{+}_{j \to i,\, t}$, we invoke a simple vector decomposition. Let Φ denote the angle between the wind direction and the direction from city j to city i. We assume that the amount of pollutant

[23] Williams and Phaneuf (2016) construct their IV for air pollution using pollutants 60 km away (or 120 km away) without exploiting wind speed or direction.

carried toward city i is $\cos(\Phi)p_{jt}$ and the speed at which the pollutant travels is $\cos(\Phi)S_{jt}$, where S_{jt} is the wind speed. In addition, $\cos(\Phi)$ is set to zero when it is negative: when wind blows away from city i, pollution from the source city j should not contribute to that of the receptor city.

Assuming a constant travel speed, the minimum days s that is required for pollutants to travel from city j to city i can be calculated as:

$$s = \left\lfloor \frac{d_{ij}}{\cos(\Phi)S_{jt}} \right\rfloor. \tag{15}$$

where d_{ij} is the distance from city j to city i, and the floor function $\lfloor \cdot \rfloor$ returns the minimum integer.

$PM_{2.5}$ concentration level decays as it travels. Let $\delta(d_{ij}, s; \theta_2)$ denote the decay function that depends on distance traveled d_{ij} and days in the air s, the amount of pollutant that reaches city i on day $t+s$ is:

$$p^{+}_{j\to i,\, t+s} = \cos(\Phi)p_{jt}\delta(d_{ij}, s; \theta_2). \tag{16}$$

Putting things together, $PM_{2.5}$ in city i at time t could be specified as:

$$\begin{aligned} p_{it} &= \theta_1 p_{i,t-1} + \sum_{j\neq i} p^{+}_{j\to i,\, t} + \mu_{it} \\ &= \theta_1 p_{i,t-1} + \sum_{j\neq i}\sum_{s} \cos(\Phi)p_{jt}\delta(d_{ij}, s; \theta_2) + \mu_{it} \\ &= \mathbf{P}(\theta) \end{aligned} \tag{17}$$

Figure 4a shows the wind-pollution vectors from over 300 cities on Dec. 5, 2013 (denoted as Day 0). The arrows' lengths and directions indicate wind speeds and directions, with the lengths rescaled to match the exact distances the arrows can cover in a day. The widths of arrows indicates the level of $PM_{2.5}$ concentration from the origination city. To illustrate how we predict city-day $PM_{2.5}$ using Beijing as an example, Figure 4b shows the subvectors of pollutants that are blown towards Beijing after the vector decomposition. The pollution level of the receptor city, or Beijing in this example, can be predicted by pollutant carried through the subvectors that reach Beijing at time t, together with local pollution level of previous days, as stated in Equation (17).

We first estimate equation (17) using Nonlinear Least Squares by minimizing the sum of squared prediction errors between the actual air quality p_{it} and the predicted air quality $\mathbf{P}$. With the estimated parameters, we then construct $\hat{p}^{far}_{it}$ only using pollutants from cities that are farther away than 150 km (see Figure 5),

$$\hat{p}^{far}_{it} = \sum_{d_{ij}>150} p^{+}_{j\to i,\, t} \tag{18}$$

This predicted *non-locally-generated* pollution serves as our instrument. According to our esti-

mates, 30% of pollutants remain after traveling 100 km.

Although the model is parsimonious, it yields very reasonable predictions. In the top panel of Figure 6, we plot both the predicted and the observed daily level of pollution for Beijing. Other than under-predicting some peaks, the predicted air quality follows data very closely. In the bottom panel, we plot the observed level and our instrument. The instrumental variable is smaller in scale, but follows a similar trend.

To examine the strength of our instruments, we regress city-daily $PM_{2.5}$ on our instrument, the predicted *non-locally-generated* $PM_{2.5}$, together with the other controls (city fixed effects, week fixed effects, etc.). There are a total of 192 thousand observations and the R-squared is 47%. The coefficient of our IV is 0.63 and highly significant, with a t-stat of 15. The cluster-robust F-statistic is 253.

We have also compared the results of this model with those from a Community Multi-scale Air Quality (CMAQ) model for the Beijing-Tianjin-Hebei area. In Appendix E, we present source-receptor matrices both as predicted by our model (Table E6) and by the CMAQ model (Table E5). The two models are broadly consistent with one another: our model implies that 56.4% of Beijing's PM 2.5 comes from other cities, whereas the corresponding number in the CMAQ model is 61.9%.

Finally, it is also important to note that in contrast to the CMAQ model, the goal of our first-stage model is not to maximize the accuracy of air quality predictions. Instead, our goal is to create an instrumental variable that is both predictive of local air pollution and at the same time exogenous to shocks to health spending. This is why we have chosen to base most of our analysis on a relatively conservative definition of the buffer zone, with only $PM_{2.5}$ coming from cities that are more than 150 km away used in constructing the IV (although we also test how robust our results are to this choice of buffer radius).

5 Empirical Results

5.1 Short-Term Impact

Our empirical analysis begins with the contemporaneous effect of air pollution on health. In the discussion below, we use the log number of transaction as the dependent variable rather than the value of transactions as in the literature using transaction-level purchase data Einav et al. (2014). The distribution of health spending rightly skewed with many large transactions (e.g., surgeries) that are unlikely caused by air pollution in the short run. In Appendix C, we report results using the value of transactions as the dependent variable. They are very similar in magnitude to those based on the number of transactions but less precise.

In all of the regressions, we include city fixed effects to control for time-invariant unobservables

and week fixed effects to control for nationwide shocks. City-specific time trend and city-specific seasonality (i.e., interactions of city fixed effects and quarterly dummies) are added to the regression to control for trends in card adoption and seasonal diseases. We also add fixed effects for major holidays, working weekend, and day-of-week, as well as weather variables to control for their direct effects on spending. For example, people may reduce non-emergent hospital visits during holidays or on days with bad weather. All standard errors are clustered at the city level.

Table 2 summarizes the OLS estimates for the short-term impacts. A 10 μg/m^3 increase in the daily $PM_{2.5}$ concentration is associated with a 0.1% increase in the total number of transactions on health care on the same day. Transactions in pharmacies and especially in Children's hospitals are more sensitive to air pollution, with an impact of 0.13% and 0.18% from a 10 μg/m^3 increase in $PM_{2.5}$. The larger impact on pharmacies and Children's hospitals make intuitive sense: children are more vulnerable to air pollution. Similarly, when elevated air pollution aggravates symptoms for people with respiratory problems, they may go to pharmacies to purchase drugs without visiting hospitals.[24]

In contrast, a temporary increase in $PM_{2.5}$ reduces transactions in daily necessities and supermarkets. This could be due to two possibilities. The first is the effect of the budget constraint: if consumers have to spend more in heath care to mitigate the negative health impact of air pollution, they may have less to spend on non-health related categories. The second possibility is avoidance behavior: consumers postpone or reduces shopping trips in response to poor air quality to reduce pollution exposure. We test these two possibilities in Section 5.4.

To graphically illustrate the relationship between pollution and spending, we plot the log number of transactions against $PM_{2.5}$ in Figure 7. All other controls (weather, city trend, etc.) are partialled out, so the plot displays the net effect of pollution on spending. For ease of presentation, we group $PM_{2.5}$ by percentiles and plot the in-group average of log number of transactions against each percentile of $PM_{2.5}$. In addition to the aggregate number of health transactions (top left corner), we also plot the relationship separately for Renmin hospital, Children's hospital, and pharmacy and two non-health categories (necessities and supermarkets). $PM_{2.5}$ has a positive relationship with spending in all health categories across *all* quantiles of $PM_{2.5}$. The cloud spots tightly center around the fitted curve, which is consistent with the fact that our standard errors are small.

To address the issue of endogeneity and measurement errors, we instrument $PM_{2.5}$ using the predicted $PM_{2.5}$ from pollutants originated from outside of the 150 km buffer zone as discussed in Section 4.2. Table 3 reports results from IV regressions. The F-statistics on the IV from the first-stage (reported in the last row of the table) varies from 184 to 258, suggesting a strong first stage.

[24]There is no distinction between prescription and over-the-counter medicines in China and medicines can be purchased without prescription from pharmacies. The Ministry of Human Resources and Social Security maintains the National Reimbursement Drug List (NRDL) and only the drugs on the list are covered by the national medical insurance programs, some in full (type A drugs) and others partially (type B).

The IV estimates are considerably larger than the OLS estimates, with most coefficients 3 to 10 times as large as their OLS counterparts. A 10 μg/m^3 increase in $PM_{2.5}$ in a day is associated with a 0.41% increase in transactions in the aggregate health care sector. The effect of air pollution on spending at Children's hospitals is the largest among different health care categories, and is more than three times as large as that for the overall healthcare spending.

The large difference between OLS and 2SLS results on the health impact of air pollution is common in this literature (Knittel et al., 2015; Schlenker and Walker, 2015). The bias toward zero in OLS estimates for both health and non-health spending is consistent with attenuation bias due to (classical) measurement error in $PM_{2.5}$ as an imperfect proxy for population pollution exposure. The downward bias could also be driven by temporary local shocks that are positively correlated with air pollution such as economic activities or big events, but these shocks need to reduce health spending but increase non-health spending to give rise to observed difference between OLS and 2SLS.

There were more than two billion transactions in hospitals and the total out-of-pocket private health spending exceeded one trillion *yuan* ($145 billion) in 2015. According to China's National Health Commission, the aggregate health expenditure, including both private and public spending, was more than three trillion *yuan* ($435 billion) in the same year. Therefore, a 0.41% increase would amounts to 8.2 million more trips to healthcare facilities and 12 billion *yuan* added healthcare cost. As discussed below in more detail in Section 5.5, our estimates on the short-term impact reflect both the true (positive) impact on health spending and the avoidance behavior suggested by the results on the two control groups which reduces health spending (in the short-term from intertemporal substitution). Therefore, the estimated impact is likely a lower bound on the true impact of air pollution on health spending.

5.2 Longer-Term Impact

Exposure to $PM_{2.5}$ could have longer-term health impacts. Directly estimating the coefficients on lagged $PM_{2.5}$ suffers from the high serial correlation in $PM_{2.5}$ and imprecision. Instead, we employ the Almon technique that is discussed in Section 4.1 by allowing pollution impacts to follow a polynomial path of decay. The classic specification of the Almon technique entails three components: the number of lags k, the order of polynomials q, and additional end-point restrictions. In our model, we do not impose any end-point restrictions so as to allow for a flexible functional form. We use model-selection criteria (AIC and BIC) as guidelines to choose the lag length and the degree of polynomial.

Table 4 reports the cumulative effects for different time periods across categories from the OLS regressions with 90 lags and 3rd order polynomials as the benchmark specification. The

standard errors are clustered at the city level and are reported in parenthesis. The first column shows that a 10 μg/m^3 increase in the $PM_{2.5}$ concentration increases today's number of transactions in all healthcare facilities by 0.03%. A permanent elevation of 10 μg/m^3 increase in the $PM_{2.5}$ concentration raises the number of transactions by 0.86%, which is eight times as large as the OLS estimate that only controls for contemporary $PM_{2.5}$ concentration. The long-term impact on pharmacies is the largest while the impact on children's hospital is not statistically significant. The last two columns show a statistically significant negative impact on necessities and supermarket spending in the short term but not in the long run.

Figure 9 shows the time path of impacts for different categories from OLS.[25] The dotted part of each line indicates the impact being statistically insignificant. Two findings emerge from the plots of coefficient estimates. First, $PM_{2.5}$ has a positive impact on health spending in the short term across all health categories. The impact diminishes over time and becomes small and imprecise after three months. Second, air pollution has a negative impact on the spending of the two control groups. The impact is short-lived and lasts for about two weeks. This provides some support for the hypothesis of avoidance behavior , rather than that of the budget constraint, since under a fixed budget, a permanent increase in health spending would lead to a permanent reduction in necessities and supermarkets. We return to this issue in the next section.

To deal with the endogeneity in $PM_{2.5}$, we extend the Almon approach to allow for the IV method. Specifically, we use lagged predicted non-local pollution in a city $\hat{p}^{far}_{c,s}, s < t$ as the instrument for its lagged pollution $p_{c,s}$, and create the Almon sums in the same way as the original regressor as shown in equation (13). In other words, the instrument for $p_{ct} + p_{c,t-1} + p_{c,t-2} + ... + p_{c,t-k}$ is $\hat{p}^{far}_{c,t} + \hat{p}^{far}_{c,t-1} + \hat{p}^{far}_{c,t-2} + ... + \hat{p}^{far}_{c,t-k}$. The contemporary and cumulative effects across different time spans are presented in Table 5.

Several important findings emerge from Table 5. First, the estimated long-term impacts of $PM_{2.5}$ on health spending across all categories from 2SLS are positive and much larger than those from the OLS results, consistent with the comparison from the short-term impact discussed in the previous section. Specifically, a permanent increase of 10 μg/m^3 in the $PM_{2.5}$ concentration raises the number of transactions in the health sector by 2.2%. Second, the largest impact is in Children's hospitals and the impact is more than four times as large as the impact on aggregate health spending. This is consistent with the fact that children are among the most vulnerable group from $PM_{2.5}$. This finding is also consistent with the results from Tables 3. Third, the effects on aggregate non-health spending as well as on daily necessities and supermarket spending are all negative and much shorter-lived and does not seem to persist longer than two weeks: the aggregate impact becomes smaller and statistically insignificant. The time paths of impact for different categories

[25] The optimal specification of the oder of polynomials and lags may differ. For example, spending in children's hospital should be estimated with fewer lags based on our estimation results. To keep the results comparable, we impose the same lag-order structure on all categories.

are presented in Figure 10.

The estimation results suggest that a 10 μg/m^3 increase in $PM_{2.5}$ would lead to an increase in the number of health-related transactions in the long term by 2.2% from 2SLS and and 0.8% from OLS. In terms of the impact on the value of transactions, the effect is about 0.5% from OLS (Table C2 in Appendix C) and 2.0% from 2SLS (Table C3 in Appendix C). The estimates are slightly less precise than those based on the number of transactions. This is likely due to the larger noise inherent in the value of health spending. For example, some of largest incidences of health transactions are likely to be surgeries which are not related to air pollution.[26] The smaller impact on the value of transactions than that on the number of transaction makes intuitive sense in that elevated pollution could reduce the desire to go to hospitals for minor illnesses (and other outdoor activities), leading to a larger impact on transaction frequency. The heterogeneity across different types of healthcare facilities and the impact on non-health spending are all very similar to the the estimates on the number of transactions.

5.3 Robustness Checks

We conduct a variety of robustness checks on model specifications. First, we test several possible combinations of the number of lags and the order of polynomials. Figure 8 plots the estimates for several combinations from OLS. The dotted part in each line indicates that the estimated lagged impact for the corresponding period is statistically insignificant at 5% level. The results show that the effect of air pollution diminishes over time and is often imprecisely estimated after a few months. Specifications with too many lags (over 150) or too few lags (fewer than 60) lead to large swings in parameter estimates. Once the length of lags is set at 90 (three months), we choose the 3rd order of polynomials that maximizes the adjusted R^2 as our benchmark specification. Table B1 in Appendix B presents the results for different order-lag combinations (4 & 90, 5 & 90, 4 & 60, and 4 & 120) from both OLS and 2SLS. The key estimates of cumulative impacts are robust across different specifications: (1) there is a positive and statistically significant impact on health spending across health categories and the effect is observed over a period of three months; (2) the estimates from 2SLS are much larger in magnitude than those from OLS, although the direction of impact is the same; and (3) the impact on the two control groups is negative and large in the immediate short term and tends to be short-lived.

Table 6 reports the cumulative impact for overall health spending under three different orders of polynomials (3, 4, and 5) and five different numbers of lags (60, 90, 120 and 150). The estimates across the three different orders are very close. The results are slightly more sensitive to the lag structure. The cumulative impact tends to be smaller with 60 days of lags and larger with 150 days

[26] Our analysis focus on transactions with amount less than 200,000 *yuan*. Among this sample, the 95 percentile of the transaction value is 6,000 *yuan* and the 99 percentile is 10,000.

of lags but the results are similar across the three other lag structures, suggesting the impact does not persist longer than three months. In addition, the estimates become imprecise with lags longer than 90 days, although the magnitude stays the same.

Our second set of robustness checks is with regard to the choice of buffer zone in constructing the IV. In the benchmark specification, we use 150 km as the buffer zone for a given city assuming that unobservables outside of the buffer zone of the city would not affect health spending in that city. There is an inherent tradeoff in the choice of the radius. On the one hand, the larger the buffer zone is, the exclusion restriction is more likely to hold; on the other hand, the correlation between the predicted $PM_{2.5}$ based on air pollution and wind conditions outside of the buffer zone and the real $PM_{2.5}$ in a given city would be weaker. Table presents several choices of the buffer zone from 50 km to 300 km with an increment of 50 km. The top panel in the table report the first-stage results and the R^2 becomes smaller as the buffer zone becomes larger, suggesting a weaker correlation between the IV and the endogenous variable. The bottom panel shows the aggregate impact during the three months. Overall, the magnitude of the estimated impacts is similar with the estimate from the 300 km buffer zone being the largest and having the largest standard error as well.

The third set of robustness checks control for other pollutants including O_3, SO_2, NO_2 and CO. These pollutants have harmful health impact themselves. Emissions sources such as electricity generation and transportation would produce both particulate matters and other pollutants. Therefore, the estimated health impact from $PM_{2.5}$ could be confounded by other pollutants especially in OLS regressions. The IV strategy should address this issue to some extent in that it leverages the long-range transport property of $PM_{2.5}$ which is different for other pollutants especially O_3 and CO. That is, the predicted $PM_{2.5}$ should be less correlated with other pollutants. Table D4 in Appendix D reports the results for regressions that include these four pollutants as additional controls. The results for both health spending and non-health spending are very similar to those reported in Table 5 without controlling for other pollutants.

5.4 Avoidance Behavior

The analysis of both the short-term and longer-term impact shows that elevated $PM_{2.5}$ leads to increased health spending and reduced non-health spending. This negative impact on non-health spending could be driven by two underlying mechanisms: budget constraint and avoidance behavior. Our analysis below helps discern the underlying cause. If consumers engage in avoidance behavior, future predictions of the pollution level should affect current consumption behavior. For example, if consumers expect pollution to get better in future days, they may postpone their consumption to avoid exposure today. On the other hand, an expectation of worse air may encourage

them to make the consumption in advance.

To capture this, we try to mimic consumers' predictions of future air quality with an OLS model. We regress the next-day pollution level P_{+1} on the variables reflecting the information we believe is available to the consumers when they make the (perhaps unconscious) prediction. The regressors include current-day variables, such as current-day pollution level P_0, weather, special day dummies (holiday, spring festival, working weekend), city FEs and weekly FEs, city-specific seasonality, linear trend, and day-of-week FEs. Additionally, we included variables reflecting next-day information, such as next-day weather expectation[27], next-day special day dummies[28], and day-of-week dummies.

The predicted next-day air quality $\widehat{P}_{+1}$ from the OLS model has a correlation of 0.788 with actual next-day air quality P_{+1}. We further generate a dummy variable indicating if the predicted next-day air $\widehat{P}_{+1}$ is worse (denoted value one) than the average air quality of current day and past 7 days. If we take this dummy variable as the criteria, the OLS prediction has a total accuracy of 71%. The prediction accuracies are summarized in Table 8.

We add a dummy variable $1\{\widehat{P}_{+1} > \bar{P}_{7\text{d}}\}$ into our baseline Almon IV estimation respectively, and the results are presented below in Table 9. The estimated health impacts based on current and lagged $PM_{2.5}$ are robust to the inclusion of this additional variable. The coefficient estimates on this dummy variable itself indicate a 0.79% increase in healthcare transactions on average when consumers anticipate that air quality will deteriorate on the following day.

The regressions based on non-health spending also show a positive correlation between the anticipated next-day pollution level and current-day spending. The positive effect from the anticipated rise in pollution next day on both health and non-health spending provides evidence of avoidance behavior. It suggests that the budget constraint is unlikely to be the channel behind the negative contemporaneous impact of pollution on non-health spending found in both short- and long-term analysis.

5.5 Discussion

Our preferred specifications show that a 10 μg/m^3 increase in $PM_{2.5}$ would lead to a 2.2% increase in the number of health-related transactions in the long term from Table 5 and a 2.0% increase in the value of transactions from Table C3 in Appendix C). Credit/debit transactions account for about half of the total spending in the health care industry with the rest from cash transactions and government transfers. Assuming that the health impact is the same for non-credit/debit card spending, the 2.0% impact translates to more than 70 billion *yuan* ($11 billion) from a 10 μg/m^3

[27] Given the high accuracy of weather forecast in the very short term, we use next-day weather variables in place of their expectations.

[28] Consumers should know, for example, if tomorrow is a holiday.

(about 18%) increase in $PM_{2.5}$. This estimated impact on health spending can directly inform the overall cost estimate of $PM_{2.5}$ and related policy discussions. For example, a 2016 report by OECD based on the benefit-transfer approach estimates that $PM_{2.5}$ and ground level ozone are associated with $20 billion in direct cost on health expenditure (due to morbidity) worldwide with half of them accounted for by non-OECD countries.[29] With the average $PM_{2.5}$ level of 56 $\mu g/m^3$ and the recommended level of 10 by WHO, a simple linear interpolation would imply a $50.6 billion in added health spending due to elevated $PM_{2.5}$ relative to the WHO recommendation in China. Even with the assumption that all increases in health spending from air pollution occur through credit or debit transactions, the interpolation based on our estimates would still suggest a $25.3 billion impact on health spending. Our estimates suggest that the OECD report significantly underestimates the health cost from outdoor air pollution, potentially up to an order of magnitude for developing countries.

To better understand the size of our estimates, we compare our results with the findings in related literature in Table 10. In a study on preventive expenditure, Mu and Zhang (2016) estimate that face mask purchases increase by 5.45% for a 10-point increase in AQI, and 7.06% for anti-$PM_{2.5}$ masks. Given that the translation from $PM_{2.5}$ concentration to AQI is piecewise linear, a 10-point increase in AQI is equivalent to an increase of 7.5 $\mu g/m^3$ to 15 $\mu g/m^3$ in $PM_{2.5}$ concentration. This means that exposure to 10 $\mu g/m^3$ more $PM_{2.5}$ leads to an increase ranging between 3.6% and 7.3% in preventive spending.

For Williams and Phaneuf (2016), whose estimation methods and data are more similar to ours, they find that a one-standard-deviation change in $PM_{2.5}$ (roughly 3.78 $\mu g/m^3$ for their data) leads to 8.3% more spending on asthma and COPD (or 22% for 10 $\mu g/m^3$ more, equivalently). To make our results comparable, we refer to statistics recorded by the National Health Commission that spending on respiratory diseases accounts for 8% of total health expenditure.[30] Assuming all additional spending induced by air pollution is for respiratory diseases, our estimates translate to about a 31% to 53% increase in respiratory-related spending, more than twice as large as the estimate from Williams and Phaneuf (2016).

Our estimates of the effect of air pollution on total health spending provide a lower bound of the social WTP for improved air quality in that the health spending impact does not take into account mortality, the quality of life, the loss of productivity as well as (costly) avoidance behavior which we find from non-health spending. An increase in overall health spending of 70 billion *yuan* (or $11 billion) from a 10 $\mu g/m^3$ increase in $PM_{2.5}$ suggests that a lower bound for the social WTP is 170 *yuan* (or $27) per household for a reduction of $PM_{2.5}$ by 10 $\mu g/m^3$.

[29] The report is titled "The Economic Consequences of Air Pollution", available at http://www.oecd.org/env/air-pollution-to-cause-6-9-million-premature-deaths-and-cost-1-gdp-by-2060.htm.

[30] For 2012, the most recent year available, and for hospital stays only.

To estimate the consumers' private WTP for air quality, we need to account for the fact that the vast majority of consumers have health insurance and do not bear the full costs of their treatment. As the theoretical model in Section 3 illustrates, the consumer WTP is bounded below by the change in out-of-pocket spending resulting from a change in air quality. For urban residents, the proportion of health spending that has to be paid for out-of-pocket equals 32% (for employees) and 52% (for non-employees).[31] An increase in overall health spending of 70 billion *yuan* (or \$11 billion) from a 10 μg/m^3 increase in $PM_{2.5}$ translates into a 29 billion *yuan* (or \$4.5 billion) increase in out-of-pocket spending.[32] This implies a lower bound for consumer WTP of 70 *yuan* (or \$11) per household for a reduction of $PM_{2.5}$ by 10 μg/m^3.

Using a discrete choice framework to estimate the demand of indoor air purifiers in China, Ito and Zhang (2016) estimate a WTP of \$1.1 for a one unit reduction in PM_{10} based on the trade-off between the high price and the better quality (ability to remove more PM_{10}). Our analysis is in line with their estimates despite that we use different empirical frameworks.Ito and Zhang (2016)'s WTP estimate for PM_{10} reduction could capture consumers' WTP for $PM_{2.5}$ as $PM_{2.5}$ constitutes the majority of harmful PM10 pollutants and the higher-quality air purifiers (with HEPA filters) in their study can also effectively remove $PM_{2.5}$.

Using the hedonic model for the housing market in the U.S., Chay and Greenstone (2005) find that consumers are willing to pay \$450 - \$1,050 more in housing price (in 1982-84 dollars) for a one μg/m^3 reduction in TSP. With a 30-year time span and 5% annual discount rate, this translate to an annual WTP of \$72.9 - \$168.5 (in 2015 dollars). Based on the discrete-choice framework for the U.S. housing market, Bayer et al. (2009) estimated the annual household WTP to be \$23.9 - \$29.5 (in 2015 dollars) for one unit reduction in PM_{10}. The WTP estimates from (Chay and Greenstone, 2005; Bayer et al., 2009) for the U.S. are substantially (at least 10 times) larger than our estimates for at least three reasons. First, during our data period, the average household income in China was about one-eighth of that in the U.S. and the environmental quality is shown to be a luxury good Kahn and Matsusaka (1997). Second, our estimate of WTP reflects the lower bound in that it does not account for the impacts on mortality and the quality of life etc. as discussed above while the WTP estimates using the revealed preference approach in both studies should in theory reflect those impacts to the extent that consumers are well-informed and rational. Third, the difference could be partly due to the potential nonlinearity in the WTP schedule since the level of air pollution is drastically different between these two countries.

[31] Due to the low penetration of credit and debit cards in rural areas, our sample primarily consists of urban residents.

[32] In 2011, there were 252 million urban employees and 221 million urban non-employees enrolled in China's public insurance programs (Yu (2015)). The population-weighted average proportion of health spending that must be covered out-of-pocket thus equals 41%.

6 Conclusion

World Health Organization's global air pollution database shows that world's most polluted cities in terms of $PM_{2.5}$ in 2016 were all from developing countries such as China, India, Iran, Pakistan, Philippines, and Saudi Arabia. The database also shows that 98% of cities in low- and middle-income countries with more than 100,000 residents do not meet WHO air quality guidelines. However, developed countries have been the focus of several decades of research on the impacts of air pollution on human health (particularly mortality) from epidemiology and economics. This analysis examines the direct health cost from $PM_{2.5}$ based on the universe of credit and debit card transactions in China and provides a lower bound estimate of consumer WTP for improved air quality that can be used as an input for the cost-benefit analysis of environmental regulations.

To address the potential endogeneity in the air pollution measure, we develop an air quality prediction model in the spirit of the US EPA's source-receptor matrix whereby we consider each city as a source and a receptor at the same time. The model allows us to tease out plausibly exogenous variations in local air quality from the spatial spillovers due to the long-range transport of $PM_{2.5}$. The IV results, three to four times larger than those from OLS, suggest that a 10 $\mu g/m^3$ decrease in $PM_{2.5}$ (or about 20% from the current level) would lead to at least 70 billion *yuan* ($11 billion) reduction in health spending annually, amounting to 2.0% of total annual healthcare expenditure nationally. Our estimate of direct health cost from $PM_{2.5}$ in China suggests that the recent report by OECD (2016) drastically underestimates worldwide impact of air pollution on health expenditure ($10 billion for non-OCED countries). Our findings based on the large-scale consumer spending data in China should inform similar studies in the future.

In many major urban centers in Northern China, the annual average concentration of $PM_{2.5}$ is close to or even exceeds 100 $\mu g/m^3$, compared to the WHO recommended level of 10 $\mu g/m^3$. The National Plan on Air Pollution Control developed by the State Council in 2013, for the first time as a national policy, set a goal of reducing $PM_{2.5}$ by 25%, 20% and 15% in 2017 relative to the 2012 levels in Beijing-Tianjin-HeBei, Yangtze River Delta, and Pearl River Delta regions. The findings from this study imply that the targeted reductions could lead to significant economic benefit.

We offer to our knowledge the first national-level analysis of the impact of air pollution on health spending in a developing country context. The air pollution level in many large urban centers in developing countries is often an order of magnitude higher than that observed in developed countries. As urbanization continues and development pressure rises, air pollution could be further exacerbated before it can get better. The full impacts of air pollution on economic growth through channels such as human capital accumulation, productivity, talent loss due to migration, and foreign direct investments are interesting and important areas for future research.

References

Almon, Shirley, "The distributed lag between capital appropriations and expenditures," *Econometrica*, 1965, pp. 178–196.

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

Bagnall, John, David Bounie, Kim P Huynh, Anneke Kosse, Tobias Schmidt, Scott D Schuh, and Helmut Stix, "Consumer cash usage: A cross-country comparison with payment diary survey data," *ECB Working Paper*, 2014.

Bank, World, *Cost of Pollution in China: Economic Estimates of Physical Damages*, The World Bank, 2007.

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

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

Brunekreef, Bert and Stephen T Holgate, "Air Pollution and Health," *Lancet*, 2002, *360*, 1233–1242.

Central Bank of China, "General Report on the Operations of Chinese Payment System, 2015 Q3," Technical Report, Central Bank of China 2015.

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

_ and _, "Does air quality matter? Evidence from the housing market," *Journal of political Economy*, 2005, *113* (2), 376–424.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li, "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy," *Proceedings of the National Academy of Sciences*, 2013, *110*, 12936–12941.

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

_ and W. Reed Walker, "Traffic Congestion and Infant Health: Evidence from E-ZPass," *Ameri-*

can Economic Journal: Applied Economics, 2011, *3*, 65–90.

Deryugina, Tatyana, Garth Heutel, Nolan Miller, David Molitor, and Julian Reif, "The Effect of Pollution on Health and Health Care Utilization: Evidence from Changes in Wind Direction," 2017. Working Paper.

Deschenes, Olivier, Michael Greenstone, and Jonathan Guryan, "Climate Change and Birth Weight," *American Economic Review*, May 2009, *99* (2), 211–217.

_ , _ , and Joseph Shapiro, "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program," forthcoming. American Economic Review.

Dockery, Douglas W., "Health Effects of Particulate Air Pollution," *Annals of Epidemiology*, 2009, *19*, 257–263.

Einav, Liran, Dan Knoepfle an Jonathan Levin, and Neel Sundaresan, "Sales Taxes and Internet Commerce," *American Economic Review*, 2014, *104* (1), 1–26.

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

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

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

Ito, Koichiro and Shuang Zhang, "Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China," 2016. Working Paper.

Kahn, Matthew E. and John G. Matsusaka, "Measuring Environmental Demand: Evidence from California Voting Initiatives," *Journal of Law and Economics*, 1997, *40* (1), 137–173.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders, "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health," *Review of Economics and Statistics*, 2015, *98*, 350–366.

Lelieveld, J., J. S. Evan, M. Fnais, D. Giannadaki, and A. Pozzer, "The contribution of outdoor air pollution sources to premature mortality on a global scale," *Nature*, 2015, *525*, 367?371.

Levinson, Arik, "Valuing public goods using happiness data: The case of air quality," *Journal of Public Economics*, 2012, *96* (9), 869–880.

Moretti, Enrico and Matthew Neidell, "Pollution, Health, and Avoidance Behavior: Evidence

from the Ports of Los Angeles," *Journal of Human Resources*, 2011, *46*, 154–175.

Mu, Quan and Junjie Zhang, "Air pollution and defensive expenditures: evidence from particulate-filtering face masks," *Journal of Environmental Economics and Management*, 2016.

Neidell, Matthew, "Air Pollution, Health, and Social-economic Status: the Effect of Outdoor Air Quality on Childhood Asthma," *Journal of Health Economics*, 2004, *23*, 1209–1236.

OECD, "The Economic Consequences of Air Pollution," 2016. http://www.oecd.org/env/air-pollution-to-cause-6-9-million-premature-deaths-and-cost-1-gdp-by-2060.htm.

Pope, C. Arden and Douglas W. Dockery, "Health Effects of Fine Particular Air Pollution: Lines that Connect," *Journal of the Air and Waste Management Association*, 2012, *56*, 709–742.

Schlenker, Wolfram and W. Reed Walker, "Airports, Air Pollution, and Contemporaneous Health," *Review of Economic Studies*, 2015, *0*, 1–43.

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

Viscusi, W. Kip and Joseph Aldy, "The Value of a Statistical Life: A Critial Review of Market Estimates Throughout the World," *Journal of Law and Economics*, 2003, *27* (1), 5–76.

Williams, Austin M. and Daniel J. Phaneuf, "The Impact of Air Pollution on Medical Expenditures: Evidence from Spending on Chronic Respiratory Conditions," 2016. Working Paper.

Yu, Hao, "Universal health insurance coverage for 1.3 billion people: What accounts for China's success?," *Health Policy*, 2015, *119* (1), 1145–1152.

Zhang, Lin, Licheng Liu, Yuanhong Zhao, Sunling Gong, and Xiaoye Zhang, "Source attribution of particulate matter pollution over North China with the adjoint method," *Environmental Research Letters*, 2015, *10* (8).

Figure 1: Three-Year Average $PM_{2.5}$ Concentration

2013 - 2015, $\mu g/m^3$

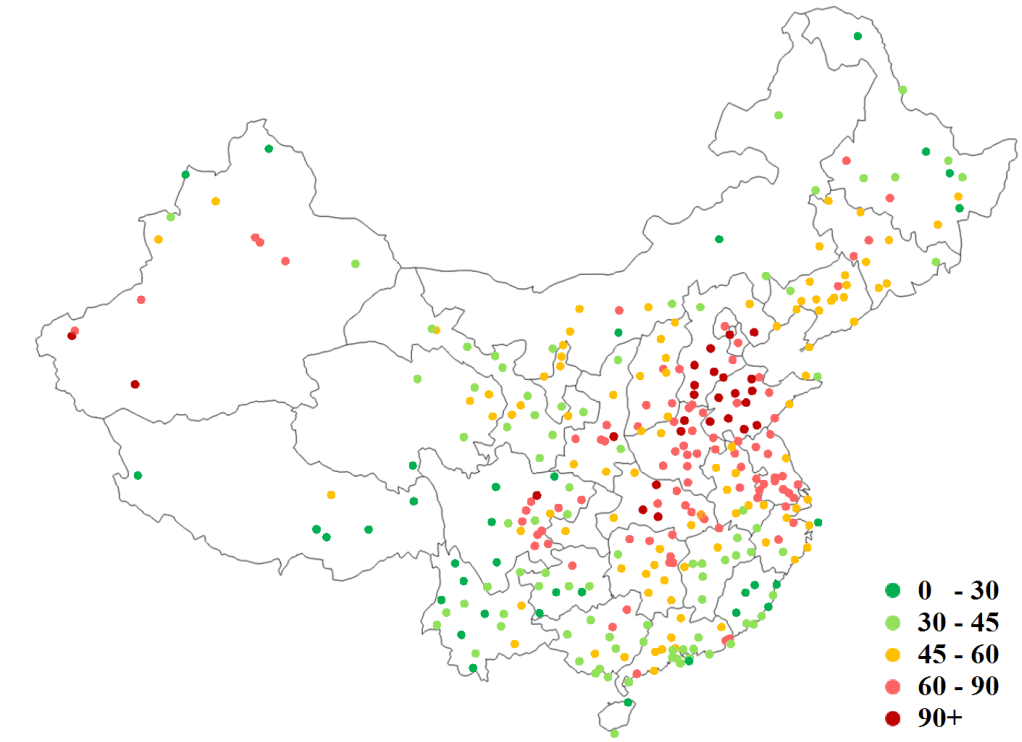
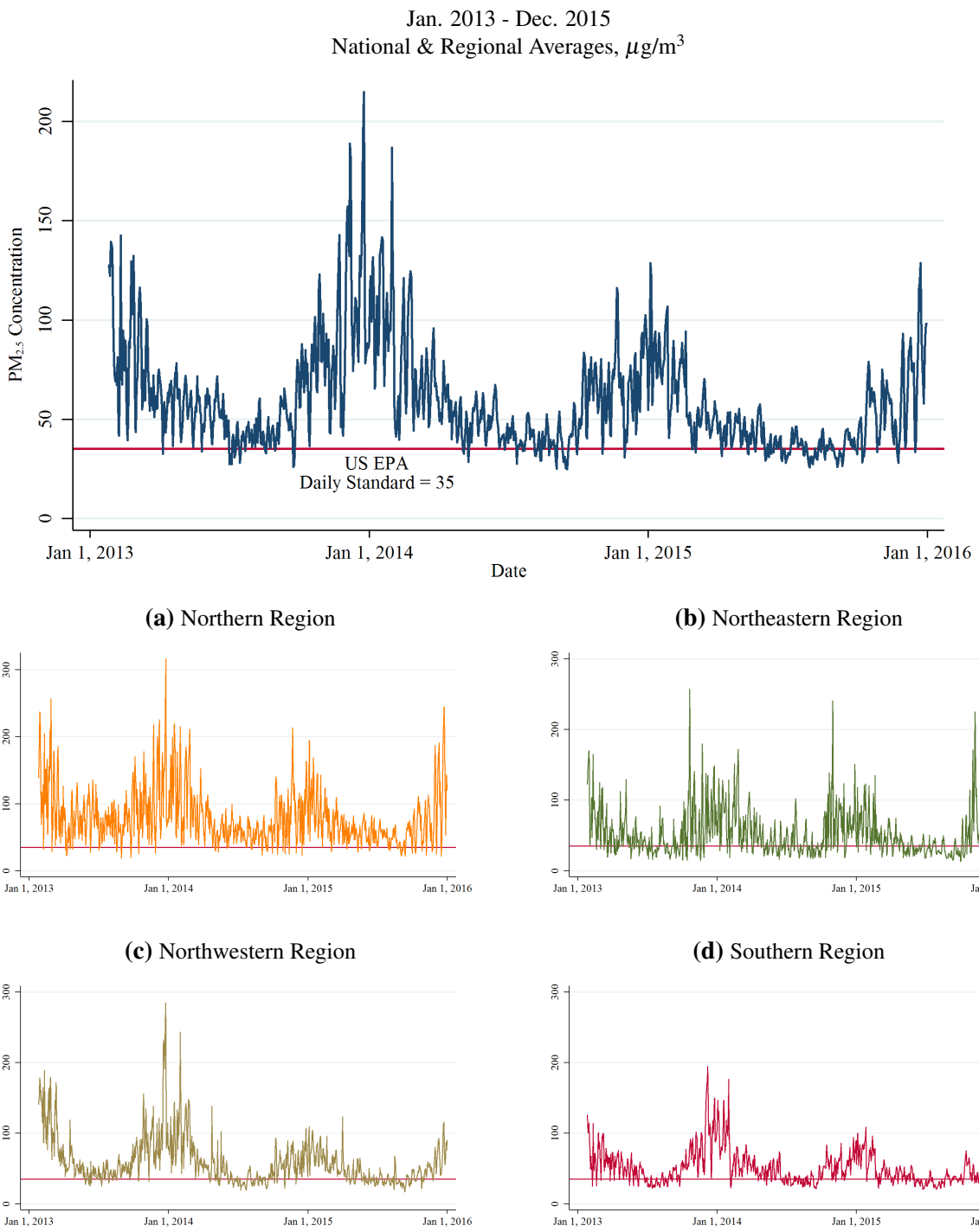


Figure 2: Daily $PM_{2.5}$ Concentration



Notes: Red line in all subfigures indicates the daily standard set by US EPA: 35 μg/m^3.

Figure 3: National Weekly Healthcare Spending, 2013 - 2015

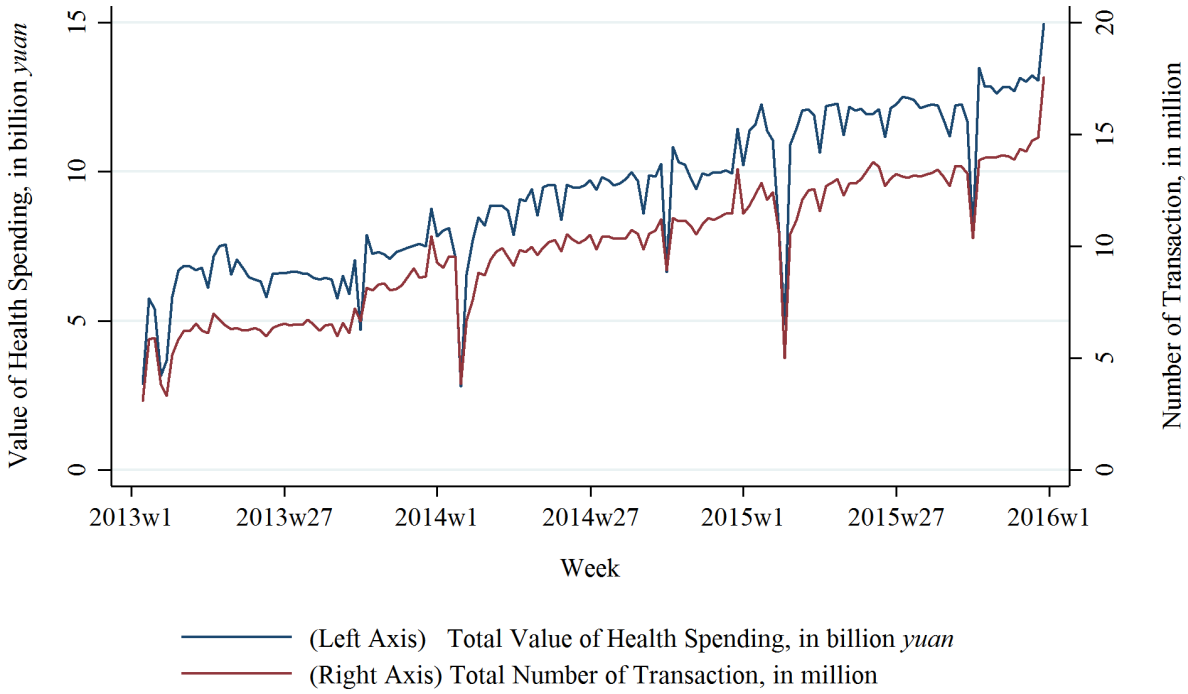
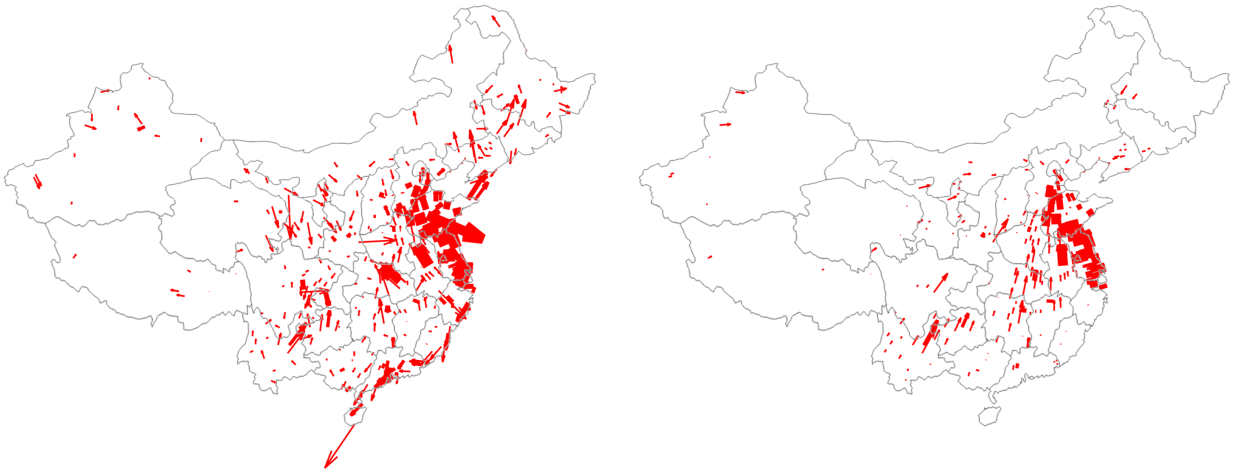


Figure 4: Wind-Pollution Vector Decomposition

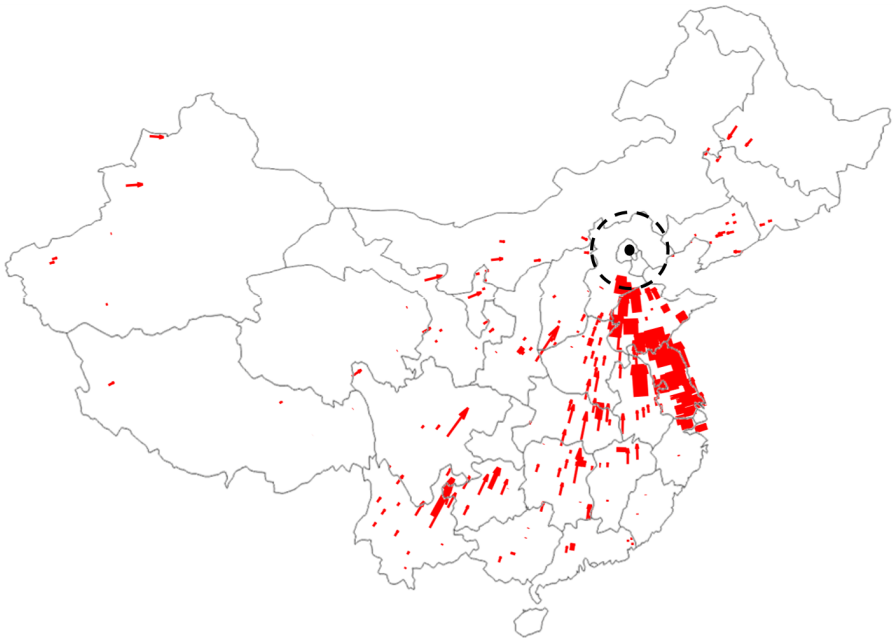
(a) Day 0: Wind-Pollution Vectors

(b) Day 0: Wind-Pollution Vectors (Decomposed)



Notes: Day 0 = Dec. 5, 2013. Subfigure 4a depicts the wind-pollution vector field on Day 0 from raw data, with vector's length indicating wind speed (rescaled to match distance traveled per day), pointing towards the direction where the wind blows to, and width indicating pollution level ($PM_{2.5}$ concentration) in the source city. In Subfigure 4b, the vectors are decomposed, and subvectors pointing towards Beijing are visualized.

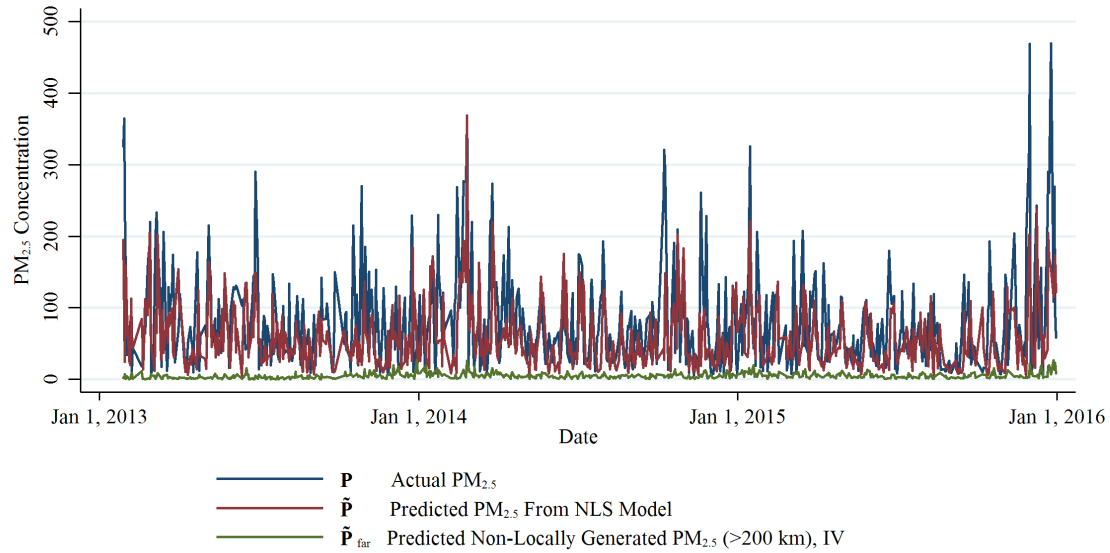
Figure 5: Wind-Pollution Vector Decomposition
From Cities Further Than 150 km



Notes: This graph is modified from Subfigure 4b, by excluding all subvectors originated from cities within a 150 km radius of Beijing. Black dot indicates the location of Beijing, and the dashed line circles the area with a radius of 150 km.

Figure 6: Actual and Predicted Level of Daily Air Pollution

Beijing, Jan. 2013 - Dec. 2015



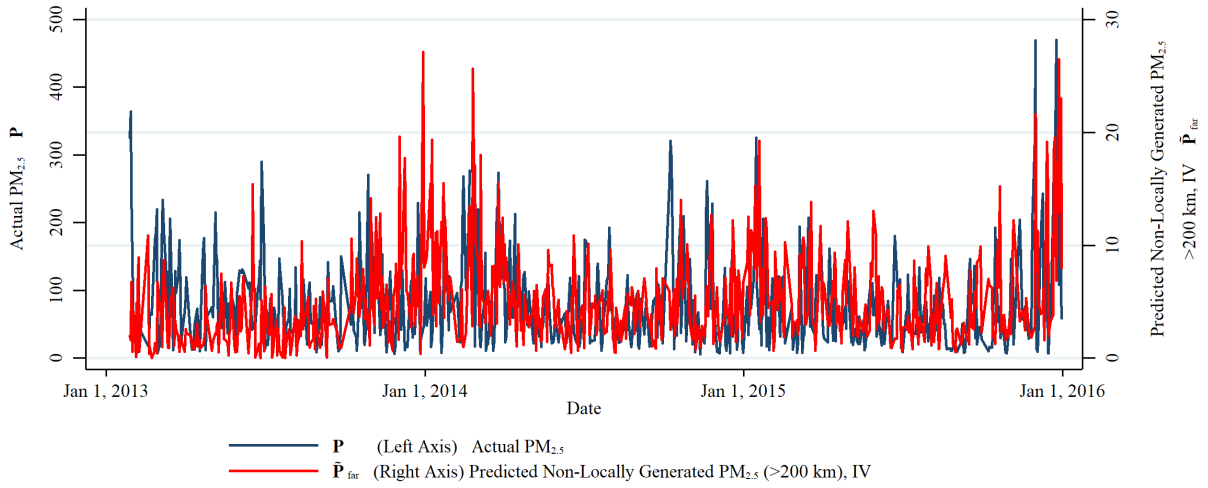
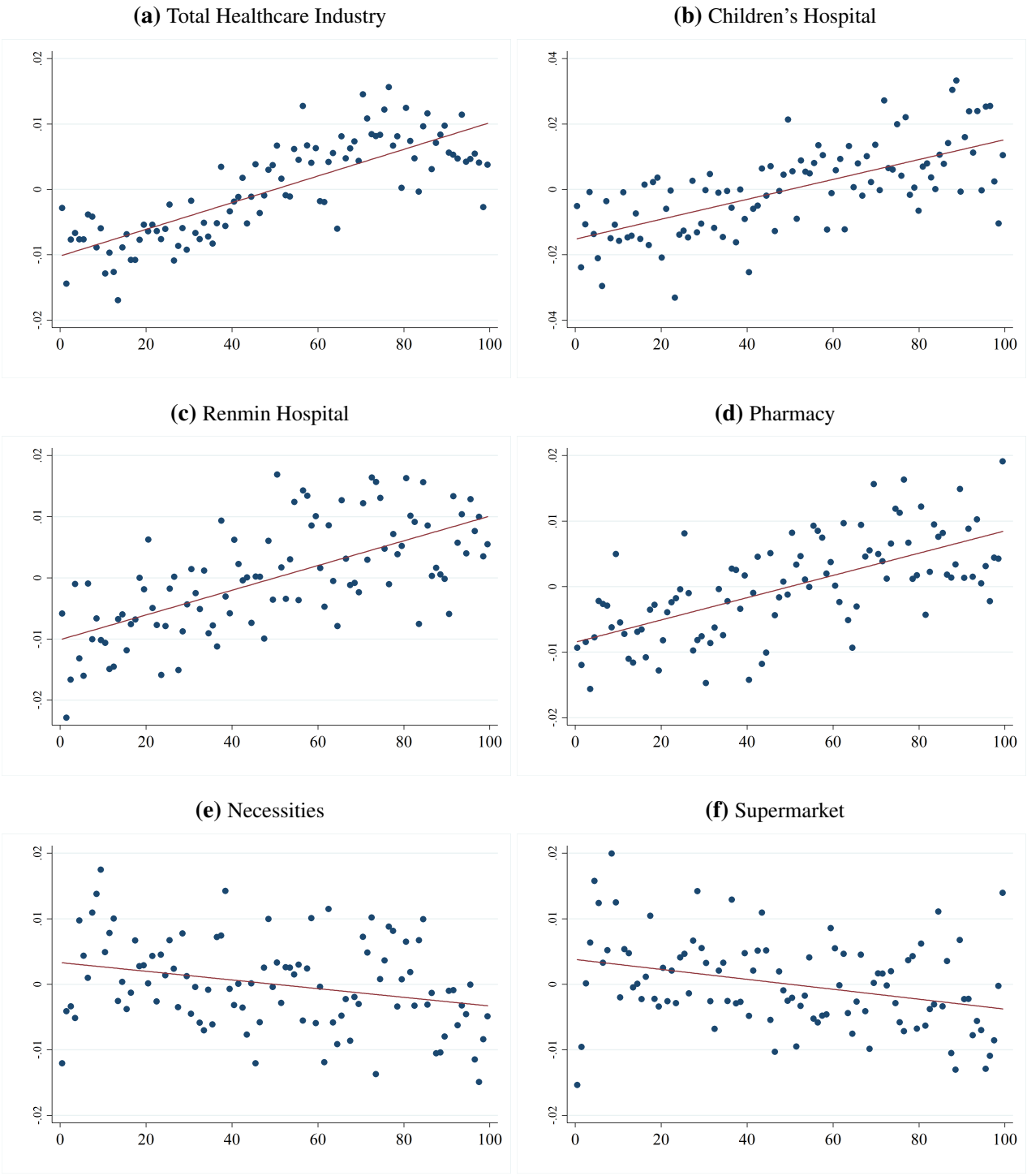


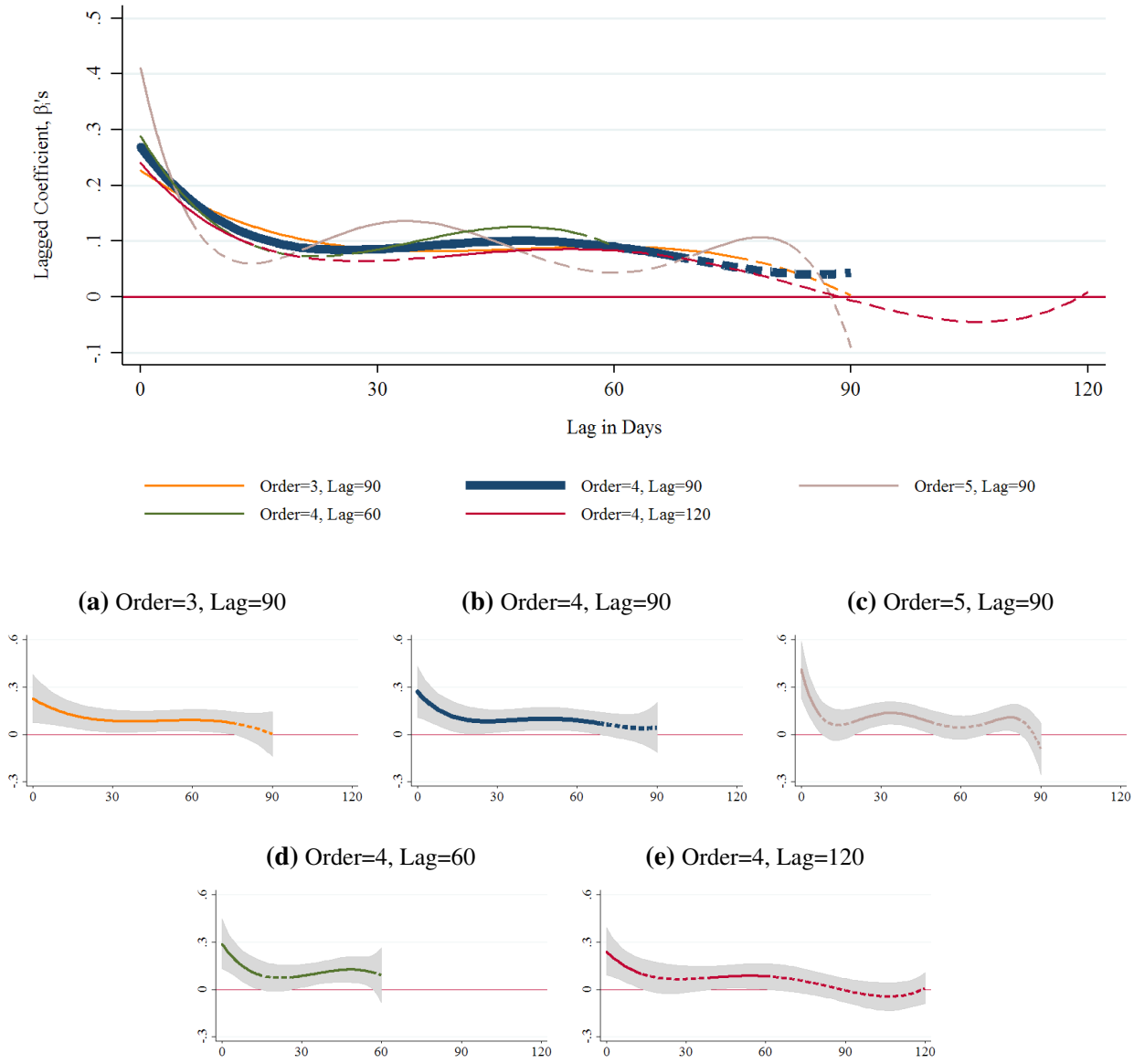
Figure 7: Residuals: Log Number of Transactions v. $PM_{2.5}$ Concentration, by Industry



Notes: Each dot denotes the in-group average residuals, partialling out city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). Groups are binned by percentiles of the x-axis variable.

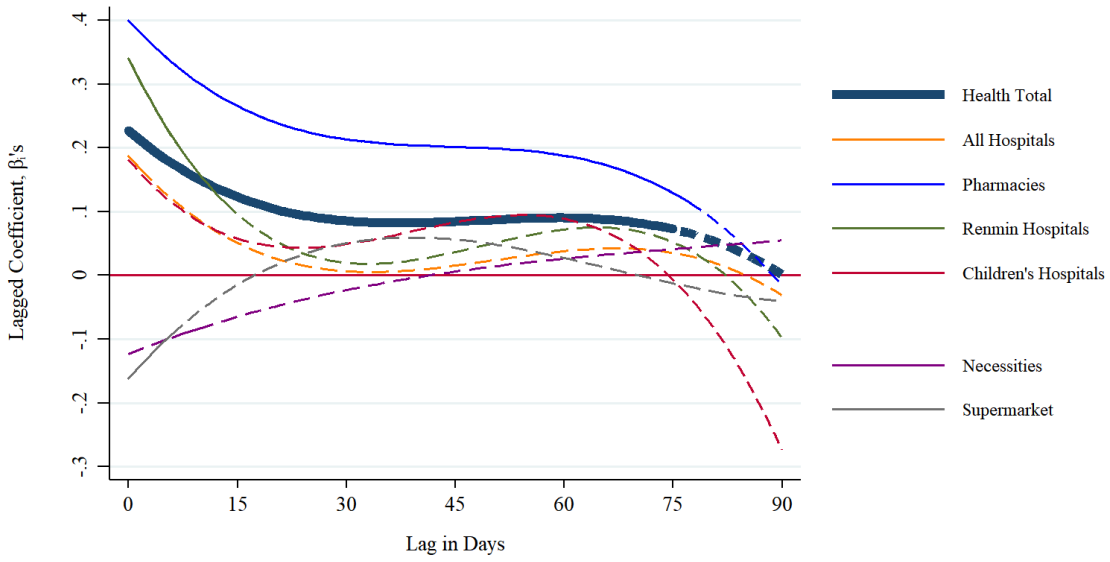
Figure 8: Impact of Air Pollution on Number of Transaction from OLS, Total Healthcare Industry

Various Order-Lag Specifications



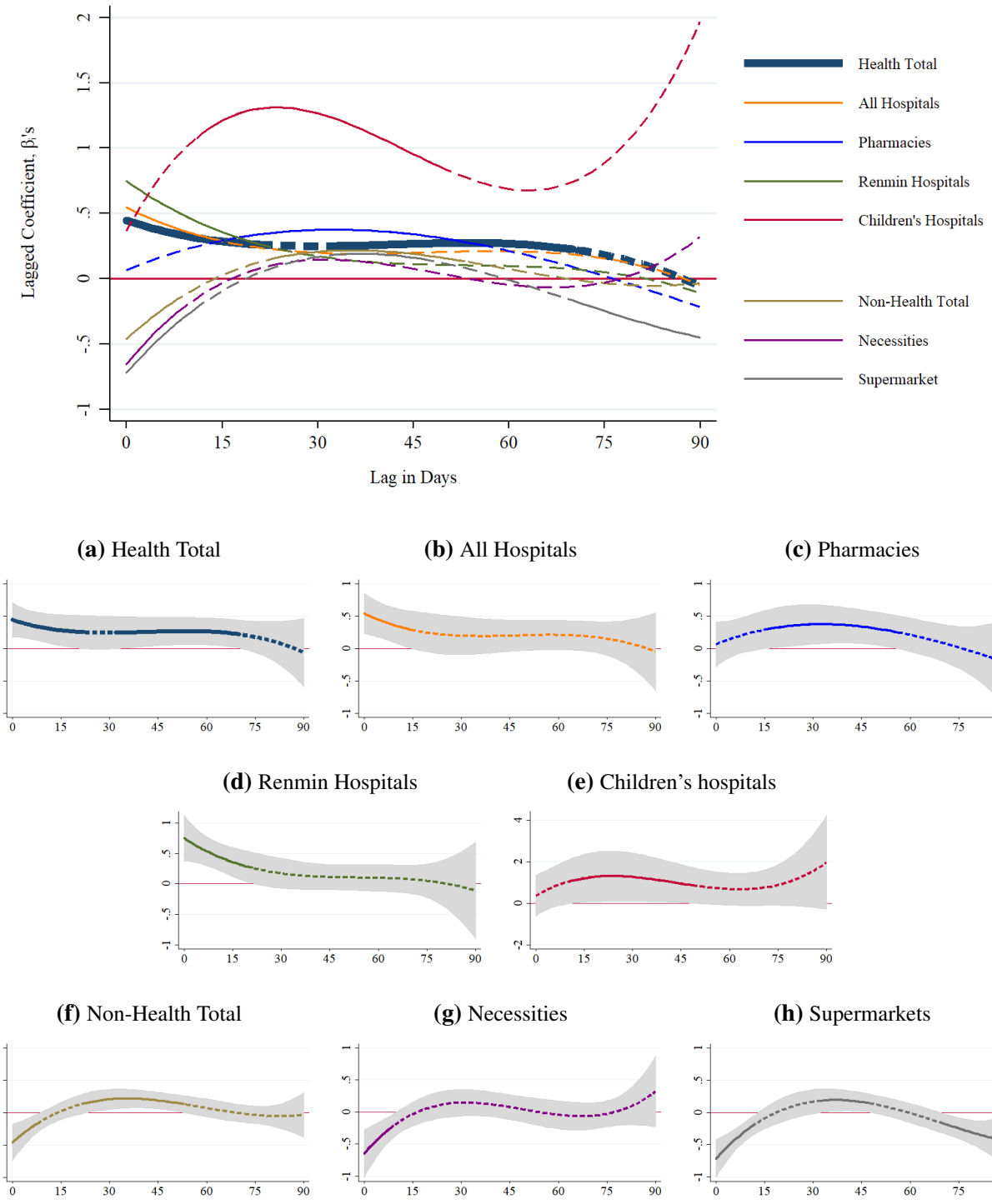
Notes: The coefficients indicate the percentage change in number of transactions per 100 μg/m^3 increase in $PM_{2.5}$ concentration. Solid line indicates $p < 0.05$ for β_i. Gray areas are 95% confidence intervals.

Figure 9: Impact of Air Pollution on Number of Transactions from OLS (Order=3, Lag=90)



Notes: The coefficients indicate the percentage change in number of transactions due to per 100 μg/m^3 increase in $PM_{2.5}$ concentration.

Figure 10: Impact of Air Pollution on Number of Transactions from IV (Order=3, Lag=90)



Notes: The coefficients indicate the percentage change in the number of transactions per 100 μg/m^3 increase in $PM_{2.5}$ concentration. Solid line indicates $p < 0.05$ for β_i. Gray areas are 95% confidence intervals. Coefficients for Children's Hospital in (e) are scaled differently from other subfigures.

Table 1: Summary Statistics

	Mean	Std. Dev.	Min.	Max.	N
Pollution					
$PM_{2.5}$ Concentration, $\mu g/m^3$	56.33	46.37	0	985.18	198,246
Number of Transactions, Daily					
Healthcare Industry, Total	7,229.22	21,308.68	0	330,974	211,318
All Hospitals	4,122.72	14,503.94	0	237,525	210,539
Renmin Hospitals	1,060.60	2,800.44	0	40,332	203,407
Children's hospitals	464.77	1,290.53	0	18,227	158,637
Pharmacies	2,245.33	7,063.37	0	96,336	210,001
Control Groups, *from 1% card sample*					
Total Non-health	778.06	1,956.94	0	27,060	211,318
Daily Necessities	233.37	628.63	0	10,865	211,318
Supermarkets	393.44	990.30	0	15,224	210,493
Total Value of Transactions, Daily, *thousand yuan*					
Healthcare Industry, Total	6,701.80	17,818.97	0	301,108.70	211,318
All Hospitals	5,556.58	15,066.80	0	275,883.00	210,539
Renmin Hospitals	1,588.13	3,401.29	0	56,856.93	203,407
Children's hospitals	363.93	843.39	0	10,324.36	158,637
Pharmacies	407.40	1,109.59	0	16,735.16	210,001
Control Groups, *from 1% card sample*					
Non-Health Total	467.67	1,554.28	0	33,457.73	390, 941
Daily Necessities	236.93	551.38	0	9,532.49	211,318
Supermarkets	232.88	643.43	0	14,404.73	210,493
Weather					
Mean Temperature, $^\circ F$	60.11	18.92	-27.50	101.63	211,317
Percipitation, *inch*	0.13	0.42	0	15.60	211,318
Rain Dummy	0.36	0.48	0	1	211,318
Mean Wind Speed, *mph*	5.50	3.11	0	48.71	211,296
Wind Direction, *navigational bearing*	-	-	0	360	211,263

Notes: Sources of the data are Ministry of Environmental Protection, P. R. China, Integrated Surface Database (ISD), and Global Surface Summary of the Day (GSOD) Database. Data for control groups are calculated for the 1% card sample. Units for variables are displayed in *italics*, if possible. While calculating total value of transaction, transactions with value larger than 200,000 *yuan* ($29,000) are excluded. For wind direction, arithmetic mean and standard deviation do not have statistical meaning, and are thus left out in the table.

Table 2: OLS Estimates of Pollution Impact on Transaction Frequency

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PM_{2.5}$ (10,000 $\mu g/m^3$)	1.00***	0.87***	1.33***	0.98***	1.79***	-0.28	-0.52**	-0.21
Current Day	(0.18)	(0.21)	(0.25)	(0.21)	(0.55)	(0.17)	(0.24)	(0.20)
N	189641	189075	188568	183214	144508	192498	189228	188995

Notes: The dependent variable is log(number of transactions). The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day FEs (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). $PM_{2.5}$ is measured in 10000 $\mu g/m^3$. The coefficient on that variable indicates the percentage change in the number of transactions per 100 $\mu g/m^3$ increase in $PM_{2.5}$ concentration. Standard errors in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3: IV Estimates of Pollution Impact on Transaction Frequency

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PM_{2.5}$, Current Day	4.11***	4.82***	3.40*	5.37***	14.96**	-0.62	-5.16***	-1.95
Current Day	(1.16)	(1.44)	(1.77)	(1.43)	(5.81)	(1.33)	(1.70)	(1.46)
N	189641	189075	188568	183214	144508	189588	189228	188995
First-stage F	258.2	257.2	256.6	248.1	184.2	258.1	258.1	257.5

Notes: The dependent variables are log(number of transactions). The instrumental variable is predicted non-locally generated $PM_{2.5}$, from cities more than 150 km away. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). The coefficient on $PM_{2.5}$ indicates the percentage change in the number of transactions per 100 $\mu g/m^3$ increase in $PM_{2.5}$ concentration. Standard errors in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and it is cluster-robust at the city level.

Table 4: Cumulative Effect of Pollution, Almon Estimation from OLS (Order=3, Lag=90)

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Current Day	0.27***	0.22**	0.44***	0.30***	0.56***	-0.20***	-0.32***	-0.21***
	(0.08)	(0.09)	(0.11)	(0.11)	(0.21)	(0.07)	(0.10)	(0.08)
Current + Past 3d	0.97***	0.78**	1.65***	1.11***	1.72**	-0.68***	-1.03***	-0.70***
	(0.29)	(0.33)	(0.38)	(0.37)	(0.79)	(0.24)	(0.33)	(0.26)
Current + Past 7d	1.71***	1.29**	3.02***	2.01***	2.36	-1.06***	-1.51***	-1.11***
	(0.51)	(0.59)	(0.65)	(0.59)	(1.51)	(0.38)	(0.52)	(0.43)
Current + Past 14d	2.63***	1.76*	4.98***	3.09***	2.11	-1.28**	-1.65**	-1.34**
	(0.80)	(0.96)	(1.00)	(0.84)	(2.77)	(0.53)	(0.72)	(0.64)
Current + Past 28d	3.88***	1.97	8.14***	3.98***	1.23	-1.08	-1.49	-0.88
	(1.25)	(1.51)	(1.64)	(1.38)	(4.95)	(0.88)	(1.15)	(1.10)
Current + Past 56d	6.54***	2.62	14.10***	4.70**	5.74	-0.68	-2.74	0.31
	(1.87)	(1.96)	(2.94)	(2.36)	(7.43)	(1.52)	(2.10)	(1.80)
Current + All Lags	8.63***	3.37	18.08***	5.92*	3.77	-0.01	-0.82	0.22
	(2.73)	(2.76)	(4.24)	(3.59)	(11.44)	(1.85)	(2.69)	(2.14)

Notes: The dependent variable is log(number of transactions). The effect of air pollution for the past 90 days is estimated allowing polynomial decay with an order of 3. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). The coefficients indicate cumulative effect on the number of transactions in percentage, from an exposure to 100 μg/m^3 more $PM_{2.5}$ for the corresponding length of time listed in the first column. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5: Cumulative Effect of Pollution from IV (Order=3, Lag=90)

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Current Day	0.45***	0.55***	0.07	0.75***	0.37	-0.46***	-0.65***	-0.72***
	(0.14)	(0.16)	(0.18)	(0.19)	(0.50)	(0.14)	(0.19)	(0.15)
Current + Past 3d	1.70***	2.04***	0.39	2.79***	1.98	-1.57***	-2.26***	-2.56***
	(0.50)	(0.60)	(0.65)	(0.67)	(1.95)	(0.50)	(0.66)	(0.51)
Current + Past 7d	3.17***	3.76***	1.06	5.10***	5.15	-2.53***	-3.71***	-4.32***
	(0.93)	(1.13)	(1.19)	(1.15)	(3.85)	(0.83)	(1.11)	(0.85)
Current + Past 14d	5.38***	6.14***	2.79	8.18***	12.65*	-3.03**	-4.77***	-5.89***
	(1.65)	(2.02)	(2.07)	(1.80)	(7.43)	(1.18)	(1.61)	(1.24)
Current + Past 28d	9.04***	9.49**	7.54*	11.88***	30.64**	-1.25	-3.73	-5.22**
	(3.23)	(3.91)	(3.95)	(3.07)	(15.37)	(1.77)	(2.49)	(2.05)
Current + Past 56d	16.37***	15.11**	17.05**	15.41***	58.94**	3.97	-1.33	-0.99
	(6.12)	(7.15)	(7.32)	(5.09)	(28.40)	(3.05)	(4.54)	(3.59)
Current + All Lags	22.01**	19.78*	18.16	16.70*	93.44**	3.85	-0.44	-8.34
	(10.16)	(11.45)	(12.74)	(8.83)	(46.15)	(4.42)	(7.16)	(5.15)
First-stage F	60.33	60.42	60.22	59.00	50.07	60.33	60.30	60.28

Notes: The dependent variable is log(number of transactions). The effect of air pollution for current day and the past k days is estimated allowing polynomial decay with an order of 3. The instrumental variable is predicted non-locally generated $PM_{2.5}$, from cities more than 150 km away, transformed by the same method as the regressor. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). The coefficients indicate the cumulative effect on the number of transactions in percentage, from an exposure to 100 $\mu g/m^3$ more $PM_{2.5}$ for the corresponding length of time listed in the first column. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported for cluster-robustness.

Table 6: Cumulative Effects of Pollution on Transaction Frequency from IV

	Lag k				
Order q	30 days	60 days	90 days	120 days	150 days
3	12.11***	20.79***	22.01**	19.86	20.78
	(3.58)	(7.23)	(10.16)	(15.46)	(21.52)
4	13.49***	20.86***	21.97**	19.60	20.50
	(3.98)	(7.62)	(10.27)	(15.31)	(21.41)
5	13.43***	19.66***	21.94**	19.28	20.54
	(3.54)	(7.12)	(9.99)	(14.97)	(21.33)

Notes: The dependent variable log(number of transactions). The effect of air pollution for current day and the past k days is estimated allowing polynomial decay with an order of q. The instrumental variable is predicted non-locally generated $PM_{2.5}$, from cities more than 150 km away, transformed by the same method as the regressor. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). Coefficients indicates total cumulative effect on the number of transactions in percentage, from an exposure to 100 $\mu g/m^3$ more $PM_{2.5}$ for the corresponding length of time listed in the given specification. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7: Health Impact Estimates with Various Buffer Radii

	Buffer Radius for IV					
	> 50 km	> 100 km	> 150 km	> 200 km	> 250 km	> 300 km
Pairwise Correlation Between IV and Regressor (Actual $PM_{2.5}$)						
	0.43	0.40	0.34	0.31	0.28	0.24
One-to-One First Stage Regression						
Coef. on IV	0.42***	0.49***	0.63***	0.81***	1.10***	1.71***
	(0.02)	(0.02)	(0.04)	(0.06)	(0.09)	(0.18)
N	192586	192586	192586	192586	192586	192586
R^2	0.494	0.477	0.466	0.461	0.458	0.457
F	606.3	513.7	253.1	158.1	130.8	88.13
IV Regression						
Total Long-Term Effect	25.88**	21.15**	22.01**	27.70**	32.75**	38.12**
on Number of Trans.	(12.73)	(8.28)	(10.16)	(12.63)	(15.49)	(15.29)
F	20.35	54.31	60.33	34.30	47.93	41.04

Notes: The dependent variable is log(number of transactions) for the healthcare industry. The effect of air pollution for current day and the past 90 days is estimated allowing polynomial decay with an order of 3. The instrument variable is predicted non-locally generated $PM_{2.5}$, from cities outside the buffer zone with each radius respectively, transformed by the same method as the regressor. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). Coefficients indicate the total cumulative effect over 3 months on the number of transactions in percentage, from an exposure to 100 $\mu g/m^3$ more $PM_{2.5}$ for the corresponding length of time listed in the given specification. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8: Accuracy of Prediction from OLS Model, Next-Day Pollution Level

	Actually Better	Actually Worse	Total
Predicted Better	72,342	23,338	95,680
Predicted Worse	33,253	68,275	101,528
Total	105,595	91,613	197,208
Accuracy	68.5%	74.5%	71.3%

Notes: The predictions generate a dummy variable being one when the predicted next-day air quality is worse than the mean pollution level of current day and past 7 days. The regressors for the prediction include current-day information, such as pollution level, weather, special day dummies (holiday, spring festival, working weekend), city FEs and weekly FEs, city-specific seasonality, linear trend, and day-of-week FEs; next-day information, such as next-day weather expectation, special day dummies, and day-of-week dummies. We use next-day weather variables in place of their expectations.

Table 9: Cumulative Effect of Pollution from IV, with Dummy $1_{\text{Predicted Worse}}$ (Order=3, Lag=90)

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$1\{\widehat{P}_{+1} > \bar{P}_{7d}\}$	0.79***	0.52*	0.63**	0.66***	1.43	0.42***	0.40*	0.12
	(0.23)	(0.27)	(0.27)	(0.23)	(1.01)	(0.15)	(0.22)	(0.18)
Current Day	0.56***	0.64***	0.15	0.87***	0.66	-0.42***	-0.61***	-0.73***
	(0.15)	(0.19)	(0.20)	(0.20)	(0.61)	(0.15)	(0.20)	(0.16)
Current Day + Past 3d	2.11***	2.38***	0.67	3.22***	3.05	-1.44***	-2.13***	-2.58***
	(0.57)	(0.70)	(0.72)	(0.73)	(2.37)	(0.53)	(0.71)	(0.55)
Current Day + Past 7d	3.89***	4.34***	1.55	5.83***	7.02	-2.31***	-3.48***	-4.36***
	(1.07)	(1.31)	(1.33)	(1.27)	(4.64)	(0.89)	(1.19)	(0.92)
Current Day + Past 14d	6.41***	6.98***	3.49	9.23***	15.43*	-2.73**	-4.45**	-5.94***
	(1.89)	(2.32)	(2.30)	(1.99)	(8.64)	(1.28)	(1.74)	(1.36)
Current Day + Past 28d	10.23***	10.46**	8.34**	13.05***	34.02**	-0.94	-3.40	-5.25**
	(3.54)	(4.28)	(4.24)	(3.26)	(16.82)	(1.89)	(2.64)	(2.20)
Current Day + Past 56d	17.60***	16.16**	17.83**	16.64***	62.36**	4.21	-1.09	-0.98
	(6.45)	(7.52)	(7.63)	(5.23)	(29.84)	(3.16)	(4.67)	(3.73)
Current Day + All Lags	23.42**	21.06*	18.90	18.16**	97.52**	4.01	-0.27	-8.37
	(10.55)	(11.90)	(13.09)	(8.94)	(47.85)	(4.56)	(7.32)	(5.32)
First-stage F	59.70	59.79	59.60	58.28	49.43	59.70	59.67	59.66

Notes: The dependent variable is log(number of transactions). The effect of air pollution for the past 90 days is estimated allowing polynomial decay with an order of 3. The instrument is predicted non-locally generated $PM_{2.5}$, from cities more than 150 km away, transformed by the same method as the regressor. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). The coefficients indicate the cumulative effect on the number of transaction in percentage, from an exposure to 100 $\mu g/m^3$ more $PM_{2.5}$ for the corresponding length of time listed in the first column. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported for cluster-robustness.

Table 10: Summary of Dose-Response Relationships from Literature

Source	Dose, additional	Response
Mu and Zhang (2016)	100-point AQI	54.5% increase in masks purchases, 70.6% in anti-$PM_{2.5}$ masks
Williams and Phaneuf (2016)	1 std. dev. $PM_{2.5}$ (3.78 $\mu g/m^3$)	8.3% more spending on asthma and COPD
Schlenker and Walker (2015)	1 std. dev. pollution	17% more asthma and total respiratory problems, 9% heart problems
Arceo et al. (2015)	1 $\mu g/m^3$ PM_{10} 1 ppb CO	0.23 per 100,000 increase in infant mortality 0.0046 per 100,000 increase in infant mortality
He et al. (2016)	10 $\mu g/m^3$ PM_{10} (roughly 10%)	8.36% in all-cause mortality rate 285,000 premature deaths each year
Chay and Greenstone (2003)	1% TSP	0.35% in infant mortality rate nationwide
Chay and Greenstone (2005)	1 $\mu g/m^3$ TSP	WTP: \$450-\$1,050 in housing price
Bayer et al. (2009)	1 $\mu g/m^3$ PM_{10}	WTP: \$149-\$185 in housing price
Ito and Zhang (2016)	1 $\mu g/m^3$ PM_{10}	WTP: \$1.1 per household
Our estimation		
OLS	10 $\mu g/m^3$ $PM_{2.5}$	0.8% in hospital visits and pharmacy purchases, 0.5% in total health expenditure
IV	10 $\mu g/m^3$ $PM_{2.5}$	2.7% in hospital visits and pharmacy purchases, 2.3% in total health expenditure WTP: \$10 per household annually

Appendices

A Derivation of Marginal Willingness-to-Pay for Clean Air

In this section, we show how to derive the expressions for the marginal willingness to pay for pollution reductions described in Section 3. Recall that individual i's maximization problem can be written as:

$$\begin{aligned} &\max_{\{m_i,c_i,s_i\}} U[h_i,c_i,s_i,e(a,m_i+c_i)], \\ &\quad \text{s.t. } y(h_i) \equiv y_0 + w(h_i) = \pi + pm_i + c_i + s_i, \\ &\quad \text{and } h_i = h_0 + m_i - g_i(e(a,m_i+c_i)), \end{aligned} \tag{19}$$

The Lagrangian can be written as:

$$L_i = U[h_i,c_i,s_i,e(a,m_i+c_i)] + \lambda_i[y(h_i) - \pi - pm_i - c_i - s_i],$$

The first-order conditions are:

$$\begin{aligned} \frac{\partial L_i^*}{\partial m_i} &= U_h(1-g_i'(e_i)e_m) + U_e e_m + \lambda_i(y_h(1-g_i'(e_i)e_m) - p) = 0, \\ \frac{\partial L_i^*}{\partial c_i} &= -U_h g_i'(e_i)e_c + U_c + U_e e_c - \lambda_i(y_h g_i'(e_i)e_c + 1) = 0, \\ \frac{\partial L_i^*}{\partial s_i} &= U_s - \lambda_i = 0, \\ \frac{\partial L_i^*}{\partial \lambda_i} &= y(h_i^*) - \pi - pm_i^* - c_i^* - s_i^* = 0. \end{aligned} \tag{20}$$

where U_h, U_c, U_s are the derivatives of the utility function with respect to each component of the utility function, and $y_h = \frac{\partial y_i}{\partial h_i}$ is the marginal effect of health stock on income.

How does an increase in air pollution a affect the consumer's health and non-health spending decisions? Because air pollution a affects a consumer's wage income, a effectively governs the relative prices of non-health consumption c_i and health consumption m_i with respect to online spending s. The intuition is that when air pollution is high, the effective price of consuming one unit of c is not just the amount spent on the good, but also the additional income lost from the increased exposure to air pollution.[33] An increase in a therefore corresponds to an increase in the relative price of c_i and a decrease in the relative price of m_i. The "price" effect causes c_i to increase

[33] To see this more formally, notice that the consumer's net income left over after purchasing c_i is equal to $y(h_i) - c_i$. Differentiating that with respect to c_i, we see that a 1-unit increase in consumption reduces the consumer's net income available for spending on other goods by $y_h e_c + 1$, which is therefore the effective price of consumption with respect to online spending. Since e_c is increasing in a, it follows that the price of consumption is increasing in a.

and m_i to decrease.

There is also an income effect caused by the reduction in income, which causes both c_i and m_i to decrease. Finally, a directly enters into the utility function through the effect on health stock, h_i, and the effect on exposure e_i. Assuming the marginal utility of consumption is non-decreasing in h_i and non-increasing in e_i, these effects also imply that c_i is decreasing in a. Thus, an increase in air pollution unambiguously causes c_i to decrease.

The effect of an increase in a on m_i is theoretically ambiguous, both because the income and price effects work in opposite directions, and because the consumer is trading off improvements in health stock h_i (which increases utility) against increased exposure to pollution e_i (which lowers utility). As long as income effects are not too large and the effect of medical spending on health stock h_i dominates the increased exposure e_i from going out to visit the doctor, m_i should be increasing in a.

We now derive the marginal willingness to pay for pollution reduction. Denote $V_i(a,h_0,y_0)$ as the indirect utility function and $L_i^*(a,h_0,y_0)$ as the optimal value of the Lagrangian. The marginal WTP for reduction in air pollution can be obtained as:

$$MWTP_i = -\frac{\frac{\partial V_i}{\partial a}}{\frac{\partial V_i}{\partial y_0}} = -\frac{\frac{\partial L_i^*}{\partial a}}{\frac{\partial L_i^*}{\partial y_0}} \tag{21}$$

By the Envelope Theorem,

$$\begin{aligned}\frac{\partial L_i^*}{\partial a} &= -U_h g_i'(e_i)e_a + U_e e_a - \lambda_i y_h g_i'(e_i)e_a - \lambda_i \pi'(a) \\ &= U_e e_a - g_i'(e_i)e_a(U_h + \lambda_i y_h) - \lambda_i \pi'(a) \qquad (22)\\ \frac{\partial L_i^*}{\partial y_0} &= \lambda_i \qquad (23)\end{aligned}$$

Taking the total derivatives of both the health stock h_i and exposure e_i with respect to a, we obtain the following equations:

$$\frac{dh_i^*}{da} = \frac{\partial m_i^*}{\partial a} - g_i'(e_i)e_a - g_i'(e_i)e_c(\frac{\partial m_i^*}{\partial a} + \frac{\partial c_i^*}{\partial a}) \tag{24}$$

$$\frac{de_i^*}{da} = e_a + e_c(\frac{\partial m_i^*}{\partial a} + \frac{\partial c_i^*}{\partial a}) \tag{25}$$

Manipulating the FOCs, we obtain the following relations:

$$\lambda_i p = U_e e c_c + (U_h + \lambda_i y_h)(1 - g_i'(e_i)e_c) \tag{26}$$

$$U_c - U_s = -U_e e_c + (U_h + \lambda_i y_h) g_i'(e_i)e_c \tag{27}$$

Plugging these equations into equation (22), we obtain[34]:

$$\begin{aligned}\frac{\partial L_i^*}{\partial a} &= U_e\big(\frac{de_i^*}{da} - e_c(\frac{\partial m_i^*}{\partial a} + \frac{\partial c_i^*}{\partial a})\big) + \big(\frac{dh_i^*}{da} - \frac{\partial m_i^*}{\partial a} + g_i'(e_i)e_c(\frac{\partial m_i^*}{\partial a} + \frac{\partial c_i^*}{\partial a})\big)\big(U_h + \lambda_i y_h\big) - \lambda_i \pi'(a) \\ &= U_e\frac{de_i^*}{da} + (U_h + \lambda_i y_h)\frac{dh_i^*}{da} - \frac{\partial m_i^*}{\partial a}\big[-U_e e_c + (U_h + \lambda_i y_h)(1 - g_i'(e_i)e_c)\big] + \\ &\quad \frac{\partial c_i^*}{\partial a}\big[-U_e e_c + (U_h + \lambda_i y_h)g_i'(e_i)e_c\big] - \lambda_i \pi'(a) \\ &= -\lambda_i p\frac{\partial m_i^*}{\partial a} + \frac{dh_i^*}{da}\big(U_h + \lambda_i y_h\big) + U_e\frac{de_i^*}{da} + (U_c - U_s)\frac{\partial c_i^*}{\partial a} - \lambda_i \pi'(a) \end{aligned} \tag{28}$$

The marginal WTP for individual i is then equal to:

$$\begin{aligned} MWTP_i &= -\frac{\frac{\partial L_i^*}{\partial a}}{\frac{\partial L_i^*}{\partial y_0}} \\ &= -\frac{U_e e_a - g_i'(a)(U_h + \lambda_i y_h) - \lambda_i \pi'(a)}{\lambda_i} \\ &= p\frac{\partial m_i^*}{\partial a} + \frac{d\pi}{da} + y_h(-\frac{dh_i^*}{da}) + \frac{U_h}{\lambda_i}(-\frac{dh_i^*}{da}) + (-\frac{U_e}{\lambda_i})\frac{de_i^*}{da} + \frac{U_c - U_s}{\lambda_i}(-\frac{\partial c_i^*}{\partial a}) \end{aligned} \tag{29}$$

Assume that premiums cannot adjust in response to pollution: $\frac{d\pi}{da} = 0$. This seems a reasonable assumption to make in the context of China, where the insurance reimbursement rates for the 3 major public insurance programs are fixed in the medium term. In that case, the MWTP for individual i can be written as:

$$MWTP_i = p\frac{\partial m_i^*}{\partial a} + y_h(-\frac{dh_i^*}{da}) + \frac{U_h}{\lambda_i}(-\frac{dh_i^*}{da}) + (-\frac{U_e}{\lambda_i})\frac{de_i^*}{da} + \frac{U_c - U_s}{\lambda_i}(-\frac{\partial c_i^*}{\partial a}) \tag{30}$$

To derive the social WTP, we additionally need to calculate how much the cost of providing insurance changes in response to air pollution. For any level of air pollution a, the net payment that the insurance program has to make to individual i equals $(1-p)m_i$. Each individual i also pays a fixed premium π into the insurance program. The insurance program's net cost therefore equals:

$$C = (1-p)\int m_i dF_i - \pi \tag{31}$$

where $\int m_i dF_i$ is aggregate medical spending in the economy. When premiums are fixed, the change in the cost of providing insurance due to a change in pollution equals:

$$\frac{dC}{da} = (1-p)\int \frac{\partial m_i^*}{\partial a} dF_i$$

[34] In the first line, we plug in (24) and (25). In the second line, we re-arrange and collect terms. To get to the third line, we plug in (26) and (27).

B Almon Estimation, Alternative Specifications, OLS & IV

Table B1: OLS and IV Estimates of Cumulative Pollution Impact

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Order=4, Lag=90, OLS	8.63***	3.37	18.08***	5.92*	3.77	-0.01	-0.82	0.22
	(2.73)	(2.76)	(4.24)	(3.59)	(11.44)	(1.85)	(2.69)	(2.14)
Order=5, Lag=90, OLS	8.65***	3.40	18.10***	5.94*	3.81	0.00	-0.82	0.22
	(2.73)	(2.76)	(4.24)	(3.59)	(11.45)	(1.85)	(2.69)	(2.14)
Order=4, Lag=60, OLS	7.17***	2.69	15.12***	5.48**	2.61	-0.70	-2.56	-0.17
	(1.98)	(2.16)	(2.82)	(2.51)	(8.06)	(1.45)	(1.99)	(1.73)
Order=4, Lag=120, OLS	6.30*	1.62	15.30***	1.97	-2.70	-1.27	-0.36	-1.08
	(3.82)	(4.09)	(5.67)	(5.98)	(16.50)	(2.65)	(3.69)	(3.17)
Order=4, Lag=90, IV	27.17**	26.03*	23.59	23.81*	109.79**	5.43	0.48	-6.21
	(12.88)	(15.01)	(15.25)	(12.74)	(52.66)	(5.40)	(9.41)	(5.61)
Order=5, Lag=90, IV	28.42**	28.71**	23.38	25.56**	107.17**	6.88	0.24	-4.80
	(12.38)	(14.46)	(14.51)	(12.22)	(49.98)	(5.17)	(9.17)	(5.26)
Order=4, Lag=60, IV	27.61***	28.23**	25.34**	18.83**	97.95**	7.68	-2.20	2.13
	(10.26)	(12.10)	(11.44)	(8.70)	(45.18)	(4.81)	(8.41)	(4.62)
Order=4, Lag=120, IV	27.89	35.17	15.30	24.46	156.45	6.40	15.70	-8.60
	(20.49)	(23.93)	(24.84)	(20.79)	(98.35)	(8.84)	(14.31)	(9.28)

Notes: The dependent variable is log(number of transactions). The instrumental variable is predicted non-locally generated $PM_{2.5}$, from cities more than 150 km away, transformed by the same method as the regressor. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). Coefficients indicates cumulative effect on the number of transaction in percentage, from an exposure to 100 $\mu g/m^3$ more $PM_{2.5}$ for the corresponding length of time listed in the first column. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported for cluster-robustness.

C Additional Regression Results with the Value of Transactions

Table C2: OLS Estimates of Pollution Impact on Value of Transactions (Order=3, Lag=90)

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Current Day	0.14*	0.16*	0.24*	0.22**	-0.40	0.01	-0.02	0.00
	(0.08)	(0.09)	(0.13)	(0.11)	(0.30)	(0.11)	(0.14)	(0.15)
Current + Past 3d	0.52*	0.54	0.90*	0.74*	-1.44	-0.04	-0.14	-0.05
	(0.30)	(0.34)	(0.46)	(0.39)	(1.11)	(0.40)	(0.51)	(0.53)
Current + Past 7d	0.89*	0.89	1.62**	1.16*	-2.48	-0.23	-0.44	-0.25
	(0.53)	(0.60)	(0.81)	(0.65)	(2.07)	(0.69)	(0.89)	(0.92)
Current + Past 14d	1.30	1.14	2.64**	1.34	-3.47	-0.75	-1.19	-0.75
	(0.82)	(0.96)	(1.27)	(0.98)	(3.52)	(1.00)	(1.36)	(1.42)
Current + Past 28d	1.72	0.98	4.55**	0.73	-2.91	-1.58	-2.61	-1.51
	(1.22)	(1.43)	(2.07)	(1.63)	(5.92)	(1.46)	(2.05)	(2.21)
Current + Past 56d	3.31*	1.27	10.30***	0.64	3.04	0.87	-1.53	0.93
	(1.73)	(1.92)	(3.67)	(2.71)	(9.08)	(2.44)	(3.40)	(3.56)
Current + All Lags	4.92**	2.18	11.40**	-0.24	4.13	1.88	1.25	2.19
	(2.51)	(2.60)	(5.11)	(4.05)	(13.15)	(3.57)	(4.66)	(4.77)

Notes: The dependent variable is log(value of transactions) instead of log(number of transactions) used in the benchmark specification. The effect of air pollution for the past 90 days are estimated allowing polynomial decay with an order of 3. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). Coefficients indicate the cumulative effect on the number of transactions in percentage, from an exposure to 100 μg/m^3 more $PM_{2.5}$ for the corresponding length of time listed in the first column. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported for cluster-robustness.

Table C3: IV Estimates of Pollution Impacts on Value of Transactions (Order=3, Lag=90)

	Health-Related Consumption					Control Groups		
	Health	Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Current Day	0.43***	0.48***	-0.12	0.69***	-0.46	-0.44	-0.78**	-1.06***
	(0.16)	(0.18)	(0.24)	(0.24)	(0.61)	(0.32)	(0.35)	(0.38)
Current Day + Past 3d	1.57***	1.76***	-0.30	2.54***	-0.99	-1.54	-2.70**	-3.90***
	(0.60)	(0.68)	(0.89)	(0.84)	(2.27)	(1.09)	(1.21)	(1.34)
Current Day + Past 7d	2.81**	3.10**	-0.16	4.54***	-0.02	-2.54	-4.42**	-6.95***
	(1.15)	(1.30)	(1.63)	(1.45)	(4.27)	(1.78)	(2.02)	(2.25)
Current Day + Past 14d	4.47**	4.74**	1.09	7.01***	5.18	-3.16	-5.57**	-10.46***
	(2.10)	(2.38)	(2.80)	(2.22)	(7.76)	(2.42)	(2.84)	(3.21)
Current Day + Past 28d	6.99*	6.70	6.84	9.52***	22.51	-1.14	-3.55	-11.83**
	(4.13)	(4.69)	(5.22)	(3.61)	(15.66)	(3.27)	(3.98)	(4.72)
Current Day + Past 56d	13.93*	11.80	24.54***	12.60**	54.68*	9.15	3.99	-3.26
	(7.60)	(8.49)	(9.21)	(6.10)	(29.35)	(5.77)	(6.98)	(8.42)
Current Day + All Lags	20.10*	16.63	27.97*	17.76	92.87*	18.59**	15.67	-5.60
	(11.97)	(13.32)	(14.48)	(11.19)	(49.24)	(9.46)	(11.68)	(12.84)
First-Stage F	60.33	60.41	60.22	58.99	50.07	60.33	60.28	60.26

Notes: The dependent variable is log(value of transactions) instead of log(number of transactions) used in the benchmark specification. The effect of air pollution for the past 90 days are estimated allowing polynomial decay with an order of 3. The instruments are predicted non-locally generated $PM_{2.5}$, from cities more than 150 km away, transformed by the same method as the regressor. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). Coefficients indicate the cumulative effect on the number of transactions in percentage, from an exposure to 100 $\mu g/m^3$ more $PM_{2.5}$ for the corresponding length of time listed in the first column. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported for cluster-robustness.

D Almon Estimation, Controlling Other Pollutants

Table D4: IV Estimates of Pollution Impact, Controlling for O_3, SO_2, NO_2 and CO (Order=3, Lag=90)

	Health-Related Consumption					Control Groups		
	Health	All Hospital	Pharmacy	Renmin	Children's	Non-Health	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Current Day	0.42***	0.52***	0.03	0.73***	0.33	-0.47***	-0.68***	-0.72***
	(0.14)	(0.16)	(0.18)	(0.19)	(0.51)	(0.15)	(0.19)	(0.15)
Current Day + Past 3d	1.59***	1.95***	0.26	2.73***	1.87	-1.63***	-2.36***	-2.57***
	(0.51)	(0.61)	(0.65)	(0.68)	(1.99)	(0.51)	(0.67)	(0.52)
Current Day + Past 7d	2.99***	3.60***	0.84	4.98***	4.96	-2.63***	-3.88***	-4.36***
	(0.94)	(1.15)	(1.20)	(1.17)	(3.92)	(0.84)	(1.12)	(0.86)
Current Day + Past 14d	5.12***	5.91***	2.47	8.01***	12.42*	-3.17***	-5.01***	-5.96***
	(1.67)	(2.04)	(2.07)	(1.82)	(7.50)	(1.19)	(1.62)	(1.25)
Current Day + Past 28d	8.84***	9.29**	7.28*	11.67***	30.55**	-1.42	-3.94	-5.39***
	(3.26)	(3.95)	(3.96)	(3.08)	(15.41)	(1.76)	(2.49)	(2.04)
Current Day + Past 56d	16.61***	15.24**	17.30**	15.31***	59.39**	3.91	-1.18	-1.35
	(6.17)	(7.20)	(7.38)	(5.14)	(28.38)	(3.03)	(4.54)	(3.58)
Current Day + All Lags	22.29**	19.87*	18.40	16.40*	93.96**	3.66	-0.33	-8.80*
	(10.23)	(11.52)	(12.80)	(8.90)	(46.15)	(4.39)	(7.12)	(5.14)
First-stage F	59.82	59.93	59.70	58.75	49.42	59.82	59.78	59.76

Notes: The dependent variable is log(number of transactions). The effect of air pollution for the past 90 days are estimated allowing the polynomial decay with an order of 3. The instrument is predicted non-locally generated $PM_{2.5}$, from cities more than 150 km away, transformed by the same method as the regressor. The controls are city FEs, weekly FEs, city-specific time trend, city-specific seasonality, day-of-week FEs, special day dummies (holiday, spring festival, working weekend), current-day weather controls (mean temperature, rain dummy, mean wind speed). Additional controls are daily average concentration level of O_3, SO_2, NO_2 and CO. Coefficients indicate the cumulative effect on the number of transactions in percentage, from an exposure to 100 $\mu g/m^3$ more $PM_{2.5}$ for the corresponding length of time listed in the first column. Standard errors in parentheses, clustered at city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported for cluster-robustness.

E $PM_{2.5}$ Source Apportionment

Table E6: $PM_{2.5}$ Source Apportionment: Model for IV Construction

	Receptor			Hebei										
	Source	Beijing	Tianjin	Shijiazhuang	Tangshan	Qinhuangdao	Handan	Xingtai	Baoding	Zhangjiakou	Chengde	Cangzhou	Langfang	Hengshui
	Beijing	43.6	3.9	0.6	0.6	0.3	0.1	0.2	3.9	21.7	2.5	1.7	8.3	0.8
	Tianjin	8.4	42.2	0.2	6.2	1.9	0.0	0.1	1.3	2.6	3.3	4.1	18.4	0.8
Hebei	Shijiazhuang	0.8	0.8	63.4	0.1	0.0	1.9	4.5	5.7	2.2	0.1	1.6	0.9	6.7
Hebei	Tangshan	4.4	8.2	0.2	77.0	13.5	0.0	0.1	1.0	2.0	6.9	1.7	6.4	0.3
Hebei	Qinhuangdao	0.2	0.7	0.0	2.2	64.1	0.0	0.0	0.1	0.1	0.8	0.2	0.4	0.0
Hebei	Handan	0.2	0.3	5.0	0.0	0.0	47.4	29.9	1.2	0.2	0.0	0.8	0.2	4.7
Hebei	Xingtai	0.3	0.4	10.2	0.1	0.0	12.4	33.8	2.0	0.4	0.0	0.9	0.4	6.1
Hebei	Baoding	8.1	4.4	3.0	1.0	0.3	0.4	1.2	62.8	8.8	0.8	1.5	8.8	2.4
Hebei	Zhangjiakou	0.8	0.2	0.1	0.1	0.0	0.0	0.0	0.2	40.9	0.3	0.1	0.4	0.0
Hebei	Chengde	0.6	0.3	0.0	1.1	2.0	0.0	0.0	0.1	1.3	77.1	0.1	0.6	0.0
Hebei	Cangzhou	3.2	17.0	0.9	3.1	1.4	0.3	0.5	3.9	1.4	0.9	49.2	8.0	3.2
Hebei	Langfang	25.3	7.8	0.3	3.0	1.1	0.0	0.1	1.9	7.8	4.5	2.2	39.3	0.7
Hebei	Hengshui	1.4	3.4	3.5	0.6	0.2	1.4	2.6	7.6	1.0	0.2	9.9	2.5	44.2
	Shanxi	0.4	0.2	5.3	0.0	0.0	3.4	4.9	1.4	7.8	0.1	0.3	0.4	1.5
	Inner Mongolia	0.1	0.0	0.0	0.1	0.4	0.0	0.0	0.0	0.6	0.8	0.0	0.1	0.0
	Shandong	2.1	10.0	5.1	3.9	4.2	10.4	10.4	6.4	1.1	0.7	25.3	4.7	26.5
	Henan	0.1	0.1	2.2	0.0	0.0	22.1	11.7	0.5	0.1	0.0	0.3	0.1	1.8
	Others	0.1	0.3	0.0	0.9	10.5	0.2	0.2	0.0	0.0	1.1	0.2	0.2	0.2

Table E5: $PM_{2.5}$ Source Apportionment: CMAQ Model

Source		Receptor: Beijing	Tianjin	Hebei: Shijiazhuang	Tangshan	Qinhuangdao	Handan	Xingtai	Baoding	Zhangjiakou	Chengde	Cangzhou	Langfang	Hengshui
	Beijing	38.1	1.9	0.9	2.0	2.0	1.0	0.9	1.7	3.3	3.5	1.7	4.3	1.7
	Tianjin	8.7	47.5	1.2	7.7	3.8	1.3	1.2	2.6	1.4	7.4	3.8	20.4	2.4
Hebei	Shijiazhuang	1.9	1.3	60.4	1.5	1.6	2.4	3.2	3.5	3.8	2.1	1.7	1.6	3.8
Hebei	Tangshan	2.1	2.7	1.2	42.8	7.9	1.1	1.0	1.3	1.0	7.8	1.5	2.7	1.5
Hebei	Qinhuangdao	1.2	1.1	0.8	1.8	23.3	0.8	0.7	1.1	0.8	1.5	1.2	1.2	1.3
Hebei	Handan	2.4	1.8	4.0	2.1	2.1	46.1	17.9	2.9	2.9	2.4	2.3	2.0	2.9
Hebei	Xingtai	1.6	1.2	4.7	1.4	1.5	3.8	43.0	2.2	1.9	1.8	1.4	1.5	3.4
Hebei	Baoding	5.8	2.2	3.4	2.0	1.7	1.2	1.1	50.8	5.5	2.5	2.1	4.6	2.7
Hebei	Zhangjiakou	1.9	1.2	0.9	1.4	1.6	0.9	0.8	1.3	52.0	1.4	1.2	1.5	1.5
Hebei	Chengde	1.2	1.1	0.9	1.5	1.8	0.8	0.7	1.1	0.9	32.2	1.2	1.3	1.3
Hebei	Cangzhou	2.7	6.1	1.0	2.3	2.2	1.0	0.8	3.0	1.0	2.9	43.3	4.4	3.4
Hebei	Langfang	6.0	2.4	0.9	1.9	1.7	0.8	0.7	2.2	1.1	3.3	1.6	26.0	1.5
Hebei	Hengshui	2.2	1.8	1.5	1.9	2.0	1.2	1.1	3.7	1.4	2.6	2.4	2.4	30.9
	Shanxi	2.7	1.9	4.0	1.8	1.8	5.9	5.7	3.2	5.8	2.7	2.3	2.2	3.1
	Inner Mongolia	2.2	1.6	1.8	1.8	2.2	1.8	1.7	2.2	3.6	2.7	1.7	2.0	2.1
	Shandong	9.9	15.6	3.7	16.7	30.1	3.8	3.4	7.9	3.6	12.7	20.2	12.5	23.3
	Henan	3.2	2.9	3.5	2.3	2.2	20.1	10.4	3.4	3.8	3.1	4.0	2.8	5.8
	Others	6.3	5.9	5.1	7.1	10.5	6.1	5.6	5.9	6.3	7.5	6.5	6.5	7.7

Notes: The CMAQ model for Beijing-Tianjin-Hebei area is developed by Xing Chang, Ph.D. candidate at School of Environment, Tsinghua University (Beijing 100084, P. R. China, email: changxing92@sina.com).