NBER WORKING PAPER SERIES

THE MORBIDITY COST OF AIR POLLUTION:
EVIDENCE FROM CONSUMER SPENDING IN CHINA

Panle Jia Barwick
Shanjun Li
Deyu Rao
Nahim Bin Zahur

Working Paper 24688
http://www.nber.org/papers/w24688

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2018

The authors thank Yongmiao Hong, Solomon Hsiang, Matthew E. Kahn, Koichiro Ito, Doug Miller, Lucija Muehlenbachs, Matthew Turner, Shuang Zhang and seminar participants at Beijing University, Beihang University, Carnegie Mellon University, Chinese University of Hong Kong, Clemson University, Jinan University, MIT, National University of Singapore, Penn State, University of Arizona, University of Chicago, 2017 China India Insights Conference, the 10th Annual Conference on China's Economic Development at George Washington University, and 2018 NBER Environmental and Energy Economics Workshop for helpful comments. Partial financial support from CDOT79841-CTECH is acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Morbidity Cost of Air Pollution: Evidence from Consumer Spending in China
Panle Jia Barwick, Shanjun Li, Deyu Rao, and Nahim Bin Zahur
NBER Working Paper No. 24688
June 2018
JEL No. I15,Q51,Q53

ABSTRACT

Developing and fast-growing economies have some of the worse air pollution in the world, but there is a lack of systematic evidence on the health especially morbidity impact of air pollution in these countries. Based on the universe of credit and debit card transactions in China from 2013 to 2015, this paper provides to our knowledge the first analysis of the morbidity cost of PM2.5 for the entire population of a developing country. To address potential endogeneity in pollution exposure, we construct an instrumental variable by modeling the spatial spillovers of PM2.5 due to long-range transport. We propose a flexible distributed-lag model that incorporates the IV approach to capture the dynamic response to past pollution exposure. Our analysis shows that PM2.5 has a significant impact on healthcare spending in both the short and medium terms that survives an array of robustness checks. The annual reduction in national healthcare spending from complying with the World Health Organization's annual standard of 10 mg/m3 would amount to $42 billion, or nearly 7% of China's total healthcare spending in 2015. In contrast to the common perception that the morbidity impact is modest relative to the mortality impact, our estimated morbidity cost of air pollution is about two-thirds of the mortality cost from the recent literature.

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

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

Deyu Rao
Warren Hall
Dyson School of Applied Economics
and Management
Cornell University
dr526@cornell.edu

Nahim Bin Zahur
Cornell University
107 East State Street, Apt 510,
Commons West
Ithaca, NY 14850
nz96@cornell.edu

1 Introduction

The health impact of air pollution is an important component of the overall benefit of environmental regulations. A rich literature from epidemiology and more recently from 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 the establishment and improvement of air quality regulations. For example, research on the health impacts of particulate matter 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 has shown significant impacts of air pollution on mortality (Chay and Greenstone, 2003; Currie and Neidell, 2005; Currie and Walker, 2011; Knittel et al., 2015; Clay et al., 2016) and contemporaneous health (Neidell, 2004; Moretti and Neidell, 2011; Schlenker and Walker, 2015). The literature has mainly focused on mortality risk, in particular for infants, in the U.S. and Europe and lacks a commonly agreed method to measure the cost of morbidity (WHO, 2015). This gap in the literature is likely driven by the fact that while data on mortality is routinely collected, morbidity outcomes are harder to measure and collect on a large scale.[1]

Due to increased pressure from economic development and lax environmental regulations, developing countries and especially emerging economies such as China and India are currently experiencing the worst air pollution in the world. This is especially concerning given the size of the 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 quality of life, there is a lack of comprehensive data and rigorous studies on the benefit of pollution reduction in these countries. As a result, the dose-response relationships (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 these two gaps in the literature by estimating the morbidity cost of $PM_{2.5}$ in China. To do so, we combine hourly air pollution readings from all monitoring stations from 2013 to 2015 with the universe of credit and 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.[2] The causal impact

[1] Among the research commendations on the economic cost of pollution, Landrigan et al. (2018) argue that further research is needed to improve the cost estimate of morbidity, which is challenging due to its diverse endpoints.

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

of air pollution on out-of-pocket healthcare spending is one of the key components in consumer Willingness-to-pay (WTP) for improved air quality. The reliance on healthcare spending and health outcomes to directly bound WTP is in contrast to the revealed preference literature that relies on the implicit trade-off between risk factors and prices in product choices. The estimates based on these different approaches can be used for cross-validation purposes.

There are a couple of key empirical challenges in identifying the causal effect of air pollution on healthcare spending. The first challenge is the potential endogeneity in contemporaneous and lagged $PM_{2.5}$ that we use 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 varies across locations within a city. Ideally residents' pollution exposure should be measured by the population weighted local 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 that measurement errors are classical, they would attenuate the estimates toward zero.

To deal with this challenge, we construct instrumental variables by modeling the spatial spillovers of $PM_{2.5}$ due to the property of long-range transport of fine particles. Our IV approach is similar to the identification strategy used in Bayer et al. (2009), Williams and Phaneuf (2016), and Deryugina et al. (2017). The first two studies construct the IV based on air quality predictions from the EPA's source-receptor matrix using distant polluting facilities as inputs, while the latter uses changes in daily wind direction in a county as exogenous shocks to local air pollution.[3] Specifically, we use a parsimonious and transparent model of $PM_{2.5}$ concentration that allows us to disentangle the contribution of local and non-local sources. The model uses wind patterns and other weather conditions, lagged pollution levels in other cities, and geographic information as inputs to generate $PM_{2.5}$ predictions from both local and non-local sources. Our instruments can be considered as various weighted sums of lagged $PM_{2.5}$ levels in other cities where the weights are a function of the distance between the origin and destination city, wind direction and speed, and other weather conditions in these two 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 of the buffer zone in generating these instruments. Our results are robust to reasonable choices of the buffer zone and a host of robustness checks to control for unobservables and spatial correlations in economic activities as discussed in detail in Section 4.2.2. Our identification strategy is different from the regression discontinuity (RD)

[3]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.

approach based on the Huai River heating policy used in Chen et al. (2013), Ito and Zhang (2016), and Ebenstein et al. (2017). The RD design is better suited to study long-term impacts, such as that on mortality, by relying on the long-term cross-sectional variation in the data. This study focuses on the short- and medium-term impacts, and the IV approach allows us to leverage rich spatial and temporal variations in our data.

The second challenge in estimating the causal effect of pollution on healthcare spending arises from the nature of the high-frequency data at the daily level. On the one hand, the data environment allows us to characterize the dynamic impacts of past pollution exposure. On the other hand, daily pollution measures exhibit a high serial correlation. A direct OLS or IV estimation that includes many lagged terms leads to oscillating estimates that are imprecise. We propose a flexible distributed-lag model that extends the Almon technique (Almon, 1965) and uses finite-order B-splines (Corradi, 1977) to flexibly capture the effects of long lags. We combine this framework with the IV method to address endogeneity in contemporaneous and lagged air pollution measures.

Based on the OLS analysis of city-level daily healthcare spending with a rich set of temporal and location fixed effects, a temporary increase of 10 μg/m^3 in $PM_{2.5}$ concentration that lasts for a week is associated with an increase of 0.11% in the total number of hospital and pharmacy transactions. A permanent elevation of 10 μg/m^3 in $PM_{2.5}$ concentration would raise the number of healthcare transactions by 0.86%. The results from IV analysis indicate impacts that are three times as large: a temporary increase of 10 μg/m^3 in $PM_{2.5}$ would lead to a 0.65% increase in healthcare transactions, while a permanent increase of the same magnitude would lead to a 2.65% increase in the number of healthcare transactions. The impact of $PM_{2.5}$ differs across health facilities: spending in Children's hospitals is more than twice as responsive as spending in other types of health facilities. For non-healthcare spending, we find a negative impact of $PM_{2.5}$ in the short-term but no significant impact beyond two weeks. In addition, a projected worsening of air quality the next day increases the current day's spending in both health and non-healthcare categories. Taken together, these results provide evidence of avoidance behavior whereby consumers reduce outdoor activities (such as shopping) to mitigate pollution exposure.

The estimates of health impacts of $PM_{2.5}$ survive a variety of robustness checks including various parametric specifications of the medium-term impact, different buffer zones in constructing the IV, and the inclusion of other pollutants such as CO, SO_2, and average $PM_{2.5}$ in nearby cities. In monetary terms, a permanent reduction of 10 μg/m^3 in daily $PM_{2.5}$ would lead to total annual savings of 59.6 billion *yuan* ($9.2 billion in 2015 terms) in healthcare spending, implying a saving of $22.4 per household per year.[4] Bringing down China's $PM_{2.5}$ to the World Health Organization's (WHO) annual standard of 10 μg/m^3 could lead to savings exceeding $42 billion, nearly 7% of China's national healthcare spending or 0.4% of China's GDP in 2015.

[4]We use an exchange rate of 1$ = 6.5 *yuan* throughout this analysis.

How does the estimated morbidity cost from this study compare to the mortality cost estimates in the literature? Ebenstein et al. (2017) examine the mortality impact of PM_{10} in China for different age groups and find that a 10 unit increase in PM_{10} raises cardiorespiratory mortalities by 8% on average. The monetized mortality cost based on the Value of a Statistical Life (VSL) is \$13.4 billion from a 10 unit increase in PM_{10}. Our estimated morbidity cost is therefore about two-thirds of the mortality cost estimated from the literature. This comparison is similar to that in Deschenes et al. (2017) from reductions in NO_x emissions in the U.S. These findings contribute to a better understanding of the morbidity cost of air pollution and contrast with the common perception that morbidity is a minor part of the overall health impact of air pollution.[5]

Our analysis on healthcare spending provides a lower bound of consumer WTP for improved air quality, a key input in the cost-benefit analysis of environmental regulations. Through a simple theoretical framework, we show that consumer WTP for clean air includes several components, one of which is the impact of air pollution on healthcare spending. Our results suggest that the annual household WTP for improved air quality due to savings in out-of-pocket healthcare cost alone is \$11.3 for a 10 μg/m^3 reduction in $PM_{2.5}$. Taking into account the mortality cost estimated from Ebenstein et al. (2017), the WTP for a 10 μg/m^3 reduction in $PM_{2.5}$ would be \$40.6 per household, or \$16.6 billion annually for the whole country.

Our study makes four contributions to the literature. First, to our knowledge, this is the first comprehensive study that analyzes the effect of pollution on the healthcare spending of the entire population of a developing country. Our analysis is made possible by a unique and rich data set that is composed of the universe of credit card and debit card transactions in China from 2013 to 2015. There are 2.7 billion credit and debit cards that contribute to 34 trillion *yuan* of economic transactions annually. Besides covering fifty percent of private healthcare spending in China, this data set also includes spending in over three hundred non-healthcare categories.

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 and interpolate the mortality or morbidity benefit 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 air pollution levels, 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 that uses the same approach (Arceo et al., 2015; Chen et al., 2013; Greenstone and Hanna, 2014; He et al., 2016; Ebenstein et al., 2017). Different from other studies in this literature, which focus on mortality, the high-frequency nature of our data allows us to identify the short- and medium-term impacts on healthcare spending.

[5] EPA (2011) estimates that the mobility benefit from the Clean Air Act from 1990 to 2020 is about 8% of the mortality benefit. WHO (2015) applies an additional 10% of the overall mortality cost as an estimate for the morbidity cost.

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 housing and other consumer goods (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 a lack of awareness), the estimated WTP would be biased toward zero. Different from the revealed preference literature, this study uses realized healthcare spending data and contributes to the growing literature that estimates WTP for improved air quality using medical expenditures (Deschenes et al., 2017; Williams and Phaneuf, 2016; Deryugina et al., 2017). This approach does not rely on the informational assumption: the estimates are derived from the fact that elevated pollution leads to illnesses that are treated through healthcare spending. Whether or not consumers know about the underlying causes for their illnesses is irrelevant for our estimates. The disadvantage of this approach is that the components of WTP such as impacts on morbidity, mortality, labor productivity, and quality of life would need to be estimated from different data sources.

Fourth, the rich spatial and temporal variations in our data allow us to examine both the short- and medium-term impacts of air pollution on healthcare spending. The aforementioned studies using health insurance data all focus on the contemporaneous impact by using daily or quarterly data. We are interested in both the contemporaneous and future health consequences of pollution. However, as mentioned above, directly controlling for lagged daily measures leads to unstable estimates. Our flexible distributed-lag model with IVs is computationally light and has several advantages over existing methods such as VARs or local projection methods. It delivers a smooth impulse response function, allows researchers to estimate both the short-term and long-run effects, and can easily incorporate instrumental variables. To our knowledge, our study is the first analysis in the environmental literature that uses this technique to estimate the short- and medium-term impacts with high frequency data.

The rest of the paper is organized as follows. Section 2 describes the data and the air pollution challenges facing China. Section 3 provides a stylized model to illustrate that the estimated impact on healthcare spending can be used as a 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.

2 Data

Our analysis is based on three comprehensive, nation-wide, micro-level datasets of air pollution, consumer spending by category, and meteorology conditions from 2013 to 2015, aggregated to daily and city-level. These datasets enable us to evaluate the impact of air pollution on spending in both the short- and medium-term, as well as heterogeneous impacts across pollution levels.

2.1 Air Pollution

For nearly four decades, China has maintained its GDP growth at an annual rate of nearly 10% and 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 world's largest energy consumer, accounting for roughly a quarter of world total energy consumption and half of world coal consumption.

Fast economic growth and rising energy consumption have put enormous pressure on the environment, with air, water, and soil pollution becoming serious challenges that adversely affect human health, ecosystems, and the quality of life.[6] Improving air quality has become an important policy goal for the central government, which extensively revised the Environmental Protection Law in 2014 and defined goals of pollution abatement in both the 12th (2011 - 2015) and 13th (2016 - 2020) five-year plans.

Fine-scale air quality data at monitoring stations 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 increased steadily from 1003 stations in 159 cities in 2013 to 1582 stations in 367 cities in 2015. We calculate the daily concentration of $PM_{2.5}$ and other pollutants at the city level by averaging data across monitoring stations within a city.

Figure 1 plots the three-year average of $PM_{2.5}$ from 2013 to 2015 across cities. The nationwide average during this period is 56 μg/m^3 (with a standard error of 46 μg/m^3), which is much higher than the annual standard of 12 μg/m^3 that is set by the U.S. Environmental Protection Agency and the standard of 35 μg/m^3 by the China MEP.[7] Notably, there is considerable regional disparity. Cities in northern and central China with a high concentration of manufacturing industries suffer from the most severe pollution, with many of them experiencing a three-year average $PM_{2.5}$ concentration of 90 μg/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 noticeable improvement in air quality as a result of shutting down or relocating polluting industries and

[6]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 China's GDP in 2003.

[7]The EPA's daily standard is 35 μg/m^3 and annual standard is 12 μg/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 μg/m^3 and annual standard is 35 μg/m^3.

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). In all regions of the country, the daily $PM_{2.5}$ concentration is higher than 35 μg/m^3, the official MEP standard, for most days. The northern regions have more pronounced peaks in the winter than the southern region, largely because of the coal-fired central heating systems north of the 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

The second main database for our analysis is the universe of credit and debit card (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. It is the largest network in the world in terms of both the number and value of transactions, ahead of Visa and Mastercard. There were 2.7 billion cards in use from 2013 to 2015 with transactions covering over 300 merchant categories.[8] The database includes 34 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. 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 that are similar to the Medicare in the US. Medical expenses that are covered by the Chinese government programs are often directly billed on medicare cards, most of which are settled through 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 by filing claims. If consumers pay for these expenses via their bank cards, then these transactions will be included in our database.[9]

Our data account for 31% of aggregate private healthcare spending in 2013, and as card penetra-

[8]There are seven major categories and 300 subcategories. The major categories are: retail; wholesale; direct sales; real estate and finance; residential and commercial service; hotel, restaurant, and entertainment; and education, health, and government service. Merchants are classified by these categories.

[9]The healthcare system and the insurance market in China have been improving with significant government support. In 2009-2011, China's central government provided 850 billion *yuan* to overhaul its healthcare system and increase the basic health insurance coverage. In 2011, the insurance coverage through three major government supported insurance programs reached nearly 95% from 65% in 2009 (Yu, 2015).

tion grew, the coverage rose 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 the late 2000s that consumers began to adopt bank cards). According to official statistics from the Central Bank of China (2015), bank card transactions accounted for 48% of overall spending in 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).

Figure 3 shows the spatial pattern of card adoption by plotting the number of active cards per registered resident by city in 2015. We assign each card to one primary city based on the location of its most frequent usage. Card adoption is higher in coastal or high-income cities. Table C1 in Appendix C correlates cross-sectional card adoption with city demographics. It shows that cities with a higher household income and education and a younger population are associated with higher adoption.

Despite the richness and uniqueness of the credit and debit card transactions, they cover only about half of private healthcare spending in 2015. In order to interpret our health impacts as the population impacts, we need to assume that the health impacts are not correlated with the method of payment. To the extent that the elderly are more vulnerable to air pollution but less likely to use credit and debit cards, our estimates provide a lower bound of the population impacts, though elderly Chinese tend to be cared for by their children who likely accompany them to hospital visits and pay the bill. In addition, low-income residents might have a lower baseline health status. If this implies that air pollution has a more severe health impact on them, then our analysis would underestimate the population impacts.

Healthcare spending includes transactions at hospitals, pharmacies, and other healthcare facilities (e.g. small health clinics). In 2015, hospitals account for 83.5% of healthcare spending in our data, and 56.8% of transactions. Different from pharmacies in the U.S., such as CVS or Walgreens, 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 People's hospitals and Children's hospitals from other hospitals. People's hospitals are state-owned general hospitals and tend to be the largest health care facilities in a city. Each city has at least one People's hospital. Children's hospitals accept mostly child patients. Birth centers and infant health centers are grouped into Children's hospitals. People's and Children's hospitals account for 24.1% and 4.2% of total healthcare spending respectively, and 26.2% and 9.0% of the total number of transactions in 2015.[10]

In addition to healthcare spending, we also analyze spending in non-healthcare categories, such

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

as daily necessities. We closely follow the United Nations' Classification of Individual Consumption According to Purpose (COICOP) in defining necessity goods.[11] Relative to healthcare spending, spending on daily necessities is three times as large and transactions three times as frequent. A unique feature of Chinese consumers' shopping behavior is their frequent trips to supermarkets for groceries (often on a daily basis) . We therefore use supermarket spending as another proxy for daily consumption, in addition to spending on necessities.[12] Spending in supermarkets is over four times as large as healthcare spending in value and five times as frequent in 2015.

To illustrate inter-temporal spending patterns, Figure 4 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.[13] 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.[14] Daily wind direction is calculated by adding up twenty-four hourly vectors of wind direction, 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, with the inter-quartile range being from 27 to 69 μg/m^3 and the maximum being 985 μg/m^3 . 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. For healthcare spending, the average daily number of transactions is 7,229 per city, and the average daily spending is 6.7 million *yuan*.

[11] United Nations' 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.

[12] We exclude supermarkets from necessity spending because they sell a large variety of goods other than necessities.

[13] These stations cover most major Chinese cities from as early as the 1940s to the present.

[14] 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 use this measure as our daily precipitation.

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 in 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 healthcare 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. (2017) and Williams and Phaneuf (2016) show that the marginal effect of air pollution exposure on total healthcare spending is one of the components of consumers' WTP for improved air quality, with other components being the mortality impact, the loss of productivity, and quality of life. In addition, the defense spending such as purchases of air purifiers or face masks and avoidance behavior such as staying indoors should also be accounted for. While the literature has largely neglected the role of avoidance behavior and reduction in 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 healthcare spending (m_i), non-healthcare offline spending (c_i), and non-healthcare online spending (o_i), subject to his budget constraint. 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 air pollution level a (which is exogenous to consumer i's spending) and spending activities $m_i + c_i$, but is not affected by online spending o_i.[15]

Consumer i has an endowed health stock h_0, which evolves as a result of exposure to air pollution and his own healthcare spending that mitigates the negative consequences of pollution. Individuals differ in how sick they become when exposed. This is captured by $g_i(e_i)$, where $g_i \sim F_i$ is a non-decreasing function that is individual-specific and represents how much the individual's health

[15] We combine all non-healthcare 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.

stock changes with respect to e_i.[16] Thus the health stock equation can be written as:

$$h_i = h_0 + m_i - g_i(e_i)$$

Consumers have health insurance, with π denoting the premium and p the proportion of healthcare spending that needs to be paid out-of-pocket.[17] Thus, if the consumer undergoes hospital treatments that cost 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:

$$y(h_i) \equiv y_0 + w(h_i) = \pi + pm_i + c_i + o_i$$

Consumer utility $U(h_i, c_i, o_i, e_i)$ depends on health stock (h_i), offline consumption (c_i), online spending (o_i), and pollution exposure (e_i). We allow utility to be both directly and indirectly affected by pollution exposure. The indirect effect comes through reduction in health stock which could capture the mortality impact. The direct mechanism arises because consumers value quality of life, which decreases with air pollution. Heavy haze and smoky air reduce consumers' utility even if their health stock is restored (i.e. held constant). For example, Levinson (2012) finds that people report lower levels of happiness on days with worse local air pollution.

Consumer i chooses spending to maximize utility, subject to his budget constraint and the rule of health stock evolution:

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

Our specification of pollution exposure $e(a, m_i + c_i)$ makes it explicit that all offline spending, whether health-related or not, affects pollution exposure because it involves time spent outdoors.[18] This is the key difference between our model and those in Deschenes et al. (2017) and Williams and Phaneuf (2016). In addition, our model incorporates utility from health (for example, through

[16] An example is $g_i(e_i) = \alpha_i e_i$, 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.

[17] 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).

[18] 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.

morbidity) 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:

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

The first-order conditions are:

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

where U_h, U_e, U_c, U_o are partial derivatives of the utility function with respect to health stock, pollution, consumption, and online spending, respectively. We assume $U_h > 0, U_c > 0, U_o > 0, U_e < 0$, since health and consumption are desirable but pollution is not. As consumers are exposed to air pollution whether buying food or seeing a doctor, we set $e_m = e_c > 0$ to be 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 of pollution exposure from hospital visits.[19] Exposure increases with pollution ($e_a > 0$). Finally, y_h is the effect of health on income and is assumed to be positive.

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:

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

As shown in Appendix A, individual i' marginal WTP can be expressed as:

$$MWTP_i = p\frac{\partial m_i^*}{\partial a} + y_h(-\frac{dh_i^*}{da}) + \frac{U_h}{\lambda_i}(-\frac{dh_i^*}{da}) + (-\frac{U_e}{\lambda_i})\frac{de_i^*}{da} + \frac{U_c - U_o}{\lambda_i}(-\frac{\partial c_i^*}{\partial a}) \tag{1}$$

Equation (1) illustrates the relationship between the impact of air pollution on healthcare spending (the morbidity effect), given by $\frac{\partial m_i^*}{\partial a}$, and MWTP for improved air quality. Changes in an

[19]Optimal healthcare spending is 0 if $1 - e_m < 0$.

individual's out-of-pocket healthcare spending is one of the determinants of his MWTP. 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 a 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 which could capture the mortality impacts. 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. The last term $\frac{U_c - U_o}{\lambda_i}(-\frac{\partial c_i^*}{\partial a})$ denotes reduction in monetized utility due to the sub-optimal level of consumption caused by pollution exposure (e.g., avoidance behavior). As shown in Appendix A, these four terms are all positive under fairly weak assumptions.

Our model encompasses that of Deschenes et al. (2017), which abstracts away the exposure associated with consumption ($e_c = 0$), as well as the utility loss of reduced quality of life ($U_e = 0$).[20] When $e_o = U_e = 0$, the FOCs indicates $U_c = U_o = \lambda$, and

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

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 our expression for the marginal willingness-to-pay collapses to theirs:[21]

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

To summarize, the utility maximization framework illustrates that consumer WTP for clean air can be estimated by adding up different components of the impact of the air pollution on population health and behavior. In the empirical analysis, we focus on quantifying the impact of air pollution on healthcare spending ($\frac{\partial m_i^*}{\partial a}$), and use changes in non-healthcare spending ($\frac{\partial c_i^*}{\partial a}$) to assess the importance of avoidance behavior. Then we contrast the morbidity component based on our parameter estimates with the mortality component based on the recent literature for China.

[20] In Deschenes et al. (2017), $MWTP = w\frac{ds}{dc} + p_a\frac{\partial a}{dc} - \frac{U_s}{\lambda}\frac{ds}{dc}$, 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 air pollution a in our framework.)

[21] 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 the equilibrium insurance premiums will adjust in response to expected pollution. 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 the large variance in pollution across cities, we find it more reasonable to assume that $\frac{\partial \pi}{\partial a} = 0$.

4 Empirical Framework

In this section, we first present a flexible econometric model that allows us to estimate the short- and medium-term impacts of air pollution on healthcare spending. Then we discuss our estimation strategy and the construction of instrumental variables.

4.1 Flexible Distributed-Lag Model

Air pollution has both short- and long-term consequences on healthcare spending. Different from quarterly or annual data commonly used in the literature, our daily level data allow us to characterize the path of health impacts from both contemporaneous and past air pollution exposure. We use the following distributed lag model (DL) to capture this relationship:

$$y_{ct} = \sum_{i=0}^{k} \beta_i p_{c,t-i} + \mathbf{x}_{ct}\alpha + \kappa_c t + \xi_c + \eta_w + \varepsilon_{ct} \tag{2}$$

where y_{ct} is daily healthcare 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 longer-term causal impacts of pollution exposure on healthcare 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 variables (three important issues that we will return to in the next section). Then the DL model can be estimated using OLS. But the linear estimation with a large number of lags is undesirable due to the high serial correlation among the lag terms $p_{c,t-i}$. The parameter estimates tend to be imprecise with artificial oscillations. To reduce the number of parameters that need to be estimated while allowing for flexible and smooth longer-term impacts, we follow Almon (1965) and Corradi (1977) and specify β_i's as cubic B-spline functions of time with z segments, where z is a constant chosen by econometricians.[22] The intuition is that 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 basis splines. Take $z = 1$ as an example, in which case the B-splines amount to a simple 3rd order polynomial:

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

[22] Almon (1965) first proposed approximating the lag coefficients with polynomial functions. Poirier (1975), Corradi and Gambetta (1976) and Corradi (1977) suggested using spline functions, which impose weaker restrictions on the lag coefficients than polynomials while maintaining the ability to estimate the model using a relatively small number of parameters.

where the contemporaneous effect of pollution on spending is captured by γ_0, the effect of yesterday's pollution is $\beta_1 = \gamma_0 + \gamma_1 + \gamma_2 + \gamma_3$, while the effect of pollution from i days' in the past is $\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2 + \gamma_3 i^3$. Appendix B describes how to extend this to the more general case where there are multiple segments and the coefficients β_i are piecewise polynomials in i.

Plug (3) into (2) and rearrange terms, we have:

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

With this reformulation, we only need to estimate four coefficients γ's rather than $k+1$ coefficients (the number of lags plus current day). The four key regressors are:

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

where the first term is the sum of past pollution exposure, and the others are a 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 (4) 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 are either 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 healthcare spending, then $\beta_{-1} = 0$. If pollution exposure prior to k lags should not affect current healthcare spending, then $\beta_{k+\tau} = 0, \forall \tau \in \mathbb{N}$ and $\tau > 0$. These assumptions can be imposed individually or jointly as constraints in the estimation and 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 an arbitrary correlation between the contemporaneous error term and past error terms, which is difficult in geometric decay models.

Once we choose the number of lags k, the order of B-spline polynomials q, the number of segments z, and additional restrictions on γ's, the estimation can be carried out in (constrained) OLS and β's can be recovered from the parameter estimates.

4.2 Identification

4.2.1 Sources of Endogeneity

There are multiple sources of potential endogeneity in the key variable of interest, pollution exposure. As is common in the literature on estimating the health impact of air pollution, our measure of pollution exposure likely suffers from measurement errors. This arises from the fact that pollution levels vary across locations within a city and that we average the pollution data from monitoring stations to the city level. For example, among the 9 monitoring stations in the urban core of Beijing, the average difference between the maximum and minimum pollution level in a day is about 35 $\mu g/m^3$ in 2014 while the daily average at the city level is 87 $\mu g/m^3$. Since population is not evenly distributed within a city and the spatial distribution 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 the population-weighted average of local 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) and the fact that many monitoring stations are located outside of population centers. In addition, our daily pollution measure is a simple average of hourly measurements and abstracts away the temporal variation. To the extent that the measurement errors are classical, our OLS estimates would suffer from the attenuation bias.[23]

Second, pollution exposure is potentially endogenous due to avoidance behavior in both the short- and longer-term. Chinese consumers now have a high awareness of air quality and its impact on health. $PM_{2.5}$ readings are becoming readily accessible through cell phone apps and from government websites in recent years.[24] During days of severe air pollution, consumers may reduce 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 healthcare spending and ren-

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

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

der the pollution measure endogenous. In the long run, 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 polluted cities. As a result, air pollution could be correlated with the error term (such as the health stock of local residents).

In a short or medium time frame, such as the one used in our analysis, location-specific time trend help 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 directly responds to air pollution (and hence will be correlated with shocks that affect air pollution). We use spending on daily necessities and at supermarkets to quantify avoidance behavior. Our results indicate that avoidance behavior reduces spending in the short term (i.e., up to two weeks) through inter-temporal substitutions, but there is no significant aggregate impact over a longer period (a month or longer).

Another source of endogeneity in pollution measures 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 healthcare spending, such as income shocks, could be correlated with economic activities and thus with air quality. Temporary local shocks, such as major sport and political events, could affect both the air pollution level and healthcare spending (and consumer activities in general). These unobservables that are not absorbed by our location and trend/seasonality interactions could render the air quality variable endogenous.

4.2.2 IV Construction

To address the concern of endogeneity, we exploit the spatial spillovers of $PM_{2.5}$ due to its long-range transportability to construct instruments. $PM_{2.5}$ particles are light, can travel at a speed of 10 mph, and often reside in the atmosphere for 3-4 days. Their region of influence is determined by wind speed and direction. Based on atmospheric modeling, Zhang et al. (2015) document significant regional pollutant 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 other cities could serve as exogenous shocks to the pollution level 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 parsimonious model to predict the air pollution level of a given city based on lagged pollution levels in the same city and other cities, wind patterns (direction and speed), other weather conditions (precipitation and temperature), and distances between city pairs.[25] This model allows us to estimate the contribution to the $PM_{2.5}$ level in a given city from non-local sources. We construct a buffer zone to minimize the correlation in unobserved regional economic shocks and only use pollution from cities outside of the buffer zone to construct the instruments.

Our identification assumption is that pollution shocks (e.g., economic activities) in regions outside of the buffer zone are uncorrelated with local shocks to spending. This assumption would be violated if spending shocks (e.g., high temperature that leads to more hospital visits as well as increased demand for electricity) in a given city affect production activities in other cities (e.g., electricity generation) outside of the buffer zone, which in turn affect the pollution level in those cities. We address this concern in four ways. First, we test the robustness of our results to the buffer-zone radius in section 5 and show that the results are robust to different radii. Second, our instruments are weighted sums of *lagged* pollution levels in other cities, with the weights being a function of wind patterns and other weather conditions as well as the distance between cities. To the extent that economic shocks in a given city affect production and hence pollution in other cities, this should induce correlation between the error term and future pollution levels rather than lagged pollution levels in other cities. In addition, the exogenous variation in wind speed and direction should reduce such correlations. Third, in one of the robustness analysis, we include the average $PM_{2.5}$ in other cities outside of the buffer zone but within the same region as an additional regressor to control for regional spillovers in economic activities. The parameter estimates on local $PM_{2.5}$ levels are very similar to those in the benchmark analysis. Fourth, we construct an alternative IV based on the weather variables and the *average* $PM_{2.5}$ in other cities. The within-city variation of this IV is solely driven by the wind pattern and other weather conditions (rather than time-varying pollution levels in other cities), hence should not be correlated with unobserved economic shocks. The results from this specification are very similar to the benchmark results.

In principle, our identification assumption implies that *any* function of pollution and weather conditions in cities outside the buffer zone is a valid instrument for pollution in city i. The set of such instruments, however, is very large and includes many weak instruments. We use a simple model of air pollution transmission to guide our construction of instrumental variables.

Denote 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:

$$p_{it} = \theta_1 p_{i,t-1} + \sum_{j \neq i} p^{+}_{j \to i,\, t} + \mu_{it}, \tag{5}$$

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

where θ_1 captures the amount of pollution that is carried over from the previous day, which can be affected by local meteorological conditions. $p^+_{j\to i,\ t}$ denotes the amount of $PM_{2.5}$ pollutants in city i at time t that is originated from city j. μ_{it} is the error term. The contribution of non-local sources to the pollution level of a given city could be affected by a host of weather and topography conditions and is the subject of sophisticated air quality modeling. We use the following parsimonious model to capture the key feature that $PM_{2.5}$ pollutants dissipate over time and across space as they move:

$$p^+_{j\to i,\ t+s_{ijt}} = \begin{cases} \cos\Phi\ p_{jt}\ f(d_{ij}, w_{jt}, w_{i,t+s_{ijt}}), & \text{if } \cos\Phi > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

$p^+_{j\to i,\ t+s_{ijt}}$ is the amount of pollution that enters city i on day $t+s_{ijt}$, having originated from city j on day t. Φ denotes the angle between the wind direction and the direction from city j to city i. We invoke a simple vector decomposition and assume that the amount of pollutants carried toward city i from city j is $\cos(\Phi)p_{jt}$ at speed $\cos(\Phi)S_{jt}$, where S_{jt} is the wind speed in city j. Pollution decays over time as it travels and only part of the pollution from city j enters the atmosphere of city i. This is represented by $f(d_{ij}, w_{jt}, w_{i,t+s_{ijt}}) \in [0,1]$, a function of the distance between the two cities (d_{ij}), weather conditions in the source city when the pollution is generated (w_{jt}), and weather conditions in the destination city when the pollution enters its atmosphere ($w_{i,t+s_{ijt}}$). The number of days it takes pollutants to travel from city j to city i, s_{ijt}, is calculated as the following and rounded to the next smallest integer:

$$s_{ijt} = \left\lfloor \frac{d_{ij}}{\cos(\Phi)S_{jt}} \right\rfloor .$$

Figure 5a shows the wind-pollution vectors from over 300 cities on Dec. 5, 2013 (denoted as Day 0). Each arrow's length indicates the wind speed, rescaled to match the exact distance the arrow can travel in a day. The arrow width indicates the level of $PM_{2.5}$ concentration at the source city. To illustrate how we predict city-day $PM_{2.5}$, Figure 5b shows all subvectors of pollutants that are blown towards Beijing on the same day. The pollution level of the receptor city, Beijing in this example, is predicted by pollutants carried through the subvectors that reach Beijing at time t, together with the lagged local pollution levels, as stated in equation (5).

The decay function $f(d_{ij}, w_{jt}, w_{i,t+s_{ijt}})$ in equation (6) is unknown. We approximate it by polynomial functions $u_l(1/d_{ij}, w_{jt}, w_{i,t+s_{ijt}})$:

$$p^+_{j\to i,\ t+s_{ijt}} = max(0, \cos\Phi) p_{jt} \sum_l \gamma_l u_l(1/d_{ij}, w_{jt}, w_{i,t+s_{ijt}}) \tag{7}$$

We now describe how to construct instruments using the above model . Let r denote the radius of the buffer zone. For most of our results we assume a buffer zone of 150 km, but we also check that results are robust to different values of r. The total amount of pollution imported from cities outside of the buffer zone, $\hat{p}_{it}^{far}$, is the following (interchanging the summation signs in the third equation):

$$
\begin{aligned}
\hat{p}_{it}^{far} &= \sum_{j:\, d_{ij}>r} p_{j\to i,\ t}^{+} \\
&= \sum_{j:\, d_{ij}>r} max(0, \cos\Phi) p_{j,t-s_{ijt}} \sum_{l} \gamma_l u_l(1/d_{ij}, w_{j,t-s_{ijt}}, w_{i,t}) \\
&= \sum_{l} \gamma_l \sum_{j:\, d_{ij}>r} max(0, \cos\Phi) p_{j,t-s_{ijt}} u_l(1/d_{ij}, w_{j,t-s_{ijt}}, w_{i,t}) \\
&= \sum_{l} \gamma_l Z_{it}^{l}
\end{aligned}
$$

where $Z_{it}^{l} = \sum_{j:\, d_{ij}>r} max(0, \cos\Phi) p_{j,t-s_{ijt}} u_l(1/d_{ij}, w_{j,t-s_{ijt}}, w_{i,t})$.

A natural strategy is to use Z_{it}^{l}, $l = 1, ..., L$ as instruments for p_{it}. These are valid instruments since they depend only on weather within city i, which we control for in our regressions, and on pollution and weather variables in cities outside of the buffer zone, which are uncorrelated with local shocks to spending by our identification assumption. We assume second-order polynomials and include wind speed, precipitation, and temperature as the weather variables in $u_l(.)$, leading to 15 instruments in Z_l. An alternative approach is to estimate the unknown parameters γ_l in equation (7), construct $\hat{p}_{it}^{far}$, and then use $\hat{p}_{it}^{far}$ as an instrument for p_{it}. The benefit of our approach of using Z_l directly as instruments is that we avoid having to make functional form assumptions in order to estimate γ_l.[26]

Notice that although we do not estimate the air pollution transmission model directly, we exploit a number of model restrictions to construct more powerful IVs. For example, if the prevailing wind conditions are such that it takes a couple of days for pollution generated in city j to reach city i, we would expect $p_{j,t}$ to affect $p_{i,t+2}$ instead of $p_{i,t}$. Our IVs take into account these considerations and thus out-perform naive approaches such as the sum of pollution levels in all cities outside the buffer zone. As documented below, the first-stage cluster-robust F-stats of excluded instruments vary from 38 to 62 across specifications, indicating a strong predictive power of the endogenous variables. It is important to note that the goal of our first-stage model is not to maximize the accuracy of air quality predictions. Instead, we want to create instruments that are both predictive of local air pollution and at the same time exogenous to shocks to healthcare spending. This is why we base our analysis on a relatively conservative definition of the buffer zone and exclude $PM_{2.5}$ from cities within 150 km (although our results are robust to the choice of buffer radius).

[26] As a robustness check, we tried the alternative approach of estimating γ_l and constructing $\hat{p}_{it}^{far}$ as the IV. We also constructed $\hat{p}_{it}^{far}$ using an exponential decay function to proxy for $f(d_{ij}, w_{jt}, w_{i,t+s_{ijt}})$. The results are similar to what we report in the paper, though the first-stage is slightly weaker.

5 Empirical Results

5.1 Short-Term Impacts

Our empirical analysis begins with the contemporaneous effect of air pollution on health. In the discussion below, we use the logarithm of the number of transactions as the dependent variable rather than the value of transactions as in the literature using similar transaction-level purchase data (Einav et al., 2014). The distribution of healthcare spending is right-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 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 credit/debit card adoption and seasonal diseases. We also add fixed effects for state holidays, working weekend,[27] day of the week, as well as weather variables to control for their direct effects on spending. For example, people may reduce non-urgent hospital visits during holidays or on days with bad weather. All standard errors are clustered at the city level.

Table 2 summarizes the short-term impacts estimated with OLS. A 10 μg/m^3 increase in the daily $PM_{2.5}$ concentration is associated with a 0.11% increase in the total number of healthcare transactions. Transactions in Children's hospitals are more sensitive to air pollution, with an impact of 0.18% from a 10 μg/m^3 increase in $PM_{2.5}$. This is intuitive since children are more vulnerable to air pollution. 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 is avoidance behavior: consumers postpone or reduce shopping trips in response to poor air quality to reduce pollution exposure. We test these two possibilities in Section 5.5.

To graphically illustrate the relationship between pollution and spending, we plot the log number of transactions against $PM_{2.5}$ in Figure 6. All other controls (weather, city trend, etc.) are partialled out, so the figure 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 healthcare transactions (top left), we also plot the relationship separately for pharmacies, People's hospitals, Children's

[27]In mainland China, weekends near multi-day holidays are usually swapped with weekdays next to actual holidays (if possible) to create a longer holiday period. As a result, businesses and schools would treat that weekend as a *working weekend*.

hospitals, and two non-healthcare categories (necessities and supermarkets). $PM_{2.5}$ has a positive relationship with spending in all health categories across nearly all quantiles of $PM_{2.5}$. The data points 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 variables constructed from pollutants outside of the 150 km buffer zone as discussed in Section 4.2. Table 3 reports IV results. The first-stage cluster-robust F-statistics on the instruments (reported in the last row of the table) vary from 52 to 62, suggesting a strong correlation between the instrument and the endogenous variable. The IV estimates are considerably larger than the OLS estimates, with most coefficients 3 to 7 times as large as their OLS counterparts. A 10 μg/m^3 increase in $PM_{2.5}$ in a day is associated with a 0.65% contemporaneous 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 nearly twice as large as that for 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 health spending is consistent with the attenuation bias due to (classical) measurement errors 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, which reduce healthcare spending but increase non-healthcare spending (more outdoor activities and fewer hospital visits).

As discussed below in more detail in Section 5.6, our estimated short-term impact includes both the direct positive effect on healthcare spending and the indirect negative effect through avoidance behavior. Therefore, the direct effect of air pollution on healthcare spending is likely larger.

5.2 Longer-Term Impacts

Exposure to $PM_{2.5}$ could have dynamic longer-term health impacts that are unlikely to be linear. Directly estimating the coefficients of a large number of lagged $PM_{2.5}$ in equation (2) suffers from high serial correlation and imprecise estimates. Instead, we employ the flexible Distributed-Lag model discussed in Section 4.1 and allow pollution impacts to follow a smooth path of decay.

Table 4 reports the cumulative effects for different time periods across categories from the OLS regressions. Our benchmark specification incorporates 90 lags (daily pollution exposure for the previous three months) and three segments for the cubic B-splines, where the decaying pattern for each month is characterized by separate cubic polynomials. Effects beyond 90 days are modest and often imprecisely estimated. The standard errors are clustered at the city level and are reported in parentheses. We examine the robustness of our results to the choice of lags and B-spline segments in section 5.4.

The first column of Table 4 shows that a temporary surge of 10 μg/m^3 in $PM_{2.5}$ concentration increases today's transactions in all healthcare facilities by 0.03%. A permanent elevation of 10 μg/m^3 raises the number of transactions by 0.86%, eight times as large as the effect reported in Table 2 when only the contemporary $PM_{2.5}$ concentration is controlled. The last two columns report a statistically significant negative impact on necessities and supermarket spending within two weeks, but not in the long run.

To deal with the endogeneity in $PM_{2.5}$, we use the instruments discussed in Section 4.2. Specifically, we instrument for the local pollution on day s, p_{is}, using the instruments Z^l_{cs}, which are functions of pollution in faraway sources that reach city i on day s. The contemporary and cumulative effects across different time spans are presented in Table 5.

Several important findings emerge from Table 5. First, the estimated longer-term impacts of $PM_{2.5}$ on healthcare spending across all categories from 2SLS are positive and much larger than their OLS counterparts, consistent with the comparison for the short-term impact discussed in Section 5.1. Specifically, a permanent increase of 10 μg/m^3 in $PM_{2.5}$ concentration raises the number of transactions in the health sector by 2.65%. Second, the impact on Children's hospitals is the largest and more than twice as large as the impact on aggregate healthcare spending, consistent with the fact that children are among the most vulnerable groups. Pharmacy is the second most responsive category among the four healthcare categories. When elevated air pollution aggravates symptoms for people with respiratory problems, they may go to pharmacies to purchase drugs without visiting hospitals.[28] Third, the effects on daily necessities and supermarket spending are all negative and appear to be short-lived.

To examine how the impact on spending changes over time, Figure 7 plots the estimates of both current and past 90 days of pollution exposures for different categories.[29] The (dotted) solid part of each line indicates the impact being statistically (in)significant. There are several noticeable patterns. First, $PM_{2.5}$ has a positive impact on healthcare 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 spending on necessities and in supermarkets in day zero, but the effect disappears after two weeks. This temporal reduction is inconsistent with the budget constraint hypothesis, since a permanent increase in healthcare spending would lead to a permanent reduction in necessities and supermarkets with a fixed budget. Instead, our result lends

[28]The Ministry of Human Resources and Social Security maintains the National Reimbursement Drug List (NRDL). Only the drugs on this list are covered by China's national medical insurance programs, some in full (type A drugs) and others partially (type B).

[29]The optimal number of lags should in theory differ across categories. For example, the effect of pollution on non-healthcare categories appears to be short-lived, while for children's hospitals it could last for more than 3 months. To keep the results comparable, we impose the same lag structure on all categories.

support to the hypothesis of avoidance behavior. We return to this issue in Section 5.5.

Our results so far suggest that a 10 μg/m^3 increase in $PM_{2.5}$ would lead to an increase in the number of health-related transactions by 2.65% in the long term. In terms of the value of transactions, the effect is 1.5% (Table C2 in Appendix C). The estimates are somewhat less precise than those based on the number of transactions. This is likely due to the larger noise inherent in the value of healthcare spending. For example, some of largest incidences of healthcare transactions are surgeries which might not be related to air pollution.[30] The smaller impact on the value of transactions 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-healthcare spending are similar to results using the number of transactions, but less precise.

5.3 Nonlinearity

Among the underlying concerns for the external validity of the benefit-transfer approach is the potential nonlinearity of the dose-response function. The pollution level observed in developing countries such as China and India is far greater than the prevailing level studied in the literature. The potential nonlinearity could lead to under- or over-estimation of the health costs of air pollution in developing countries based on the linear projections in the benefit-transfer approach (Lelieveld et al. (2015) and World Bank (2007)). Despite of its important implications, there is a lack of empirical evidence on the nonlinearity of the dose-response function (Lelieveld and Poschl, 2017). The rich spatial and temporal variation in our data allows us to examine the health impacts of $PM_{2.5}$ for a wide range of pollution levels.

To capture nonlinearity, we include the quadratic term of $PM_{2.5}$ in addition to its linear form.[31] The top panel of Figure 8 plots the estimated surface of the marginal response for varying levels of $PM_{2.5}$ and along the time path for up to 100 days. For each value of $PM_{2.5}$, the slice of the surface along the p-axis is the estimated dynamic response as in Figure 7a. The surface is slightly tilted upwards with a higher marginal response for a higher pollution level, indicating an increasing marginal impact of $PM_{2.5}$ on healthcare spending. To further illustrate this, we plot the cumulated marginal effect over three months ($\sum_t \beta_t$) against pollution level in the bottom panel of Figure 8. Interestingly, the marginal impact on healthcare spending increases in $PM_{2.5}$ but at a diminishing rate. The cumulative effect ranges from 2.16% when $PM_{2.5}$ is near zero to 2.25% when the concen-

[30] Our analysis focuses on transactions that cost less than 200,000 *yuan*. Among this sample, the 95th percentile of the transaction value is 6,000 *yuan* and the 99th percentile is 10,000 *yuan*.

[31] To conserve the number of parameters, we use one-segment instead of three-segment B-splines, since results are very similar across different segments (see section 5.4).

tration reaches 150 μg/m^3 (i.e., the 90 percentile of the daily average).[32] Overall, the nonlinearity of the health impact appears modest. Based on this finding, we extrapolate our estimates across a wide range of pollution levels in the discussion below.

Figure 9 examines the impact of air pollution across cities with different per capita income. In 2015, China's city average annual disposable income per capita varied from 12,000 *yuan* to 53,000 *yuan*, with an average of 25,530 *yuan*. The top panel depicts the marginal response across income levels and over time and the bottom panel plots the cumulative long-term impact against income. Healthcare response appears largest in cities with the lowest income. This may be driven by the limited avoidance behavior (e.g., use of air purifier) among low income households. To the extent that low-income cities have a lower penetration of UnionPay cards, as shown in Table C1, this finding suggests that we might underestimate the overall health impact of air pollution based on bank cards rather than all payment methods, though it is unlikely a major concern given the limited heterogeneity across income levels.

5.4 Robustness Checks

We conduct a variety of robustness checks. Table 6 reports the cumulative impact for overall healthcare spending under three different numbers of B-spline segments (1, 2, and 3) and five different numbers of lags (60, 90, 120 and 150). The estimates across different numbers of segments are very similar. We choose three segments for our base specification since it provides a good balance between flexibility and precision.[33] The cumulative impact tends to be smaller with 60 days of lags and larger with 120 days of lags than that with 90 days, but the difference is small. The cumulative impact using 30-day lags is considerably smaller.[34] We choose 90 lags because the estimated effects for many of the lagged pollution measures are significant until around 90 days and start to lose significance for earlier periods.

Our second set of robustness checks is with regard to the radius size of the buffer zone in constructing the IV. We fix the radius at 150 km in the benchmark specification and assume that unobservables outside of the buffer zone of a city would not affect healthcare spending in that city. There is an inherent trade-off in the choice of the radius. On the one hand, the larger the buffer zone, the easier it is for the exclusion restriction to hold. On the other hand, the bigger the radius, the weaker the correlation between the predicted $PM_{2.5}$ using non-local pollution and the observed $PM_{2.5}$ in a given city. Table 7 presents several choices of the buffer zone from 50 km to 300 km with an increment of 50 km. The top panel reports the first-stage results. Generally, both the R^2 and the

[32]Our preferred estimate of the cumulative effect is 2.65%, at the high end of this range, because the estimation constraints are nonlinear.

[33]Results from more than three segments suffer from the over-fitting problem and exhibit large swings over time.

[34]Cross-validation results indicate that models with long lags are preferred to the model with 30 days of lags.

F-statistics decrease with the radius of the buffer zone, suggesting a weaker correlation between the IV and the endogenous variable as the buffer zone gets larger. The bottom panel shows the cumulative long-term impact on healthcare spending, which varies from 2.42% to 2.88% across different radii when $PM_{2.5}$ increases by 10 μg/m^3 permanently. Our preferred specification with a 150 km radius delivers an estimate that is in the middle of this range.

The third set of robustness checks controls for other pollutants including O_3, SO_2, NO_2 and CO. Emission sources such as electricity generation and transportation produce both particulate matters and other pollutants, which also have harmful health impacts. 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 the observed level of local pollutants. Table 8 reports estimates with these four pollutants as additional controls. The results for both healthcare spending and non-healthcare spending categories are very similar to those in Table 5 without controlling for other pollutants.[35]

The fourth set of robustness checks further addresses the concern of regional economic spillovers by controlling for the average level of $PM_{2.5}$ of nearby cities in the same region outside of the buffer zone. If regional economic activities have systematic spillover effects beyond the buffer zone, one might be concerned with the exogeneity of our IVs: local unobservables could be correlated with economic activities in other cities, which are in turn correlated with pollution levels in other cities. Including the average level of $PM_{2.5}$ of nearby cities in the regressions could help control for economic activities in other cities. Table 9 presents estimation results with this additional control, and the results are very close to the benchmark specifications.

To further address the concerns on the exogeneity of the instruments, we create an alternative IV where all of the variation comes from changing wind patterns. In our main specification, the IVs are functions of both the level of pollution in source cities and prevailing wind directions and speeds. In this robustness check, we use the historical average (time-invariant) level of air pollution in the source city rather than the actual observed pollution level that could be subject to regional economic spillovers. All of the within-city variation in these IVs comes from variation in wind and weather patterns, and thus, after controlling for city fixed effects, the IVs should be uncorrelated with any unobserved economic shocks. Table 10 presents results from this specification. Though the IVs are not as strong as those in the main specification, as indicated by a decrease in the F-statistic, the estimated effect of pollution on healthcare spending is very similar to the benchmark specification.

As a final robustness check, we drop the following large cities from the analysis: Beijing,

[35]The correlation coefficient between $PM_{2.5}$ and O_3, SO_2, NO_2 and CO is -0.13, 0.55, 0.66, 0.03, respectively.

Shanghai, Guangzhou, Shenzhen, Wuhan, Chongqing, Chengdu, and Nanjing. Due to having superior medical facilities and being large transportation hubs, these cities receive a large number of hospital patients from other areas. If some of these out-of-town patients come from cities that export pollution to these major cities, this could lead to correlation between the instruments and unobserved healthcare spending shocks. Table 11 reports the estimated long-term effect of air pollution when these cities are removed from the estimation sample. The estimates are similar to those in the benchmark specification.

5.5 Avoidance Behavior

The analyses of both the short-term and longer-term impact suggest that elevated $PM_{2.5}$ leads to increased healthcare spending and reduced non-healthcare spending. This negative impact on non-healthcare spending could be driven by two underlying mechanisms: the budget constraint and avoidance behavior. As we argued in Section 5.2, the short-lived nature of the negative consequences is inconsistent with the budget constraint hypothesis. In this section, we examine whether households engage in avoidance behavior to mitigate their pollution exposure.

A key insight of our analysis is that when consumers engage in avoidance behavior, expectations of *future* pollution levels should affect current consumption. For example, if consumers expect pollution to improve in the near future, they may postpone their consumption to avoid exposure today. On the other hand, an expectation of worse air tomorrow may encourage them to make the consumption in advance. To investigate this, we assume that the consumers are aware of whether the next day's air quality is better or worse than the current day's.[36]

We add the dummy variable $\mathbb{1}\{p_{i,t+1} > p_{i,t}\}$ in our baseline specification and report the results in Table 12. The coefficient on this dummy variable indicates a 0.41% increase in healthcare transactions when consumers anticipate worse air quality the next day. In addition, spending in necessities and supermarkets *increases* when next-day pollution is expected to deteriorate. The coefficient is also found to be larger for pharmacies than hospitals, consistent with the fact that hospital visits are often scheduled in advance and are less substitutable intertemporally. The estimated cumulative impact on healthcare spending that is associated with a permanent reduction of 10 μg/m^3 of $PM_{2.5}$ is 2.71%, slightly higher than when we do not control for avoidance.

5.6 Morbidity Cost

Our preferred specifications show that a 10 μg/m^3 increase in $PM_{2.5}$ would lead to a 2.65% increase in the number of health-related transactions in the long term (Table 5) and a 1.5% increase in the

[36]Given the ubiquitous forecasts and apps on $PM_{2.5}$, this appears a reasonable assumption. Note that it is much weaker than assuming consumers know about the actual *level* of tomorrow's pollution.

value of transactions (Table C2). To better understand the size of our morbidity estimates, we compare our results with the findings in the related literature in Table 13. In a study on preventive expenditure, Mu and Zhang (2016) estimate that face mask purchases in China increase by 5.45% for a 10-point increase in Air Quality Index (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 anywhere between 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 from 3.6% to 7.3% in preventive spending.

Williams and Phaneuf (2016) use similar estimation methods and data in the U.S. context and 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, which is equivalent to a 22% increase for 10 $\mu g/m^3$ more $PM_{2.5}$. According to China's National Health Commission, spending on respiratory diseases accounts for 8% of total health expenditure in 2012. Assuming all additional spending induced by air pollution is for respiratory diseases, our estimates translate to a 33% increase in respiratory-related spending, about one and a half times as large as the estimate from Williams and Phaneuf (2016).

Based on the parameter estimates, we now conduct back-of-the-envelope calculations to estimate the morbidity cost from elevated $PM_{2.5}$. Credit and debit card transactions (i.e., bank card transactions) account for about half of the total private 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-bank-card spending, the 1.5% impact translates to 59.6 billion Yuan ($9.2 billion) from a 10 $\mu g/m^3$ (about 18%) increase in $PM_{2.5}$. These numbers can directly inform the overall welfare cost 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 a $20 billion direct cost on health expenditures (due to morbidity) worldwide, with half of these costs coming from non-OECD countries.[37] A simple linear interpolation based on our estimates implies $42 billion in added healthcare spending in China owing to elevated $PM_{2.5}$ (56 $\mu g/m^3$ on average) relative to WHO's recommended level of 10 $\mu g/m^3$.[38]

These results indicate that the OECD report significantly underestimates the health cost from outdoor air pollution, potentially up to an order of magnitude for developing countries. The underestimation from the OECD report could be due to: (1) the downward bias from endogeneity in the dose-response function; (2) the inherent differences in the dose-response function across countries;

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

[38] According to China's National Health Commission, aggregate health expenditure, including both private and public spending, was more than four trillion *yuan* ($614 billion) in 2015.

and (3) the monetization of the disease incidences. The discrepancy highlights the importance of empirical studies using data on health spending from developing countries.

In addition to the direct healthcare cost, the morbidity cost should also include the value of lost time from the illnesses including hospital visits. Based on the fact that our database recorded 670 million health-related transactions in 2015 which accounted for 50% of private health spending, our estimate implies 35.5 million additional trips to healthcare facilities from a 10 μg/m^3 increase in $PM_{2.5}$. To monetize the lost time, we assume that each trip takes three hours and the value of time (VOT) is 100% of the hourly wage, which is an upper-end estimate of VOT in the literature (Small, 2012; Wolff, 2014). The total value of the lost time from additional trips to healthcare facilities amounts to 2.3 billion Yuan in 2015, compared to 59.6 billion Yuan in additional healthcare costs from a 10 μg/m^3 increase in $PM_{2.5}$. This suggests that the direct healthcare cost is the major component of the overall morbidity cost, recognizing that we are not accounting for the lost productivity and loss in quality of life.

5.7 WTP for Clean Air

As illustrated in Section 3, consumer WTP for clean air is composed of several components including the impact on healthcare spending, mortality, labor productivity and the quality of life. While there is a growing literature on the impact of air pollution on labor productivity and the quality of life (Levinson, 2012; Zivin and Neidell, 2012; Chang et al., 2016; He et al., 2018), there is a need for comprehensive country-level analysis in China. In the following analysis, we focus on the morbidity and mortality cost in order to provide a lower bound for consumer WTP for clean air, recognizing that the full WTP calculation would necessitate additional empirical evidence on other components.

Our analysis suggests that a 10 μg/m^3 increase in $PM_{2.5}$ would lead to an increase of 59.6 billion *yuan* (or $9.2 billion) in overall healthcare spending. This can be considered as the social cost of air pollution due to increased spending in healthcare. To examine the private cost of air pollution and consumers' private WTP for air quality, we need to take into consideration that the vast majority of consumers have health insurance and do not bear the full cost of their treatment. For urban residents, the proportion of healthcare spending that is paid for out-of-pocket equals 32% for employees and 52% for non-employees. For the rural population, the share of out-of-pocket spending is 56%. An increase in overall healthcare spending of 59.6 billion *yuan* translates into 30.2 billion *yuan* (or $4.6 billion) additional out-of-pocket spending.[39] This implies that the morbidity component of consumer WTP is 74 *yuan* per household (or $11.3) for a 10 μg/m^3 reduction of

[39] In 2011, there were 252 million urban employees, 221 million urban non-employees, and 832 million rural residents enrolled in China's public insurance programs (Yu, 2015). The population-weighted average proportion of healthcare spending that must be covered out-of-pocket thus equals 50.7%.

$PM_{2.5}$.

We estimate the mortality cost based on Ebenstein et al. (2017). Using detailed mortality data by gender, age cohort, and disease types in 161 representative counties across China, they estimate that a 10 $\mu g/m^3$ increase of PM_{10} would increase the cardiorespiratory mortality rate by 8% on average and the impact varies across age cohorts by not across gender. We take two steps to monetize the mortality impact. First, we rely on the benefit transfer approach to estimate the VSL for the Chinese population due to the lack of a national-level estimate of VSL for China. Ashenfelter and Greenstone (2004) provides an upper-bound VSL estimate of $2.27 million (in 2015) for the U.S. population. This value is used in Deschenes et al. (2017) to estimate the mortality cost of NO_x emissions reductions.[40] To adjust for income differences between the U.S. and China, we use the transfer elasticity of 1.2 as suggested by Narain and Sall (2016). This leads to an estimate of $0.21 million (in 2015) for VSL in China.[41] Second, we develop estimates of the VSL for each age group based on the age-adjustments in Murphy and Topel (2006). The VSL is at the full value for people less than 40 years old but reduced to 40% of its full value by age 65 and 15% by age 80. Similar to what Deschenes et al. (2017) find, this adjustment is important because although the age groups with age 65 and above account for less than 9% of the total population, they account for nearly 75% of the avoided mortalities from an improvement in air quality.

This back-of-the-envelope analysis implies that a 10 $\mu g/m^3$ increase of $PM_{2.5}$ would generate a mortality impact in the neighborhood of $13.4 billion in 2015 in China (Table C3 in Appendix C).[42] In comparison, the corresponding healthcare cost is $9.2 billion, which is 69% of the mortality cost. In terms of consumer WTP for clean air, the mortality component amounts to $29.3 per household for a 10 $\mu g/m^3$ reduction of $PM_{2.5}$. Taking into account both morbidity and mortality impacts, consumer WTP for a 10 $\mu g/m^3$ reduction of $PM_{2.5}$ would be $40.6 per household.

Using a discrete choice framework to estimate the demand of indoor air purifiers in China, Ito and Zhang (2016) estimate a consumer WTP of $11 for a 10-unit reduction in PM_{10} based on the trade-off between price and quality (purifiers' ability to remove PM_{10}). Note that the benefit of an air purifier is confined to the location where the equipment is installed (e.g., a room) and does not apply elsewhere. Consumers' WTP for a uniform reduction of ambient pollution (e.g., indoors and outdoors) is likely bigger. Though the context and approach are different, our estimate of the morbidity component of WTP is broadly in line with their findings.

Using the hedonic model for the U.S. housing market, Chay and Greenstone (2005) find that

[40] Primarily based on studies of compensating wage differentials in the labor market, U.S. EPA recommends a central estimate of $8.7 million (in 2015) as the VSL value for their benefit and cost analysis. Other studies have estimated a lower value in the neighborhood of $2 million (Alberini et al., 2004; Ashenfelter and Greenstone, 2004).

[41] The transfer elasticity measures the income elasticity of VSL. Narain and Sall (2016) suggest a transfer elasticity of 1.2 for transferring the VSL from U.S. or a developed country to a developing country with low or middle income.

[42] $PM_{2.5}$ accounts for about 60% of PM_{10} during our data period. We assume that the mortality cost of a 10 $\mu g/m^3$ increase of $PM_{2.5}$ is the same as that of a 10 $\mu g/m^3$ increase of PM_{10}.

consumers are willing to pay \$450 - \$1,050 more in housing price (in 1982-84 dollars) for a *one* μg/m^3 reduction in total suspended particles (TSP). With a 30-year time span and 5% annual discount rate, this translates to an annual WTP of \$72.9 - \$168.5 in 2015 dollars. Based on the discrete-choice framework that is also applied to the U.S. housing market, Bayer et al. (2009) estimate the annual household WTP to be \$23.9 - \$29.5 in 2015 dollars for a one-unit reduction in PM_{10}. These estimates are larger than ours for at least two reasons. First, the average household income in China in 2015 is about one-seventh 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 a lower bound and does not account for the impacts on the loss of productivity and quality of life, etc. as discussed above, while the WTP estimates using the revealed preference approach should in theory reflect those impacts (provided that consumers are well-informed and rational).

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

6 Conclusion

WHO's global air pollution database shows that the 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, past research from epidemiology and economics going back several decades has focused on the impacts of air pollution on human health, particularly mortality, in developed countries. This study provides the first comprehensive analysis on the direct healthcare cost from $PM_{2.5}$ based on the universe of credit and debit card transactions in China.

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 that allows us to isolate exogenous variations in local air quality using the spatial spillovers of $PM_{2.5}$. We propose a flexible distributed-lag model to estimate the temporal effect on healthcare spending. Our IV results, three to four times larger than those from OLS, suggest that a 10 μg/m^3 decrease in $PM_{2.5}$ would lead to at least \$9.2 billion reduction in healthcare spending annually, or 1.5% of national annual healthcare expenditure. The estimated healthcare cost exceeds two-thirds of the mortality cost based on the recent literature. China's elevated $PM_{2.5}$ level relative to the WHO's annual standards entails $42 billion additional healthcare expenditure in 2015. Together, these results indicate that the recent report by OECD (2016) drastically underestimates the worldwide impact of air pollution on health expenditure ($10 billion for all non-OCED countries including China).

We offer to our knowledge the first national-level analysis of the impact of air pollution on healthcare spending in a developing country context. The air pollution level in 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 get worse before it gets better. The aggregate impact of air pollution on economic growth, including factors such as human capital accumulation, productivity, talent loss due to migration, and foreign direct investment, is an interesting and important area for future research.

References

Alberini, Anna, Maureen Cropper, Alan Krupnick, and Nathalie B. Simond, "Does the value of a statistical life vary with age and health status? Evidence from the US and Canada," *Journal of Environmental Economics and Management*, 2004, *48* (1), 769–792.

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.

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

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.

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.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell, "The effect of pollution on office workers: Evidence from call centers in china," 2016. Working Paper.

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.

Clay, Karen, Joshua Lewis, and Edson Severnini, "Canary in a Coal Mine: Impact of Mid-20th Century Air Pollution on Infant Mortality and Property Values," 2016. Working Paper.

Corradi, Corrado, "Smooth Distributed Lag Estimators and Smoothing Spline Functions in Hilbert Spaces," *Journal of Econometrics*, 1977, *5* (1), 211–219.

_ and G. Gambetta, "The Estimation of Distributed Lags by Spline Functions," *Empirical Economics*, 1976, *1* (1), 41–51.

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," *American 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," *American Economic Review*, 2017, *107* (10), 2958–89.

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

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

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

EPA, *The benefits and costs of the Clean Air Act from 1990 to 2020*, Washington, DC: United States Environmental Protection Agency, 2011.

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.

He, Jiaxiu, Haoming Liu, and Alberto Salvo, "Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China," *American Economic Journal: Applied Economics*, 2018.

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.

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

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.

Lelieveld, Jos and Ulrich Poschl, "Chemists can help to solve the air-pollution health crisis," *Nature*, 2017, *551*, 291–293.

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.

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

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

Poirier, Dale J., "On the Use Of Bilinear Splines in Economics," *Journal of Econometrics*, 1975, *3* (1), 23–34.

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.

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

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.

WHO, *Economic cost of the health impact of air pollution in Europe: Clean air, health and wealth*, Copenhagen: WHO Regional Office for Europe, 2015.

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.

Wolff, Hendrik, "Value of time: Speeding behavior and gasoline prices," *Journal of Environmental Economics and Management*, 2014, *67* (1), 71–88.

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

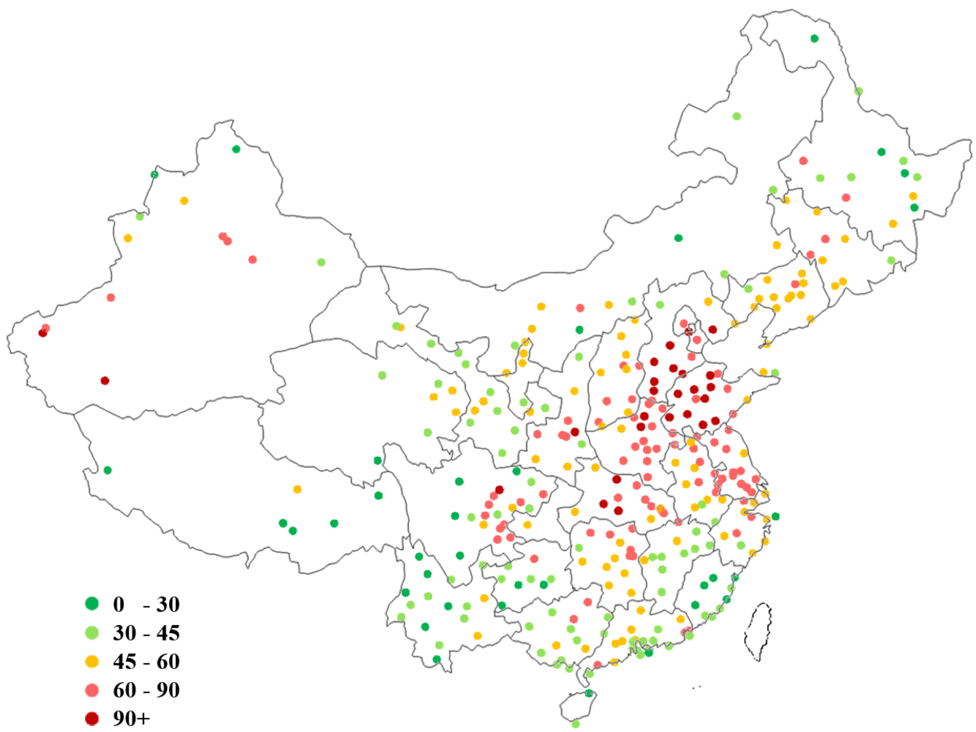
Yu, Hao, "Universal health insurance coverage for 1.3 billion people: What accounts for China's success?," *Health Policy*, 2015, *119* (1), 1145–1152.

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).

Zivin, Joshua Graff and Matthew Neidell, "The impact of pollution on worker productivity," *American Economic Review*, 2012, *102* (7), 3652–73.

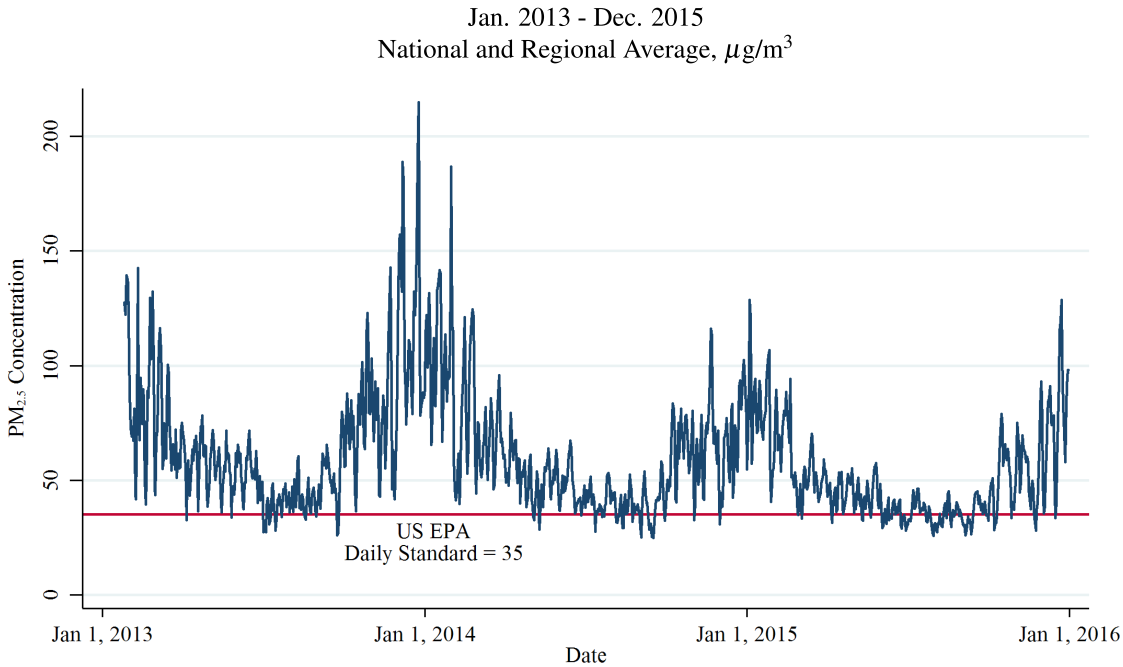
Figure 1: Three-Year Average $PM_{2.5}$ Concentration

Jan. 2013 - Dec. 2015, μg/m^3

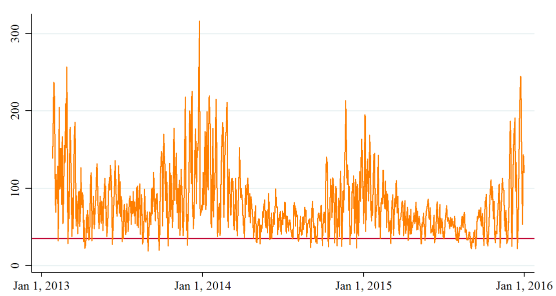


Notes: Each dot represents a city. There are 329 cities in total.

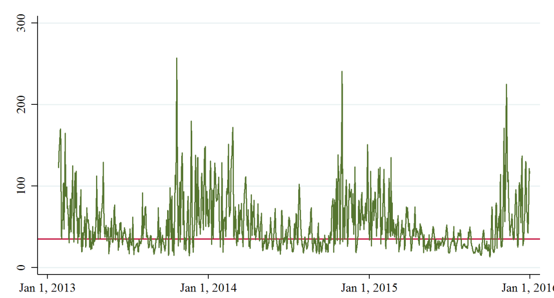
Figure 2: Daily $PM_{2.5}$ Concentration



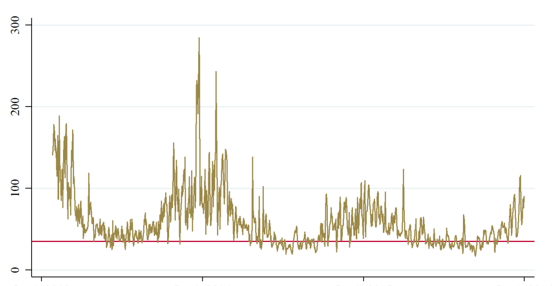
(a) Northern Region



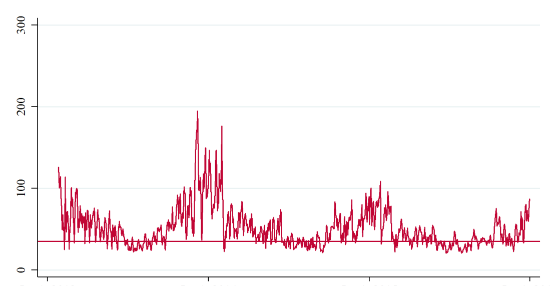
(b) Northeastern Region



(c) Northwestern Region

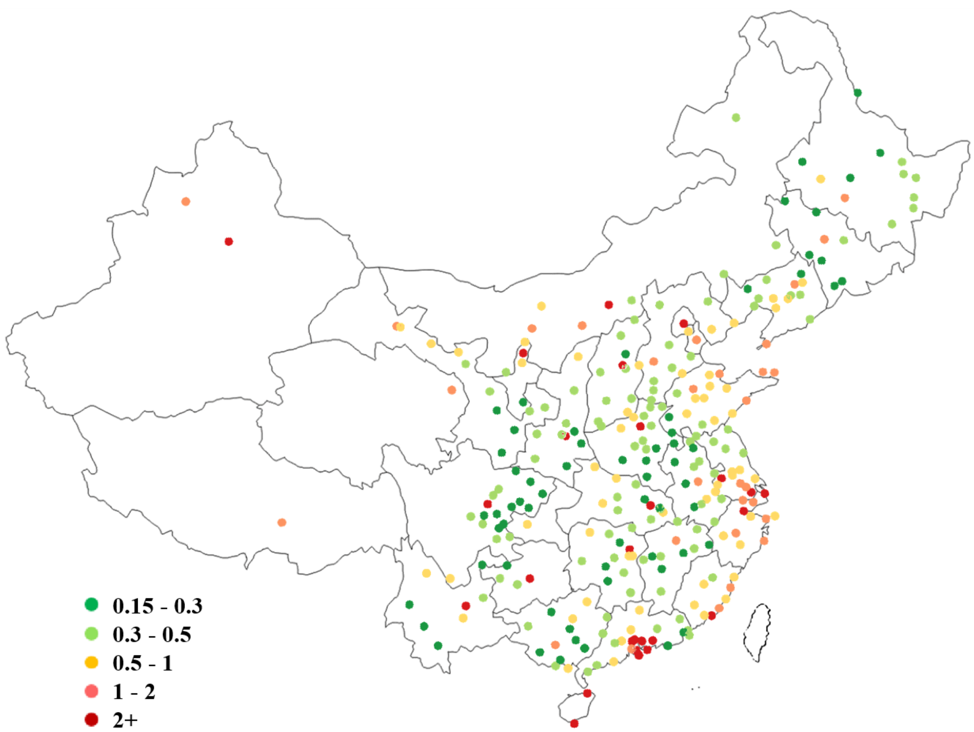


(d) Southern Region



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

Figure 3: The Number of Active Bank Cards per Capita, 2015



Notes: Bank cards include debit and credit cards. Active bank cards are defined as cards that have been used at least once in the given year. Each card is assigned to one primary city based on the location of its most frequent usage. Population measure is year-end registered population of each city.

Figure 4: National Weekly Healthcare Spending, 2013 - 2015

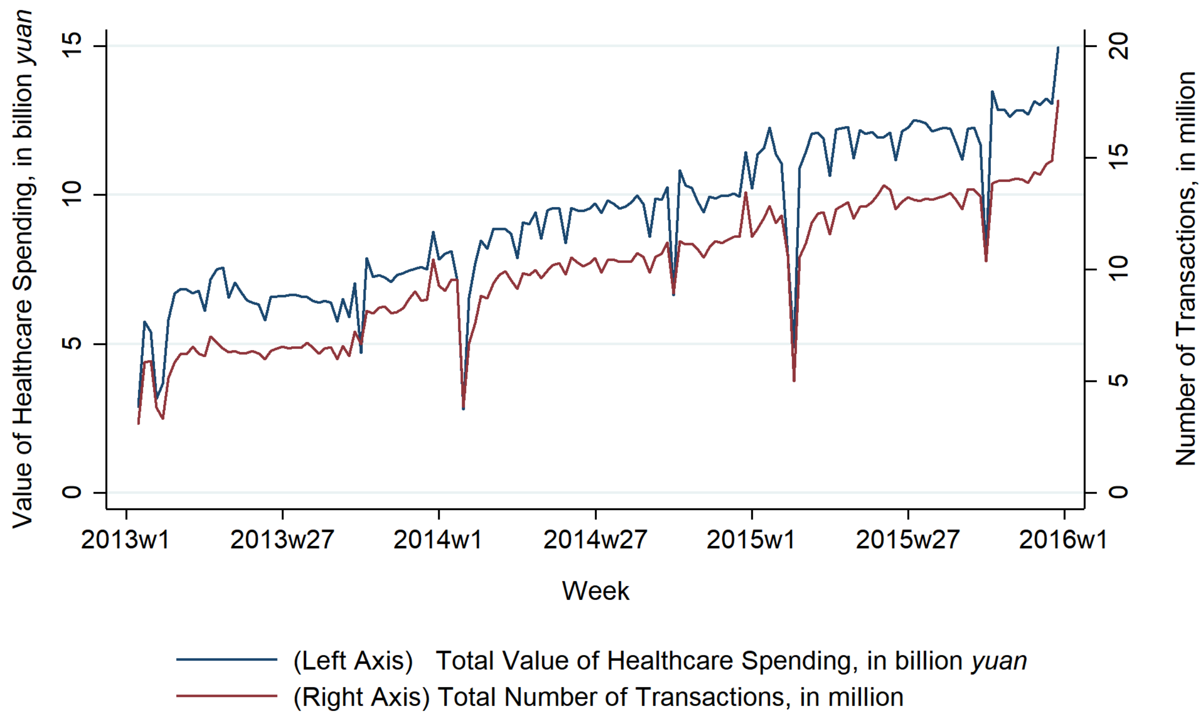


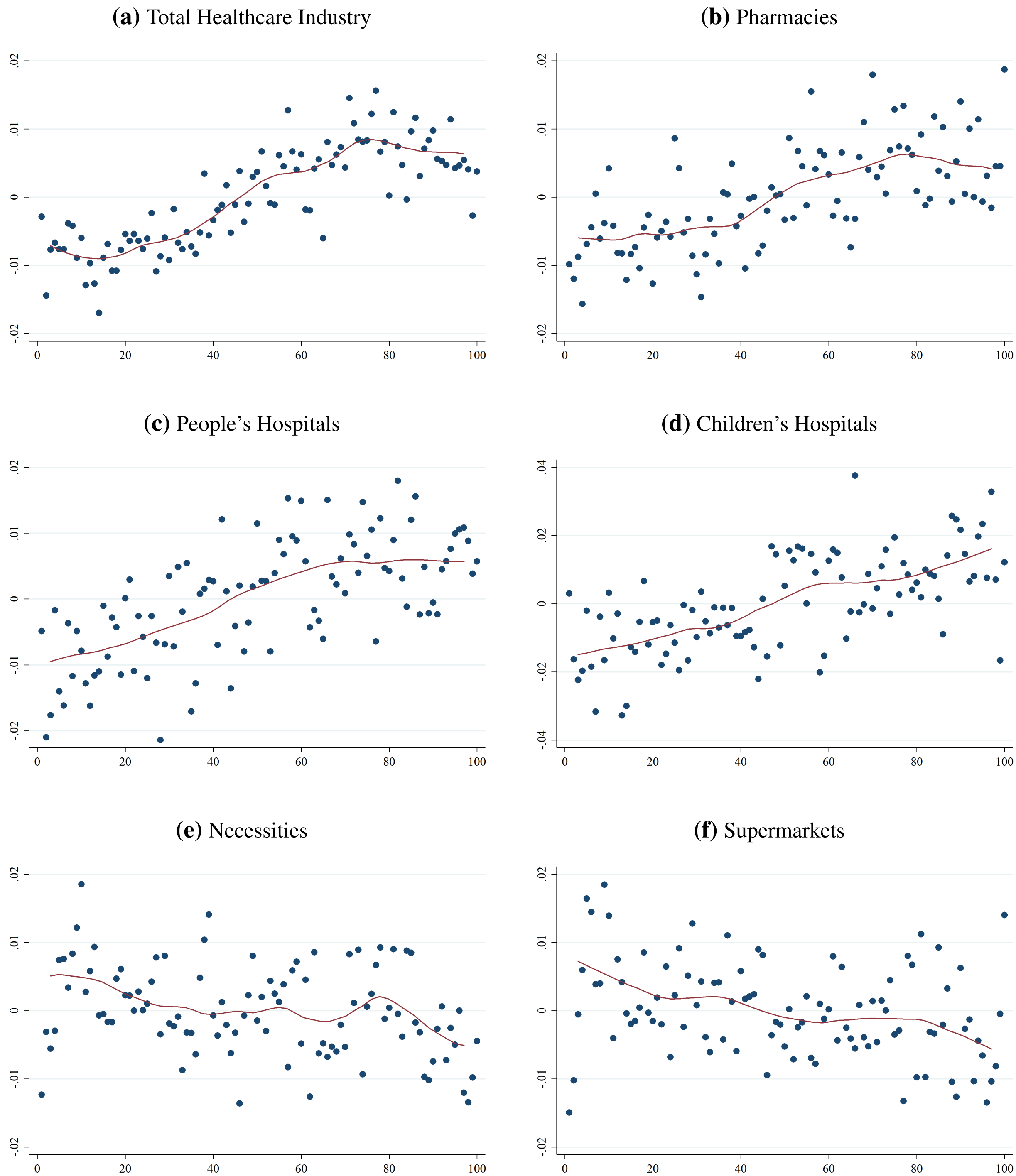
Figure 5: Wind-Pollution Vector Decomposition

(a) Day 0: Wind-Pollution Vectors

(b) Day 0: Wind-Pollution Vectors toward Beijing

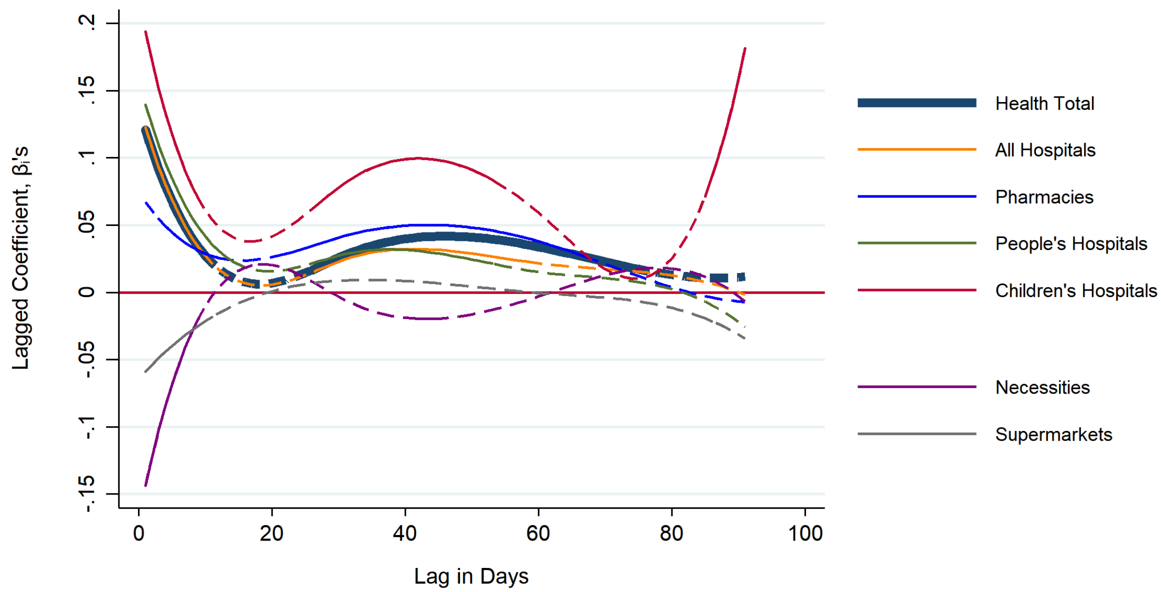
Notes: Day 0 refers to Dec. 5, 2013. Subfigure 5a depicts the wind-pollution vector fields on Day 0, with each vector's length indicating wind speed (rescaled to match the distance traveled per day) and width indicating $PM_{2.5}$ concentration level in the source city. Subfigure 5b plots the decomposed subvectors pointing towards Beijing.

Figure 6: Residuals of Log Number of Transactions v. $PM_{2.5}$ Concentration, by Category

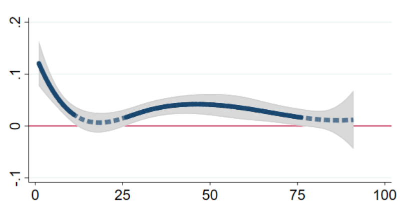


Notes: Each dot denotes the in-group average residuals, partialing out city FEs, weekly FEs, city-specific time trends, city-specific seasonality, day-of-week FEs, dummies for holidays and working weekends, and weather controls (temperature, precipitation, wind speed). Groups are binned by percentiles of $PM_{2.5}$, which is depicted by the x-axis.

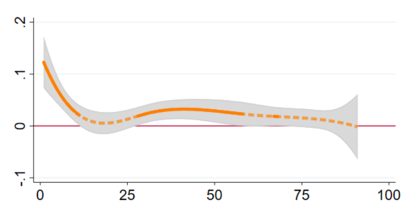
Figure 7: Impact of Air Pollution on Number of Transactions from IV Regressions with 90 Lags



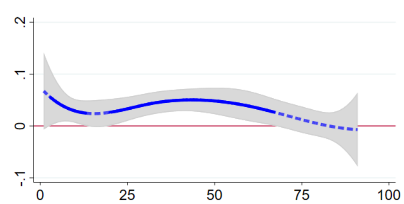
(a) Health Total



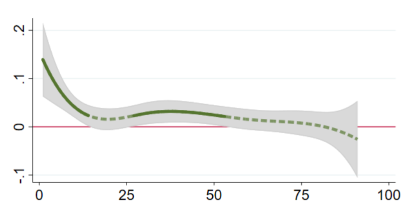
(b) All Hospitals



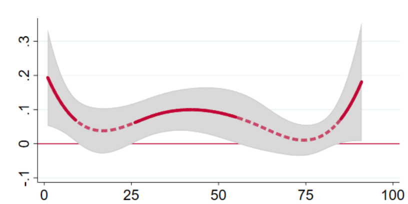
(c) Pharmacies



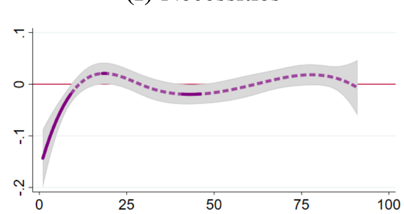
(d) People's Hospitals



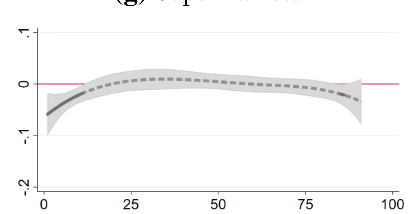
(e) Children's hospitals



(f) Necessities

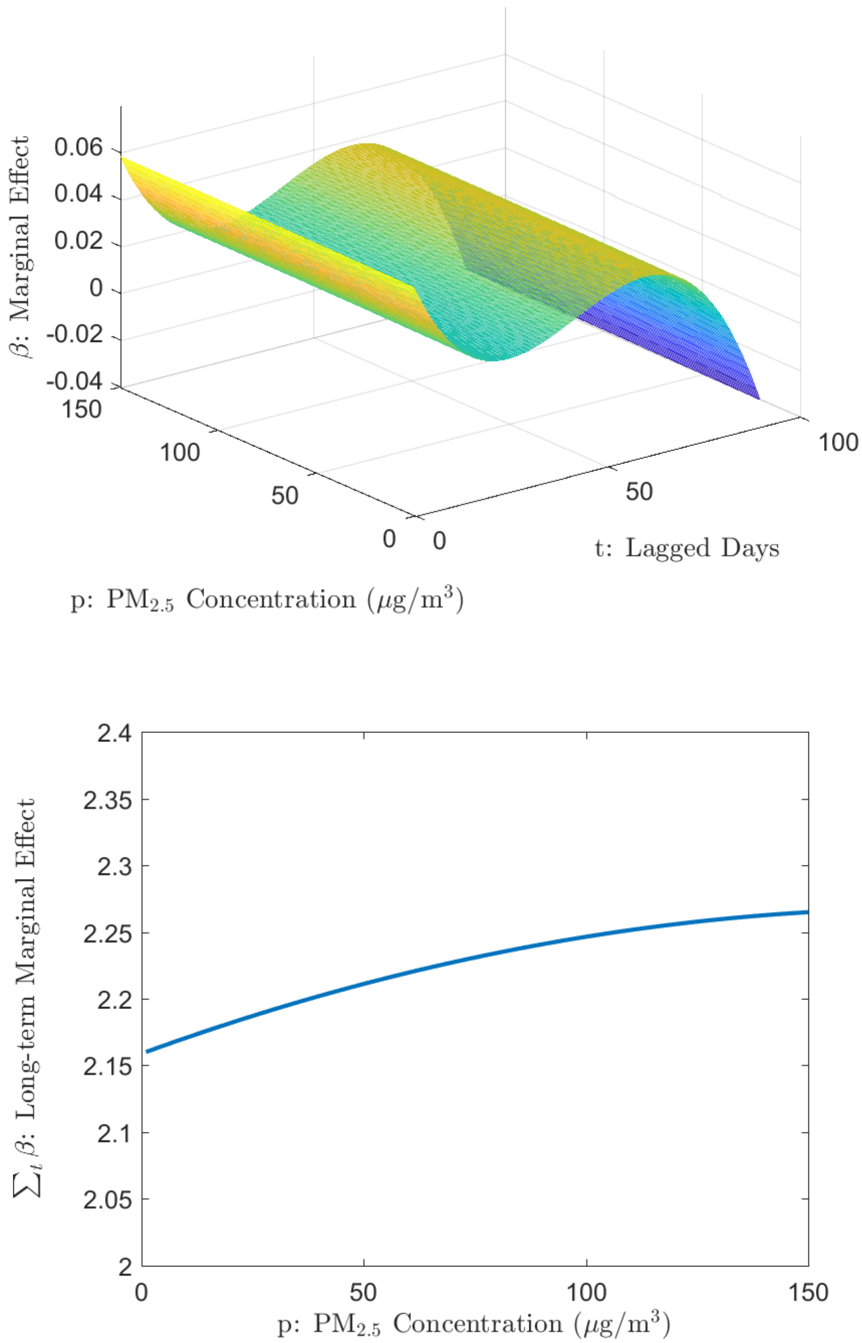


(g) Supermarkets



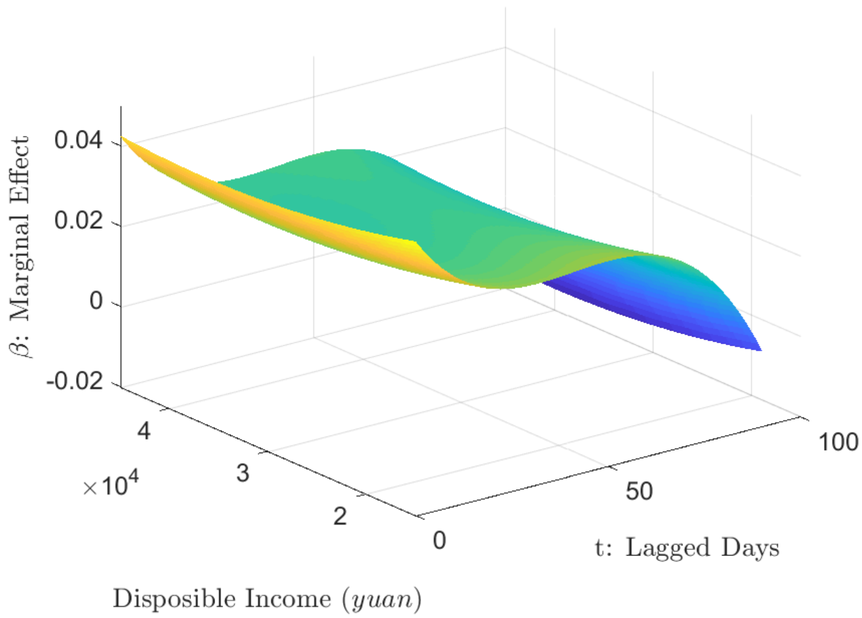
Notes: The y-axis indicates the percentage change in the number of transactions per 10 μg/m^3 increase in $PM_{2.5}$ concentration. On the x-axis, 0 refers to current pollution, 50 refers to pollution 50 days ago, etc. Solid line indicates significance at the 5% level. Gray areas are 95% confidence intervals.

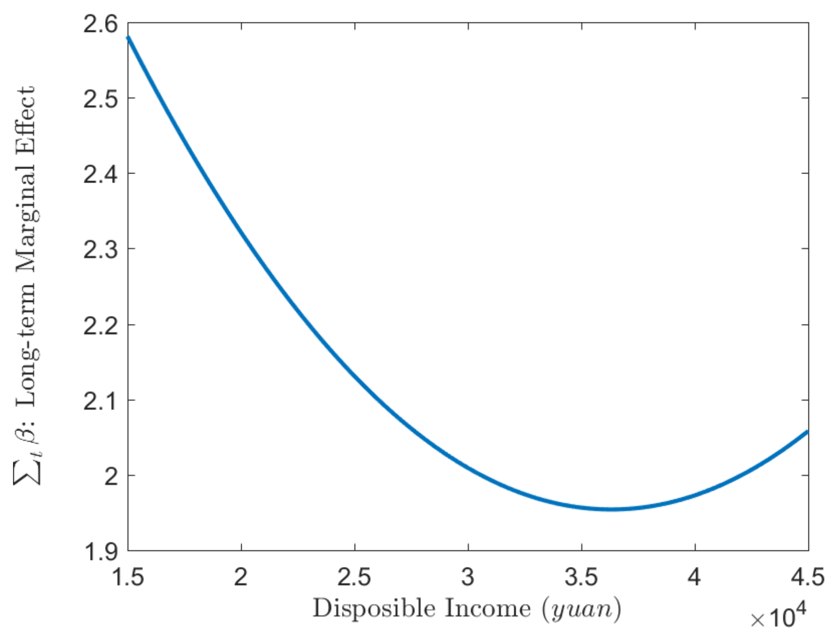
Figure 8: Nonlinear Impacts of Air Pollution



Notes: The y-axis in the top panel indicates the percentage change in the number of transactions for a 10 μg/m^3 increase in $PM_{2.5}$ at a specific level of pollution concentration and on a given day. On the t-axis, 0 refers to current pollution, 50 refers to pollution 50 days ago, etc. The y-axis in the bottom panel indicates the percentage change in the number of transactions for a permanent 10 μg/m^3 increase in $PM_{2.5}$. The x-axis denotes different levels of pollution. Results are from the IV estimation.

Figure 9: Impacts of Air Pollution Across Income Levels





Notes: The y-axis in (a) indicates the percentage change in the number of transactions for a 10 μg/m^3 increase in $PM_{2.5}$ on a given day. The t-axis (from 0 to 90) refers to current pollution, pollution in the previous day, pollution t days before, etc. The y-axis in (b) indicates the percentage change in the number of transactions for a persisting 10 μg/m^3 increase in $PM_{2.5}$ during the last three months. The x-axis denotes different levels of disposable income. Results are from the IV estimation.

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	7,229.2	21,308.6	0	330,974	211,318
All Hospitals	4,122.7	14,503.9	0	237,525	210,539
People's Hospitals	1,060.6	2,800.4	0	40,332	203,407
Children's Hospitals	464.7	1,290.5	0	18,227	158,637
Pharmacies	2,245.3	7,063.3	0	96,336	210,001
Comparison Groups, *from 1% card sample*					
Daily Necessities	233.3	628.6	0	10,865	211,318
Supermarkets	393.4	990.3	0	15,224	210,493
Total Value of Transactions, Daily, *thousand yuan*					
Healthcare Industry, Total	6,701.8	17,818.9	0	301,108.7	211,318
All Hospitals	5,556.5	15,066.8	0	275,883.0	210,539
People's Hospitals	1,588.1	3,401.2	0	56,856.9	203,407
Children's Hospitals	363.9	843.3	0	10,324.3	158,637
Pharmacies	407.4	1,109.5	0	16,735.1	210,001
Comparison Groups, *from 1% card sample*					
Daily Necessities	236.9	551.3	0	9,532.4	211,318
Supermarkets	232.8	643.4	0	14,404.7	210,493
Weather					
Mean Temperature, $^\circ F$	60.11	18.92	-27.50	101.6	211,317
Precipitation, *inch*	0.13	0.42	0	15.6	211,318
Mean Wind Speed, *mph*	5.50	3.11	0	48.7	211,296
Wind Direction, *navigational bearing*	-	-	0	360	211,263

Notes: Data sources include China's Ministry of Environmental Protection, Integrated Surface Database (ISD), and Global Surface Summary of the Day (GSOD) Database. Data for comparison groups are calculated from a subsample that uses a randomly selected 1% of bank cards. Transactions larger than 200,000 *yuan* ($29,000) are excluded from the total value of transactions. The arithmetic mean and standard deviation of wind directions do not have statistical meaning and are left out in the table.

Table 2: OLS Estimates of the Pollution Impact on Health Spending: Contemporaneous Effects

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$PM_{2.5}$, Current Day	0.11***	0.11***	0.12***	0.13***	0.18***	-0.06***	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)
N	192,586	191,814	191,277	185,773	146,224	192,035	191,766

Notes: The dependent variable is log(number of transactions). The controls are city FEs, weekly FEs, city-specific time trends, city-specific seasonality, day-of-week FEs, dummies for holidays and working weekends, and weather controls (temperature, precipitation, wind speed). Each column reports the percentage change in the number of transactions per 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

Table 3: IV Estimates of the Pollution Impact on Health Spending: Contemporaneous Effects

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$PM_{2.5}$, Current Day	0.65***	0.73***	0.60***	0.77***	1.13***	-0.09	-0.10
	(0.09)	(0.11)	(0.15)	(0.13)	(0.37)	(0.15)	(0.12)
N	192,586	191,814	191,277	185,773	146,224	192,035	191,766
First-stage F	61.93	61.77	61.78	59.47	52.32	61.92	61.97

Notes: The dependent variable is log(number of transactions). The IVs are various interactions of (both current and lagged) wind patterns, weather conditions and $PM_{2.5}$ in cities more than 150 km away, which capture non-local $PM_{2.5}$. Same controls as in Table 2. Each column reports the percentage change in the number of transactions per 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 4: Cumulative Effect of Pollution, OLS with 90 Lags

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Current Day	0.03***	0.04***	0.05***	0.04***	0.06***	-0.03***	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Current + Past 3d	0.12***	0.11***	0.18***	0.13***	0.19**	-0.11***	-0.07**
	(0.03)	(0.03)	(0.04)	(0.04)	(0.08)	(0.03)	(0.03)
Current + Past 7d	0.19***	0.16***	0.32***	0.21***	0.25*	-0.16***	-0.11***
	(0.05)	(0.06)	(0.07)	(0.06)	(0.15)	(0.05)	(0.04)
Current + Past 14d	0.25***	0.16	0.49***	0.30***	0.20	-0.16**	-0.13**
	(0.08)	(0.10)	(0.10)	(0.08)	(0.28)	(0.07)	(0.06)
Current + Past 28d	0.38***	0.18	0.80***	0.39***	0.12	-0.15	-0.09
	(0.13)	(0.15)	(0.16)	(0.14)	(0.50)	(0.12)	(0.11)
Current + Past 56d	0.66***	0.27	1.42***	0.47**	0.57	-0.27	0.03
	(0.19)	(0.20)	(0.29)	(0.24)	(0.74)	(0.21)	(0.18)
Current + All Lags	0.86***	0.34	1.81***	0.59*	0.38	-0.08	0.02
	(0.27)	(0.28)	(0.42)	(0.36)	(1.14)	(0.27)	(0.21)
N	141,794	141,657	141,567	137,853	110,259	141,770	141,652

Notes: The dependent variable is log(number of transactions). The effect of current and past air pollution is estimated using the Flexible Distributed-Lag Model with 90 lags and 3 evenly-split segments. Same controls as in Table 2. Each row reports the cumulative percentage change in the dependent variable in response to a 10 μg/m^3 increase in $PM_{2.5}$ for the corresponding period. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

Table 5: Cumulative Effect of Pollution, IV with 90 Lags

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Current Day	0.12***	0.12***	0.07*	0.14***	0.19***	-0.14***	-0.06***
	(0.02)	(0.03)	(0.04)	(0.04)	(0.07)	(0.03)	(0.02)
Current + Past 3d	0.40***	0.40***	0.23*	0.47***	0.65***	-0.45***	-0.21***
	(0.07)	(0.08)	(0.12)	(0.13)	(0.23)	(0.09)	(0.07)
Current + Past 7d	0.61***	0.62***	0.39**	0.75***	1.04***	-0.64***	-0.34***
	(0.10)	(0.12)	(0.18)	(0.19)	(0.36)	(0.13)	(0.10)
Current + Past 14d	0.74***	0.75***	0.57***	0.97***	1.40***	-0.63***	-0.45***
	(0.14)	(0.16)	(0.21)	(0.22)	(0.50)	(0.16)	(0.12)
Current + Past 28d	0.91***	0.90***	0.99***	1.24***	2.12***	-0.44*	-0.41**
	(0.22)	(0.25)	(0.30)	(0.27)	(0.79)	(0.23)	(0.21)
Current + Past 56d	1.97***	1.71***	2.31***	2.01***	4.65***	-0.85**	-0.23
	(0.42)	(0.47)	(0.54)	(0.46)	(1.56)	(0.41)	(0.36)
Current + All Lags	2.65***	2.18***	2.80***	2.13***	6.37***	-0.55	-0.57
	(0.68)	(0.71)	(0.89)	(0.75)	(2.33)	(0.58)	(0.47)
N	141,794	141,657	141,567	137,853	110,259	141,770	141,652
First-stage F	38.35	38.36	38.37	39.69	47.79	38.29	38.29

Notes: The dependent variable is log(number of transactions). The effect of current and past air pollution is estimated using the Flexible Distributed-Lag Model with 90 lags and 3 evenly-split segments. The IVs are various interactions of (both current and lagged) wind patterns, weather conditions and $PM_{2.5}$ in cities more than 150 km away, which capture non-local $PM_{2.5}$. Same controls as in Table 2. Each row reports the cumulative percentage change in the dependent variable in response to a 10 $\mu g/m^3$ increase in $PM_{2.5}$ for the corresponding period. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 6: IV Cumulative Effects of Pollution: Different Number of Lags and Segments

	Lag k				
Segments z	30 days	60 days	90 days	120 days	150 days
1	1.18***	2.12***	2.42***	2.60***	2.58*
	(0.25)	(0.51)	(0.69)	(0.98)	(1.48)
2	1.41***	2.26***	2.67***	2.80***	2.62*
	(0.25)	(0.52)	(0.69)	(0.95)	(1.43)
3	1.28***	2.16***	2.65***	2.74***	2.41*
	(0.25)	(0.49)	(0.68)	(0.93)	(1.40)

Notes: The dependent variable is log(number of transactions). Each row indicates the number of segments for the cubic B-splines. Each column reports the cumulative percentage change in the dependent variable in response to a 10 μg/m^3 increase in $PM_{2.5}$, over different numbers of days. Same IVs and controls as in Table 5. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

Table 7: IV Cumulative Effects of Pollution: Different Buffer Zone Radii

	Radius for the Buffer Zone					
	50 km	100 km	150 km	200 km	250 km	300 km
First Stage Regression						
N	192,586	192,586	192,586	192,586	192,586	192,586
R^2	0.502	0.486	0.474	0.467	0.464	0.462
First-stage F	34.48	46.69	38.35	34.14	35.36	35.33
IV Regression						
Total Long-Term Effect	2.56***	2.42***	2.65***	2.86***	2.86***	2.88***
	(0.78)	(0.60)	(0.68)	(0.71)	(0.72)	(0.70)

Notes: The dependent variable is log(number of transactions). Each column uses a different buffer zone radius in constructing the instruments and reports the cumulative percentage change in the dependent variable in response to a permanent 10 μg/m^3 increase in $PM_{2.5}$. Same controls as in Table 5. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 8: IV Cumulative Effects of Pollution: Controlling for O_3, SO_2, NO_2 and CO

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Current Day	0.09***	0.09***	0.05	0.12***	0.16**	-0.16***	-0.05**
	(0.02)	(0.03)	(0.04)	(0.04)	(0.08)	(0.03)	(0.02)
Current + Past 3d	0.32***	0.31***	0.16	0.40***	0.56**	-0.51***	-0.18***
	(0.07)	(0.08)	(0.12)	(0.13)	(0.24)	(0.09)	(0.07)
Current + Past 7d	0.50***	0.49***	0.29	0.65***	0.91**	-0.73***	-0.31***
	(0.11)	(0.12)	(0.18)	(0.19)	(0.37)	(0.14)	(0.10)
Current + Past 14d	0.64***	0.63***	0.48**	0.89***	1.27**	-0.71***	-0.44***
	(0.14)	(0.16)	(0.22)	(0.23)	(0.50)	(0.16)	(0.13)
Current + Past 28d	0.84***	0.82***	0.93***	1.19***	2.01**	-0.49**	-0.44**
	(0.22)	(0.25)	(0.31)	(0.27)	(0.79)	(0.23)	(0.20)
Current + Past 56d	1.87***	1.60***	2.24***	1.91***	4.51***	-0.88**	-0.30
	(0.43)	(0.47)	(0.55)	(0.46)	(1.55)	(0.41)	(0.36)
Current + All Lags	2.55***	2.07***	2.73***	2.01***	6.21***	-0.55	-0.69
	(0.69)	(0.72)	(0.91)	(0.76)	(2.34)	(0.58)	(0.46)
N	141,779	141,642	141,552	137,838	110,244	141,755	141,637
First-stage F	39.76	39.85	39.75	41.61	50.98	39.71	39.71

Notes: The dependent variable is log(number of transactions). Same IVs as in Table 5. In addition to controls in Table 5, daily average concentration levels of O_3, SO_2, NO_2 and CO are included. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 9: IV Cumulative Effects of Pollution: Controlling for Economic Spillovers

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Current Day	0.11***	0.10***	0.05	0.13***	0.19***	-0.15***	-0.05***
	(0.02)	(0.03)	(0.04)	(0.04)	(0.07)	(0.03)	(0.02)
Current + Past 3d	0.35***	0.34***	0.17	0.44***	0.64***	-0.46***	-0.19***
	(0.07)	(0.08)	(0.12)	(0.13)	(0.23)	(0.09)	(0.07)
Current + Past 7d	0.55***	0.54***	0.31*	0.71***	1.03***	-0.66***	-0.32***
	(0.11)	(0.13)	(0.18)	(0.19)	(0.37)	(0.14)	(0.10)
Current + Past 14d	0.70***	0.68***	0.50**	0.95***	1.39***	-0.65***	-0.43***
	(0.14)	(0.17)	(0.21)	(0.23)	(0.51)	(0.16)	(0.13)
Current + Past 28d	0.89***	0.87***	0.96***	1.23***	2.13***	-0.46**	-0.41*
	(0.22)	(0.25)	(0.30)	(0.27)	(0.80)	(0.23)	(0.21)
Current + Past 56d	1.94***	1.66***	2.27***	1.99***	4.66***	-0.86**	-0.22
	(0.42)	(0.47)	(0.54)	(0.46)	(1.57)	(0.41)	(0.36)
Current + All Lags	2.62***	2.15***	2.76***	2.12***	6.37***	-0.56	-0.56
	(0.68)	(0.72)	(0.89)	(0.76)	(2.34)	(0.59)	(0.47)
N	138,390	138,254	138,164	134,544	107,345	138,366	138,250
First-stage F	37.53	37.49	37.54	38.91	45.28	37.49	37.48

Notes: The dependent variable is log(number of transactions). Same IVs as in Table 5. In addition to controls in Table 5, we include the average pollution level in cities outside of the buffer zone but within the same region to control for regional economic spillovers. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 10: Cumulative Effects of Pollution: IV constructed using source cities' time-invariant pollution

	Health-Related Consumption					Control Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
Current Day	0.11***	0.18***	0.01	0.21***	0.27**	-0.22***	-0.10***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.12)	(0.05)	(0.03)
Current + Past 3d	0.37***	0.55***	0.07	0.68***	0.83**	-0.70***	-0.35***
	(0.09)	(0.11)	(0.15)	(0.18)	(0.37)	(0.14)	(0.11)
Current + Past 7d	0.54***	0.77***	0.19	0.99***	1.13**	-1.04***	-0.54***
	(0.14)	(0.16)	(0.22)	(0.27)	(0.54)	(0.21)	(0.16)
Current + Past 14d	0.62***	0.77***	0.50*	1.07***	1.01	-1.13***	-0.64***
	(0.18)	(0.20)	(0.28)	(0.34)	(0.69)	(0.26)	(0.20)
Current + Past 28d	0.83***	0.84***	1.25**	1.19***	0.86	-0.90**	-0.57*
	(0.29)	(0.29)	(0.51)	(0.40)	(0.96)	(0.36)	(0.31)
Current + Past 56d	2.34***	2.20***	3.25***	2.49***	3.19***	-1.65***	-0.44
	(0.40)	(0.38)	(0.76)	(0.55)	(1.21)	(0.49)	(0.44)
Current + All Lags	2.69***	2.17***	3.83***	2.16**	2.41	-1.74***	-0.56
	(0.64)	(0.58)	(1.22)	(0.90)	(1.99)	(0.67)	(0.57)
N	141794	141657	141567	137853	110259	141770	141652
First-stage F	26.96	27.02	26.95	26.88	29.34	26.79	26.93

Notes: The dependent variable is log(number of transactions). The effect of current and past air pollution is estimated using Flexible Distributed-Lag Model with 90 lags. The IVs are various functions (both current and lagged) of wind patterns, weather conditions, and the *average* (time-invariant) level of $PM_{2.5}$ in cities more than 150 km away. Same controls as in Table 2. Each row reports the cumulative percentage change in the dependent variable in response to a 10 $\mu g/m^3$ increase in $PM_{2.5}$ for the corresponding period. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Cluster-robust Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 11: IV Cumulative Effects of Pollution: Drop Major Cities

	Health-Related Consumption					Control Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
Current Day	0.12***	0.12***	0.06	0.14***	0.18**	-0.15***	-0.06***
	(0.02)	(0.03)	(0.04)	(0.04)	(0.08)	(0.03)	(0.02)
Current + Past 3d	0.39***	0.39***	0.22*	0.46***	0.60**	-0.46***	-0.22***
	(0.07)	(0.08)	(0.13)	(0.13)	(0.24)	(0.09)	(0.07)
Current + Past 7d	0.59***	0.60***	0.37**	0.74***	0.93**	-0.65***	-0.35***
	(0.11)	(0.12)	(0.19)	(0.19)	(0.37)	(0.14)	(0.10)
Current + Past 14d	0.70***	0.70***	0.56**	0.98***	1.20***	-0.62***	-0.45***
	(0.12)	(0.14)	(0.22)	(0.23)	(0.46)	(0.16)	(0.13)
Current + Past 28d	0.80***	0.77***	0.99***	1.28***	1.79***	-0.41*	-0.40*
	(0.17)	(0.18)	(0.31)	(0.27)	(0.65)	(0.23)	(0.21)
Current + Past 56d	1.72***	1.42***	2.27***	2.09***	4.05***	-0.82**	-0.20
	(0.33)	(0.32)	(0.56)	(0.46)	(1.26)	(0.42)	(0.37)
Current + All Lags	2.25***	1.75***	2.75***	2.21***	5.50***	-0.48	-0.51
	(0.55)	(0.51)	(0.92)	(0.76)	(2.03)	(0.60)	(0.48)
N	136067	135930	135840	132126	104532	136043	135925
First-stage F	37.00	37.03	37.06	38.16	47.25	36.98	36.93

Notes: The dependent variable is log(number of transactions). The effect of current and past air pollution is estimated using Flexible Distributed-Lag Model with 90 lags. Same IVs and controls as in Table 5. The following large cities are dropped from the analysis: Beijing, Shanghai, Guangzhou, Shenzhen, Wuhan, Chongqing, Chengdu, and Nanjing. Each row reports the cumulative percentage change in the dependent variable in response to a 10 μg/m^3 increase in PM$_{2.5}$ for the corresponding period. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Cluster-robust Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 12: IV Cumulative Effects of Pollution: Avoidance Behavior

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\{P_{+1} > P_0\}$	0.41***	-0.15	0.95***	-0.11	0.29	0.36**	0.95***
	(0.11)	(0.13)	(0.19)	(0.20)	(0.32)	(0.15)	(0.14)
Current Day	0.13***	0.13***	0.07*	0.15***	0.20***	-0.14***	-0.05***
	(0.02)	(0.03)	(0.04)	(0.04)	(0.07)	(0.03)	(0.02)
Current + Past 3d	0.41***	0.41***	0.25**	0.49***	0.68***	-0.44***	-0.18***
	(0.07)	(0.08)	(0.12)	(0.13)	(0.24)	(0.09)	(0.07)
Current + Past 7d	0.63***	0.63***	0.42**	0.78***	1.09***	-0.63***	-0.31***
	(0.11)	(0.12)	(0.18)	(0.19)	(0.37)	(0.13)	(0.10)
Current + Past 14d	0.77***	0.77***	0.60***	1.01***	1.48***	-0.63***	-0.43***
	(0.14)	(0.16)	(0.22)	(0.23)	(0.51)	(0.16)	(0.13)
Current + Past 28d	0.95***	0.94***	1.03***	1.29***	2.25***	-0.45*	-0.41*
	(0.22)	(0.25)	(0.31)	(0.27)	(0.80)	(0.23)	(0.21)
Current + Past 56d	2.02***	1.77***	2.34***	2.07***	4.82***	-0.88**	-0.24
	(0.43)	(0.47)	(0.54)	(0.46)	(1.59)	(0.42)	(0.37)
Current + All Lags	2.71***	2.27***	2.83***	2.22***	6.57***	-0.60	-0.60
	(0.69)	(0.72)	(0.90)	(0.75)	(2.38)	(0.59)	(0.48)
N	141,272	141,136	141,046	137,347	109,862	141,248	141,132
First-stage F	37.76	37.79	37.77	38.88	45.27	37.72	37.70

Notes: The dependent variable is log(number of transactions). Same IVs as in Table 5. Besides controls used in Table 5, a flag of whether pollution level the next day is worse than current day is included to control for avoidance behavior. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table 13: Summary of the Dose-Response Relationship from Literature

Source	Dose, additional	Response
Mu and Zhang (2016)	100-point AQI	54.5% increase in mask purchases, 70.6% increase in anti-$PM_{2.5}$ mask purchases
Williams and Phaneuf (2016)	1 std. dev. of $PM_{2.5}$ (3.78 $\mu g/m^3$)	8.3% more spending on asthma and COPD
Schlenker and Walker (2015)	1 std. dev. of pollution	17% more asthma and other respiratory incidences, 9% more heart incidences
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% increase in all-cause mortality rate 285,000 more premature deaths each year
Chay and Greenstone (2003)	1% TSP	0.35% increase 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 per year
Our estimation		
IV	10 $\mu g/m^3$ $PM_{2.5}$	2.6% increase in hospital visits and pharmacy purchases, 1.5% increase in total health expenditure WTP: \$11.3 per household per year from morbidity reduction

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:

$$\max_{\{m_i, c_i, o_i\}} U[h_i, c_i, o_i, e(a, m_i + c_i)],$$
$$\text{s.t. } y(h_i) \equiv y_0 + w(h_i) = \pi + pm_i + c_i + o_i,$$
$$\text{and } h_i = h_0 + m_i - g_i(e(a, m_i + c_i)),$$

The Lagrangian can be written as:

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

The first-order conditions are:

$$\frac{\partial L_i^*}{\partial m_i} = U_h(1 - g_i'(e_i)e_m) + U_e e_m + \lambda_i(y_h(1 - g_i'(e_i)e_m) - p) = 0,$$
$$\frac{\partial L_i^*}{\partial c_i} = -U_h g_i'(e_i)e_c + U_c + U_e e_c - \lambda_i(y_h g_i'(e_i)e_c + 1) = 0,$$
$$\frac{\partial L_i^*}{\partial o_i} = U_o - \lambda_i = 0,$$
$$\frac{\partial L_i^*}{\partial \lambda_i} = y(h_i^*) - \pi - pm_i^* - c_i^* - o_i^* = 0.$$

where U_h, U_c, U_o 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-healthcare spending decisions? Because air pollution a affects a consumer's wage income, a effectively governs the relative prices of non-healthcare consumption c_i and health consumption m_i with respect to online spending o. 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.[43] 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 decrease

[43] 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 $y_h g_i'(e)e_c + 1$, 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 increase.

There is also an income effect: the reduction in income causes both c_i and m_i to decrease. Finally, a directly lowers the consumer's utility by increasing exposure e_i and decreasing the health stock h_i. Assuming that consumption and health are complements, the marginal utility of consumption is non-decreasing in h_i and non-increasing in e_i. Thus a decrease in h_i and an increase in e_i caused by an increase in a lead to a further decrease in c_i. Taking into account of all these effects, an increase in air pollution unambiguously causes c_i to decrease.

The effect of an increase in a on m_i is theoretically ambiguous, 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:

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

By the Envelope Theorem,

$$\begin{aligned}\frac{\partial L_i^*}{\partial a} &= -U_h g_i'(e_i)e_a + U_e e_a - \lambda_i y_h g_i'(e_i)e_a - \lambda_i \pi'(a) \\ &= U_e e_a - g_i'(e_i)e_a(U_h + \lambda_i y_h) - \lambda_i \pi'(a) \qquad (1) \\ \frac{\partial L_i^*}{\partial y_0} &= \lambda_i\end{aligned}$$

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

$$\frac{dh_i^*}{da} = \frac{\partial m_i^*}{\partial a} - g_i'(e_i)e_a - g_i'(e_i)e_c\left(\frac{\partial m_i^*}{\partial a} + \frac{\partial c_i^*}{\partial a}\right) \qquad (2)$$

$$\frac{de_i^*}{da} = e_a + e_c\left(\frac{\partial m_i^*}{\partial a} + \frac{\partial c_i^*}{\partial a}\right) \qquad (3)$$

Rearranging terms, we obtain the following relations:

$$\lambda_i p = U_e e c_c + (U_h + \lambda_i y_h)(1 - g_i'(e_i)e_c) \qquad (4)$$

$$U_c - U_o = -U_e e_c + (U_h + \lambda_i y_h) g_i'(e_i) e_c \qquad (5)$$

Plugging these equations into equation (1), we obtain[44]:

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

The marginal WTP for individual i is then equal to:

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

We assume that premiums cannot adjust in response to pollution: $\frac{d\pi}{da} = 0$. This seems reasonable in the context of China, where the insurance reimbursement rates for the 3 major public insurance programs are fixed by the government and don't depend on individuals' pollution exposure. The MWTP for individual i can be simplified as:

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

[44] In the first line, we plug in (2) and (3). In the second line, we re-arrange and collect terms. To get to the third line, we plug in (4) and (5).

B B-Spline Construction

Let β_i denote the impact of pollution exposure i days in the past on today's spending. We assume that β_i can be approximated by a set of basis functions B_j in i:

$$\beta_i = \sum_j \gamma_j B_j(i)$$

Section 4.1 discusses the case where the basis function $B_j(i)$ is a polynomial function of i up to the 3rd order: $\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2 + \gamma_3 i^3$. We now describe how to extend this to the more general case where $B_j(i)$ is an r-th order B-spline in i with z segments.

To do so we introduce some new notation. Let the support of i be $[0, \bar{s}]$. We divide the support into z sub-intervals by a vector of $z+1$ knots $\mathbf{t} = [t_0, t_1 \cdots, t_z]$, where $t_0 = 0$ and $t_z = \bar{s}$. The r-th order B-spline, which is equivalent to a piecewise polynomial of order $r-1$ (enforcing C^{r-2} continuity), can be constructed from a set of basis functions:

$$B_{j,r}(i|\mathbf{t}) = (t_{j+r} - t_j) \sum_{k=0}^{r} \left[\prod_{0 \le h \le r, h \ne k} (t_{j+h} - t_{j+k}) \right]^{-1} (i - t_{j+k})_+^{r-1}$$

where

$$(i - t_{j+k})_+^{r-1} = \mathbb{1}(i > t_{j+k}) \cdot (i - t_{j+k})^{r-1}$$

Since there are z subintervals and the order of the spline is r, there will be $z+r-1$ such B-splines. We can now define β_i as a linear combination of these B-splines, where γ_j are unknown parameters to be estimated:

$$\beta_i = \sum_{j=1-r}^{z-1} \gamma_j B_{j,r}(i|t_j, \ldots, t_{j+r})$$

To implement this, the econometrician has to determine both the order of the spline, $r-1$, and the number of segments, z. In our benchmark estimates, we set $z = 3$ and $r = 4$. By choosing $r = 4$, we approximate β_i with a linear combination of cubic B-splines. The cubic B-spline is equivalent to a piecewise cubic polynomial with smoothness constraints at each knot up to the 2nd order derivative, which makes the spline twice continuously differentiable.

C Additional Results

Table C1: Coverage of UnionPay Cards in 2015

	Log(No. of cards per capita)	
	(1)	(2)
log(household income)	1.556***	1.362***
	(0.093)	(0.126)
Years of education	0.156***	0.327***
	(0.041)	(0.055)
Average age	-0.040***	0.005
	(0.012)	(0.014)
Constant	-13.00***	-13.82***
	(0.662)	(0.983)
Province fixed effects	No	Yes
No. of obs.	287	287
R^2	0.682	0.831

Notes: The unit of observation is a city. The dependent variable is the log of number of active bank cards per capita in 2015 as shown in Figure 3. The city-level demographics (income, education, and age) are from the 2005 Census.

Table C2: IV Estimates of Pollution Impacts on Value of Transactions with 90 Lags

	Health-Related Consumption					Comparison Groups	
	Health	All Hospital	Pharmacy	People's	Children's	Necessities	Supermarket
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Current Day	0.07***	0.07**	0.01	0.10**	0.01	-0.12**	-0.04
	(0.02)	(0.03)	(0.05)	(0.05)	(0.09)	(0.05)	(0.05)
Current + Past 3d	0.23***	0.25***	0.04	0.34**	0.05	-0.38**	-0.16
	(0.08)	(0.08)	(0.15)	(0.15)	(0.30)	(0.17)	(0.17)
Current + Past 7d	0.36***	0.38***	0.04	0.54**	0.13	-0.55**	-0.36
	(0.11)	(0.13)	(0.21)	(0.23)	(0.45)	(0.24)	(0.25)
Current + Past 14d	0.42***	0.46***	0.03	0.69**	0.33	-0.56**	-0.68**
	(0.15)	(0.17)	(0.24)	(0.27)	(0.59)	(0.27)	(0.31)
Current + Past 28d	0.43*	0.44	0.29	0.79***	1.09	-0.34	-0.99**
	(0.25)	(0.29)	(0.34)	(0.30)	(0.88)	(0.34)	(0.43)
Current + Past 56d	1.04**	0.83	1.64***	1.15**	4.07**	-0.32	-0.95
	(0.47)	(0.54)	(0.61)	(0.47)	(1.72)	(0.58)	(0.75)
Current + All Lags	1.47**	1.08	1.96**	1.20	6.12**	0.54	-0.66
	(0.70)	(0.78)	(0.96)	(0.83)	(2.61)	(0.87)	(1.04)
N	141,794	141,656	141,566	137,854	110,257	141,757	141,641
First-stage F	38.35	38.38	38.37	39.68	47.79	38.26	38.30

Notes: The dependent variable is log(value of transactions). Same IVs and controls as in Table 5. Each row reports the cumulative percentage change in the dependent variable in response to a 10 μg/m^3 increase in $PM_{2.5}$ for the corresponding period. Standard errors are in parentheses, clustered at the city level. Significance levels are indicated by *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Kleibergen-Paap Wald rk F-statistics are reported in the last row and are cluster-robust at the city level.

Table C3: Mortality Cost Calculation

Age group	Urban population	Rural Population	Urban mortality rate (per 100,000)	Rural mortality rate (per 100,000)	VSL in million (2015$)	Mortality impact in percentage
20-24	73,195,616	58,048,857	4.31	3.81	0.2106	10
25-29	59,414,692	44,637,171	5.47	5.54	0.2106	11
30-34	57,695,497	42,364,156	8.07	7.07	0.2106	14
35-39	66,981,015	54,594,597	13.71	12.48	0.2106	10
40-44	65,704,887	62,801,076	26.02	23.90	0.1895	12
45-49	55,242,460	53,527,870	42.25	46.27	0.1684	13
50-54	40,364,926	40,756,761	65.87	71.27	0.1474	13
55-59	38,563,476	45,194,486	105.52	125.79	0.1263	12
60-64	26,819,982	33,611,729	209.62	255.81	0.1053	12
65-69	18,448,986	23,900,786	402.25	459.16	0.0842	11
70-74	15,221,689	18,742,359	880.11	1092.46	0.0632	9
75-79	10,848,240	13,721,250	1744.92	1998.33	0.0421	7
80-84	5,936,146	7,839,253	3632.06	4316.95	0.0316	5
85 and above	3,370,721	4,474,484	9685.26	13128.58	0.0211	3

Notes: The population data are for 2015. The mortality rates per 100,000 are only for cardiorespiratory diseases and are from the 2015 National Health Statistics. The VSL is calculated using the transfer elasticity of 1.2 based on the VSL of 2.27 million (in 2015$) estimated for the U.S. population from Ashenfelter and Greenstone (2004). The estimated VSL for the Chinese population is $0.2106 million for a prime age person. The age adjustment for VSL is based on Murphy and Topel (2006). The mortality calculation follows closely with Deschenes et al. (2017). The mortality impact (in percentage) is the estimated impact of a 10 μg/m^3 increase in PM_{10} on cardiorespiratory mortality during life cycle from Table S6 in Ebenstein et al. (2017).