The Human Impacts of Air Pollution: Three Studies Using Internet Metrics

Mingying Zhu

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfilment of the requirements for the degree: Doctorate in Philosophy Economics

Department of Economics

Faculty of Social Sciences

University of Ottawa

Abstract

Chapter 1.—We provide first evidence of a link from daily air pollution exposure to sleep loss in a panel of Chinese cities. We develop a social media-based, city-level metric for sleeplessness, and bolster causal claims by instrumenting for pollution with plausibly exogenous variations in wind patterns. Estimates of effect sizes are substantial and robust. In our preferred specification, a one standard deviation increase in AQI causes an 11.6% increase in sleeplessness. The results sustain qualitatively under OLS estimation but are attenuated. The analysis provides a previously unaccounted-for benefit of more stringent air quality regulation. It also offers a candidate mechanism in support of recent research that links daily air quality to diminished workplace productivity, cognitive performance, school absence, traffic accidents, and other detrimental outcomes.

Chapter 2.—We provide linear and non-parametric estimates of the causal impact of short-term exposure to polluted air on the prevalence of cough in a panel of a hundred Chinese cities. In our central estimate, which exploits plausibly-exogenous variations in the number of agricultural fires burning in the vicinity as an instrument, we find that a one standard deviation increase in airborne pollution causes a roughly 5% increase in the prevalence of cough in the affected city. Amongst pollutants the effect can be tied specifically to particulate matter ($PM_{2.5}$). The results prove resilient in a series of robustness tests and falsification exercises.

Chapter 3.—We provide the first study of the relationship between air pollution and students' migration intentions for higher education. Young people's interest in local study is proxied by their Baidu search index for local universities. The IV method is supplemented to identify the causal link by instrumenting for particular matter with plausibly exogenous variations in temperature inversion strength. The estimates of effect sizes are substantial and robust. When air quality in Beijing moves from good-day level to moderately-polluted level, people's search for local education decreases by 3.8% under OLS and 11.8% under IV. The results release the signal that people lost their interest in local universities due to the elevated air pollution. There could be future out-migration to cleaner cities for higher education.

Declaration

I acknowledge the contribution of Professor Anthony Heyes for the first chapter. His contribution is equal to my own. The article in this chapter has been revised and resubmitted to the *Journal of Environmental Economics and Management*.

Acknowledgements

I would like to express my great thanks to my supervisor Professor Anthony Heyes. Thanks for his patient guidance and insightful comments. PhD research work is long and sometimes arduous. He always generously shares his new idea to us and encourages us to keep improving. I still remember his words at the beginning when I became his student in 2016, “Time goes fast.”, “If I did not work for a week, the world totally changed to me.”, “You should have novel ideas.”, “You need to really think.”, “Thinking accounts for 70% of the research, and 30% is writing.”, etc. All he said will have a profound impact on my later research work. Professor Heyes is a hard worker and very responsive to his students. I learned a lot from him about how to do the research, and how to guide the students.

Two anonymous referees from *Journal of Environmental Economics and Management* provided excellent suggestions for the work in my first chapter, not only helping me solve the major concerns in the identification strategy, but also helping me to develop the better design in the later research.

I would also like to thank my family. They supported me a lot in my life and even travelled far away from home to Canada to take care of me. They are the most important people to me. Many thanks to my parents, my husband, and my parents-in-law.

General Introduction

In this thesis we investigate the causal link from short-term air pollution to a series of novel outcome variables. In particular, (1) sleep quality, (2) cough and (3), plans to migrate. Each study is based on China, where poor air quality is widely-understood to be an important social issue, focussing on fine particulate matter ($PM_{2.5}$) and the multi-pollutant Air Quality Index (AQI). For dependent variables in each case, we exploit data derived from internet behavior: In particular search behavior on the very widely-used Baidu search engine and posts on the ubiquitous social media site Weibo.[1] Data is scraped and used to develop pertinent proxies in each case. In each chapter causal claims are bolstered by the use of instrumental variables (IV) methods, in order to address endogeneity concerns. Results in each chapter are highly statistically significant, prove to be resilient in a wide array of robustness tests and falsification exercises, and are substantial in size. Our analysis further broadens and deepens our understanding of the various ways in which the air pollution levels seen in Chinese cities, and in many other parts of the world, can damage human well-being, and identifies previously unaccounted-for benefits of more stringent regulation of air quality.

Chapter one explores the causal impact of air pollution on sleeplessness in the "first-tier" Chinese cities (this is an official designation of the largest and most-developed 19 cities). We build a nightly, city-level measure for sleeplessness, based on a count of the intensity with which inhabitants in each city are searching for keywords meaning 'can't sleep', 'sleepless',

[1] These are often referred to as the "Chinese Google" and "Chinese Twitter" respectively. Most of the websites commonly used in North America and Europe are barred in China, including Google, Twitter, Facebook, Youtube, Instagram, Snapchat and many others.

etc. in Chinese characters. Ambient pollutant and weather covariates are all integrated from hourly data into daily averages by city. We first use the ordinary-least squares (OLS) method to estimate the association between air pollution and sleeplessness in a straightforward panel fixed effects setting, finding a strong positive association. Non-linear effects are also discussed with the inclusion of different pollutant bins.

Due to the potential measurement error in air pollution and omitted unobservables, we supplement the OLS results with IV estimation, instrumenting for the air pollution in the target cities with the weighted average pollution from surrounding cities, similar to the approach used in Bayer et al. (2009) and Schlenker and Walker (2016). Variations in pollution here are driven by plausibly-exogenous variations in wind direction. The validity of the instrument is discussed in detail in the first chapter.

The IV estimates are the same in sign as those generated by OLS, but the estimated sizes are larger, consistent with the OLS estimates being attenuated. They are robust to a wide set of checks, including alternative constructions of the instrument, inclusion of fixed effects in many different permutations, the conduct of a number of sub-sample exercises, and so on. Our preferred specification implies that, for the composite air-quality index, moving from a median clean decile day to a median dirty day (that is from the 5th to the 95th percentile when days are ranked from clean to dirty) increases city-level sleeplessness by 36.3% relative to the mean. For $PM_{2.5}$ the number is 37.3%.

The second chapter analyses the short-term exposure to polluted air on the prevalence of cough in a panel of the hundred largest cities in China. Of course there is a very extensive existing literature on the relationship between air pollution and serious health outcomes such as mortality, heart attacks, and asthma attacks that manifest themselves in hospitalization or other interaction with health-care professionals. Our focus is different. We target a sub-clinical health condition - the phenomenon of cough - which is a very common ailment, and one which substantially damages human well-being, but which does not typically show up in hospitalization or in other medical data.

As a daily, city-level metric for coughing, we exploit people's web-search behavior on the Baidu platform. Considering plausible lags among pollution exposure, cough symptom and web-search behavior, we integrate our daily sample into three-day time bands. We first use the OLS method to search for evidence of a positive association between the city-level intensity of searches for cough keywords and air pollution levels ($PM_{2.5}$ and AQI) in the study cities. We also explore the possibility of non-linear effects. In the complementary IV analysis, we construct a second instrument for air pollution based on the daily average number of agricultural fires burning at any given time within 150km of the city in question. Instrument validity requires that the fire points within the neighborhood of a particular city impact the propensity to cough by inhabitants of the city only through its contribution to increased air pollution in the city. This is our primary identifying assumption. We also conduct a set of robustness checks, including strategy of altering the time unit, using an alternative fire count radius, accounting for different wind directions, considering confidence-unweighted fires, restricting the sample to a high fire month, and using alternative time fixed effects. Our central estimate means that a one standard deviation increase in $PM_{2.5}$ in the air causes a roughly 5% increase in the prevalence of cough.

The third chapter explores the causal effect of air pollution in a single city, Beijing, on the mobility aspirations of young people. We rely on a measure that combines the intensity with which people in the city engage in searches for universities in Beijing, compared to those in other, cleaner cities. We compute the daily share of Baidu search index for universities in Beijing over the total search for all universities, and link it to the variation of $PM_{2.5}$ in Beijing. The result from OLS method displays the negative relation, implying that people have less search interest in local education, when local air pollution gets worse. The plausible linear relationship is identified by including the categorical bins for $PM_{2.5}$, as well as using the spline function with corresponding knots.

Beijing has very frequent temperature inversions due to its special topography, with the occurrence almost every one or two days. On days when temperature is inverted, pollutants generated in the city are kept pinned to the ground, rather than the air ventilating, and pollution levels are substantially higher. In this chapter, temperature inversion is used as an instrument for local $PM_{2.5}$, following a number of recent papers, including Jans et al. (2018), Chen et al. (2017) and Sager (2016). The underlying assumption for the validity of our instrument is that people's search behavior is not affected directly by temperature inversion, only through its detrimental implications for air quality.

The estimated effect sizes from our IV exercise are substantial and robust to a battery of tests. In our preferred specifications, moving from the 5th to the 95th percentile in terms of air quality reduces the propensity of people in Beijing to search for information on local universities to fall by 3.6% under OLS and 11.2% under IV. There is a corresponding increase in searches for universities outside of Beijing, with the effect particularly pronounced for universities in cleaner cities. The study adds to the nascent literature on how environmental factors might motivate intention to migrate.

Contents

List of Tables

List of Figures

Chapter 1

Air Pollution as a Cause of Sleeplessness: Social Media Evidence from a Panel of Chinese Cities

1.1 Introduction

Our objective in this paper is to investigate a possible causal effect of urban air pollution on the sleep of city inhabitants. Air quality - particularly in cities - is one of the great policy challenges of our time. Understanding the full range of negative impacts of pollution is an essential prerequisite for welfare evaluation of policy interventions.

Sleep is an essential input to human well-being. Loss of sleep reduces mental function along various dimensions, such as learning (Huber et al., 2004), memory (Diekelmann and Born, 2010), judgement (Killgore et al., 2006), speed of reflex (Maquet, 2001) and emotional balance (Ireland and Culpin, 2006). It is correlated with lower self-reported well-being (Hamilton et al., 2007; Steptoe et al., 2008). Tiredness - the inevitable consequence of sleeplessness - has been causally linked to various negative outcomes, including road traffic accidents (Valent et al., 2010), reduced workplace productivity (Zammit et al., 2010; Rosekind et al., 2010), industrial injuries (Barnes and Wagner, 2009), absenteeism (Daley et al., 2009), deteriorated relationship quality (Gordon and Chen, 2014), domestic violence (Meijer et al., 2010), and compromised school performance (Chung and Cheung, 2008). In terms of health outcomes, shortage of sleep over various time scales has been linked to reduced functioning of immune systems and subsequent increased susceptibility to disease, increased risk of hypertension, cardiac and breathing problems, increased adiposity, and negative mental health outcomes.[1]

It is not a surprise that both individuals and governments invest in protecting sleep, and that individuals when asked express a substantial willingness-to-pay to avoid sleep loss (Pollinger, 2014; Delfino et al., 2008).[2] In summary, given that the typical adult in most societies spends between 7 and 8 hours of each day engaged in the activity of sleep (and children longer): "If sleep does *not* serve an absolutely vital function, then it is the biggest mistake the evolutionary process has ever made." (Rechtschaffen, 1971).

Despite the centrality of sleep to humans, and the diverse contributions that it makes to individual and societal well-being, economic analysis of it has been cursory. Biddle and Hamermesh (1990) treat sleep choice as a time allocation problem. Similarly, Asgeirsdottir and Zoega (2011) provide a model of sleep behavior as an investment that an individual makes in the level of alertness he then enjoys during the day, in the spirit of the approach taken to health as human capital.

While the channels that might link pollution exposure to lower quantity or quality of sleep are obvious (shortness of breath, elevated heart-rate, irritation of upper airways, eyes etc.), research linking pollution exposure to sleep outcomes is limited. (1) Strøm-Tejsen et al. (2016) manipulate indoor air quality in the campus bedrooms of 16 students and find

[1] There is a large literature on the health implications of both short-term and chronic sleep loss (Altevogt and Colten, 2006; Cappuccio et al., 2010).

[2] For example, individuals spend on good mattresses and other aids to healthful sleep, worry about the noise environment when they buy a home, etc.. Governments spend on sleep research, impose regulation on night-time noise around airports, etc.. Employers are also aware of the benefits of sleep. See for example the lead article *Why Companies are Willing to Pay to Make Sure You Get a Good Night's Sleep* in Executive Style Magazine (21 April 2016) on the productivity benefits of well-rested employees.

that indoor air quality impacts both sleep quality (as measured by subject-worn actigraphs) and next-day performance on math and language tests. (2) Using measures of outdoor air quality in a small number of US cities, Zanobetti et al. (2010) show that the same-night AQI in the city in which the subject resides correlates with likelihood of episodes of sleep apnea (pauses in breathing during sleep) and other physiological correlates of sleep health. This is an important paper to which ours is complementary. The advantage of their methods is that they deliver very precise individual-level metrics of sleep. The downsides relate to its focus on sleep-illness, and the observation of subjects via a polysomnograph (sensors at nose, fingers, face and scalp) - not more natural sleeping circumstances - and at much lower levels of ambient pollution than we see in the Chinese cities that we study. (3) Focusing on long-term exposure, and without using tools that would allow for causal inference, Billings et al. (2017) find a negative association between sleep efficiency amongst a sample of older people and 5-year and 1-year measures of $PM_{2.5}$ in the neighborhoods of the six US cities in which they live.

Sleep loss is a significant problem in China (Luo et al., 2013) and elsewhere. For the 19 largest Chinese cities, we construct a nightly, population-level measure of sleeplessness using frequency of use of the Chinese characters meaning 'can't sleep', 'sleepless' etc. on the very widely-used social media site Weibo.[3] We estimate an equation using OLS to characterize a positive association between that measure and same-day local air quality. To reinforce our causal interpretation of this relationship, we apply IV methods, using plausible exogenous

[3]We will be careful to qualify our use of the term "population level" in the data section. Population-level behavior on various internet platforms is increasingly being exploited by social scientists. Choi and Varian (2012) show that Google search data can be used to predict demand for automobiles, home sales and travel behavior. Several papers demonstrate the efficacy of using internet search metrics to predict health outcomes - especially flu - and Google itself established the Google Flu Trends tool in 2008. Goel et al. (2010) show that searches can predict the success of movies, songs and video games. In an environmental application, Herrnstadt and Muehlegger (2014) show that searches for "climate change" and "global warming" in a particular US city are sensitive to short-term deviations of weather from normal. Much recent work has been devoted to Twitter-driven predictive analytics. For three examples among many: Bollen et al. (2011) show that Twitter mood can be used to add explanatory power to stock market forecasts, Gerber (2014) uses Twitter key words to predict crime patterns, and Gayo-Avello et al. (2011) are among several papers using Twitter to predict elections. A central way in which our methods depart from this literature is that we will use measures from social media as our dependent variable. In that regard the paper relates to Baylis (2015), who shows the effect on unusual temperature on Twitter-sentiment.

variations in short-term wind patterns to instrument for air quality. In our preferred specifications, we find that a one standard deviation increase in $PM_{2.5}$ causes an 12.8% increase in sleeplessness relative to mean. The statistical significance and estimated effect size prove to be remarkably robust to a battery of alternative specifications and tests.

We are cautious not to over-interpret the results. Monetizing the sleep loss caused by diminished air quality is beyond the scope of this paper, though it is worth noting that previous research does provide WTP estimates that could be exploited in a back-of-the envelope exercise. The results are instructive in two ways. First, the loss of sleep plausibly impacts the well-being of the affected individual him or herself through a variety of channels. Second, as noted, the results provide a mechanism consistent with recent research linking short-term variations in air quality to reduced workplace productivity (Zivin and Neidell, 2012; Chang et al., 2016), school absence (Currie et al., 2009), exam performance Mendell and Heath (2005), motor vehicle accidents (Sager, 2016) etc..

Section 1.2 details data sources. Section 1.3 describes methods. Section 1.4 and Section 1.5 present main and robustness results. Section 1.6 conducts the placebo test. Section 1.7 summarizes the results from joint estimation. Section 1.8 concludes.

1.2 Data

We investigate the effect of air pollution on sleep in the "first-tier" Chinese cities (19 cities). To do this we develop a nightly, city-level measure of sleep quality derived from posts on social media and connect it to high frequency data on air quality. We also include detailed meteorological data both to control for the likely confounding influences of weather on sleep and for the construction of our instrument.

1.2.1 Sleep

A challenge in this research is to develop a defensible measure of sleeplessness that is a nightly index for how badly (or well) the inhabitants of a particular city are sleeping.

A number of surveys have asked questions about sleep.[4] However none of these provide the temporal granularity that we require (the exact date of interview and some question about short-term, ideally daily, sleep experience). Even if such questions were asked, the resulting responses would be threatened by imperfect recall of respondents, and other shortcomings typical of retrospective survey-derived data.

We exploit what people are saying on the Chinese micro-blogging Weibo. Weibo was launched in August 2009, and growth in its use was explosive, not least because most of the key social media platforms familiar to those living elsewhere (including Twitter, Facebook, Instagram and Youtube) are blocked in China. It is the biggest social media site in China, and by 2016 it had more than 503 million registered and 313 million regular users amongst the 720 million internet users in that country (DeLuca et al., 2016). As with Twitter, messages were - at least during the period that we analyse - subject to a tight word limit. In comparison to Twitter, it has a greater personal than professional orientation in the way it is used (Sullivan, 2012), with substantially more posts outside standard office hours (Gao et al., 2012). Users typically post what they see, hear and think (Cain K, 2015, September 21) and, while it needs to be mined with caution, the content of posts provides the researcher with a potential 'window' into the mind of users and a rich data source.

Keywords

Written Chinese is not alphabetic but rather comprises self-standing characters or glyphs. It is logo-syllabic, which means that a character represents a whole word (physical object, concept, *etc.*). Literacy requires the memorization of a large number of such characters, and a well-educated Chinese person knows more than 4000, while between 2000 and 3000 are needed to read a newspaper (Norman, 1988). This characteristic is helpful to us. By its nature there are many fewer duplicative ways to express concepts than is common in alphabetic languages such as English. "Shimian" and "Shuibuzhao" are the two characters

[4] For example Chen et al. (2004), Yu et al. (2007), and Sun et al. (2015).

that have meaning equivalent to that covered by English words and expressions such as "sleepless", "can't fall asleep", "losing sleep", "insomnia", *etc.*. A further advantage of Chinese is that these are used in the affirmative, so we avoid complications arising from conventions for negation that would arise in most other languages.

We search for the hourly use of these keywords in Weibo posts from users located within each of the "first-tier" cities in China (these are Shanghai, Beijing, Shenzhen, Guangzhou, Chengdu, Hangzhou, Chongqing, Wuhan, Suzhou, Xi'an, Tianjin, Nanjing, Zhengzhou, Changsha, Shenyang, Qingdao, Ningbo, Dongguan, and Wuxi). Weibo offers advanced search tools that enable users to obtain the public posts filtered by keyword, date, time period (minimum duration 1 hour), and location (city). We use these to construct a panel of the number of posts featuring the keywords of interest for each hour of each night (11pm through 7am) for each city for the two year period 2014 and 2015.

It is worth reflecting on this as a dependent variable. The question is not whether keyword use on Weibo is a perfect measure of the thing that we want to measure (the extent to which inhabitants of a particular city are sleeping on a particular night) - of course it is not. Rather, is it a good enough measure, and is it better than others available?

There are two main challenges to our claim that intensity of use of the words "shimian" and "shuibuzhao" provides a valid proxy for city-level sleeplessness. First, perhaps other terms exist that might be used to express the difficulty sleeping that we fail to consider. Inspection of Chinese thesauri and discussion with Chinese speakers make us doubt that this is the case. However, even if it were, it is unlikely to disturb our conclusions. (1) The correlation between use of "shimian" and "shuibuzhao" in our sample is very high (0.96), and the ratio between use of one and the other proves to be insensitive to air quality conditions. We use the word counts as an index, rather than focus on absolute levels. If an additional synonym exists that we have ignored, then provided that focussing its use is closely correlated with these two, then its exclusion is not a concern.[5] (2) Measurement error in the dependent

[5] A problem would arise for us if there was an excluded means of expression whose comparative intensity of use varied systematically with air quality conditions. This seems implausible.

variable that such an oversight would imply should not bias OLS or 2SLS estimates, only reduce their efficiency. We also investigate and refute the possibility that what we are picking up is a simple proxy for overall Weibo use by showing that the sleep metric is uncorrelated with the use of a series of sleep-neutral words (table, cat, etc.), with appearances of the latter not systematically sensitive to air pollution conditions.

Second, Weibo users are not representative of the Chinese population in general. In particular users are younger, more educated, and earn higher income than the broader population (Chan et al., 2012; Chiu al., 2012). While results should most properly be seen as reflecting a treatment effect in the Weibo-using part of the community, we do not see this as problematic. These are likely the high-value workers in Chinese urban society, and disturbance of their sleep can be expected to have correspondingly important economic impacts. Further, there is no reason to think that effects observed in this group would not be observed in the non-Weibo-using part of the population. Indeed, it is plausible to think that those effects could be larger for at least two reasons: (1) In terms of self-protection from pollutants, those with internet access are disproportionately likely to own both air conditioners and air purifiers. (2) Weibo-users are younger than the general population, and most physical effects of pollution are more pronounced among the old.

An additional point to note is that when interpreting coefficients, we assume that the propensity to report sleeplessness is not itself sensitive to pollution conditions. In other words, if we observe a 5% rise in messages about sleeplessness, we take that to imply a 5% increase in sleeplessness in the Weibo-using population. This is similar to the approach in which many researchers have interpreted changes in people reporting health symptoms of pollution, or attending a physician with health symptoms, for example, to reflect changes in the prevalence of those symptoms in the population. If there were some change in the propensity to report - for example via pollution-induced changes in emotional state (as suggested by Zeidner and Shechter (1988) - then the observed change in the proxy would have to be calculated by the actual prevalence multiplied by the propensity. If it happened that the propensity was *increased* by high pollution, then our estimates would over-state the true effect size. However, recent evidence linking short-run exposure to a depressive mood and risk aversion might lead us to speculate that the propensity to message would be reduced. While we are unable to address definitively either possibility, we have no reason to expect any such effect would be significant. However, in the interest of caution, it is worth keeping this in mind when interpreting results.

1.2.2 Pollution

Data on pollution at our locations of interest were collected from www.aqistudy.cn. This website compiles real-time data on pollutants from the Chinese Ministry of Environmental Protection (MEP) and converts it into daily average measures. The pollutants for which we have data are $PM_{2.5}$, CO, NO_2, SO_2, and O_3 (in addition to AQI).[6] Summary statistics for daily ambient measures in our whole sample are included in Table 1.1 (and by city in the Appendix Table A1.1).[7]

Table 1.2 lists the categories of air quality days as defined by the Chinese government and - in the right hand column - the percentage of days in our sample that fall within each category on the AQI measure. Table 1.3 summarizes the correlations between daily city-level measures of the individual pollutants in our sample. In a number of cases, the correlations are quite high, often exceeding 0.6. Most of our analysis will be conducted pollutant by pollutant; only later do we include all pollutants in the same regressions. This follows Schlenker and Walker (2016).

Our analysis is conducted at the city-level, and we calculate air quality measures by taking a simple arithmetic mean of data from all monitors within a city (the number of monitors within our 19 cities varies between 9 and 17). While we know that a user is based in a particular city, we do not know precisely where, nor his or her movements during the

[6] PM_{10} is not reported or studied due to its high correlation with $PM_{2.5}$.

[7] Historically the quality of official data on air quality in China has been questioned. In particular there has been evidence of manipulation around key thresholds (Chen et al., 2012). Stoerk (2016) tests the consistency of official data with Benford's Law, and with US Embassy data, and concludes that it is reliable from 2013.

day. To allay concerns about intra-city variations in pollution conditions, we calculate the correlations between readings at each pair of monitors in each city. The results are reported in Appendix Table A1.2 (and for illustration in detail for Beijing in Appendix Table A1.3 through A1.8). With the exception of CO - a more localized pollutant - pairwise correlations are very high, especially for AQI, $PM_{2.5}$ and O_3 which are close to or above 0.9. In other words pollution measured at any particular monitor is a good indicator of levels across the city.[8]

1.2.3 Weather

Disentangling the potentially confounding effects of weather is important. Weather conditions (in particular temperature, humidity, precipitation) can influence sleep activity directly (Okamoto-Mizuno and Mizuno, 2012; Van, 2006). For example, in recent interesting analysis, Obradovich et al. (2017) identify an effect of external ambient temperature on sleep.

Meteorological data are obtained from the weather stations registered by the World Meteorological Organization (WMO) that are collated by the National Oceanic and Atmospheric Administration (NOAA). The weather variables involved in the study comprise average temperature ($^\circ C$), average humidity (%), sea-level pressure (hPa), wind speed (Km/h), wind direction (°) and precipitation (mm). We combine the hourly weather data into daily mean levels corresponding to the daily average air pollution levels of each city. Summary statistics for the dataset appear in Table 1.1 (and for each city separately in the Appendix Table A1.1).

1.3 Methods

We investigate a link from air pollution in city i on day t to our city-level metric for sleeplessness in that city on that night. In simple terms: if the air in Nanjing is highly

[8] Insofar as measurement error exists in this regressor, we expect it to attenuate OLS estimates, implying that the effect sizes identified under OLS should be interpreted as *under*-stating true effects. The coefficients from the IV exercise will not be subject to such bias.

polluted today, does that damage the quality of sleep in Nanjing tonight?

1.3.1 Ordinary least squares

We first use OLS to estimate the association between air quality and sleeplessness in a straight-forward panel fixed effects setting. We estimate the following specification

$$lnS_{it} = \alpha_0 + P_{it}\beta + W_{it}\gamma + \theta_i + \lambda_t + \epsilon_{it}. \tag{1}$$

S_{it} is the sleeplessness index in city i on the night following calendar date t. lnS_{it} denotes that the outcome variable is logged. P_{it} is the daily average pollutant concentration in city i on date t. The primary pollutants that we consider in turn are $PM_{2.5}$ and the composite AQI measure.

We control for a wide set of potential confounders. W_{it} is a vector of weather controls containing average temperature, average humidity, precipitation, wind speed, and sea-level pressure. The temperature and humidity measures enter as indicators or 'bins' (5 $°C$ indicators for average temperature, 20 % indicators for average humidity) to accommodate possible non-linear effects.[9] θ_i is a city fixed effect that controls for time-invariant city characteristics. λ_t is a vector of time fixed effects, comprising year by month fixed effects, city by year fixed effects, city by quarter fixed effects, day of week and a dummy for holiday dates. ϵ_{it} is the error term.

Our coefficient of interest is β, which relates air pollution to sleeplessness. It can be interpreted as 100*β% increase in sleeplessness due to additional unit of pollutant. Most of the estimated effect sizes that we will report are based on the percentage change due to one standard deviation change in pollutant, which could be computed by multiplying 100*β by one standard deviation (44.993 for $PM_{2.5}$ and 52.202 for AQI).

[9] Results prove to be similar under quadratic estimation, a popular alternative approach to non-linearity.

1.3.2 Single pollutant versus joint estimation

Our initial results are derived from single pollutant models in which regressions are run that incorporate $PM_{2.5}$ without co-emission. There is also an AQI variant, where AQI is a composite measure that captures the 'binding' pollutant on any particular date. We report the joint estimation exercise in Section 1.7. Note that research in this area is plagued by the difficulties of disentangling the effects of *particular* pollutants from the overall cocktail of pollutants that an individual will typically be inhaling on a 'bad air' day.

Some settings do allow for a clean route around this problem. A nice recent example is Lavaine and Neidell (2017). Helpfully for them, the oil refinery strikes that they exploit as exogenous events that temporarily improved air quality in a set of French towns acted on sulphur dioxide in particular, leaving ambient levels of other key pollutants undisturbed. But often the inclusion or exclusion of pollutants is driven by data availability in particular settings. Papers typically report results of regressions that include a single (or limited subset) of pollutants. For example, among well-known investigations of the effect of short term air quality variations on various outcomes; (1) Zivin and Neidell (2012), who look at productivity of agricultural works, select ozone as their pollutant of interest and control only for $PM_{2.5}$. (2) Ransom and Pope (1992), looking at school absences, exploit data only on PM_{10}, finding negative effects.[10] (3) Ebenstein, Lavy and Roth (2016), studying the effect of daily pollution levels on the exam performance of Israeli children, consider only $PM_{2.5}$.[11] (4) Schlenker and Walker (2016), looking at the health impacts of pollution, exploit data on only CO, NO_2 and ozone, and their main results are derived from specifications in which each pollutant is used as explanatory variable sequentially, without controls for the other two (indeed all but one of the eight tables in Schlenker and Walker (2016) report results of single pollutant exercises). They later insert the three pollutants in the same regression,

[10] In the pursuant literature various authors have considered varying permutations of the major pollutants. For example, Gilliland et al. (2001) add ozone and NO_2 and find *beneficial* effects of PM_{10} on absences. Currie et al. (2009) study three of the main pollutants, CO, PM_{10} and ozone.

[11] While in an earlier version (Ebenstein et al., 2016) they also investigate CO, they did not do so simultaneously, and were unable to account for other major pollutants.

which generated a qualitative loss of results.[12]

We are to some extent insulated from these problems because our main estimates derive from IV methods. However, given the (sometimes strong) covariance between pollutants, we will follow Schlenker and Walker's caution in tying effects to particular individual pollutants. As it turns out, our results all work in the same direction - more pollution causing greater sleeplessness. But we are more confident interpreting this as a story about 'dirty air', and circumspect as far as pollutant by pollutant inferences are concerned.

1.3.3 Instrumental variable estimation

There are several challenges to the validity of OLS estimation here. First, there is likely measurement error in pollution. Our theoretical foundation is predicated on the possibility, founded on plausible physiological foundations, that exposure of an individual to elevated levels of pollution increases the chance of disturbed sleep. However, we observe ambient air quality (which we have shown to be comparatively uniform across monitor sites within a particular city on a particular date) rather than individual exposure. For example, we do not observe self-defensive behavior, such as closing of windows and use of air purifiers, which can reduce effective exposure.[13] The measurement error that would be present in the independent variable would lead to attenuated OLS estimates of our coefficient of interest.[14] Second, while we included a rich set of controls for potential confounders - taking particular care with weather - we cannot rule out the presence of omitted variables. For example, air pollution may be positively correlated with unobserved variations in city-level economic

[12]They are explicit in "...acknowledging that we may be picking up the health effects of other pollutants" (page 787). The omission of $PM_{2.5}$ and PM_{10} - with clear links to a variety of cardiovascular and other health outcomes - is a challenge for the interpretation of their results. In an Appendix exercise, they note that this is due to the absence of data. As such they conclude that: "We believe that some amount of caution is warranted in interpreting CO as the unique pollution-related causal channel leading to adverse health outcomes; there may be in fact other unobserved sources of ambient air pollution that covary with CO that may also effect health" (page 800).

[13]In some sense this doesn't matter. What we end up with is not an individual level sleep 'production function' but a population-level effect from ambient conditions to sleep. In terms of defensive behaviors, our results should be interpreted as incorporating such margins of adjustment.

[14]In their investigation of the effects of short-term exposure to health, Moretti and Neidell (2011) provide evidence and insightful discussion of the problems associated with measurement error in this context.

activity, which may in turn influence sleeplessness through other channels. For these reasons we supplement our OLS analysis using two-stage least squares (2SLS), with an instrument based on wind direction.

Instrumental variable

Air pollution in Chinese cities is known to be highly sensitive to wind direction and speed, as pollutants are carried from neighboring cities (Fu et al., 2017). Ambient pollutants, especially fine particles, can travel over long distances by wind, ranging from hundreds to thousands of kilometers (EPA, 1996). The fact that airborne particles can be transported by wind and affect the places on the downwind side has been used in linking air pollution to health outcomes, for example by Schlenker and Walker (2016) in their study of adverse health effects downwind of airports. Bayer et al. (2009) use pollution levels in nearby (but further than 80km) cities to instrument for local pollutant levels. There are also studies that focus on estimating movement of air pollutants between cities (for example Chen and Ye (2015)). We develop an instrument based on plausibly exogenous day-to-day variations in wind patterns which, consistent with the existing literature, proves to have strong relevance (satisfies the first stage). The method is similar to that applied by Schlenker and Walker (2016), but whereas they exploited a single source of emission of pollution (an airport) to any particular neighborhood, our study's cities typically import wind-borne emissions from multiple neighboring cities, requiring that we apply an intuitive weighting scheme.

For each study or target city i - recall that we consider the 19 most populous in China - we identify other smaller cities located (centre to centre) within between 100km and 200 km. These are likely sources of pollution imported to city i if the wind happens to blow in the 'right' direction. We refer to these as 'source' cities for city i. Neighboring cities within

100km are excluded to minimize risk of endogeneity (Bayer et al., 2009; Zheng et al., 2014).[15] Source cities and their coordinates are listed in Appendix Table A1.9.

We deliberately take a 'standard' approach to constructing our first stage equation, which is

$$P_{it} = \eta_0 + \psi P_{source_{it}} + W_{it}\gamma + \theta_i + \lambda_t + \epsilon_{it} \tag{2}$$

where

$$P_{source_{it}} = \sum_{j}^{J} \omega_{ijt} \overline{P_{jtmonth}}$$

P_{it} is actual pollution in target city i on date t. The coefficient of interest is ψ and captures the effect of pollution from upwind source cities on the target city. $P_{source_{it}}$ is an index that proxies the amount of pollution expected imported into target city i from source cities on a particular day. It is important that the construction of this index is fully understood, so that we will describe its components in some detail. Validity of the instrument will require that the only way in which wind directions influence sleep patterns in the target city is through induced changes in target city air quality.

$\overline{P_{jtmonth}}$ is the mean level of pollution in source city j in the associated month. In other words a measure of how 'potent' a particular source is as a supplier of pollution. As is well known, transport of pollution from source to target city on a particular day depends upon wind direction and speed. In particular, other things being equal, imports of pollution from city j by air to city i are greater when: (a) the city is close, (b) windspeed is high on a particular day, and (c) the angle between wind direction and an imaginary line joining the two cities is narrow (Zahran et al., 2017; Anderson, 2015; Schlenker and Walker, 2016). The vector of weights ω_{ijt} capture this. In particular we weight the source cities by inverse-

[15] Bayer et al. (2009) exclude the distant sources within 80km, Zheng et al. (2014) within 120km. In their study of medium-term health effects of $PM_{2.5}$ and SO_2, Barreca, Neidell and Sanders (2017) allow for the transport of pollution from a single power station up to 100 miles (161 km). We also tried the different cut-off distances with 100 to 300km, but the instrument is not strong enough because of the uncertainty embedded in long-distance travel.

distance (Equation 3), where geographical distance is adjusted to allow for windspeed and angle (Equation 4).

$$\omega_{ijt} = \frac{\frac{1}{trans_{jt}}}{\sum_j^J \frac{1}{trans_{jt}}} = \sum_j^J \frac{\frac{1}{trans_{jt}}}{\frac{1}{trans_{1t}} + \frac{1}{trans_{2t}} + \frac{1}{trans_{3t}} + ... + \frac{1}{trans_{Jt}}}, \tag{3}$$

where

$$trans_j = \frac{dj}{windspeed_i * \cos|\phi_i - \phi_j|_{>0}} \tag{4}$$

Wind direction can vary during the course of a day. We use daily average direction constructed from hourly data, consistent with first principles and most existing studies (including Schlenker and Walker (2016) and Herrnstadt et al. (2016)). Only positive values of $\cos|\phi_i - \phi_j|$ are included when the index is calculated, *i.e.* attention is limited to source cities that are (not necessarily directly) downwind on any particular day.[16] This occurs where the difference between wind direction and the direction of the vector between cities j and i is less than 90 degrees. In a robustness check we find that results are largely undisturbed if we instead limit to those where the difference is no greater than 60 degrees. The complexities of pollution transport by wind do not allow us to specify fully the process whereby pollution from one city influences air quality in another, but the functional form here is a simplified version standard in modelling of this sort. For a recent application, the analysis here coincides with Schlenker and Walker (2016), who account for the cosine of variation of wind direction from point source (airport) to the centre point of zipcode. Importantly, it is

[16]The angle between wind direction and the line joining the central points of cities i and j is $|\phi_i - \phi_j|$. All angles are measured in degrees clockwise from due North (0° and 360° equal North). The cosine transformation implies a particular weighting to sources at different angles. Recall that the cosine of zero degrees is 1, cosine of 20 degrees is 0.93, cosine of 60 degrees is 0.5 and so on. So other things being equal, a source 60 degrees off the wind line carries half the weight as a source that is directly upwind. The weighting is consistent with first principles (Anderson, 2015). In our unreported analysis, the results are also qualitatively robust to dropping the weighting scheme altogether. As would be expected the precision of estimates is compromised, though significance of results is maintained.

unlikely that the precise functional form adopted here would influence the defensibility of the exclusion restriction. Moreover, we will try some alternatives for the purposes of robustness later. The relevance of the instrument is assessed statistically at the first stage.

Lagged instrumental variable

As noted, in our base specifications, we limit attention to source cities located 100 to 200 km from the target city ($100km < d_{ij} < 200km$). Airborne pollutants leaving one city take more time to transport over a greater distance, which points to a delayed impact on the target. Our primary measure of pollution is average ambient concentrations from midnight to midnight, and the outcome of interest is sleeplessness in pursuant night (11 pm to 7 am). With average wind speed in the sample at around 8 km/h transport of air from a city at distance of 100 km would take over 12 hours, from 200km over 24 hours. To capture this lagged effect in some specifications, we include a one-day lag,

$$P_{it} = \eta_0 + \psi_{it-1} \sum_{j}^{J} \omega_{ij(t-1)} \overline{P_{j(t-1)month}} + \psi_{it} \sum_{j}^{J} \omega_{ijt} \overline{P_{jtmonth}} + W_{it}\gamma + \theta_i + \lambda_t + \epsilon_{it} \qquad (5)$$

We expect each of the coefficients ψ_{it-1} and ψ_{it} to be positive and similar in order of magnitude. In unreported analysis we have tried alternative specifications with additional lags without disturbing results discernibly.

1.4 Results

1.4.1 Ordinary least squares

Table 1.4 reports the coefficients from estimating equation (1) using OLS regression model for AQI (Panel A) and $PM_{2.5}$ (Panel B), where the dependent variable is log form of sleeplessness, and the independent variable of interest is daily pollution. Each of the 14 coefficients reported in Table 1.4 is derived from a separate regression. We will talk for now about coefficient magnitudes, and return to interpret the effect size that they imply later.

Standard errors are clustered at city level. As there are only 19 clusters (cities), we use wild cluster bootstrap method (Cameron et al., 2008), one of the most versatile remedies for small numbers of clusters.[17] The likely alternative approaches would have been cluster-adjustment of the t-statistics (Bakirov and Székely, 2006) and pairs cluster bootstrap (Cameron et al., 2008; Harden, 2011).[18]

Column (1) is the sparsest specification and includes only city fixed effects, netting out any unobserved, time-invariant city characteristics (size, Weibo-penetration, building characteristics, etc.). Reading down this column, we see positive coefficients for each pollutant, in most cases significantly stronger than 5%.

From Column (2) to Column (6), we add time controls (year by month fixed effects, city by year fixed effects, city by quarter fixed effects, day of week and holiday fixed effects one by one). As expected, monthly effects have an important impact on sleep. Cities have different characteristics that vary by year and season. Besides, sleep behavior may be expected to be different on weekdays versus weekends, and on holidays versus non-holidays. The inclusion of these has little impact on the estimated coefficients on the pollution regressors in Column (6).

In Column (7) we control for weather effects. The weather controls include bins for average temperature and humidity, and linear measures for precipitation, sea-level pressure, and wind speed. Weather effects are known to have a meaningful impact both on sleep (estimates not presented in this table) but, more importantly for us, may affect the strength of the relationship between air quality and sleeplessness. However, after inclusion of time fixed effects and city by time fixed effects, the inclusion of weather controls does not disturb substantially our coefficient estimate of interest. For the part of the empirical analysis based on OLS estimation, Column (7) summarizes the preferred specification.

[17] We report P-values based on wild cluster-bootstrap (1000 replications) in brackets. Robust standard errors clustered at city level are reported in parentheses.

[18] Unfortunately, our data are insufficient to generate the estimates under pairs wild bootstrap due to the inclusion of multiple fixed effects.

While the sign and significance obtained for coefficients on all pollutants in this section provide valuable insight, earlier we identified concerns - in particular measurement error related to effective pollution exposure levels - that led us to expect attenuation in estimated coefficient values. Insofar as these concerns are valid we would expect the effects summarized in the last paragraph to *under*-state true effect sizes. To address this concern we will report IV estimates below.

Non-linear effects

In order to check for the possible non-linear effects of air pollution on sleeplessness, we incorporate the categorical variables for different levels of pollution in the regression model. AQI and $PM_{2.5}$ are classified into five bins: less than 50, 50 - 100, 100 - 150, 150 - 200 and larger than 200, with the first group serving as the reference group. In Table 1.5, the coefficients display evidence of an increasing effect at higher pollution levels. Compared to the reference group, a realization $PM_{2.5}$ between 100 and 150 $\mu g/m^3$ causes an increase in sleeplessness of 3.5%, 150 to 200 an increase of 6.2% and 200 plus an increase of 8.4%. Notice that the effects are close to their linear counterpart, which motivates our use of the linear specifications elsehwere in the paper. They are also comparable with the effects obtained by linear OLS, namely 4.3%, 6.5% and 8.6% respectively.[19]

We also plot the points estimates and the corresponding 95% confidence intervals in Figure 1.1. Graphs (a) and (b) depict the estimates reported in Table 1.5 with 50 unit as the width of each bin. Graphs (c) and (d) repeat the exercise but with bins 25 units in width.

1.4.2 Instrumental variable

The main IV results are reported in Table 1.6. From Columns (1) to (4), all the regressions include the full suite of controls which are shown under the preferred specification in Column

[19] 4.3% is calculated by multiplying the coefficient under the preferred estimation in Column (7) by 100.

(7) of Table 1.4. Each column reports the outcome of a separate regression, and for $PM_{2.5}$ and AQI, we run alternatives without and with the lagged instrument included in the first stage (odd and even numbered columns respectively).

The dependent variable in the first stage is daily-mean pollutant in target city i, and the IV is the weighted average pollution of surrounding source cities. Recall that cities are included if they are between 100km and 200km in the upwind direction, where upwind is defined as within 90 degrees of the average within-day wind direction.

The first stage estimation works well. We find a strong effect of variations in pollution in source (upwind) cities on the target city. In each case significance is achieved at better than the 1% level. The lagged pollutant measure is also significant in both cases, as anticipated. The Kleibergen-Paap-Wald-F statistic in each of the four first-stages are high enough compared to the Stock-Yogo weak ID test critical values (10% maximal IV size) listed below. So we have no concerns about weak instruments.

The second stage replicates the preferred OLS specification, regressing the daily sleeplessness measure on the predicted level of pollution obtained from the first stage. Comparing the coefficients in the odd and even columns, the lagged pollution measure 'matters' in the first stage; its inclusion has a little decease impact on coefficient of interest in the second.

The estimates under stronger instruments with larger F-statistics are chosen to be our preferred specifications, which are listed in odd column of Table 1.6.[20] In each case the estimated coefficients are four times larger in absolute size than those derived from OLS, consistent with our expectation that the estimates from the latter were attenuated.

A one standard deviation increase in $PM_{2.5}$ causes an increase in sleeplessness equal to 12.8% of the daily mean. For AQI a one standard deviation increase causes a similar increase in sleeplessness, amounting to 11.6% of mean level.

[20]In addition, and following Schlenker and Walker (2016) Table 1.1, we explore the possibility that pollution may be dispersed by high winds by adding an interaction term $P_{source_{it}} * windspeed_{it}$ to our preferred first-stage specification. This has little impact on results - summarized in the Appendix Table A1.10.

1.5 Robustness and falsification

1.5.1 Wind direction

In developing the instrument, to be considered 'upwind', the angle between the wind line and a straight line drawn between source and target city had to be less than 90 degrees (*i.e.* $|\phi_i - \phi_j| < 90°$). Since source cities are described by their monthly average pollutant characteristics, and locations do not move, the only variation in source city across dates comes from plausibly exogenous day-to-day variations in wind direction. As such it is important to check that alternative definitions of 'upwind' would not deliver a substantially different result.

In Table 1.7 we report the results of re-estimating the preferred IV specification but with cities selected as sources if they lie within a narrower, 60 degree angle of the wind line (*i.e.*, $|\phi_i - \phi_j| < 60°$). The results from the first stage maintain significance at the 1% level, but decrease a bit in magnitude for both odd and even columns due to the loss of that part of the information at $60° < |\phi_i - \phi_j| < 90°$. The second stage regression looks very similar to those reported in Table 1.6. Significance and coefficient values are little disturbed.[21]

1.5.2 Reduced form

Table 1.8 reports the results of the reduced form estimation corresponding to our preferred IV specification. Columns (1) and (2) reproduce the OLS and IV results respectively. The coefficients in Column (1) coincide with Column (7) from the OLS regressions in Table 1.4. Column (2) repeats the second stage results under the odd columns in Table 1.6. Column (3) reports the reduced form exercise in which $P_{source_{it}}$ is the regressor of interest in an OLS regression with lnS_{it} as the dependent variable. In other words from:

[21] The F-statistics from the first stages are somewhat smaller, though still good. This reflects the fact that building the instrument on a basis that excludes source cities at $60° < |\phi_i - \phi_j| < 90°$ means that we lose part of correlation of the instrument with target city pollution.

$$lnS_{it} = \alpha_0^{'} + P_{source_{it}}\beta_{up} + W_{it}\gamma^{'} + \theta_i + \lambda_t + \epsilon_{it}^{'} \tag{6}$$

Again, each coefficient presented in this table comes from a different regression. As expected, the estimates from the upwind variant remain significant - the usual reduced form ‘works’.

1.5.3 Alternative fixed effects

Our preferred estimation accounts for a suite of fixed effects, including year by month, city by year, city by quarter, day of week and holiday. Table 1.9 re-conducts the OLS regression using alternative fixed effects. Each of the eight coefficients comes from a separate regression. Column (1) reprints the outcomes reported in Column (7) of Table 1.4. Column (2) adds date-of-observation fixed effects, which helps to account for likely daily heterogeneity in economic activity. Column (3) further controls the variations from the differential trends by week of year, replacing year by month fixed effects in Column (1) with year by week fixed effects. Column (4) replaces city by quarter fixed effects with city by month fixed effects. Results are consistent across specifications. Similarly, we re-estimate the IV regressions in table 10.[22] Again, second stage results vary only a little between columns.

1.5.4 Precipitation

The confounding role of rain is a potentially important challenge to our inference. Introspection suggests that rainfall - either contemporaneously, or lagged through effect on mood etc. - might plausibly inhibit sleep. While we include controls for daily-average precipitation amongst our weather controls, we further probe this possibility by conducting two sub-sample exercises.

First, we re-estimate our preferred specifications on that sub-sample of days on which

[22] City by month fixed effects are not considered under IV because they render our instrument weak.

recorded night-time precipitation (from 11 pm to 7 am) in the target city is zero. This causes us to lose around 21% of the sample. The results of this exercise are reported in Column (2) of Table 1.11 and in Column (2) of Table 1.12 for OLS and IV respectively. Results are little disturbed. This implies that the effects observed are not driven by contemporaneous rainfall.

Second, we re-estimate our preferred specifications on that sub-sample of days on which recorded precipitation during the night in question *and* the whole of the preceding calendar day in the target city is zero. This causes us to lose around 34% of the sample. The results of this exercise are reported in Column (3) of Table 1.11, and Column (3) of Table 1.12 for OLS and IV respectively. The signs and magnitudes of the coefficients are in all cases quite similar (the IV estimates in each case in fact become somewhat larger than those derived from the whole sample). The level of statistical significance obtained is sustained in almost all cases - better than might have been anticipated given the considerable erosion of sample size.

1.5.5 Beijing and environs, Shanghai and environs, Guangzhou and environs

While we derive results from a panel of the 19 most populous cities in China, a further concern might be that the results are driven by a small subset of the cities. In an unreported exercise we re-estimate our preferred specifications on restricted samples of cities, dropping each individually in turn, and in no case do we observe more that slight disturbances in our results. However, in this section, we report the impact of dropping clusters of cities that may exhibit particular features that might be driving the results. In particular, first, we exclude the cities of Beijing and Tianjin (the Beijing-Tianjin corridor is the country's most heavily industrialized 'rust-belt' area (Shao et al., 2006); second, separately we exclude the cluster of the south-eastern coastal cities of Shanghai, Hangzhou and Suzhou, as well as southern coastal cities of Guangzhou, Shenzhen and Dongguan (these are less polluted, less industrialized, and more influenced by coastal effects).

The results of these exercises are summarized from Columns (2) to (4) of Appendix Tables A14 and A15 for OLS and IV respectively. Again, the results are little-disturbed. The first stage regressions continue to work well, and the second stage estimates are largely robust.

It is also concerned that whether air pollution remains its health effect across the 19 cities in the sample. Both OLS and IV estimators of individual city are reported in Appendix Tables A16 and A17 for AQI, and A18 and A19 for $PM_{2.5}$ respectively. Although the magnitude of the effects varies across the cities, most of them still have a significant impact on city sleeplessness.

1.5.6 Alternative standard errors

In the calculation of standard errors in the main tables, we chose to cluster at the city level, judging this to account for the potential correlations among regressors and errors within clusters. However, this approach delivers only nineteen clusters (each with around 730 observations), which Angrist and Pischke (2008) suggest may be too few. Cameron et al. (2008) show that small cluster numbers can bias downwards cluster-robust standard errors, leading researchers to overstate the statistical significance of results. To address this we use the wild cluster bootstrap technique for the results in our main tables (Cameron et al., 2008; Esarey and Menger, 2015).

We also supplement the analysis with more clusters by using cluster-robust standard errors at city by year by season (152 clusters), city by year by month (456 clusters) and city by year by week (1976 clusters) to explore whether our conclusions would have been changed substantially by such alternative approaches. The results are displayed in Appendix Tables A1.11, A1.12 and A1.13 for OLS and IV respectively. All the estimators retain significance at conventional levels (in almost all cases at better than 1%).

A separate concern related to standard errors is that spatial correlation can in some circumstances bias standard errors and so invalidate inference (Hoechle, 2007). To investigate this possibility in our setting, we apply the methods of Driscoll and Kraay (1998). They introduce a non-parametric covariance matrix estimator for which standard errors are assumed to be heteroscedastic, auto-correlated with $MA(q)$ within panel (each city), and potentially correlated among panels. The method is appropriate for panels with small numbers of panels (in our case 19) but many observations per panel (730). The results of this exercise (for $q = 7$, though very similar results emerge with different values) are reported in Column (2) of Tables A1.11 to A1.13. Again statistical significance is maintained at conventional levels.

1.6 Placebos

The air pollution of our target city is instrumented by the daily weighted average pollutants of source cities upwind. The exogeneity of our instrument would be threatened if there were unobserved (and therefore uncontrolled for), daily-varying factors that are correlated between source and target cities and cause pollution in both places. For example, daily variations in traffic density could be correlated across cities source and target cities.

A number of elements of our design, however, point to this not being a big problem. First, we exclude potential source cities within 100 km of the target, and these sorts of correlations are likely to be less pronounced over longer distances (think of the traffic spike caused by a major sporting event, for example). Second, the potency or 'dirtiness' of a particular city as a source of exported pollution is based on a long-run, *monthly* average measure of air pollution in that city. As such daily variations in the instrument are - by construction - *not* driven by daily or short-term variations in pollution levels in the source city. As such the variation in daily contribution of distant sources to air quality in a particular target city should be driven only by variations in wind direction.

However, to further test the instrument we conduct two placebo tests, replacing wind directions in the vicinity of each target city with placebo series of irrelevant wind directions. The tests differ in how we generate the irrelevant or placebo wind data. First, we scramble the wind directions within our sample of Chinese cities by using a reverse-alphabetical assignment. For example Beijing, the first city in our sample when listed alphabetically, is falsely-assigned the wind direction series from Zhengzhou, which appears last, and so on.[23] Second, we conduct an out-of-sample placebo. To do this we draw wind directions from US cities (19 largest cities by population), which we matched based on alphabetical assignment. For example, the first city in our sample (Beijing) is falsely-assigned the wind direction in the first-alphabetical US city (Austin). The fifth city, Dongguan, has the wind direction from Dallas in US sample, and so on.

The results of these exercises are reported in Table 1.13. We can see that in each case, the first-stage regression breaks down comprehensively. This provides compelling evidence that the variation in target-city pollution in our IV specifications is driven by (plausibly exogenous) variations in wind direction, not by correlated day-to-day variations in local air quality conditions in source and target cities.

1.7 Joint estimation

Disentangling the independent effects of particular pollutants is a challenge for research on both health and non-health outcomes. Various authors have addressed the problem in different ways; typically this involves excluding a subset of the potentially confounding substances altogether (often due to data limitations). If pollutants tend to positively covary then this leads to effects being loaded onto that pollutant or subset of pollutants that are included.[24]

Ambient levels of the various pollutants (with the exception of ozone) covary positively.[25] Some of the pollutants are precursors in the production of ozone. Furthermore the overall impact of a particular cocktail of pollutants may depend upon their mixture in complex ways. This leads us to be cautious in interpreting the results reported thus far. Taken collectively we believe that the results presented in Tables 1.4 though 1.13 provide a compelling case

[23] For the middle one, we replace that using wind direction in Beijing.

[24] A different approach taken in some recent work (for example Gendron-Carrier et al. (2017)) exploits data from NASA satellites that measures Aerosol Optical Depth (AOD). AOD in effect measures how optically 'thick' the air is over a particular GIS point, but does not allow for pollutant-by-pollutant inference.

[25] In our dataset, the correlation between $PM_{2.5}$ and CO is 0.709, between $PM_{2.5}$ and NO_2 is 0.669, and between CO and NO_2 is 0.584.

that polluted air has a causal impact on city-level sleep quality. While we have focussed on $PM_{2.5}$ and the multi-pollutant AQI measure, for completeness we summarize in Table 1.14 the results of additional joint estimation exercises.

Columns (1) and (3) report the OLS and IV results from estimation of our preferred specifications on each single pollutant ($PM_{2.5}$, CO, NO_2 and O_3) in turn. Column (2) reports the results of including the four pollutants in an OLS regression simultaneously - the so-called 'horse-race' regression. As in Schlenker and Walker (2016), signs become mixed. $PM_{2.5}$ remain its positive sign and retains significance at the 3.7% level.

Column (4) follows the method proposed by Schlenker and Walker (2016), Knittel et al. (2016) and Sager (2016).[26] In our case, different pollutants are instrumented by their corresponding levels in source cities, and the instrumented pollution levels are then included simultaneously in the same regression.[27] The coefficient on $PM_{2.5}$ remains positive and is comparable in magnitude to those from the single pollutant exercises; significance is maintained at better than 1%.

An alternative approach - adopted by Moretti and Neidell (2011) - is to instrument for one pollutant at a time, in each case including the other pollutants, uninstrumented, as linear controls in both the first and second-stage regressions.[28] In that case the coefficient on the instrumented pollutant is unbiassed, though those on the control pollutants are not. Column (5) reports the results of conducting that exercise repeatedly, with each pollutant in turn being the one that is subject to instrumentation.[29] Under this alternative approach, $PM_{2.5}$ remains significant at the 5% level with an associated coefficient estimate that is somewhat larger than in our preferred specification.

[26] Only pollutants with strong instrument are included in the joint estimation, otherwise it is presented by "-". See the first stage results in Appendix Table A1.22 from Column (1) to Column (4).

[27] To be clear, while each coefficient in Columns (1), (3) and (5) is derived from a separate regression, Column (2) and (4) each report a single regression.

[28] More concretely, Moretti and Neidell (2011) instrument for ozone, and include controls (uninstrumented) for CO, O_3 and NO_2. They do not include measures for particulate matter.

[29] Only pollutants with strong instrument are included in the joint estimation, otherwise it is presented by "-". See the first stage results in Appendix Table A1.22, Columns (5) through (8).

1.8 Conclusions

Sleep is a central contributor to human well-being, and its disturbance has been linked to a wide set of negative outcomes. If pollution in a city has a significant detrimental impact on how the inhabitants of that city sleep, this would imply a hitherto unaccounted for social cost of air pollution. Understanding the full range of channels through which pollution effects welfare - and by implication the benefits of clean air - is a prerequisite for the design of welfare-maximising policy interventions in this area.

We provide what we believe to be the first evidence that air pollution on a particular day has a causal impact on sleep quality in a city on the following night. The estimated effect is substantial. For the composite air quality index (AQI), notionally moving from a median clean decile day to a median dirty day (in other words from the 5th to the 95th percentile when days are ranked from clean to dirty) increases city-level sleeplessness by 36.3% of its mean value. For $PM_{2.5}$ that number is 37.3%. The estimates prove to be robust to a wide set of checks.

The analysis provides further evidence of the susceptibility of individual and social outcomes to anthropogenic pollution. We have argued that sleep loss is an important outcome in its own right, but also that it can provide a mechanism to underpin a suite of less proximate outcomes identified in recent research. Further validation of the results, using alternative metrics and instruments, is planned in future research.

Table 1.1: Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
Sleeplessness Index	13870	798.306	5226.540	12	301569
AQI	13620	92.598	52.202	12	486.5
PM2.5 ($\mu g/m^3$)	13620	61.810	44.993	0	884
CO (mg/m^3)	13620	1.117	0.555	0	12.6
NO2 ($\mu g/m^3$)	13620	44.706	18.717	0	171
SO2 ($\mu g/m^3$)	13620	25.050	25.332	0	335
O3 ($\mu g/m^3$)	13620	91.101	49.431	0	294
Temperature ($^{\circ}C$)	13870	16.869	9.814	-21.9	34.8
Humidity (%)	13870	69.481	16.450	7.875	99.5
Sea-level Pressure (hPa)	13870	1016.262	9.378	983.2	1054.5
Wind Speed (Km/h)	13870	7.944	3.598	1	33
Precipitation (mm)	13870	3.35	11.253	0	204.8

Notes: The dataset contains daily data from 19 target cities from 2014 to 2015.

Table 1.2: Air Quality Index (AQI) and Pollutant Concentrations

	Level	Description	AQI	PM2.5 (24hr) ($\mu g/m^3$)	Number of Days (AQI)	Percent
Low	I	Excellent	0-50	0-35	2145	15.75%
	II	Good	51-100	36-75	7090	52.06%
Medium	III	Light Polluted	101-150	76-115	2801	20.57%
		Moderately Polluted	151-200	115-150	912	6.70%
High	IV	Heavily Polluted	201-300	151-250	578	4.24%
	V	Severely Polluted	301-500	251-500	94	0.69%

Notes: The table maps $PM_{2.5}$ to AQI categories. Classification principles are taken from the *Technical Regulation on Ambient Air Quality Index HJ 633-2012*. Levels I and II do not have health implications, and are thus suitable for outdoor activities. Higher levels of pollutants leads to higher risk of breathing or heart problems. Outdoor exercise should be reduced. Level V may induce respiratory diseases, and outdoor exposure is to be avoided for elderly and sick people. The last two columns report the number of days and corresponding percentage of days falling into each category in the sample.

Table 1.3: Correlations between Pollutants

	AQI	PM2.5	PM10	CO	NO2	SO2	O3
AQI	1.000						
PM2.5	0.957	1.000					
PM10	0.896	0.875	1.000				
CO	0.672	0.709	0.672	1.000			
NO2	0.640	0.669	0.658	0.584	1.000		
SO2	0.490	0.498	0.522	0.496	0.508	1.000	
O3	0.008	-0.125	-0.064	-0.313	-0.155	-0.223	1.000

Notes: The table displays correlation matrix of pollutants in the dataset.

Table 1.4: Air Quality and Sleeplessness — OLS

Independent Variable Daily Pollutant	Dependent Variable Ln(Sleepless)						
	City FEs	Temporal Controls					Weather Covariates
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: AQI	0.050***	0.056**	0.055***	0.041**	0.040**	0.038**	0.037***
	(0.017)	(0.022)	(0.018)	(0.014)	(0.014)	(0.014)	(0.012)
	[0.010]	[0.029]	[0.008]	[0.021]	[0.025]	[0.034]	[0.001]
Panel B: PM2.5	0.043*	0.056**	0.057***	0.049***	0.046**	0.044**	0.043***
	(0.021)	(0.018)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)
	[0.072]	[0.019]	[0.007]	[0.011]	[0.021]	[0.030]	[0.012]
Observations	13617	13617	13617	13617	13617	13617	13617
Additional Controls							
City FEs	Y	Y	Y	Y	Y	Y	Y
Year by month FEs	N	Y	Y	Y	Y	Y	Y
City by year FEs	N	N	Y	Y	Y	Y	Y
City by quarter FEs	N	N	N	Y	Y	Y	Y
Day of Week FEs	N	N	N	N	Y	Y	Y
Holiday FEs	N	N	N	N	N	Y	Y
Weather Covariates	N	N	N	N	N	N	Y

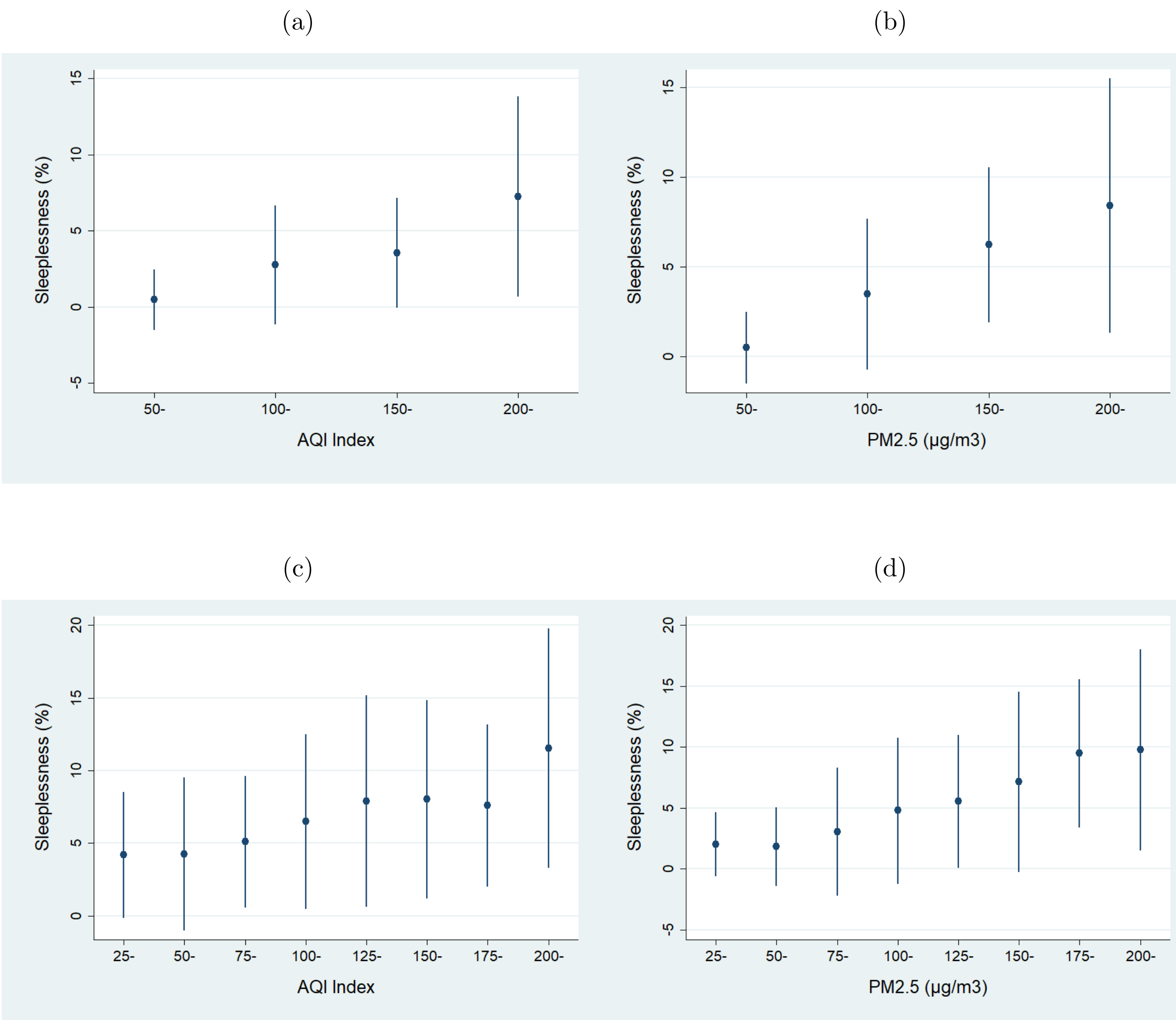
Notes: Dependent variable is log form of Sleeplessness Index. Data collection period runs from 11pm to 7am. Independent variable of interest is daily average measure of specific pollutant. All estimators have been adjusted into percentage by multiplying 100. Temporal controls include year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects. Weather controls contain temperature, humidity, precipitation, wind speed and sea-level pressure. Temperature and humidity are measured by the way of bins (5 degree C indicators for average temperature, 20 percent indicators for average humidity). Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.5: Non-linear Effects

	AQI	PM2.5
	(1)	(2)
AQI/PM2.5 <50 Omitted		
[50, 100)	0.483	0.504
	(0.932)	(0.942)
	[0.749]	[0.622]
[100, 150)	2.766*	3.493*
	(1.839)	(1.992)
	[0.094]	[0.060]
[150, 200)	3.561**	6.231***
	(1.702)	(2.043)
	[0.044]	[0.002]
>200	7.264**	8.411**
	(3.112)	(3.368)
	[0.021]	[0.051]
Observations	13617	13617
Additional Controls		
City FEs	Y	Y
Year by month FEs	Y	Y
City by year FEs	Y	Y
City by quarter FEs	Y	Y
Day of week FEs	Y	Y
Holiday FEs	Y	Y
Weather Covariates	Y	Y

Notes: This table reports the non-linear effects of pollutants on sleeplessness, with AQI and $PM_{2.5}$ incorporated in the form of bins (50 units in each bin). All the regressions include city fixed effects, temporal controls, and weather covariates. Temperature and humidity are measured in the form of bins. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Figure 1.1: Non-linear Effects



Notes: This diagram displays the estimates of non-linear effects of AQI and $PM_{2.5}$. Graphs (a) and (b) plots the point estimates reported in Table 1.5, as well as the corresponding 95% confidence intervals. (c) and (d) break the pollutant into more bins with 25 units per bin, and re-estimate the effects in the same way as Table 1.5 did.

Table 1.6: Air Quality and Sleeplessness — IV

2SLS	AQI (1)	AQI (2)	PM2.5 (3)	PM2.5 (4)
First Stage[(a)]				
Instrumental Pollution t	0.455***	0.181***	0.463***	0.201***
	(0.062)	(0.034)	(0.061)	(0.032)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Instrumental Pollution lagged t-1		0.514***		0.50***
		(0.059)		(0.065)
		[<0.01]		[<0.01]
Kleibergen-Paap rk Wald F statistic	54.791	52.868	56.846	46.156
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	19.93	16.38	19.93
Second Stage[(b)]				
Instrumented Pollution	0.223**	0.149**	0.285**	0.213**
	(0.096)	(0.075)	(0.118)	(0.098)
	[0.039]	[0.045]	[0.033]	[0.045]
Observations	12904	12579	12989	12662
Additional Controls				
City FEs	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y
City by quarter FEs	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y

Notes: (a) Dependent variable in the first stage is daily-mean pollutant of target city, and independent variable is daily weighted average pollution of surrounding cities ($100km < d_{ij} < 200km$) from upwind direction (within 90 degree to the wind). (b) Second stage reports the results regressing the log form of Sleeplessness Index on the instrumented daily pollution with estimators being adjusted into percentages by multiplying by 100. Columns (2) and (4) incorporate the day before variable as an additional instrument. Temporal controls include year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects. Weather controls contain temperature, humidity, precipitation, wind speed and sea-level pressure. Temperature and humidity are measured by the way of bins. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.7: Robustness — IV with 60 Degree Wind Angle Inclusion

2SLS	AQI		PM2.5	
	(1)	(2)	(3)	(4)
First Stage[(a)]				
Instrumental Pollution t	0.378***	0.175***	0.392***	0.200***
	(0.077)	(0.033)	(0.078)	(0.032)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Instrumental Pollution lagged t-1		0.437***		0.433***
		(0.077)		(0.078)
		[<0.01]		[<0.01]
Kleibergen-Paap rk Wald F statistic	23.883	23.415	25.092	23.919
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	19.93	16.38	19.93
Second Stage[(b)]				
Instrumented Pollution	0.250**	0.173*	0.306**	0.234**
	(0.098)	(0.079)	(0.114)	(0.097)
	[0.038]	[0.057]	[0.030]	[0.034]
Observations	11944	11049	12021	11116
Additional Controls				
City FEs	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y
City by quarter FEs	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y

Notes: (a) Dependent variable in the first stage is daily-mean pollutant of target city, and independent variable is the daily weighted average pollution of surrounding cities ($100km < d_{ij} < 200km$) from upwind direction (within 60 degree to the wind). (b) Second stage reports the results regressing the log form of Sleeplessness Index on the instrumented daily pollution. Columns (2) and (4) incorporate the day before measure as an additional instrument. Temporal controls include year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects. Weather controls contain temperature, humidity, precipitation, wind speed, and sea-level pressure. Temperature and humidity are measured by the way of bins. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.8: Reduced Form

	OLS	IV	Reduced Form
	(1)	(2)	(3)
Panel A: AQI	0.037***	0.223**	0.102**
	(0.012)	(0.096)	(0.044)
	[0.001]	[0.039]	[0.039]
Observations	13617	12904	13093
Panel B: PM2.5	0.043***	0.285**	0.132**
	(0.017)	(0.118)	(0.055)
	[0.012]	[0.033]	[0.033]
Observations	13617	12989	13179
Additional Controls			
City FEs	Y	Y	Y
Year by month FEs	Y	Y	Y
City by year FEs	Y	Y	Y
City by quarter FEs	Y	Y	Y
Day of week FEs	Y	Y	Y
Holiday FEs	Y	Y	Y
Weather Covariates	Y	Y	Y

Notes: Column (1) repeats the OLS results of Column (7) in Table 1.4. Column (2) repeats the second stage results under Column (1) and Column (3) in Table 1.6. Column (3) presents the results of reduced form regressing the log form of daily Sleeplessness Index on daily weighted average pollutant of peripheral cities ($100km < d_{ij} < 200km$) from the upwind direction (within 90 degree to the wind). All the regressions include city fixed effects, temporal controls (year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects) and weather controls (average temperature bins, average humidity bins, precipitation, sea-level pressure, and wind speed). Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.9: Alternative Fixed Effects — OLS

	(1)	(2)	(3)	(4)
Panel A: AQI	0.037***	0.041***	0.032***	0.025***
	(0.012)	(0.013)	(0.011)	(0.008)
	[0.001]	[<0.01]	[0.001]	[0.010]
Panel B: PM2.5	0.043***	0.046***	0.036***	0.027*
	(0.017)	(0.018)	(0.016)	(0.012)
	[0.012]	[0.008]	[0.013]	[0.060]
Observations	13617	13617	13617	13617
Additional Controls				
City FEs	Y	Y	Y	Y
Year by month FEs	Y	Y	N	Y
City by year FEs	Y	Y	Y	Y
City by quarter FEs	Y	Y	N	N
City by month FEs	N	N	N	Y
Year by week FEs	N	N	Y	N
Date FEs	N	Y	N	N
Day of week FEs	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y

Notes: This table re-runs the OLS estimation for AQI and $PM_{2.5}$ under various sets of fixed effects. Column (1) replicates the preferred OLS results of Column (7) in Table 1.4. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.10: Alternative Fixed Effects — IV

	AQI			PM2.5		
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage						
Instrumental Pollution t	0.455***	0.518***	0.455***	0.463***	0.522***	0.462***
	(0.062)	(0.046)	(0.059)	(0.061)	(0.048)	(0.058)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Kleibergen-Paap rk Wald F statistic	54.791	130.600	60.317	56.846	117.546	62.509
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	16.38	16.38	16.38	16.38	16.38
Second Stage						
Instrumented Pollutant	0.223**	0.200*	0.204**	0.285**	0.264*	0.264**
	(0.096)	(0.097)	(0.092)	(0.118)	(0.121)	(0.113)
	[0.039]	[0.063]	[0.044]	[0.033]	[0.056]	[0.045]
Observations	12904	12904	12904	12989	12989	12989
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Year by month FEs	Y	Y	N	Y	Y	N
City by year FEs	Y	Y	Y	Y	Y	Y
City by quarter FEs	Y	Y	N	Y	Y	N
City by month FEs	N	N	N	N	N	N
Year by week FEs	N	N	Y	N	N	Y
Date FEs	N	Y	N	N	Y	N
Day of week FEs	Y	Y	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y	Y	Y

Notes: This table re-conducts the IV estimation under various sets of fixed effects. Column (1) and Column (4) reprint the preferred second stage results under Column (1) and Column (3) in Table 1.6. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.11: Precipitation Exclusion — OLS

	Full	Clear Nights	Zero Rain Days
	(1)	(2)	(3)
Panel A: AQI	0.037***	0.039***	0.049***
	(0.012)	(0.013)	(0.016)
	[0.001]	[0.003]	[<0.01]
Panel B: PM2.5	0.043***	0.044**	0.057***
	(0.017)	(0.017)	(0.021)
	[0.012]	[0.016]	[0.001]
Observations	13617	10774	9051
Additional Controls			
City FEs	Y	Y	Y
Year by month FEs	Y	Y	Y
City by year FEs	Y	Y	Y
City by quarter FEs	Y	Y	Y
Day of week FEs	Y	Y	Y
Holiday FEs	Y	Y	Y
Weather Covariates	Y	Y	Y

Notes: Dependent variable is log form of Sleeplessness Index. Independent variable is city daily-mean value of specific pollutant. Column (1) displays the results for all observations replicating the results under Column (7) in Table 1.4. Column (2) excludes days with precipitation from 11pm to 7am. Column (3) excludes days with precipitation from 12pm to 7am on the following day. All the regressions include city fixed effects, temporal controls and weather covariates. Temperature and humidity are measured in the form of bins. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.12: Precipitation Exclusion — IV

	Full (1)	Clear Nights (2)	Zero Rain Days (3)
	Panel A: AQI		
First Stage			
Instrumental AQI t	0.455***	0.440***	0.454***
	(0.062)	(0.080)	(0.082)
	[<0.01]	[<0.01]	[<0.01]
Kleibergen-Paap rk Wald F statistic	54.791	30.529	30.417
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	16.38	16.38
Second Stage			
Instrumented AQI	0.223**	0.225**	0.215**
	(0.096)	(0.094)	(0.084)
	[0.039]	[0.037]	[0.038]
Observations	12904	10207	8560
	Panel B: PM2.5		
First Stage			
Instrumental PM2.5 t	0.463***	0.436***	0.446***
	(0.061)	(0.082)	(0.082)
	[<0.01]	[<0.01]	[<0.01]
Kleibergen-Paap rk Wald F statistic	56.846	28.070	29.399
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	16.38	16.38
Second Stage			
Instrumented PM2.5	0.285**	0.284**	0.288**
	(0.118)	(0.116)	(0.104)
	[0.033]	[0.041]	[0.032]
Observations	12989	10280	8633
Additional Controls			
City FEs	Y	Y	Y
Temporal Controls	Y	Y	Y
Weather Covariates	Y	Y	Y

Notes: Column (1) displays IV results for all observations, replicating the results under Column (1) and Column (3) in Table 1.6. Column (2) excludes the days with snowy or rainy nights. The results in Column (3) limit the sample to clear days without rain or snow in the daytime or nighttime. All the regressions include the same city fixed effects, temporal controls and weather controls as those in Table 1.6. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.13: Placebo Test

	Preferred	Chinese Cities Reverse-alphabetic (wind direction)	US largest Cities (wind direction)
	(1)	(2)	(3)
		Panel A: AQI	
First Stage			
Instrumental AQI t	0.455***	0.063	0.045
	(0.062)	0.093	0.048
	[<0.01]	[0.658]	[0.352]
Kleibergen-Paap rk Wald F statistic	54.791	-	-
Second Stage			
Instrumented AQI	0.223**	-	-
	(0.096)	-	-
	[0.039]	-	-
Observations	12904	12810	13502
		Panel B: PM2.5	
First Stage			
Instrumental PM2.5 t	0.463***	0.081	0.023
	(0.061)	0.088	0.018
	[<0.01]	[0.537]	[0.18]
Kleibergen-Paap rk Wald F statistic	56.846	-	-
Second Stage			
Instrumented PM2.5	0.285**	-	-
	(0.118)	-	-
	[0.033]	-	-
Observations	12989	12896	13502
Additional Controls			
City FEs	Y	Y	Y
Temporal Controls	Y	Y	Y
Weather Covariates	Y	Y	Y

Notes: Column (1) reports the IV estimations from the preferred specification in Table 1.6. Column (2) and Column (3) re-construct the weights based on the scrambled wind directions in other Chinese cities (reverse-alphabetic order in the sample) and US largest cities, respectively. All the regressions include city fixed effects, temporal controls (year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects) and weather controls (average temperature bins, average humidity bins, precipitation, sea-level pressure, and wind speed). Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 1.14: Joint Estimation

	OLS		2SLS		
	Single Estimation	Joint Estimation	Single Estimation	Joint Estimation (Schlenker and Walker 2016)	Joint Estimation (Moretti and Neidell 2011)
	(1)	(2)	(3)	(4)	(5)
PM2.5	0.043***	0.041**	0.285**	0.371***	0.519**
	(0.017)	(0.016)	(0.118)	(0.143)	(0.229)
	[0.012]	[0.037]	[0.033]	[0.006]	[0.049]
CO	2.688**	1.030	-	-	-
	(1.135)	(1.570)	-	-	-
	[0.023]	[0.547]	-	-	-
NO2	0.064	-0.024	0.018	-0.551**	-
	(0.041)	(0.047)	(0.220)	(0.250)	-
	[0.171]	[0.674]	[0.944]	[0.032]	-
O3	0.022	0.012	0.072	0.052	0.071
	(0.017)	(0.015)	(0.172)	(0.168)	(0.179)
	[0.261]	[0.521]	[0.835]	[0.854]	[0.834]
Observations	13617	13617	12989	12989	12989
Additional Controls					
City FEs	Y	Y	Y	Y	Y
Temporal Controls	Y	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y	Y
City by quarter FEs	Y	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y	Y

Notes: Column (1) and Column (3) repeat the OLS and IV estimations from the preferred specification in Table 1.4 and 1.6. Joint estimations that include co-emission are reported in Column (2), Column (4) and Column (5). Column (4) regresses Sleeplessness Index on different instrumented pollutants together. Column (5) instruments for one pollutant at a time, in each case including the other pollutants, uninstrumented, as linear controls in both the first and second-stage regressions. All the regressions include city fixed effects, temporal controls and weather covariates. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.1: Summary Statistics by City

	Beijing (1)	Changsha (2)	Chengdu (3)	Chongqing (4)	Dongguan (5)	Guangzhou (6)	Hangzhou (7)
Sleeplessness Index	5676.008	258.247	458.869	347.193	127.575	3331.537	649.112
	(18024.220)	(58.018)	(92.582)	(78.808)	(26.877)	(12400.320)	(562.917)
AQI Index	119.335	92.860	97.616	85.764	71.122	68.645	93.996
	(76.912)	(52.043)	(54.903)	(45.355)	(36.409)	(31.623)	(43.868)
$PM_{2.5}$ ($\mu g/m^3$)	81.882	67.443	67.006	58.570	40.202	42.988	62.780
	(70.034)	(43.542)	(47.492)	(37.251)	(22.462)	(23.035)	(36.668)
CO (mg/m^3)	1.278	1.026	1.079	1.127	0.878	0.976	1.087
	(0.997)	(0.344)	(0.363)	(0.269)	(0.238)	(0.259)	(0.345)
NO_2 ($\mu g/m^3$)	51.290	37.999	50.815	39.849	37.702	45.168	44.422
	(24.239)	(16.290)	(15.475)	(11.703)	(16.570)	(17.459)	(16.274)
SO_2 ($\mu g/m^3$)	16.605	20.756	16.742	19.638	17.372	14.529	32.338
	(19.585)	(10.794)	(8.407)	(11.792)	(9.014)	(6.364)	(14.689)
$O_3(\mu g/m^3)$	100.085	76.858	90.064	66.825	110.322	89.897	98.223
	(65.199)	(39.527)	(51.312)	(46.810)	(53.341)	(49.914)	(48.935)
Temperature ($°C$)	14.093	18.531	17.125	17.849	22.835	22.317	17.075
	(10.693)	(8.294)	(7.152)	(7.567)	(6.196)	(6.622)	(8.535)
Humidity (%)	52.560	58.076	80.219	78.975	76.225	75.011	74.596
	(19.837)	(16.823)	(8.506)	(10.441)	(10.140)	(12.878)	(12.992)
Sea-level Pressure (hPa)	1016.923	1015.726	1014.542	1014.212	1013.424	1013.655	1016.579
	(10.044)	(9.083)	(8.963)	(9.060)	(6.995)	(7.328)	(9.221)
Wind Speed (Km/h)	7.826	8.185	4.914	6.023	8.327	7.341	6.589
	(2.930)	(4.024)	(1.581)	(1.738)	(3.276)	(3.197)	(2.814)
Precipitation (mm)	1.260	4.007	2.542	3.813	6.447	6.680	3.974
	(5.700)	(10.063)	(8.533)	(10.857)	(16.846)	(18.671)	(10.722)

continued

	Nanjing (8)	Ningbo (9)	Qingdao (10)	Shanghai (11)	Shenyang (12)	Shenzhen (13)	Suzhou (14)
Sleeplessness Index	471.386	141.430	133.243	845.545	206.784	468.010	307.849
	(109.433)	(42.431)	(26.092)	(2498.374)	(45.883)	(151.464)	(226.141)
AQI Index	96.843	71.608	88.059	81.320	104.742	51.551	90.072
	(46.782)	(33.376)	(40.671)	(39.625)	(57.086)	(20.945)	(41.381)
$PM_{2.5}$ ($\mu g/m^3$)	65.151	45.067	52.833	53.053	70.576	31.140	61.765
	(39.841)	(28.382)	(36.993)	(33.542)	(58.055)	(17.731)	(34.119)
CO (mg/m^3)	0.936	0.924	0.926	0.836	1.044	0.978	0.927
	(0.346)	(0.267)	(0.698)	(0.286)	(0.499)	(0.213)	(0.291)
NO_2 ($\mu g/m^3$)	50.200	41.779	36.529	44.465	47.634	32.823	51.744
	(19.145)	(17.972)	(17.297)	(19.910)	(17.768)	(11.581)	(18.590)
SO_2 ($\mu g/m^3$)	20.958	18.707	30.090	17.220	67.358	8.130	20.886
	(12.057)	(10.378)	(18.949)	(9.977)	(67.123)	(3.132)	(10.223)
O_3($\mu g/m^3$)	100.517	96.606	102.146	102.752	92.272	78.678	97.158
	(53.182)	(39.177)	(40.703)	(42.989)	(48.966)	(32.456)	(48.678)
Temperature ($°C$)	16.828	17.551	14.214	17.261	9.290	24.130	17.261
	(8.495)	(7.556)	(8.996)	(8.220)	(12.749)	(5.605)	(8.220)
Humidity (%)	72.626	79.627	69.600	72.682	59.307	71.594	72.682
	(14.393)	(11.447)	(16.290)	(12.602)	(15.627)	(11.839)	(12.602)
Sea-level Pressure (hPa)	1017.057	1016.565	1017.550	1017.085	1016.625	1013.252	1017.085
	(9.250)	(8.695)	(9.170)	(8.974)	(9.859)	(6.657)	(8.974)
Wind Speed (Km/h)	9.412	7.616	11.823	9.308	8.018	7.579	9.308
	(3.789)	(3.170)	(4.735)	(3.308)	(3.250)	(2.472)	(3.308)
Precipitation (mm)	3.915	4.789	1.490	4.035	1.277	4.424	4.035
	(13.777)	(13.851)	(6.889)	(11.550)	(4.298)	(15.748)	(11.550)

continued

	Tianjin (15)	Wuhan (16)	Wuxi (17)	Xian (18)	Zhengzhou (19)
Sleeplessness Index	443.933	459.155	105.373	357.637	375.643
	(369.787)	(106.965)	(34.321)	(59.897)	(78.325)
AQI Index	112.737	106.519	95.279	104.029	128.748
	(62.388)	(54.983)	(42.070)	(54.515)	(62.903)
PM2.5 ($\mu g/m^3$)	78.577	74.769	64.210	66.339	91.210
	(54.586)	(47.490)	(33.967)	(48.639)	(55.645)
CO (mg/m^3)	1.526	1.149	1.078	1.783	1.664
	(0.796)	(0.393)	(0.335)	(0.730)	(0.665)
NO2 ($\mu g/m^3$)	48.008	50.227	42.725	43.593	52.807
	(23.867)	(20.332)	(16.349)	(14.638)	(18.164)
SO2 ($\mu g/m^3$)	38.401	25.905	26.976	27.380	37.405
	(35.471)	(15.127)	(11.661)	(22.797)	(27.967)
O3($\mu g/m^3$)	80.822	96.440	100.660	71.073	79.994
	(49.732)	(49.278)	(52.992)	(44.906)	(45.317)
Temperature ($°C$)	14.333	17.376	17.075	9.084	16.287
	(10.786)	(8.676)	(8.535)	(15.190)	(9.400)
Humidity (%)	56.554	78.278	74.596	57.982	58.949
	(17.466)	(10.546)	(12.992)	(12.442)	(18.056)
Sea-level Pressure (hPa)	1017.158	1016.150	1016.579	1021.632	1017.183
	(9.992)	(9.260)	(9.221)	(12.655)	(9.726)
Wind Speed (Km/h)	9.970	5.890	6.589	9.167	7.055
	(3.665)	(3.161)	(2.814)	(3.787)	(2.661)
Precipitation (mm)	1.391	3.606	3.974	0.300	1.700
	(6.226)	(11.005)	(10.722)	(1.436)	(6.516)

Notes: The table lists the sample means at daily level. Standard deviations are shown in parentheses.

Table A1.2: Average Correlation among Monitoring Stations within Each City

City	Average Correlation AQI	Average Correlation PM2.5	Average Correlation CO	Average Correlation NO2	Average Correlation SO2	Average Correlation O3
Beijing	0.885	0.952	0.933	0.881	0.917	0.944
Changsha	0.925	0.966	0.544	0.692	0.635	0.812
Chengdu	0.830	0.932	0.804	0.716	0.593	0.797
Chongqing	0.831	0.895	0.553	0.672	0.723	0.840
Dongguan	0.907	0.965	0.735	0.885	0.802	0.939
Guangzhou	0.845	0.932	0.454	0.795	0.716	0.860
Hangzhou	0.803	0.866	0.665	0.719	0.522	0.852
Nanjing	0.942	0.972	0.720	0.772	0.802	0.880
Ningbo	0.895	0.967	0.742	0.836	0.815	0.860
Qingdao	0.859	0.953	0.714	0.546	0.786	0.791
Shanghai	0.932	0.966	0.773	0.908	0.899	0.873
Shenyang	0.867	0.944	0.872	0.796	0.810	0.847
Shenzhen	0.762	0.901	0.350	0.568	0.424	0.775
Suzhou	0.942	0.977	0.794	0.819	0.819	0.914
Tianjin	0.881	0.940	0.789	0.908	0.837	0.812
Wuhan	0.897	0.947	0.708	0.802	0.649	0.724
Wuxi	0.926	0.972	0.611	0.749	0.767	0.911
Xian	0.765	0.866	0.678	0.454	0.806	0.859
Zhengzhou	0.905	0.955	0.746	0.834	0.880	0.856
Overall Average	0.874	0.940	0.694	0.755	0.747	0.850

Notes: The table reports the average pairwise correlations for daily average pollutant levels from all the monitoring stations in each city. The mean values under Beijing are the same as those average values from Table A3 to A8.

Table A1.3: Pairwise Correlations among Monitoring Stations in Beijing — AQI

Correlation	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6	Station 7	Station 8	Station 9	Station 10	Station 11	Station 12
Station 1	-											
Station 2	0.792	-										
Station 3	0.957	0.816	-									
Station 4	0.976	0.799	0.962	-								
Station 5	0.947	0.811	0.968	0.960	-							
Station 6	0.962	0.826	0.963	0.961	0.953	-						
Station 7	0.928	0.854	0.936	0.927	0.931	0.961	-					
Station 8	0.883	0.818	0.888	0.888	0.897	0.889	0.879	-				
Station 9	0.840	0.857	0.841	0.841	0.844	0.858	0.864	0.915	-			
Station 10	0.776	0.907	0.795	0.784	0.794	0.804	0.840	0.790	0.821	-		
Station 11	0.931	0.841	0.956	0.938	0.959	0.959	0.950	0.899	0.863	0.812	-	
Station 12	0.912	0.839	0.904	0.900	0.892	0.932	0.940	0.865	0.862	0.817	0.910	-
Average	0.885											
Longitude	116.366	116.170	116.434	116.434	116.473	116.361	116.315	116.720	116.644	116.230	116.407	116.225
Latitude	39.867	40.287	39.952	39.875	39.972	39.943	39.993	40.144	40.394	40.195	40.003	39.928

Notes: The table reports the correlations among daily average AQI generated from each monitoring station in Beijing.

Table A1.4: Pairwise Correlations among Monitoring Stations in Beijing — $PM_{2.5}$

Correlation	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6	Station 7	Station 8	Station 9	Station 10	Station 11	Station 12
Station 1	-											
Station 2	0.868	-										
Station 3	0.968	0.908	-									
Station 4	0.984	0.894	0.992	-								
Station 5	0.963	0.906	0.992	0.989	-							
Station 6	0.971	0.913	0.992	0.990	0.986	-						
Station 7	0.953	0.937	0.978	0.974	0.975	0.988	-					
Station 8	0.948	0.912	0.957	0.960	0.964	0.962	0.960	-				
Station 9	0.918	0.937	0.926	0.926	0.930	0.940	0.949	0.973	-			
Station 10	0.887	0.978	0.921	0.912	0.924	0.923	0.952	0.921	0.934	-		
Station 11	0.957	0.914	0.989	0.985	0.988	0.994	0.986	0.965	0.939	0.925	-	
Station 12	0.958	0.925	0.967	0.970	0.965	0.979	0.983	0.961	0.951	0.933	0.975	-
Average	0.952											
Longitude	116.366	116.170	116.434	116.434	116.473	116.361	116.315	116.720	116.644	116.230	116.407	116.225
Latitude	39.867	40.287	39.952	39.875	39.972	39.943	39.993	40.144	40.394	40.195	40.003	39.928

Notes: The table reports the correlations among daily average $PM_{2.5}$ generated from each monitoring station in Beijing.

Table A1.5: Pairwise Correlations among Monitoring Stations in Beijing — CO

Correlation	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6	Station 7	Station 8	Station 9	Station 10	Station 11	Station 12
Station 1	-											
Station 2	0.858	-										
Station 3	0.953	0.881	-									
Station 4	0.966	0.872	0.987	-								
Station 5	0.950	0.876	0.980	0.978	-							
Station 6	0.958	0.884	0.989	0.988	0.980	-						
Station 7	0.914	0.885	0.962	0.949	0.946	0.966	-					
Station 8	0.928	0.888	0.932	0.933	0.933	0.927	0.902	-				
Station 9	0.909	0.916	0.913	0.914	0.911	0.914	0.893	0.952	-			
Station 10	0.903	0.944	0.922	0.915	0.916	0.923	0.937	0.911	0.927	-		
Station 11	0.954	0.902	0.977	0.974	0.979	0.981	0.943	0.941	0.922	0.929	-	
Station 12	0.937	0.895	0.942	0.943	0.944	0.951	0.952	0.907	0.911	0.936	0.948	-
Average	0.933											
Longitude	116.366	116.170	116.434	116.434	116.473	116.361	116.315	116.720	116.644	116.230	116.407	116.225
Latitude	39.867	40.287	39.952	39.875	39.972	39.943	39.993	40.144	40.394	40.195	40.003	39.928

Notes: The table reports the correlations among daily average CO generated from each monitoring station in Beijing.

Table A1.6: Pairwise Correlations among Monitoring Stations in Beijing — NO_2

Correlation	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6	Station 7	Station 8	Station 9	Station 10	Station 11	Station 12
Station 1	-											
Station 2	0.799	-										
Station 3	0.962	0.780	-									
Station 4	0.962	0.794	0.965	-								
Station 5	0.946	0.808	0.967	0.948	-							
Station 6	0.968	0.802	0.961	0.955	0.949	-						
Station 7	0.919	0.750	0.882	0.858	0.886	0.904	-					
Station 8	0.897	0.796	0.894	0.872	0.914	0.881	0.868	-				
Station 9	0.827	0.885	0.780	0.792	0.831	0.817	0.818	0.877	-			
Station 10	0.866	0.941	0.856	0.864	0.884	0.872	0.831	0.876	0.892	-		
Station 11	0.932	0.789	0.970	0.948	0.948	0.932	0.847	0.883	0.773	0.844	-	
Station 12	0.943	0.843	0.933	0.929	0.919	0.935	0.887	0.875	0.826	0.881	0.927	-
Average	0.881											
Longitude	116.366	116.170	116.434	116.434	116.473	116.361	116.315	116.720	116.644	116.230	116.407	116.225
Latitude	39.867	40.287	39.952	39.875	39.972	39.943	39.993	40.144	40.394	40.195	40.003	39.928

Notes: The table reports the correlations among daily average NO_2 generated from each monitoring station in Beijing.

Table A1.7: Pairwise Correlations among Monitoring Stations in Beijing — SO_2

Correlation	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6	Station 7	Station 8	Station 9	Station 10	Station 11	Station 12
Station 1	-											
Station 2	0.869	-										
Station 3	0.973	0.882	-									
Station 4	0.952	0.849	0.937	-								
Station 5	0.974	0.906	0.982	0.944	-							
Station 6	0.976	0.890	0.972	0.936	0.979	-						
Station 7	0.971	0.903	0.969	0.922	0.971	0.975	-					
Station 8	0.903	0.858	0.898	0.896	0.918	0.900	0.901	-				
Station 9	0.875	0.849	0.875	0.879	0.893	0.877	0.875	0.926	-			
Station 10	0.887	0.936	0.879	0.881	0.917	0.909	0.921	0.893	0.877	-		
Station 11	0.936	0.884	0.944	0.900	0.956	0.964	0.953	0.878	0.848	0.914	-	
Station 12	0.961	0.878	0.937	0.912	0.948	0.955	0.967	0.903	0.878	0.914	0.922	-
Average	0.917											
Longitude	116.366	116.170	116.434	116.434	116.473	116.361	116.315	116.720	116.644	116.230	116.407	116.225
Latitude	39.867	40.287	39.952	39.875	39.972	39.943	39.993	40.144	40.394	40.195	40.003	39.928

Notes: The table reports the correlations among daily average SO_2 generated from each monitoring station in Beijing.

Table A1.8: Pairwise Correlations among Monitoring Stations in Beijing — O_3

Correlation	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6	Station 7	Station 8	Station 9	Station 10	Station 11	Station 12
Station 1	-											
Station 2	0.900	-										
Station 3	0.974	0.907	-									
Station 4	0.974	0.900	0.983	-								
Station 5	0.968	0.905	0.990	0.976	-							
Station 6	0.969	0.901	0.984	0.972	0.978	-						
Station 7	0.946	0.899	0.972	0.951	0.970	0.969	-					
Station 8	0.937	0.912	0.960	0.950	0.961	0.942	0.936	-				
Station 9	0.884	0.915	0.906	0.900	0.905	0.886	0.885	0.957	-			
Station 10	0.933	0.943	0.948	0.941	0.947	0.942	0.947	0.957	0.942	-		
Station 11	0.953	0.902	0.975	0.963	0.974	0.968	0.970	0.948	0.902	0.949	-	
Station 12	0.969	0.896	0.974	0.965	0.969	0.972	0.970	0.936	0.883	0.947	0.971	-
Average	0.944											
Longitude	116.366	116.170	116.434	116.434	116.473	116.361	116.315	116.720	116.644	116.230	116.407	116.225
Latitude	39.867	40.287	39.952	39.875	39.972	39.943	39.993	40.144	40.394	40.195	40.003	39.928

Notes: The table reports the correlations among daily average O_3 generated from each monitoring station in Beijing.

Table A1.9: Target City and Source Instrumental Cities

Target Cities	Instrumental Cities	Coordinate (°)	Distance (km)	Location (°)
Beijing	Chengde	40.97N 117.94E	175.93	56
39.96N 116.43E	Tangshan	39.62N 118.18E	154.4	101
	Tianjin	39.09N 117.20E	113.34	140
	Langfang	39.54N 116.68E	48.48	150
	Baoding	38.89N 115.47E	140.06	222
	Zhangjiakou	40.76N 114.88E	161.34	297
Changsha	Yueyang	29.36N 113.31E	126.24	10
28.23N 112.94E	Xinyu	27.81N 114.92E	200	102
	Yichun	27.81N 114.42E	152.5	105
	Pingxiang	27.62N 113.85E	112.43	123
	Zhuzhou	27.83N 113.13E	48.86	159
	Xiangtan	27.83N 112.94E	44.99	180
	Hengyang	26.89N 112.57E	153.68	195
	Shaoyang	27.25N 111.47E	181.64	236
	Loudi	27.69N 111.99E	110.36	240
	Yiyang	28.55N 112.36E	66.7	300
	Changde	29.03N 111.7E	149.44	303
Chengdu	Aba	31.91N 102.22E	229.74	305
30.58N 104.07E	Fanzi	30.05N 101.96E	209.96	256
	Yaan	30.00N 103.02E	119.56	241
	Leshan	29.57N 103.76E	115.40	196
	Yibin	28.73N 104.65E	208.80	162
	Zigong	29.33N 104.78E	153.36	150
	Meishan	30.08N 103.85E	58.52	203
	Ziyang	30.13N 104.63E	72.59	128
	Neijiang	29.59N 105.05E	144.97	134
	Luzhou	28.89N 105.43E	229.73	140
	Chongqing	29.56N 106.54E	301.39	112
	Suining	30.54N 105.59E	147.17	91
	Guangan	30.47N 106.63E	246.20	92
	Nanchong	30.85N 106.13E	197.34	82
	Bazhong	31.87N 106.75E	292.92	64
	Guangyuan	32.44N 105.84E	267.09	43
	Deyang	31.13N 104.40E	69.52	30
	Mianyang	31.47N 104.68E	115.30	34

continued

Target Cities	Instrumental Cities	Coordinate (°)	Distance (km)	Location (°)
Chongqing				
29.56N 106.54E	Guangan	30.45N 106.64E	99.19	6
	Dazhou	31.21N 107.47E	203.96	29
	Zunyi	27.73N 106.92E	207.33	168
	Luzhou	28.86N 105.44E	133.13	238
	Yibin	28.73N 104.65E	206.64	246
	Zigong	29.33N 104.78E	173.39	263
	Neijiang	29.58N 105.05E	145.03	271
	Ziyang	30.13N 104.63E	195.25	286
	Suining	30.54N 105.59E	142.07	316
	Nanchong	30.85N 106.13E	148.98	342
Dongguan	Huizhou	23.11N 114.41E	68.26	82
23.02N 113.75E	Heyuan	23.76N 114.7E	127.16	52
	Meizhou	24.3N 116.12E	280.21	61
	Jieyang	23.58N 116.37E	274.72	77
	Chaozhou	23.67N 116.62E	301.64	77
	Shantou	23.37N 116.68E	302.32	83
	Shanwei	22.81N 115.37E	167.44	97
	Zhuhai	22.28N 113.58E	84.11	192
	Yangjiang	21.86N 111.98E	223.00	236
	Jiangmen	22.58N 113.08E	84.49	236
	Foshan	23.03N 113.13E	63.46	270
	Yunfu	22.91N 112.04E	175.50	266
	Zhaoqing	23.02N 112.48E	129.97	0
	Guangzhou	23.12N 113.27E	50.35	281
	Wuzhou	23.46N 111.27E	258.06	280
	Hezhou	24.41N 111.57E	270.65	302
	Qingyuan	23.68N 113.06E	101.72	313
	Shaoguan	24.8N 113.6E	198.51	355
Guangzhou	Shaoguan	24.8N 113.6E	118.05	11
23.12N 113.27E	Heyuan	23.76N 114.7E	162.57	66
	Huizhou	23.11N 114.41E	117.49	91
	Dongguan	23.02N 113.75E	52.29	101
	Shenzhen	22.55N 114.06E	104.1	126
	Zhuhai	22.28N 113.58E	99.92	160
	Jiangmen	22.58N 113.08E	64.29	119
	Yangjiang	21.86N 111.98E	192.95	226
	Foshan	23.03N 113.13E	18.27	236
	Yunfu	22.91N 112.04E	127.79	260
	Zhaoqing	23.02N 112.48E	80.5	263
	Qingyuan	23.68N 113.06E	64.85	340

continued

Target Cities	Instrumental Cities	Coordinate (°)	Distance (km)	Location (°)
Hangzhou				
30.28N 120.15E	Suzhou	31.32N 120.59E	118.59	23
	Jiaxing	30.75N 120.76E	78.84	52
	Shanghai	31.23N 121.47E	164.45	54
	Zhoushan	30.02N 122.21E	200.09	97
	Ningbo	29.88N 121.54E	142.69	105
	Shaoxing	30.01N 120.61E	52.9	124
	Lishui	28.48N 119.95E	203.64	187
	Jinhua	29.06N 119.65E	140.28	203
	Quzhou	29N 118.9E	188.8	225
	Huangshan	29.72N 118.38E	183.42	253
	Xuancheng	30.94N 118.76E	152.66	295
	Wuhu	31.37N 118.42E	203.74	302
	Huzhou	30.89N 120.08E	68.97	355
	Changzhou	31.81N 119.97E	172.53	353
Nanjing	Huaian	33.6N 119.02E	172.28	8
32.06N 118.79E	Yangzhou	32.38N 119.41E	68.42	63
	Taizhou	32.45N 119.91E	114.29	71
	Zhenjiang	32.2N 119.43E	61.71	78
	Nantong	31.96N 120.89E	199.18	93
	Changzhou	31.81N 119.97E	115.35	102
	Wuxi	31.48N 120.3E	156.81	111
	Suzhou	31.32N 120.59E	189.96	113
	Huzhou	30.89N 120.08E	178.78	132
	Xuancheng	30.94N 118.76E	126.48	181
	Wuhu	31.37N 118.42E	86.46	206
	Maanshan	31.66N 118.51E	51.54	215
	Tongling	30.94N 117.82E	153.9	221
	Chizhou	30.66N 117.5E	198.53	223
	Hefei	31.83N 117.23E	149.82	262
	Huainan	32.64N 117.01E	179.49	288
	Chuzhou	32.25N 118.33E	48.49	293
	Bengbu	32.91N 117.39E	162.03	301
	Suqian	33.96N 118.28E	216.55	345
				continued

Target Cities	Instrumental Cities	Coordinate (°)	Distance (km)	Location (°)
Ningbo	Shanghai	31.23N 121.47E	150.26	357
29.88N 121.54E	Nantong	31.96N 120.89E	239.45	342
	Changzhou	31.81N 119.97E	262.01	320
	Wuxi	31.48N 120.30E	213.80	322
	Suzhou	31.32N 120.59E	183.61	326
	Xuancheng	30.94N 118.76E	291.47	290
	Hangzhou	30.28N 120.15E	141.16	286
	Shaoxing	30.01N 120.61E	90.76	277
	Huangshan	29.72N 118.38E	305.45	267
	Quzhou	29.00N 118.90E	273.73	251
	Jinhua	29.06N 119.65E	204.42	246
	Lishui	28.48N 119.95E	219.34	228
	Wenzhou	28.00N 120.69E	224.78	204
	Taizhou	28.66N 121.42E	136.17	185
	Zhoushan	30.02N 122.21E	66.40	78
Qingdao	Yantai	37.46N 121.46E	181.46	37
36.08N 120.39E	Weifang	36.72N 119.16E	131.51	297
	Jinan	36.66N 117.11E	300.80	280
	Linyi	35.11N 118.35E	213.50	244
	Lianyungang	34.61N 119.21E	194.33	218
	Weihai	37.52N 122.11E	222.18	50
	Dongying	37.45N 118.67E	216.40	308
	Binzhou	37.39N 117.97E	260.94	298
	Zibo	36.82N 118.06E	224.27	287
	Taian	36.21N 117.08E	297.57	272
	Laiwu	36.22N 117.68E	243.63	273
	Rizhao	35.43N 119.52E	106.50	233
	Suqian	33.96N 118.28E	303.02	224
	Huaian	33.56N 119.11E	303.57	206
	Zaozhuang	34.82N 117.33E	310.41	247
Shanghai	Zhoushan	30.02N 122.21E	155.64	149
31.23N 121.47E	Ningbo	29.88N 121.54E	150.87	177
	Shaoxing	30.01N 120.61E	157.34	215
	Hangzhou	30.28N 120.15E	164.45	233
	Suzhou	31.32N 120.59E	82.2	276
	Changzhou	31.81N 119.97E	154.08	292
	Taizhou	32.45N 119.91E	198.03	309
	Nantong	31.96N 120.89E	100.1	323

continued

Target Cities	Instrumental Cities	Coordinate (°)	Distance (km)	Location (°)
Shenyang	Fushun	41.87N 123.96E	44.18	83
41.81N 123.43E	Tonghua	41.73N 125.94E	208.14	91
	Baishan	41.95N 126.41E	247.10	87
	Benxi	41.50N 123.69E	40.49	140
	Anshan	41.14N 122.99E	83.61	212
	Yingkou	40.69N 122.23E	160.51	226
	Panjin	41.14N 122.06E	136.40	243
	Chaoyang	41.60N 120.43E	249.88	265
	Fuxin	42.03N 121.68E	146.73	277
	Tongliao	43.66N 122.24E	226.89	327
	Changchun	43.84N 125.32E	272.86	43
	Siping	43.18N 124.35E	170.16	33
	Tieling	42.23N 123.72E	52.06	34
Shenzhen	Heyuan	23.76N 114.70E	148.86	27
22.55N 114.06E	Meizhou	24.30N 116.12E	286.70	49
	Jieyang	23.58N 116.37E	262.38	65
	Shanwei	22.81N 115.37E	137.27	78
	Shantou	23.37N 116.68E	283.63	72
	Yangjiang	21.86N 111.98E	226.50	251
	Wuzhou	23.46N 111.27E	302.84	288
	Zhaoqing	23.02N 112.48E	174.50	288
	Guangzhou	23.12N 113.27E	106.03	306
	Dongguan	23.02N 113.75E	61.85	326
	Shaoguan	24.80N 113.60E	257.64	348
Suzhou	Taizhou	28.66N 121.42E	305.37	162
31.32N 120.59E	Jinhua	29.06N 119.65E	261.33	203
	Hangzhou	30.28N 120.15E	120.78	202
	Nanjing	32.06N 118.79E	190.64	293
	Nantong	31.96N 120.89E	80.13	25
	Shanghai	31.23N 121.47E	83.71	95

continued

Target Cities	Instrumental Cities	Coordinate (°)	Distance (km)	Location (°)
Tianjin	Tangshan	39.62N 118.18E	103.88	60.84
39.09N 117.19E	Binzhou	37.39N 117.97E	200.5	155.381
	Cangzhou	38.31N 116.84E	92.16	304.27
	Dezhou	37.44N 116.36E	198.12	206.83
	Hengshui	37.73N 115.66E	200.69	228.52
	Baoding	38.89N 115.47E	151.83	263.06
	Langfang	39.54N 116.68E	65.5	310.95
	Beijing	39.96N 116.43E	113.34	319.13
Wuhan	Huanggang	30.45N 14.88E	56.93	104.84
30.61N 114.33E	Hangshi	30.19N 115.05E	84.53	119.04
	Jiujiang	29.71N 116.00E	191.59	117.81
	Xianning	29.83N 114.33E	84.72	178.51
	Yueyang	29.36N 113.31E	178.8	222.77
	Jinzhou	30.34N 112.24E	198.96	262.81
	Xiaogan	30.91N 113.94E	49.14	310.22
	Suizhou	31.69N 113.4E	149.01	319.99
	Xinyang	32.15N 114.09E	173.69	351.97
Wuxi	Suzhou	31.32N 120.59E	32.77	118
31.48N 120.30E	Shanghai	31.23N 121.47E	114.52	102
	Jiaxing	30.75N 120.76E	92.23	147
	Ningbo	29.88N 121.54E	213.80	142
	Zhoushan	30.02N 122.21E	244.27	127
	Shaoxing	30.01N 120.61E	166.12	168
	Hangzhou	30.28N 120.15E	134.10	187
	Jinhua	29.06N 119.65E	276.24	195
	Huzhou	30.89N 120.08E	68.86	200
	Huangshan	29.72N 118.38E	268.74	227
	Chizhou	30.66N 117.50E	281.83	253
	Tongling	30.94N 117.82E	243.37	90
	Xuancheng	30.94N 118.76E	158.29	250
	Wuhu	31.37N 118.42E	178.80	266
	Heifei	31.83N 117.23E	293.17	276
	Maanshan	31.66N 118.51E	170.76	275
	Chuzhou	32.25N 118.33E	204.79	291
	Nanjing	32.06N 118.79E	156.64	291
	Zhenjiang	32.2N 119.43E	114.73	309
	Yangzhou	32.38N 119.41E	130.65	315
	Huaian	33.60N 119.02E	264.51	328
	Taizhou	32.45N 119.91E	113.96	338
	Changzhou	31.81N 119.97E	48.19	315

continued

Target Cities	Instrumental Cities	Coordinate (°)	Distance (km)	Location (°)
Wuxi	Suzhou	31.32N 120.59E	32.77	118
31.48N 120.30E	Shanghai	31.23N 121.47E	114.52	102
	Jiaxing	30.75N 120.76E	92.23	147
	Ningbo	29.88N 121.54E	213.80	142
	Zhoushan	30.02N 122.21E	244.27	127
	Shaoxing	30.01N 120.61E	166.12	168
	Hangzhou	30.28N 120.15E	134.10	187
	Jinhua	29.06N 119.65E	276.24	195
	Huzhou	30.89N 120.08E	68.86	200
	Huangshan	29.72N 118.38E	268.74	227
	Chizhou	30.66N 117.50E	281.83	253
	Tongling	30.94N 117.82E	243.37	90
	Xuancheng	30.94N 118.76E	158.29	250
	Wuhu	31.37N 118.42E	178.80	266
	Heifei	31.83N 117.23E	293.17	276
	Maanshan	31.66N 118.51E	170.76	275
	Chuzhou	32.25N 118.33E	204.79	291
	Nanjing	32.06N 118.79E	156.64	291
	Zhenjiang	32.20N 119.43E	114.73	309
	Yangzhou	32.38N 119.41E	130.65	315
	Huaian	33.60N 119.02E	264.51	328
	Taizhou	32.45N 119.91E	113.96	338
	Changzhou	31.81N 119.97E	48.19	315
Xi'an	Shiyan	32.65N 110.77E	253.39	132
34.34N 108.94E	Shangluo	33.87N 109.93E	104.93	115
	Ankang	32.69N 109.02E	183.80	177
	Hanzhong	33.06N 107.02E	227.65	236
	Tianshui	34.59N 105.70E	298.41	274
	Baoji	34.37N 107.24E	156.23	271
	Pingliang	35.54N 106.63E	249.23	297
	Guyuan	36.01N 106.24E	307.84	301
	Qingyang	35.71N 107.64E	193.39	316
	Tongchuan	34.90N 108.93E	62.41	358
	Yanan	36.59N 109.49E	255.56	13
	Weinan	34.50N 109.49E	53.67	73
	Yuncheng	35.04N 111.00E	203.28	71
	Linfen	36.10N 111.53E	305.97	55
	Sanmenxia	34.80N 111.20E	212.86	78
	Xianyang	34.34N 108.71E	21.57	270

Table A1.10: Robustness — Accounting for Daily Wind Speed in IV

	AQI		PM2.5	
	(1)	(2)	(3)	(4)
First Stage$^{\mathbf{(a)}}$				
$P_{source_{it}}$	0.529***	0.268***	0.536***	0.270***
	(0.083)	(0.078)	(0.076)	(0.058)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
$P_{source_{it}} * windspeed_{it}$	-0.010	-0.003	-0.009	-0.001
	(0.007)	(0.008)	(0.005)	(0.005)
	[0.249]	[0.721]	[0.123]	[0.873]
$P_{source_{it-1}}$		0.726***		0.778***
		(0.063)		(0.067)
		[<0.01]		[<0.01]
$P_{source_{it-1}} * windspeed_{it-1}$		-0.035***		-0.045***
		(0.005)		(0.007)
		[<0.01]		[<0.01]
Kleibergen-Paap rk Wald F statistic	40.422	95.029	49.312	103.185
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	19.93	16.38	19.93
Second Stage$^{\mathbf{(b)}}$				
Instrumented pollutant	0.219**	0.121**	0.297**	0.170**
	(0.089)	(0.049)	(0.116)	(0.061)
	[0.026]	[0.02]	[0.034]	[0.012]
Observations	12904	12579	12989	12662
Additional Control				
City FEs	Y	Y	Y	Y
Temporal Controls	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y

Notes: (a) Dependent variable in the first stage is daily-mean pollutant of target city, and independent variable is source pollutant ($100km < d_{ij} < 300km$) from upwind direction (within 90 degrees to the wind), and its interaction term with wind speed in the target city. (b) Second stage reports the results regressing log form of Sleeplessness Index on the instrumented daily pollution. Column (2) and (4) incorporate day before as an additional instrument. Temporal controls include year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects. Weather controls contain temperature, humidity, precipitation, wind speed and sea-level pressure. Temperature and humidity are measured by the way of bins. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.11: Alternative Standard Errors — OLS

	Wild Cluster Bootstrap	Driscoll-Kraay Spatial Correlation	Alternative Clusters			
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: AQI	0.037***	0.037***	0.037***	0.037***	0.037***	0.037***
	[0.001]	(0.010)	(0.013)	(0.012)	(0.010)	(0.012)
Observations	12365	12365	12365	12365	12365	12365
Panel B: PM2.5	0.043***	0.043***	0.043***	0.043***	0.043***	0.043***
	[0.012]	(0.012)	(0.013)	(0.013)	(0.011)	(0.017)
Observations	13617	13617	13617	13617	13617	13617
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y	Y	Y
City by quarter FEs	Y	Y	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y	Y	Y
Clusters	City (19)	-	City by year by season (152)	City by year by month (456)	City by year by week (1976)	City Year by month

Notes: Dependent variable is log form of Sleeplessness Index. Independent variable is city daily-mean value of specific pollutant. Column (1) implements the wild bootstrap procedure as described in Cameron et al. (2008), which replicates the results under Column (7) in Table 1.4. Column (3) follows Driscoll and Kraay (1998) to consider spatial correlation. Column (3) through Column (5) are clustered at city by year by season, city by year by month and city by year by week, respectively. Column (6) adopts the multi-way clusters at both city and year by month. Temporal controls include year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects. Weather controls contain temperature, humidity, precipitation, wind speed and sea-level pressure. Temperature and humidity are measured by the way of bins. Robust standard errors clustered at different levels are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets in Column (1). Robust standard errors clustered at alternative levels are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.12: Alternative Standard Errors — IV

AQI	Wild Cluster Bootstrap	Driscoll-Kraay Spatial Correlation	Alternative Clusters			
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage						
Instrumental AQI t	0.455***	0.455***	0.455***	0.455***	0.455***	0.455***
	[<0.01]	(0.056)	(0.058)	(0.051)	(0.047)	(0.080)
F-statistic	54.791	65.3	61.362	77.607	92.526	32.120
Second Stage						
Instrumented AQI	0.223**	0.223***	0.223**	0.223***	0.223***	0.223**
	[0.039]	(0.062)	(0.098)	(0.089)	(0.062)	(0.091)
Observations	12904	12904	12904	12904	12904	12904
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y	Y	Y
City by quarter FEs	Y	Y	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y	Y	Y
Clusters	City (19)	-	City by year by season (152)	City by year by month (456)	City by year by week (1976)	City Year by month

Notes: Dependent variable in the first stage is daily-mean AQI of local city, and independent variable is daily weighted average pollution of source cities. Second stage reports the results regressing log Sleeplessness Index on the instrumented daily pollution. Column (1) repeats the IV results under Column (1) and Column (3) in Table 1.6, in which wild bootstrap clustered at city is used to indicate the significance level. Column (3) follows Driscoll and Kraay (1998) to consider spatial correlation. Column (3) through Column (5) are clustered at city by year by season, city by year by month and city by year by week, respectively. Column (6) adopts the multi-way clusters at both city and year by month. Temporal controls include year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects. Weather controls contain temperature, humidity, precipitation, wind speed and sea-level pressure. Temperature and humidity are measured by the way of bins. Robust standard errors clustered at different levels are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets in Column (1). Robust standard errors clustered at alternative levels are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.13: Alternative Standard Errors — IV

PM2.5	Wild Cluster Bootstrap	Driscoll-Kraay Spatial Correlation	Alternative Clusters			
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage						
Instrumental PM2.5 t	0.463***	0.463***	0.463***	0.463***	0.463***	0.463***
	[<0.01]	(0.058)	(0.063)	(0.057)	(0.052)	(0.075)
F-statistic	56.846	62.820	53.071	65.903	78.358	38.450
Second Stage						
Instrumented PM2.5	0.285**	0.285***	0.285***	0.285***	0.285***	0.285**
	[0.033]	(0.068)	(0.113)	(0.100)	(0.068)	(0.117)
Observations	12989	12989	12989	12989	12989	12989
City FEs	Y	Y	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y	Y	Y
City by quarter FEs	Y	Y	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y	Y	Y
Clusters	City (19)	-	City by year by season (152)	City by year by month (456)	City by year by week (1976)	City Year by month

Notes: Dependent variable in the first stage is daily-mean $PM_{2.5}$ of local city, and independent variable is daily weighted average pollution of source cities. Second stage reports the results regressing log Sleeplessness Index on the instrumented daily pollution. Column (1) repeats the IV results under Column (1) and Column (3) in Table 1.6, in which wild bootstrap clustered at city is used to indicate the significance level. Column (3) follows Driscoll and Kraay (1998) to consider spatial correlation. Column (3) through Column (5) are clustered at city by year by season, city by year by month and city by year by week, respectively. Column (6) adopts the multi-way clusters at both city and year by month. Temporal controls include year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects. Weather controls contain temperature, humidity, precipitation, wind speed and sea-level pressure. Temperature and humidity are measured by the way of bins. Robust standard errors clustered at different levels are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets in Column (1). Robust standard errors clustered at alternative levels are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.14: City Sub-samples — OLS

	Full	Exclude Beijing and Environ	Exclude Shanghai and Environs	Exclude Guangzhou and Environs
	(1)	(2)	(3)	(4)
Panel A: AQI	0.037***	0.026**	0.041***	0.038***
	(0.012)	(0.010)	(0.013)	(0.013)
	[0.001]	[0.019]	[0.002]	[<0.01]
Panel B: PM2.5	0.043***	0.026**	0.047***	0.043***
	(0.017)	(0.011)	(0.018)	(0.017)
	[0.012]	[0.032]	[0.005]	[0.009]
Observations	13617	12173	11479	11433
Additional Controls				
City FEs	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y
City by quarter FEs	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y

Notes: Column (1) replicates the OLS results in Column (7) of Table 1.4. Column (2) excludes Beijing and its nearby city, Tianjin, both of which are situated in northern heavy industrial region. Column (3) excludes Shanghai and its nearby cities, Suzhou and Hangzhou, which are coastally located and dominated by light industry. Column (4) exludes the cleanest part in Southern China, Guangzhou and its nearby cities, Shenzhen and Dongguan. All the regressions include city fixed effects, temporal controls (year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects) and weather controls (temperature, humidity, precipitation, wind speed and sea-level pressure). Temperature and humidity are measured in the form of bins. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.15: City Sub-samples — IV

	Full	Exclude Beijing and Environ	Exclude Shanghai and Environs	Exclude Guangzhou and Environs
	(1)	(2)	(3)	(4)
	Panel A: AQI			
First Stage				
Instrumental AQI t	0.455***	0.509***	0.469***	0.436***
	(0.062)	(0.060)	(0.067)	(0.061)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Kleibergen-Paap rk Wald F statistic	54.791	72.472	48.488	51.714
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	16.38	16.38	16.38
Second Stage				
Instrumented Pollutant	0.223**	0.134*	0.213*	0.239**
	(0.096)	(0.073)	(0.098)	(0.105)
	[0.039]	[0.079]	[0.076]	[0.047]
Observations	12904	11461	11037	10720
	Panel B: PM2.5			
First Stage				
Instrumental PM2.5 t	0.463***	0.524***	0.473***	0.444***
	(0.061)	(0.051)	(0.067)	(0.060)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Kleibergen-Paap rk Wald F statistic	56.846	72.472	48.488	51.714
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	16.38	16.38	16.38
Second Stage				
	0.285**	0.177*	0.276*	0.302**
	(0.118)	(0.094)	(0.119)	(0.128)
	[0.033]	[0.078]	[0.056]	[0.052]
Observations	12989	11546	11122	10805
Additional Controls				
City FEs	Y	Y	Y	Y
Temporal Controls	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y

Notes: Column (1) repeats the IV results under Column (1) and Column (3) in Table 1.6 with full sample. Column (2) excludes Beijing and its nearby city, Tianjin, both of which are situated in northern heavy industrial region. Column (3) excludes Shanghai and its nearby cities, Suzhou and Hangzhou, which are coastally located and dominated by light industry. Column (4) exludes the cleanest part in Southern China, Guangzhou and its nearby cities, Shenzhen and Dongguan. All the regressions include city fixed effects, temporal controls (year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects) and weather controls (temperature, humidity, precipitation, wind speed and sea-level pressure). Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.16: Individual City Effect — AQI

Independent Variable	Dependent Variable (Ln(Sleepless))										
(Daily Pollutant)	Full	Beijing	Changsha	Chengdu	Chongqing	Dongguan	Guangzhou	Hangzhou	Nanjing	Ningbo	Qingdao
Individual City	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A11: AQI-OLS											
AQI (OLS)	0.037***	0.050	0.046***	0.049***	0.019	0.057**	0.095	0.089**	0.058***	-0.022	0.038**
	[0.001]	(0.046)	(0.015)	(0.017)	(0.017)	(0.024)	(0.089)	(0.037)	(0.018)	(0.037)	(0.018)
Observations	13617	730	713	714	730	724	730	714	714	713	714
Panel A21: AQI-IV											
First Stage											
Instrumental AQI t	0.455**	0.300***	0.781***	0.968***	0.835***	0.690***	0.737***	0.460***	0.937***	0.688***	0.394***
	[<0.01]	(0.112)	(0.122)	(0.165)	(0.092)	(0.098)	(0.101)	(0.182)	(0.161)	(0.126)	(0.108)
F-statistics	54.791	7.13	41.191	34.274	82.471	49.529	53.054	6.407	33.945	29.686	13.454
Second Stage											
Instrumented AQI	0.223**	-0.510	0.190***	0.302***	0.244***	0.288***	-1.573***	1.642***	0.153***	0.017	0.104
	[0.039]	(0.398)	(0.057)	(0.075)	(0.054)	(0.098)	(0.411)	(0.626)	(0.050)	(0.124)	(0.128)
Observations	12904	729	575	651	730	724	730	599	714	692	714

continued

Table A1.17: Individual City Effect — AQI

Independent Variable (Daily Pollutant) Individual City	Dependent Variable (Ln(Sleepless)) Full (1)	Shanghai (2)	Shenyang (3)	Shenzhen (4)	Suzhou (5)	Tianjin (6)	Wuhan (7)	Wuxi (8)	Xian (9)	Zhengzhou (10)
					Panel A12: AQI-OLS					
AQI (OLS)	0.037***	-0.010	0.061***	0.121**	-0.005	0.039	0.080***	0.023	-0.030	0.019*
	[0.001]	(0.022)	(0.019)	(0.051)	(0.041)	(0.025)	(0.016)	(0.028)	(0.020)	(0.011)
Observations	13617	714	707	730	710	712	713	708	713	714
					Panel A22: AQI-IV					
First Stage										
Instrumental AQI t	0.455**	0.497***	0.504***	0.526***	0.582***	0.699***	0.613***	0.373***	0.619***	0.504***
	[<0.01]	(0.110)	(0.170)	(0.059)	(0.154)	(0.139)	(0.097)	(0.132)	(0.150)	(0.111)
F-statistics	54.791	20.438	8.745	79.673	14.268	24.05	40.046	7.937	17.029	20.572
Second Stage										
Instrumented AQI	0.223**	-0.076	0.337***	0.745***	0.471**	0.037	0.281***	0.795***	-0.261***	0.452***
	[0.039]	(0.152)	(0.121)	(0.166)	(0.205)	(0.097)	(0.061)	(0.305)	(0.101)	(0.120)
Observations	12904	699	590	730	569	712	708	708	616	714

Notes: Column (1) repeats the preferred results for *AQI* in Table 1.4 and Table 1.6. Column (2) through Column (11) present individual health effect of *AQI* for each city via both OLS and IV. The instruments are as used in the Column (6) of Table 1.5. All the regressions include temporal controls (year by season fixed effect) and weather controls (average temperature bins, average humidity bins, precipitation, sea-level pressure, and wind speed). Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.18: Individual City Effect — PM2.5

Independent Variable (Daily Pollutant) Individual City	Full (1)	Beijing (2)	Changsha (3)	Chengdu (4)	Chongqing (5)	Dongguan (6)	Guangzhou (7)	Hangzhou (8)	Nanjing (9)	Ningbo (10)	Qingdao (11)
	Dependent Variable (Ln(Sleepless))										
						Panel B11: PM2.5-OLS					
PM2.5 (OLS)	0.043***	0.064	0.056***	0.058***	0.039*	0.100***	-0.071	0.100**	0.067***	-0.022	0.039*
	[0.017]	(0.050)	(0.018)	(0.021)	(0.022)	(0.040)	(0.120)	(0.044)	(0.020	(0.045)	(0.020)
Observations	13617	730	713	714	730	724	730	714	714	713	714
						Panel B21: PM2.5-IV					
First Stage											
Instrumental PM2.5 t	0.463***	0.306***	0.725***	1.040***	0.787***	0.776***	0.787***	0.360**	0.952***	0.724***	0.524***
	[<0.01]	(0.112)	(0.121)	(0.172)	(0.088)	(0.084)	(0.092)	(0.172)	(0.174)	(0.120)	(0.101)
F-statistics	56.846	7.467	36.137	36.506	79.96	84.809	73.105	4.377	29.83	36.298	26.665
Second Stage											
Instrumented PM2.5	0.285**	-0.192	0.244***	0.347***	0.332***	0.301***	-1.668***	2.374**	0.193***	0.174	0.069
	[0.033]	(0.372)	(0.073)	(0.084)	(0.071)	(0.104)	(0.430)	(1.092)	(0.061)	(0.136)	(0.104)
Observations	12989	729	579	651	730	724	730	599	714	692	714
											continued

Table A1.19: Individual City Effect — PM2.5

Independent Variable (Daily Pollutant) Individual City	Dependent Variable (Ln(Sleepless)) Full (1)	Shanghai (2)	Shenyang (3)	Shenzhen (4)	Suzhou (5)	Tianjin (6)	Wuhan (7)	Wuxi (8)	Xian (9)	Zhengzhou (10)
	Panel B12: PM2.5-OLS									
PM2.5 (OLS)	0.043***	-0.011	0.068***	0.148**	-0.012	0.036	0.095***	0.031	0.007	0.031***
	[0.017]	(0.027)	(0.011)	(0.064)	(0.052)	(0.029)	(0.019)	(0.036)	(0.023)	(0.012)
Observations	13617	714	707	730	710	712	713	708	713	714
	Panel B22: PM2.5-IV									
First Stage										
Instrumental PM2.5 t	0.463***	0.569***	0.628***	0.564***	0.617***	0.595***	0.641***	0.419***	0.790***	0.492***
	[<0.01]	(0.100)	90.2240	(0.061)	(0.147)	(0.129)	(0.101)	(0.133)	(0.154)	(0.107)
F-statistics	56.846	32.357	7.865	86.125	17.612	20.074	40.174	9.983	26.466	21.007
Second Stage										
Instrumented PM2.5	0.285**	-0.032	0.239***	0.907***	0.636***	0.072	0.360***	1.042***	-0.059	0.526***
	[0.033]	(0.155)	(0.096)	(0.194)	(0.239)	(0.125)	(0.074)	(0.352)	(0.093)	(0.136)
Observations	12989	699	290	730	569	712	708	708	701	714

Notes: Column (1) repeats the preferred results for $PM_{2.5}$ in Table 1.4 and Table 1.6. Column (2) through Column (11) present individual health effect of $PM_{2.5}$ for each city via both OLS and IV. The instruments are as used in the Column (6) of Table 1.5. All the regressions include temporal controls (year by season fixed effect) and weather controls (average temperature bins, average humidity bins, precipitation, sea-level pressure, and wind speed). Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.20: Air Quality and Sleeplessness — All Coefficients

	AQI		PM2.5	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Pollutant	0.037***	0.223**	0.043***	0.285**
	(0.012)	(0.096)	(0.017)	(0.118)
	[0.001]	[0.039]	[0.012]	[0.033]
Average Temperature (T∈ [10,15) Omitted)				
T <0	8.294	12.653**	7.996	11.573*
	(5.579)	(5.537)	(5.494)	(5.457)
	[0.203]	[0.025]	[0.218]	[0.067]
T ∈[0,5)	9.926***	11.706***	9.803***	11.205***
	(3.790)	(3.727)	(3.776)	(3.703)
	[0.005]	[<0.01]	[0.007]	[<0.01]
T ∈[5,10)	6.291***	5.344**	6.305***	5.191**
	(2.377)	(2.257)	(2.393)	(2.188)
	[0.011]	[0.026]	[0.007]	[0.029]
T ∈[15,20)	7.214***	6.111***	7.227***	6.437***
	(2.148)	(1.817)	(2.131)	(1.888)
	[0.001]	[0.002]	[0.001]	[0.002]
T ∈[20,25)	13.915***	11.635***	13.808***	12.126***
	(3.681)	(3.156)	(3.678)	(3.166)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
T ∈[25,30)	15.546***	10.916***	15.587***	12.020***
	(4.961)	(4.279)	(4.988)	(4.144)
	[<0.01]	[0.008]	[<0.01]	[0.001]
T ≥ 30	10.943**	5.425	11.311***	8.656*
	(5.872)	(5.613)	(5.940)	(5.327)
	[0.015]	[0.381]	[0.014]	[0.097]
Precipitation	-0.021	0.015	-0.023	0.008
	(0.021)	(0.012)	(0.021)	(0.013)
	[0.316]	[0.128]	[0.277]	[0.472]
Sea-level Pressure	-0.214	-0.047	-0.202	-0.009
	(0.122)	(0.177)	(0.127)	(0.185)
	[0.112]	[0.801]	[0.128]	[0.959]
Wind Speed	0.124	0.503**	0.128	0.546**
	(0.094)	(0.204)	(0.092)	(0.214)
	[0.197]	[0.022]	[0.164]	[0.016]
				Continued

	AQI		PM2.5	
	(1)	(2)	(3)	(4)
Average Humidity (H∈[40,60) Omitted)				
H <20	13.753	22.587	14.134	26.456
	(4.889)	(7.612)	(5.090)	(8.998)
	[0.264]	[0.125]	[0.277]	[0.108]
H ∈[20,40)	0.462	5.480	0.654	7.181
	(2.608)	(3.953)	(2.696)	(4.475)
	[0.849]	[0.312]	[0.808]	[0.229]
H ∈[60,80)	3.285***	2.459**	3.223***	1.311
	(1.128)	(1.078)	(1.186)	(1.226)
	[<0.01]	[0.027]	[0.002]	[0.309]
H $\geq$ 80	5.613***	7.091***	5.329***	4.802***
	(1.060)	(0.914)	(1.153)	(1.228)
	[<0.01]	[<0.01]	[<0.01]	[0.001]
Day of Week (Monday Omitted)				
Tuesday	-13.195***	-13.065***	-13.186***	-12.968***
	(0.991)	(1.024)	(0.987)	(1.044)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Wednesday	-12.798***	-12.507***	-12.782***	-12.333***
	(1.104)	(1.122)	(1.103)	(1.114)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Thursday	-13.023***	-12.310***	-13.000***	-12.223***
	(1.128)	(1.209)	(1.131)	(1.198)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Friday	-13.465***	-13.165***	-13.445***	-12.984***
	(1.075)	(1.139)	(1.076)	(1.156)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Saturday	-14.542***	-14.235***	-14.501***	-14.042***
	(1.060)	(1.125)	(1.067)	(1.138)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Sunday	-9.022***	-8.666***	-9.019***	-8.639***
	(0.750)	(0.794)	(0.750)	(0.789)
	[<0.01]	[<0.01]	[<0.01]	[<0.01]
Holiday	-9.664***	-8.576***	-9.521***	-7.924***
	(1.150)	(1.333)	(1.162)	(1.451)
	[<0.01]	[<0.01]	[0.012]	[<0.01]

Notes: The table reports detailed OLS and IV results. Each column represents a separate regression. Dependent variable is log form of Sleeplessness Index. Independent variables include daily mean level of specific pollutant, weather controls (average temperature bins, average humidity bins, precipitation, sea-level pressure, and wind speed), temporal controls (year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects) and city fixed effects. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.21: Pollution Regressed on Imported Source Pollutants All Coefficients (First Stage)

First Stage	**AQI**	**PM2.5**
	(1)	(2)
Instrumental	0.455***	0.463***
Pollutant t	(0.062)	(0.061)
	[<0.01]	[<0.01]
Average Temperature (T∈ [10,15) Omitted)		
T <0	-26.803**	-18.887*
	(7.187)	(7.094)
	[0.024]	[0.066]
T ∈[0,5)	-15.260***	-10.746***
	(3.759)	(3.196)
	[<0.01]	[0.002]
T∈ [5,10)	1.386	1.702
	(2.145)	(1.782)
	[0.500]	[0.345]
T ∈[15,20)	6.948***	4.064**
	(2.184)	(1.674)
	[0.002]	[0.017]
T∈ [20,25)	12.147***	6.735**
	(3.540)	(2.756)
	[0.002]	[0.023]
T [25,30)	24.962***	14.608***
	(5.463)	(4.169)
	[0.001]	[0.004]
T ≥ 30	30.398***	11.283***
	(5.161)	(3.957)
	[<0.01]	[0.005]
Precipitation	-0.178***	-0.114***
	(0.036)	(0.027)
	[<0.01]	[<0.01]
Sea-level Pressure	-0.937***	-0.835***
	(0.254)	(0.213)
	[0.006]	[0.002]
Wind Speed	-1.914***	-1.602***
	(0.271)	(0.202)
	[<0.01]	[<0.01]
		continued

First Stage	AQI	PM2.5
	(1)	(2)
Average Humidity (H ∈[40,60) Omitted)		
H <20	-42.911	-45.284*
	(14.122)	(11.782)
	[0.108]	[0.086]
H∈ [20,40)	-25.111***	-24.904***
	(9.139)	(7.447)
	[0.014]	[0.001]
H∈ [60,80)	5.814	8.884**
	(4.152)	(3.974)
	[0.197]	[0.026]
H $\geq$ 80	-6.629	2.927
	(4.927)	(4.473)
	[0.199]	[0.578]
Day of Week (Monday Omitted)		
Tuesday	-0.368	-0.617
	(0.802)	(0.777)
	[0.631]	[0.426]
Wednesday	-2.114*	-2.215**
	(1.126)	(0.967)
	[0.092]	[0.049]
Thursday	-3.549**	-2.863**
	(1.206)	(1.138)
	[0.019]	[0.048]
Friday	-2.103	-2.228
	(1.259)	(1.274)
	[0.129]	[0.110]
Saturday	-1.631	-1.830*
	(1.159)	(0.945)
	[0.170]	[0.078]
Sunday	-0.834	-0.679
	(0.865)	(0.727)
	[0.383]	[0.384]
Holiday	-7.642***	-7.935***
	(1.346)	(1.029)
	[<0.01]	[<0.01]

Notes: The table reports detailed results of the first stage under IV estimations. Each column represents a separate regression. Dependent variable is daily mean level of specific pollutant for each city. Independent variables include instrumented pollutant, weather controls (average temperature bins, average humidity bins, precipitation, sea-level pressure, and wind speed), temporal controls (year by month fixed effects, city by year fixed effects, city by quarter fixed effects, as well as day of week and holiday fixed effects) and city fixed effects. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.22: Joint Estimation (First Stage and Second Stage)

Individual Pollutant	Whether to Control Co-emissions in the first stage: NO				Yes			
	PM2.5	CO	NO2	O3	PM2.5	CO	NO2	O3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First Stage								
Instrumental Pollutant t	0.463***	0.184***	0.326***	0.442***	0.236***	0.105***	0.168**	0.433***
	(0.061)	(0.077)	(0.071)	(0.044)	(0.054)	(0.050)	(0.066)	(0.045)
	[<0.01]	[0.011]	[0.001]	[<0.01]	[<0.01]	[0.006]	[0.023]	[<0.01]
Kleibergen-Paap rk Wald F statistic)	56.846	5.742	20.621	101.465	18.987	4.451	6.508	93.942
Stock-Yogo weak ID test critical values: 10% maximal IV size	16.38	16.38	16.38	16.38	16.38	16.38	16.38	16.38
Second Stage								
Instrumented Pollutant	0.285**	-	0.018	0.072	0.519**	-	-	0.071
	(0.118)	-	(0.220)	(0.172)	(0.229)	-	-	(0.179)
	[0.033]	-	[0.944]	[0.835]	[0.049]	-	-	[0.834]
Observations	12989	-	12989	12989	12989	-	-	12989
Additional Controls								
City FEs	Y	Y	Y	Y	Y	Y	Y	Y
Temporal Controls	Y	Y	Y	Y	Y	Y	Y	Y
Weather Covariates	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table reports the results from both the first stage and second stage of joint estimation in Table 1.14. Column (1) through Column (4) does not include co-emissions at the first stage. The second stage regresses Sleeplessness Index on different instrumented pollutants together, which corresponds to Column (4) of Table 1.14. Column (5) through Column (8) controls co-pollution when making instrument and reports each second stage estimate one by one, which corresponds to Column (5) of Table 1.14. All the regressions include city fixed effects, temporal controls and weather covariates. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A1.23: Air Pollution and Neutral Keywords

	Sleeplessness (1)	Cat (2)	Table (3)
Panel A: AQI	0.037***	-0.004	-0.017
	(0.012)	(0.015)	(0.027)
	[0.001]	[0.798]	[0.499]
Panel B: PM2.5	0.043***	-0.009	-0.022
	(0.017)	(0.019)	(0.029)
	[0.012]	[0.600]	[0.395]
Observations	13617	13617	13617
Additional Controls			
City FEs	Y	Y	Y
Year by month FEs	Y	Y	Y
City by year FEs	Y	Y	Y
City by quarter FEs	Y	Y	Y
Day of week FEs	Y	Y	Y
Holiday FEs	Y	Y	Y
Weather Covariates	Y	Y	Y

Notes: This table compares the effects of air pollution on the neutral keywords "cat" and "table", by regressing air pollution on the log form of the keywords. Entries have been adjusted to percentage form. All the regressions include city fixed effects, temporal controls and weather covariates. Robust standard errors clustered at the city level are reported in parentheses. P-values based on wild cluster-bootstrap (1000 replications) are reported in brackets. Asterisk indicates the statistical significance according to the wild bootstrap p-values (* significant at 10%, ** significant at 5%, *** significant at 1%).

Chapter 2

Air Pollution and Morbidity: Evidence from Internet Search Behavior in a Panel of 100 Chinese Cities

2.1 Introduction

Air pollution is a pervasive feature of modern life, particularly for inhabitants of cities. The problems are particularly marked in China, which is the setting for our study. The short- and long-term impacts of exposure on human health have been extensively investigated.[1] Thousands of studies investigate the correlation or causal link from exposure to air pollution over various time frames to indicate for mortality, hospitalization and health conditions that precipitate an interaction with the health sector.

[1]The elevated air pollution has been linked to the increased mortality due to respiratory and cardiovascular diseases over both short term and long term (see summary in Brunekreef and Holgate (2009)). For example, in recent work, Ebenstein et al. (2017) identify a substantial decrease in life expectancy in Northern China as a result of its winter heating under China's Huai River Policy, which generates more coal-fired pollution (PM_{10}) compared to Southern China. Currie et al. (2009) show that exposure to air pollution, especially carbon monoxide, has a consistently negative effect on infant health. Lavaine and Neidell (2017) link a significant decrease in sulfur dioxide to increases in birth weight and gestation. In the latest study, Burnett et al. (2018) find a much larger detrimental effect of particular matter based on their risk model, indicating 8.9 million cases of global mortality attributable in 2015 to air pollution that is more than twice as much as previous outcomes.

Our focus is different. There are a number of less-serious health and body conditions that do not typically manifest in death, hospitalization or interaction with a healthcare professional, but that can nonetheless harm human well-being substantially, and that might plausibly be exacerbated by exposure to polluted air. We study the phenomenon of the cough. Persistent or episodic coughing has been shown to have important negative consequences for workplace productivity (Keech et al., 1998; Dicpinigaitis et al., 2015), mood (Chamberlain et al., 2013), ability to concentrate (Bucks et al., 2008), interpersonal relationships (Baiardini et al., 2005), and other outcomes important for human well-being. A survey-based literature has used contingent valuation and related methods to uncover individual willingness to pay (WTP) for avoidance of 1 day of cough, which are substantial, often in the range of 10 - 20 USD per cough-day avoided.[2]

A number of epidemiological papers estimate the *association* between long-term exposure to air pollution and the prevalence of cough. That people living in more polluted locales are more prone to episodes is not surprising. However such an association may be explained, in part or whole, by locational sorting and other selection effects, or by unobservable contributors to those outcomes that are correlated with air quality.[3] Omitted explanatory variables and the (inevitable) measurement of pollution exposure with error will cause standard correlational and OLS-based studies to underestimate the true sensitivity of cough to pollution levels.

While for completeness we report significant correlations estimated by ordinary-least squares (OLS), our central results use instrumental variables (IV) methods and have *causal*

[2]A good survey of some of the early studies is provided by Vassanadumrongdee et al. (2004). In the recent work, Deryugina et al. (2016) study the medical cost among the US elderly and find that each unit increase in $PM_{2.5}$ $\mu g/m^3$ leads to \$299,000 per million beneficiaries. Barwick et al. (2018) examine the morbidity cost of air pollution via 'bank card' transactions on health care spending. That analysis suggests that \$42 billion per annum would be saved on bank-card purchased medicines alone if China's $PM_{2.5}$ were to be reduced to 10 $\mu g/m^3$.

[3]For example stress from over-work might cause an individual to cough, which might correlate with pollution from businesses working at high-capacity. A notable exception to the correlational focus of most existing research is Gupta and Spears (2017). They use a differences-in-differences design to obtain evidence of a causal link from the opening of coal-fired power plants in some Indian districts but not others to the increase between 2005 and 2012 in the number of respondents to the large-scale Indian Human Development Survey reporting having experienced an episode of coughing in the 30 days before interview.

interpretation. With particular focus on fine particulate matter ($PM_{2.5}$), we uncover highly significant causal effects of daily variations in air quality on the incidence of cough. As already noted, we expect OLS to deliver attenuated results in this setting and, consistent with that, the effect sizes implied by IV estimation are much larger (typically around five times larger) than those from the former. The results prove to be robust to a suite of robustness and falsification exercises.

Back-of-the-envelope calculations suggest that there exists a substantial monetary impact. We will be careful to emphasize the shortcomings of such calculations. However, the implied dollarized values are of the same order of magnitude as those ascribed to mortality effects in papers such as Anderson (2015) and Barreca et al. (2017). Failure to recognize them means that we substantially under-state the true social benefits of policies that deliver cleaner air.

Our setting is a panel of 100 large Chinese cities, and we conduct an analysis at daily frequency from 1 January 2014 to 31 December 2017. There are two central challenges that we face in the conduct of our study. First, we need to derive a credible, high-frequency (which here means daily) measure for the number of people in a set of locations that are suffering from the sub-clinical health condition of interest to us, namely coughing, at a particular time. Second, we should also connect those locations to an instrument which generates plausibly exogenous day-to-day variations in air quality conditions in those locations.

To address the first concern, we use an index of the intensity with which the keyword 'cough' is searched for on Baidu from IP addresses in each city in our panel on each day in our period of interest. Baidu is the universally-used search engine in China (where Google is blocked by the government). In China in 2018 PC web-search users reach 688 m people, of which the Baidu Company accounts for 60% of the market share dominating the Chinese PC search market.[4] It is increasingly recognized that internet behavior - what people are posting on social media sites, their search patterns, etc. - can provide a powerful source of

[4]The statistic is reported from "2018 China PC Search Market Special Report" released by iiMedia Research Group which is a world-renowned third-party for big data analysis.

data, offering a “window into the thoughts” of users (Choi and Varian, 2012; Qin and Zhu, 2018). Metrics derived from such sources have been used as both dependent and independent variables in a number of recent studies. We will sketch some of these in Section 2.2 and in Section 2.3 to explain in detail why this approach is well-suited for use in our setting.

With regards to the second challenge, we use agricultural fires in the area surrounding each of our cities on a given time frame (three-day period) as an instrument for air quality in that city. Many small farmers in China set fire to the stubble in their fields between harvesting one crop and planting the next. On a given time in a circle of 150 km surrounding a city, there can be dozens of active fires, and we use this count as our preferred instrument. Depending on the direction and speed of the wind, particulate pollution from these fires can spread across the city. Agricultural fires have been identified as a major contributor to air pollution in many Chinese cities (Zhang et al., 2016; Chen et al., 2017), particularly at times of year when burning is at its peak, which varies across regions. Counts of fires are derived from high-frequency NASA satellite imagery parsed using specialist software developed with NASA sponsorship at University of Maryland (Giglio, 2015). Consistent with existing research, the fire counts, constructed in various ways, prove to be a strong instrument for city-level $PM_{2.5}$. Helpfully for us the fire count does not prove to be a strong instrument for the other common pollutants that we study, allowing our instrumental variables analysis to disentangle the particular effect of particulate matter from the broader effect of ‘dirty air’. We present arguments and the results of empirical exercises in attempts to convince the reader of the validity of the instrument.[5]

The rest of the paper is organized as follow. In Section 2.2 we identify some pertinent existing research. In Section 2.3 we describe our data. In Section 2.4 we outline our empirical strategy, and Section 2.5 shows the results from both the OLS and IV estimations. Section

[5]The fire count data has recently been used in a reduced form setting to estimate the effect of agricultural fires on infant health in Brazil (Rangel and Vogl, 2016). Unlike that study, our interest is not in the health impact of fires *per se*, though that is undoubtedly an important public health question in those countries where fires are common. Rather we use the fires to instrument for the air pollutant $PM_{2.5}$, which is our variable of interest.

2.6 presents the results of a battery of robustness checks and falsification exercises. Section 2.7 concludes.

2.2 Some existing research

There has been extensive literature exploring the health effects of air pollution from both short-term and long-term, especially for infant health, adult mortality, respiratory and cardiovascular disease. We do not attempt to survey those topics here. For a couple of examples using natural experimental methods, Currie and Walker (2011) find evidence that after the introduction of electronic toll collection, the reduced traffic congestion along with less emission accumulation significantly increases the gestation and birth weight in the nearby area. The increased infant health outcomes are also found from superfund cleanups in Currie et al. (2011). Anderson (2015) explores the causal link from air pollution to increased adult mortality with the instrument of the exogenous wind patterns near major highways. Schlenker and Walker (2016) show that daily air pollution increases the hospitalization rate for respiratory and heart diseases, exploiting plausibly-exogenous variations in the arrival, taxi and departure of flights in and out of a panel of Californian airports.

Population-level behavior on various internet platforms is increasingly being exploited by social scientists. While here we will mention a small number of relevant studies, more comprehensive overviews of the use of search engine and social media data in the social sciences are provided by Blazquez and Domenech (2018).

Internet metrics are increasingly used by social scientists as a source of data. To the best of our knowledge, Ettredge et al. (2005) were the first to promote the utility of using Web search data in empirical research, in particular using search behavior for the purpose of forecasting unemployment rates. Choi and Varian (2012) show that Google search data can be used to predict demand for auto-mobiles, home sales, and travel behavior.[6] Along similar lines, several papers demonstrate the efficacy of using internet search metrics to predict

[6]Hal Varian has served as Chief Economist at Google Inc. since since 2002.

health outcomes - especially flu - and Google itself established the 'Google Flu Trends' tool in 2008 (Doornik, 2009). Goel et al. (2010) show that search patterns can predict the success of movies, songs and video games. In an environmental application, Herrnstadt and Muehlegger (2014) show that searches for "climate change" and "global warming" in a particular US city are sensitive to short-term deviations of weather from normal.

Much recent work has been devoted to Twitter-driven predictive analytics. For three examples among many: Bollen et al. (2011) show that Twitter mood can be used to add explanatory power to stock market forecasts, Gerber (2014) uses Twitter key words to predict crime patterns, and Gayo-Avello (2012) is among several papers using Twitter to predict election results.

A central way in which our methods depart from these papers is that we will use measures from web-search as dependent variables. In other words variations in it are the outcome to be explained by variations in the world (here, pollution levels). In that regard the paper relates to Baylis (2015), who shows the effect of unusual temperatures on a Twitter-derived metric for sentiment. Zheng et al. (2019), who examine daily pollution and sentiment, and Heyes and Zhu (2018), who establish a negative and causal effect of air pollution on sleep quality, both through analysis of Weibo posts in panels of Chinese cities, are also relevant studies.

2.3 Data

Our study requires a population metric for the population level proxy for prevalence of cough symptoms in each city, on each day (we will be careful to qualify our use of the term "population level" in the data section). For this we use the intensity of searches originating from within each city on the Baidu. We interact this with city-level measures of air pollution and weather. We collect data on agricultural fires to underpin an instrumental variable analysis. Summary statistics are listed in Table 2.1.[7]

[7]The variables that we work with in our central specifications are daily averages aggregated over a three-day period.

2.3.1 Searches

Baidu provides a daily 'Baidu Index' which can be filtered by search time, time period and city of search based on IP address of user.[8] For each keyword the Index "... is calculated by the weighted sum of search frequency from users' query." Our focus is on the keyword "cough". For the purposes of falsification exercises presented later in the paper, we also collect data on keywords that we do not expect to be sensitive to short-term pollution (in particular; diarrhea, dislocation, food poisoning, stomach ache, lamp and desk). Although Baidu Index is not equal to the real search volume on Baidu engine, it has been verified as being closely linearly correlated with the search frequency (Qin and Zhu, 2018). Yuan et al. (2013) and Vaughan and Chen (2015) provide a detailed description and analysis of the Baidu Index and advice on how it should be used in social science research. Since it is the scaled-down measure of search volume, we directly generate our variable of interest based on Baidu Index.

In our main specifications, we work with the three-day average of the Index as our preferred dependent variable. Our study window runs from 1 January 2014 to 31 December 2017 and relates to the 100 largest cities in mainland China.[9] The time series for the Index for Beijing is plotted from Figure 2.1, a screenshot from the Baidu platform. Across the whole sample, the Index ranges from 0 to 997, with a mean value of 152.5 and a standard deviation of 84.9. In summary there is plenty of within- and between-city variation. Our city-fixed-effect specification will focus on within-city changes and how these relate to pollution patterns across our panel.

Besides the Index data, Baidu also releases information on the frequency with which other search terms that accompanied the keyword (for example, "cough"), either in the same search or in the search executed immediately before or immediately after. The most

[8] More refined geographical data on user (for example zipcode) is not disclosed. We ignore any possible complications due to use of VPNs.

[9] The top 100 cities with mature Internet development are selected based on the report "China Internet Plus and Digital Economy Index 2017" released by Tencent Research Institute (http://www.cbdio.com/image/site2/20170420/3417eb9bbd591a62741647.pdf).

popular related search were “treating a cough”, “cough syrup”, “ dry cough” and “throat”.

One concern regarding the use of web search as a city population-level proxy for cough propensity is that not all people have internet access or are regular internet users. Internet users in China are on average younger, more educated, and earn higher incomes than the broader population (Chan al., 2012). The results should be understood in that context.

2.3.2 Air pollution

In the years prior to 2014, China significantly improved air quality monitoring, with 1586 air monitors installed at the start of our study period. They report real-time hourly data for $PM_{2.5}$ $(\mu g/m^3)$, $CO(mg/m^3)$, $NO_2(\mu g/m^3)$, and $O_3(\mu g/m^3)$, as well as the Air Quality Index (AQI), to the Chinese Ministry of Environmental Protection (MEP).[10] We obtained city-level pollution measures from the website www.aqistudy.cn, which mirrors MEP data, combining hourly data from all monitors within each city into city-level output. Consistent data are available from 1 January 2014, which is informed choice of starting date for our study period.

2.3.3 Weather

The weather data are from the National Oceanic and Atmospheric Administration (NOAA) and collected from all stations within China registered by the World Meteorological Organization (WMO). All variables are converted into daily average measures. These include average temperature ($^\circ C$), maximum temperature ($^\circ C$), minimum temperature ($^\circ C$), average humidity (%), precipitation (mm), sea-level pressure (hPa), wind speed (Km/h), wind direction (°) and cloud coverage (on a 0 to 8 scale). We adopt the weather information from the nearest station to the city centre as the daily measure for each local city.

[10]The veracity of official Chinese air pollution data historically has been questioned. Stoerk (2016) tested Chinese data against US Embassy data and for consistency with Benford’s Law, and concluded that misreporting was not a substantial problem after 2013.

2.3.4 Fires

The instrument we use on our estimation is the number of agricultural fires burning in the vicinity of each city at different times. The upper panel in Figure 2.2 shows what a typical fire might look like from ground-level, with the farmer controlled-burning the stubble in his fields in preparation for the next round of cropping. What this looks like from space is shown by the satellite image of India and China in the lower panel, with each orange dot representing a fire point.

The daily number of count of fire points in the vicinity of each city in our sample is obtained from the Fire Information for Resource Management System (FIRMS). The system was developed and is managed by the University of Maryland in collaboration with the National Aeronautics and Space Administration (NASA) and serves to provide Near Real-Time (NRT) active fire data through the instrument Moderate Resolution Imaging Spectroradiometer (MODIS) that is operating on both the Terra and Aqua satellites.[11] The satellites between them provide observations up to 4 times daily (He et al., 2007). In typical conditions MODIS is able to observe the flaming and smoldering fires with size of 1000 m^2 and the smaller size of 100 m^2 in clean air (Giglio, 2015). They also provide the detection of confidence ranging from 0% to 100% indicating the quality of fire pixels.[12] In our preferred specifications we use the confidence-weighted series within 150 km to city centre as our instrument.

2.4 Methods

2.4.1 Linear OLS

The purpose of our study is to explore the causal impact of air pollution on the incidence of cough. Considering plausible lags between pollution exposure, cough symptoms, physical

[11] About FIRMS: https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms/about-firms Data source: https://firms.modaps.eosdis.nasa.gov/download/request.php

[12] Refer to MODIS Collection 6 Active Fire Product User's Guide Revision B (https://cdn.earthdata.nasa.gov/conduit/upload/10575/MODIS_C6_Fire_User_Guide_B.pdf)

discomfort, and search behavior, we adopt the three-day average as our preferred unit of analysis.[13] This plausibly captures lags overnight, or over a couple of nights, from pollution exposure to search for solutions or information. This follows the approach of Deryugina et al. (2016), who estimate the causal link from air pollution to medical expenses, also using three-day time windows as the preferred temporal unit of analysis. The three-day unit is a modelling choice, and later in the paper, we will show results which are consistent when estimated over alternative temporal resolutions.

To examine the association between pollution and cough, we first estimate the following equation using OLS.

$$\log H_{ct} = \alpha_0 + P_{ct}\beta + Co_{ct}\beta' + W_{ct}\gamma + \theta_c + \lambda_t + \sigma_{ct}. \tag{1}$$

H_{ct} is the Baidu Index for our keyword of interest (cough) in city c in time period t. P_{ct} is the level of $PM_{2.5}$ in city c in period t. (In complementary analysis, we replace $PM_{2.5}$ with the multi-pollutant AQI measure). W_{ct} is the vector of weather covariates in city c including all the factors that could affect health, in particular temperature, humidity, precipitation, air pressure, wind speed, and cloud coverage. For the purpose of consistency, each is measured as a three-day average.[14] All the weather covariates are included in squared term. Co_{ct} is a vector of co-pollutants, comprising three-day average concentration levels of CO, NO_2, SO_2, and O_3. θ_c denotes city fixed effects. λ_t represents time fixed effects including city by year, city by month, year by month, holiday, and day of week fixed effects. σ_{ct} is an error term. In our central case, standard errors are robust and clustered at city level.

The coefficient β represents the association between $PM_{2.5}$ (AQI) levels and the prevalence of cough, as measured using our proxy.

[13] In Appendix A5, the instrument at a daily frequency does not provide adequate variation. For the two day period, the F-statistics are not large enough compared to the critical value 16.38. Therefore, we choose three-day period as our preferred setting.

[14] Weather covariates in the regression incorporate squared term in maximum, minimum temperature, relative humidity, wind speed, cloud coverage and sea-level pressure.

2.4.2 OLS estimates derived from binned data

To explore the possibility of a non-linear association between $PM_{2.5}$ (AQI) levels and the incidence of cough, we re-estimate Equation (1) but replace the continuous measure of pollution with a series of indicator variables. Each indicator variable takes the value 1 if the pollution measure in a particular three-day time unit fell into the range defined by a particular 'bin'.

We estimate this specification using two alternative bin widths. First, we include bins of width 50 $\mu g/m^3$, starting from zero, with the 0 - 50 $\mu g/m^3$ as the omitted or reference category. Second, we further consider bins of width 25 $\mu g/m^3$, starting from zero, with the 0 - 25 $\mu g/m^3$ as the omitted or reference category.

2.4.3 IV estimates

Several challenges exist for inference from OLS estimation in this setting. First, air pollution is measured as daily average by integrating hourly pollution data recorded from all the monitors in a particular city. Given that we do not have precise information about either the location of searchers at time of search, or their behaviors and movements at other times during the day, exposure is measured with error. Such measurement error on the regressor of interest would imply that the coefficient β estimated by OLS would be attenuated. Second, there exists omitted variables. While we have controlled for a set of potential confounders, and the inclusion of diverse fixed effects can be expected to negate the role of others, we cannot rule out other omissions. Insofar as these are correlated with pollution levels, they would simply reduce precision of our estimates making inference from OLS conservative. However there may be other unobserved factors, like some other uncontrolled weather covariates, correlated with *both* short-term pollution and the propensity for inhabitants in a city to engage in search for the term cough, that could bias the coefficient estimate upwards or downwards. For these reasons we supplement the OLS results using IV methods, instrumenting for air pollution using the number of fire counts around.

The first stage is written as follows,

$$P_{ct} = \eta_0 + Fire_{ct}\delta + Co_{ct}\eta + W_{ct}\tau + \theta_c + \lambda_t + \epsilon_{ct} \tag{2}$$

Here $Fire_{ct}$ is daily average confidence-weighted fire counts in the three-day period within 150 km of the city centre. Again, P_{ct} is the level of $PM_{2.5}$ in city c in the 3-day period t (replaced by AQI in a secondary exercise). The other controls are as above.

The second stage analyses the health effect of $PM_{2.5}$ by regressing $logH_{ct}$ on the predicted air pollution $\hat{P_{ct}}$ obtained from first stage and controls. In other words,

$$\log H_{ct} = \tau_0 + \hat{P_{ct}}\psi + Co_{ct}\psi' + W_{ct}\xi + \theta_i + \lambda_t + \upsilon_{ct}. \tag{3}$$

The coefficient of interest is ψ which, if the instrument is valid, will provide an estimate of the causal impact of short-term $PM_{2.5}$ exposure on the outcome variable cough.

Further discussion of the instrument

The instrument is central to many of the results in the paper, and it merits more detailed discussion. The fire information provided by the FIRMS dataset includes the latitude and longitude of each fire. In addition, a detection confidence' number for each fire point on a scale from 0% to 100% is available, which is a measure of the degree of certainty that the bright spot detected from satellite imagery is indeed a fire. From this we compute, for each city and for each day the confidence-weighted number of fires detected within a circle of 150 km circumference from the city centre. The instrument $Fire_{ct}$ is then this number for city c averaged over the 3-day window at period t.

The instrument constructed in this way embeds a number of modelling choices, and we will test the robustness of results to alternative choices in Section 2.6. In particular we will consider: (1) Varying the unit of time for analysis from three days to seven days. (2) Varying the circumference of the circle within which fires are counted. (3) Counting only fires upwind of the city, determined by daily average wind direction. (4) Using raw fire count data rather than the confidence-weighted measure. (5) Estimating only on those months of the year when fires are particularly frequent. (6) Varying the combinations of time fixed effects.

Instrument relevance is established statistically. The instrument constructed as described will turn out to be a strong instrument for $PM_{2.5}$ (and AQI). Instrument validity requires that the number of fires within the 150 km neighborhood of a particular city impacts the propensity to cough by inhabitants of the city only through its contribution to increased air pollution in the city. This is our primary identifying assumption. There could be a confounder like that: some places in the city with more agricultural fires are poorer areas that have worse hospital care, where residents are more prone to cough. However, the hidden correlation will be diluted when we base the estimates to the whole range of the city, which is the setting of our study.[15] Moreover, the regression estimates are clustered at the city level, which helps to control the time-invariant variables, like internal economic factors.

2.5 Regression results

2.5.1 Linear OLS

Table 2.2 summarizes the results from linear OLS estimation, in particular estimates of the coefficient β in Equation (1). The first four columns relate to $PM_{2.5}$. The coefficient in each of columns (1) through (3) is positive and significant at higher than 1%. Column (4) includes as controls ambient levels of the other four pollutants for which we have data. The coefficient is rendered substantially smaller with inverse sign and loses statistical significance. This is a problem that recurs in research investigating the impact of air pollution on various outcomes, and arises from the positive and often strong positive correlation observed between ambient levels of the various pollutants in most settings (Moretti and Neidell, 2011; Schlenker and Walker, 2016; Knittel al., 2016; Arceo al., 2016). Some authors circumvent this problem

[15]The dependent variable is the total search for cough within a city. Moreover, The instrumental variable is the entire number of fires within the city scope.

by concentrating on only one or a restricted set of pollutants (Zivin and Neidell, 2012; Chang al., 2016), while others treat pollutants one by one. Such an approach does not satisfactorily disentangle the role of separate pollutants etc.. If pollutants are positively correlated, then inclusion of any as sole regressor could deliver a significant coefficient, even if only one of the pollutants was the true cause of the effect. Our instrument will allow for cleaner isolation of the role of $PM_{2.5}$, since the agricultural fires that we will exploit as instrument will relate only to $PM_{2.5}$.

The implied effect size in each case is reported below the coefficient estimate. It is the percentage change relative to the mean associated with a one-standard-deviation increase in $PM_{2.5}$. In the preferred specification (Column (3)), the effect is 0.75%. The analogous effect size with regards to AQI is 0.92%.[16] For reasons already outlined, we expect these coefficient estimates to be attenuated and implied effect sizes to be biassed downwards.

2.5.2 Binned OLS estimation

We next probe for possible non-linearity in the relationship between $PM_{2.5}$ pollution and cough using the binned approach described in Section 2.4.2. The results are summarized in Table 2.3 and Figure 2.3. Column (1) in Table 2.3 and the top panel in Figure 2.3 report the specification with bins of width 50 $\mu g/m^3$, with the 0 - 50 $\mu g/m^3$ as the reference category. Column (2) in table 3 and the lower panel in Figure 2.3 include $PM_{2.5}$ with bins of width 25 $\mu g/m^3$, with the 0 - 25 $\mu g/m^3$ as the reference category.

The results are best understood from the Figure. The central estimates in each bin are consistent with an upward-sloping and approximately linear relation. The highest bin in the top panel contains comparatively few data points (only 1% in the group of 200-250, 0.44% larger than 250).[17] As such the associated confidence interval is very wide and it would be inappropriate to deduce anything concrete from the apparent downward step at that bin

[16] The AQI is a composite or multi-pollutant measure of air pollution such that inclusion of controls for co-pollutants become redundant.

[17] See Appendix Table A2.4 and Appendix Figure A2.2, the distribution of $PM_{2.5}\_3$.

value.

2.5.3 IV estimation

The results of estimating the first stage of $PM_{2.5}\_3$ (Equation (2)) are presented in columns (1) and (2) of Table 2.4. The difference between the two columns relates to the inclusion or non-inclusion of the other pollutants as controls. In both columns, the coefficient is statistically significant at a level well above 1%. In other words $PM_{2.5}$ levels in city c are significantly influenced by the number of fires burning in the surrounding agricultural areas. The estimate under Column (2) with co-emissions suggests that each additional fire point is associated with a 0.208 unit increase in $PM_{2.5}$ levels. After comparing with the critical value of the Stock-Yogo weak ID test (10% maximal IV size) which is 16.38, we reject the hypothesis that instrument weak at 1% significance level.

For completeness we re-estimate the column (2) specification but replace $PM_{2.5}$ with the other five major pollutants (Column (3) through Column (6)). The corresponding F-statistics are much smaller in each case compared to the critical value of Stock-Yogo test (16.38), indicating that they are weak instruments. In other words the fire count measure that we adopt is a strong instrument for $PM_{2.5}$ with a large effect size and level of significance much higher than 1%. However it is not a suitable instrument for the other potentially-confounding pollutants. It is a substantial advantage in our setting because it implies that the effects that we uncover in our IV estimates can be tied specifically to $PM_{2.5}$ rather than to the correlated pollutants.

We report our main IV results in columns (1) and (2) of Table 2.5 with and without inclusion of co-pollutant controls. The first stage coefficients (0.310 and 0.208) replicate those from the first two columns of Table 2.4.[18] The second stage coefficients are our main interest. Coefficients in each of these columns are positive and statistically significant. The

[18] The associated reduced form exercise is summarised in Appendix Table A2.1. This is the most comparable table to the results of Rangel and Vogl (2016) who conduct a reduced form analysis of the impact of fires to birth outcomes in Sao Paulo, Brazil, rather than using fire counts to instrument for air quality.

coefficients are again interpretable in percentage terms. In each column we also express effect size in terms of the impact of a one-standard-deviation increase in $PM_{2.5}$. Such an increase causes a 4.19% (or 5.20%) increase in the cough index. The analogous effect size for AQI is 3.59%.

Note that the effect sizes in this table are around three to six times larger than those from the OLS estimation reported in Table 2.2. This is consistent with our prior expectation that measurement error relating to pollution exposure would attenuate OLS estimates. It also suggests that studies deriving inference from associations rather than causal methods can be expected to drastically understate the true impact of air pollution on cough.

2.5.4 Effect size and back of the envelope calculations of burden

Introspection suggest that the effect size identified - in the central case that a one-standard-deviation increase in particulate matter causes a 5.20% increase in likelihood to experience cough - is a substantial one.

Converting it to case numbers, or dollarized value of burden, requires a strong set of assumptions. However, for the purposes of developing at least a ballpark understanding of the implications, we develop here some back-of-the-envelope calculations. We do not want to over-interpret these numbers, and advise that the reader exercise due caution in drawing lessons from them.

Chen et al. (2006) found that in a large cluster-randomized sample of young adults in Guangzhou, one of the three largest cities in our sample, at any particular time the prevalence of cough was 10.9%. (They also found no difference between males and females in the sample.) According to Lai al. (2013) the prevalence of acute cough in China lies between 9 and 64%, for chronic cough between 7 and 33%, with numbers highest in urban areas. Combining the value of 5.20% derived from our regression with 20% typical prevalence would imply that a one-standard-deviation in pollution would cause about an extra 1% of the population to be subject to cough at a particular time.[19]

Various researchers have attempted to attach a dollar value to the monetary value of a cough-day. These typically use stated preference methods, in particular contingent valuation, to estimate the willingness-to-pay (WTP) of individuals to avoid a day with cough symptoms. Rozan (2004) provides a survey of methods, in addition to presenting estimate for France and Germany. Unfortunately the estimates from such studies in less-developed countries vary widely. WTP for avoidance of 1 day of symptoms of 'mild respiratory illness' were estimated to be 21 1995US$ in Taiwan (Alberini al., 1997) 17 1995US$ in Bangkok (Chestnut et al., 1998) and 11 1995US$ in Columbia (Ibáñez and McConnell, 2001).[20] Specific to WTP to pay to avoid a single day of 'mild cough' estimates include 31 1995US$ for Costa Rica (Barton, 1999), 18 1995US$ for Malaysia (Dubourg, 1998) but only 3 1995US$ for Iran (Meegan, 1998). In the context of a meta-analysis, Vassanadumrongdee et al. (2004) present a helpful table of estimates that include coded methods and location of study (Appendix Table A2).

Absent a specific number derived from a Chinese study, and on the basis that China belongs to the group of developing countries, we adopt 14 1995US$ which is the minimum value Vassanadumrongdee et al. (2004) list for developing countries. Adjusting for cumulative inflation 14 1995US$ equates to roughly 25.9 2018US$. The population of Beijing is 21.7 million. Multiplying these numbers together suggests that a one-standard-deviation increase in pollution in Beijing would cause approximately an additional 217,000 cough-days. Valuing each avoided cough-day at 25.9 2018US$ implies a total WTP to avoid those additional cough-days of approximately 5.6 m 2018US$, or about 0.259 2018US$ per capita. These numbers relate to the discomforts associated with sub-clinical cough symptoms and should be accounted for in addition to the various other symptoms and health (often much more serious) of short-term exposure.

[19] 20% is roughly medium in the prevalence of chronic cough.

[20] 1995US$ denotes US dollars at 1995 prices.

2.6 Robustness and falsification

Our central analyses embed a number of modelling assumptions. In this section we report the outcomes of a number robustness checks and falsification exercises. The results prove to be resilient to alternative specifications and approaches.

Alternative time unit: In Table 2.6 we report the effect of using time periods longer than 3 days for the analysis, in particular using 5 day blocks, and 7 day blocks. Columns (1) and (2) represent the preferred results from Tables 2.2 and 2.5. The results across columns (3) through (6) are broadly consistent, including the implied effect sizes. We also report the estimates under daily and two-day periods in Appendix Table A2.5. The first stage regression under the daily frequency is not strong enough compared to the critical value of 16.38. As for the two-day period, although the outcomes remain statistically significant, the F-statistics are very close to the threshold, while the three-day frame not only has better performance for the instrument, but also retain a sufficient sample size.[21]

Alternative fire count radius: With respect to the fire counts, we adopted a cut-off for distance threshold of 150km. However we might have chosen a tighter neighborhood around cities within which to have executed the count as, for example, did Rangel and Vogl (2016). Columns (3) and (4) in Table 2.7 repeat the preferred IV exercises using a fire count based on a smaller circle around the city: that of circumference 100 km. The first-stage coefficients are somewhat larger, and F-statistics are little changed. The effect sizes implied for the sensitivity of the cough outcome to changes in $PM_{2.5}$ are perturbed only slightly. A smaller threshold would not replace our preferred choice. Instrument relevance and strength are already established statistically, and any challenge to the exclusion restriction on the fires in our counts would plausibly apply *a fortiori* to fires closer to the affected city. But as a test of our design, it is reassuring to know that the thrust of results is not sensitive to this modelling choice.

[21] We choose the three-day period as our preferred frame because the instrument shows strong performance, but also retain larger sample size compared to longer period frame.

Accounting for wind direction: Our preferred approach to counting fires was the simplest we could conceive - a simple count of the fires in the whole circle surrounding the city in question. An alternative would have been to account for wind direction. Given that our simple approach delivered a strong instrument and a well-performing first stage regression, we did not pursue this. Any potential challenge to the associated exclusion restriction could be expected to be no weaker under a wind-sensitive instrument. However, a number of papers (Anderson, 2015; Rangel and Vogl, 2016) have sought to develop pollution instruments based on wind direction. For completeness, then, we repeat the exercise but (a) counting only those fires in the circle segment defined by the radii plus and minus 60 degrees from the average wind direction during the three-day time period in question and the centre of the city in question, and (b) plus and minus 90 degrees (so in the semi-circle centred on the average wind direction). The results of these exercises are reported in columns (3) through (6) of Table 2.8. The coefficients at the first stage are larger, and the F-statistics somewhat higher. This is unsurprising; importantly, the implied effect size is little disturbed. In fact it is a little larger in each case, suggesting our whole circle approach represents a conservative design choice.[22]

Confidence-unweighed fire count: As already noted, FIRMS is not always able to say that a particular bright area identified from satellite image is definitely a fire. In addition to the location of a fire point, therefore, it reports a confidence-measure on a scale from 0% to 100%. The instrument that we exploited in our preferred specification used a confidence-weighted count of fire points in the relevant zone. For example a fire reported with confidence 50% would carry only half the weight of a fire reported with confidence 100%. Table 2.9 reports the effect of re-estimating our preferred specifications but replacing the confidence-weighted fire count with the unweighted or 'raw' count. As we expect, the coefficients at

[22] It is important to recognize that the wind direction from which these instruments is built are either prevailing wind directions or average wind direction over the multi-day period. However, the average masks the fact that the wind in most cases comes in a variety of directions in any three day period. As such fires from all directions (including directly downwind form the average) can contribute to prevailing air pollution in a city (Barwick et al., 2018; Bayer al., 2009).

the first stage are somewhat smaller, but the new version of the instrument remains strong with a similar F-statistic. The effect size implied by the second-stage estimates is somewhat eroded.

Restrict to high-fire months: Fires do not occur evenly across months, as they are tied to the agriculture cycle that varies by season, with patterns that are different in the different regions of China. In the paper up to this point, we have exploited data from the whole year. Here we consider restricting attention only to those months of the year in which, at any particular location, fires are most prevalent. China is commonly divided into seven geographical regions (see Appendix Figure A2.1). Appendix Table A2.2 identifies in boldface those months of the year in which fire counts are higher than the mean for each region. Table 2.10 reports the results of re-estimating our preferred OLS and IV specifications based only on dates falling in those fire-frequent (boldface) months. Comparing the results in columns (1) and (2), estimated on the whole sample, with the results in columns (3) and (4) on the restricted sample, it can be seen that there is little qualitative difference. The effect sizes implied by the IV based on the restricted sample are slightly larger, suggesting that our preferred approach based on the full sample is conservative.

Alternative time fixed effects Our preferred approach includes fixed effects for city by year, city by month, year by month, day of week and a holiday indicator. As is often the case, there are a number of other combinations of time fixed effects that could plausibly have been used in place of these. Tables 2.11 (OLS) and 2.12 (IV) summarise the results of re-estimating our main equation but with four alternative combinations of time fixed effects. In Table 2.11 the preferred specification is column (3), while in Table 2.12 it is those in columns (3) and (8). Inspection of the results makes clear that the alternative approaches make little difference to the conclusion in either the OLS or IV exercises. In the case of the IV results, in Table 2.12, each of the four alternative specifications implies a somewhat larger effect size.

Placebos Table 2.13 reports the results of two placebo exercises in which we replace the present pollutant with a falsely-assigned pollutant half year ago and half year after. As for the first stage, we replace the fire count with that on the falsely-assigned date. This is an approach common in the literature on the health and non-health effects of air pollution (Lavy al., 2014; Deryugina et al., 2016; Archsmith al., 2018) and is a test of design replacing the main independent variable of interest with randomly falsely-assigned ones. We know the irrelevant variable should not show its effect any more. In the upper panel, the coefficient estimates from the OLS exercise are smaller in absolute value in each case, mixed in sign, and in no case achieve significance at conventional levels. In the lower panel, the estimated coefficient of the fire count variable in the first stage is not only weak, but also close to zero, suggesting that it is irrelevant to the present pollution.

Irrelevant keywords: In our final falsification, exercise we replace the Baidu search index relating to cough, which we expect to be sensitive to air pollution, with the index for a series of other words which we do not expect to be. These include four health conditions (diarrhea, food poisoning, dislocation and stomach-ache) and two non-health related (desk and lamp). In each case the coefficient estimates are smaller, in most cases more than 20 times smaller. They are mixed in sign, and in no case achieve statistical significance at conventional levels.

2.7 Conclusions

This study presents the first attempt to estimate the causal impact of short-term (three-daily) variations in particulate air pollution on sub-clinical morbidity, in particular the incidence of coughing. To do this we use internet searches for the word cough on the very widely-used search engine Baidu as a proxy for cough prevalence. Our setting consists of the 100 largest Chinese cities.

Our analysis suggests that a one-standard-deviation increase in particulate matter in the air causes a roughly 5% increase in the prevalence of cough. The causal inference is underpinned by an instrumental variables approach, using a count of agricultural fires burning at any given time within the 150km circle surrounding any city as the instrument. The estimates prove resilient in the context of a wide set of robustness tests and falsification exercises.

The analysis complements the large existing body of work which established a link from short and long term exposure to particulates on a range of more serious health outcomes (including death), by extending it to consideration of lower-level morbidity impacts. Such impacts do not typically manifest as contact with the health-care system, and therefore do not show up in administrative data drawn, for example, from hospital admissions records, but can nonetheless substantially hamper quality of life. Such costs should be incorporated properly into the evaluation of the benefits of air quality improvement in China and elsewhere.

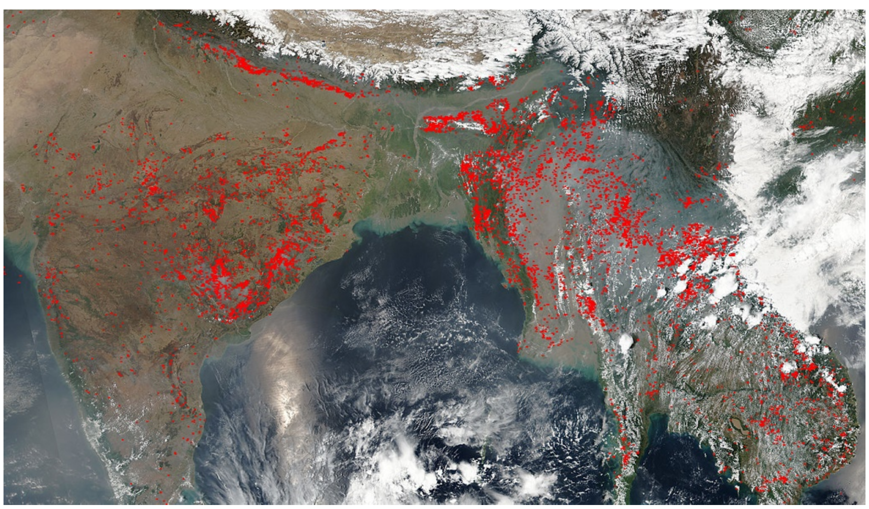
Figure 2.1: Baidu Index — Event of Searching Cough (Beijing)

Notes: This is a screenshot from online Baidu Index, which depicts the general index trend in Beijing for keyword "cough" during 2014-01-01 to 2017-12-31.

Figure 2.2: Straw Burning and Spread of Haze

Agriculture fires [Image by Manfred Sommer]
https://www.thethirdpole.net/en/2017/04/17/farm-fires-poison-air-over-south-asia/

(a)



Satellite image of agricultural fires [Photo by Jeff Schmaltz / NASA]
https://www.thethirdpole.net/en/2017/04/17/farm-fires-poison-air-over-south-asia/

(b)

Table 2.1: Summary Statistics

Three Day Average	Observations (1)	Mean (2)	Std. Dev. (3)	Min (4)	Max (5)
Cough	48635	152.454	84.934	0	997
AQI	46378	87.602	45.308	6.333	496.524
$PM_{2.5}$ $(\mu g/m^3)$	46369	62.491	43.564	0	529.667
NO_2 $(\mu g/m^3)$	46373	37.976	17.760	0	166.333
CO (mg/m^3)	46375	1.119	0.613	0	14.272
SO_2 $(\mu g/m^3)$	46374	25.701	25.575	0	365.333
O_3 $(\mu g/m^3)$	46377	92.476	42.952	0	284
Max Temperature ($^{\circ}C$)	48696	20.414	10.228	-21.433	41.033
Min Temperature ($^{\circ}C$)	48697	11.940	10.612	-33.800	31.200
Avg Humidity (%)	48696	0.685	0.163	0.120	1
Precipitation (mm)	40762	3.165	7.521	0	134.333
Sea-level Pressure (hPa)	46262	1016.302	9.172	991.167	1057.400
Wind Speed (Km/h)	48696	8.847	4.271	0.333	53.333
Cloud Cover(-/8)	48110	4.981	2.091	0	8

Notes: This table reports the summary statistics for all major variables involved in the main regression. Each unit in the dataset is a city-level observation based on the average value within three-day period. The whole period of time is from 2014 to 2017.

Table 2.2: Effect of $PM_{2.5}$ on the Incidence of Cough — OLS

OLS	$PM_{2.5}$_3				AQI_3
	(1)	(2)	(3)	(4)	(5)
Ln(cough_3)	0.084***	0.014***	0.017***	-0.001	0.020***
	(0.005)	(0.003)	(0.004)	(0.005)	(0.004)
Effect Size	3.65%	0.61%	0.75%	-	0.92%
Additional Controls					
City FEs	Y	Y	Y	Y	Y
Time FEs	N	Y	Y	Y	Y
Weather Controls	N	N	Y	Y	Y
Co-Pollutant	N	N	N	Y	-
Observations	46055	46054	36149	36149	36156

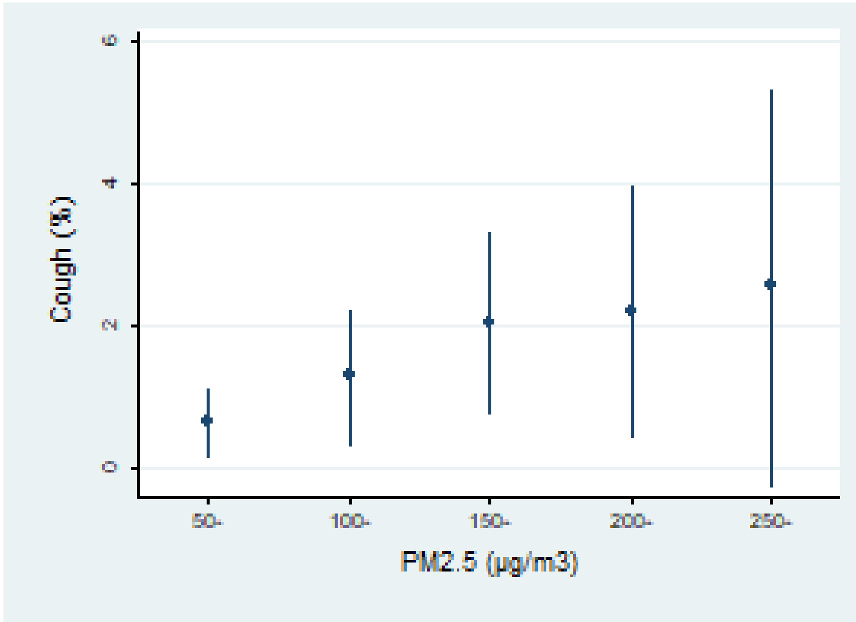
Notes: Column (1) through Column (4) estimate the effect of $PM_{2.5}$ on cough (log form) under increasing controls via OLS method. Column (5) reports the estimator for AQI under full controls as that in Column (3). All variables involved take the form of daily averages within a three-day period. Effect size indicates how much the web-search for cough increases due to a one-standard-deviation of $PM_{2.5}$ (or AQI). Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms for maximum, minimum temperature, relative humidity, wind speed, cloud coverage and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering the co-pollutants. Robust standard errors reported in parentheses are clustered at city level. (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.3: Non-linear Effect of $PM_{2.5}$ on the Incidence of Cough — OLS

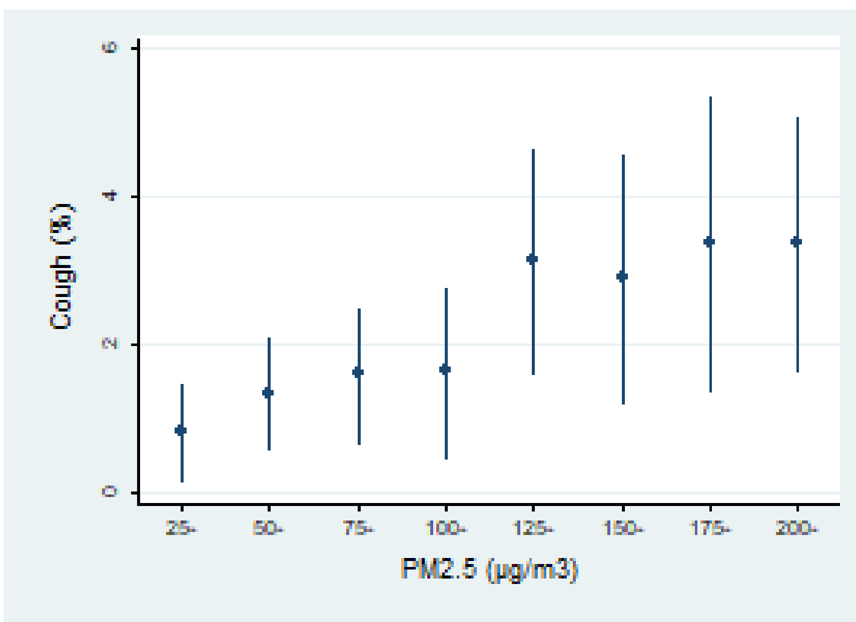
OLS	Non-linear Estimates (1)		(2)
[(0, 50]	-	[0, 25]	-
(50, 100]	0.665*** (0.245)	(25, 50]	0.804*** (0.335)
(100, 150]	1.297*** (0.480)	(50, 75]	1.339*** (0.386)
(150, 200]	2.064*** (0.651)	(75, 100]	1.581*** (0.466)
(200, 250]	2.219** (0.895)	(100, 125]	1.619*** (0.585)
>250	2.580* (1.408)	(125, 150]	3.149*** (0.772)
		(150, 175]	2.895*** (0.847)
		(175, 200]	3.367*** (1.008)
		>200	3.367*** (0.878)
Additional Controls			
City FEs	Y		Y
Time FEs	Y		Y
Weather Controls	Y		Y
Co-Pollutant	N		N
Observations	36149		36149

Notes: This table estimates the non-linear health effects of $PM_{2.5}$ measured in the form of bins. Column (1) includes $PM_{2.5}$ with 50 units per bin; Column (2) includes it with 25 units per bins. All observations involved are based on daily averages within a three-day period. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms for maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. Robust standard errors reported in parentheses are clustered at city level. (* significant at 10%, ** significant at 5%, *** significant at 1%).

Figure 2.3: Non-Linear Effects of $PM_{2.5}$ on the Incidence of Cough



(a)



(b)

Notes: Graphs (a) and (b) plot the point estimates and 95% confidence intervals listed in Column (1) and Column (2) of Table 2.3, respectively. The horizontal axis represents the level of $PM_{2.5}$ in each bin, and the vertical axis indicates the increase (in percentage form) in cough for each bin of $PM_{2.5}$ compared with the base group. The distribution of $PM_{2.5}$_3 is listed in Appendix Table A2.4 and Figure A2.2.

Table 2.4: First Stage Results for Each Pollutant Instrumented by Fire Points

Confidence-weighted Fires_3	$PM_{2.5}$_3 (1)	$PM_{2.5}$_3 (2)	CO_3 (3)	NO_2_3 (4)	SO_2_3 (5)	O_3_3 (6)	AQI_3 (7)
First Stage							
	0.310****	**0.208******	-0.0008	0.018**	-0.017	0.018	0.373***
	(0.065)	**(0.047)**	(0.0005)	(0.008)	(0.012)	(0.025)	(0.075)
F-Statistic	**22.492**	**19.752**	2.526	5.891	2.076	0.528	24.984
Additional Controls							
City FEs	Y	Y	Y	Y	Y	Y	Y
City by year FEs	Y	Y	Y	Y	Y	Y	Y
City by month FEs	Y	Y	Y	Y	Y	Y	Y
Year by month FEs	Y	Y	Y	Y	Y	Y	Y
Holiday FEs	Y	Y	Y	Y	Y	Y	Y
Day of week FEs	Y	Y	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y	Y	Y
Co-Pollutant	N	Y	Y	Y	Y	Y	-
Observations	36149	36149	36149	36149	36149	36149	36156

Notes: This table reports the first stage results by regressing each pollutant on confidence-weighted number of fires within the 150 km of the city centre. All observations are measured as daily averages within a three-day period. Kleibergen-Paap rk Wald F statistic is used to check whether the instrument is weak or not by comparing with the critical value of Stock-Yogo weak ID test (10% maximal IV size) which is 16.38 in our case. Obviously, $PM_{2.5}$ is the only pollutant that could be strongly instrumented by fire points with co-pollutant controls. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms for maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. Robust standard errors reported in parentheses are clustered at city level. Co-pollutants include $PM_{2.5}$, CO, NO_2, SO_2, and O_3. (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.5: Effect of $PM_{2.5}$ on the Incidence of Cough — IV

2SLS	$PM_{2.5}\_3$		AQI_3
	(1)	(2)	(3)
First Stage[(a)]			
	0.310***	0.208***	0.373***
	(0.065)	(0.047)	(0.075)
F-Statistics	22.492	19.752	24.984
Second Stage[(b)]			
Ln(Cough_3)	0.095***	0.118**	0.079***
	(0.037)	(0.058)	(0.031)
Effect Size	4.19%	5.20%	3.59%
Additional Controls			
City FEs	Y	Y	Y
Time FEs	Y	Y	Y
Weather Controls	Y	Y	Y
Co-Pollutant	N	Y	-
Observations	36149	36149	36156

Notes: (a) Dependent variable in the first stage is the measure of $PM_{2.5}$ (or AQI), and the independent variable is confidence-weighted number of fires within 150 km to the city centre. (b) Second stage regresses Baidu Index of cough (log form) on instrumented $PM_{2.5}$ (or AQI). All observations are measured as daily averages within a three-day period. Effect size indicates how much the web-search for cough increases due to a one-standard-deviation of $PM_{2.5}$. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level. (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.6: Robustness Check — Alternative Time Unit

Alternative Time Unit	Three Day Average (1)	Three Day Average (2)	Five Day Average (3)	Five Day Average (4)	Seven Day Average (5)	Seven Day Average (6)
			Panel A: OLS			
Ln(Cough_t)	0.017***	-0.001	0.021***	-0.006	0.021***	-0.003
	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.008)
Effect Size	0.75%	-	0.88%	-	0.88%	-
			Panel B: 2SLS			
First Stage						
	0.310***	0.208***	0.414***	0.284***	0.374***	0.275***
	(0.065)	(0.047)	(0.087)	(0.066)	(0.060)	(0.041)
F-Statistic	22.492	19.752	22.471	18.435	38.129	30.941
Second Stage						
Ln(Cough_t)	0.095***	0.118**	0.111***	0.129***	0.144**	0.171*
	(0.037)	(0.058)	(0.030)	(0.046)	(0.066)	(0.093)
Effect Size	4.19%	5.20%	4.64%	5.37%	5.76%	6.85%
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Time FEs	Y	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y	N	Y
Observation	36149	36149	18986	18986	11900	11900

Notes: This table re-conducts the preferred estimation for both OLS and IV methods within alternative periods, three-day period (Column (1) and (2)), five-day period (Column (3) and (4)) and seven-day period (Column (5) and (6)). Effect size indicates how much the web-search for cough increases due to a one-standard-deviation of $PM_{2.5}$. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.7: Robustness Check — Alternative Fire Count Radius

Alternative Fire Count Radius	Instrumented $PM_{2.5}\_3$			
	Within 150 km		**Within 100 km**	
	(1)	(2)	(3)	(4)
First Stage				
	0.310***	0.208***	0.544***	0.354***
	(0.065)	(0.047)	(0.118)	(0.078)
F-Statistic	22.492	19.752	21.327	20.722
Second Stage				
Ln(Cough_t)	0.095***	0.118**	0.101**	0.130*
	(0.037)	(0.058)	(0.046)	(0.073)
Effect Size	4.19%	5.20%	4.42%	5.71%
Additional Controls				
City FEs	Y	Y	Y	Y
Time FEs	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y
Observations	36149	36149	36149	36149

Notes: This table reports the robustness check by using alternative measure of fire counts as instruments in the first stage. Fires are counted within different scopes, the preferred 150 km (Column (1) and Column (2)), and the shorter one of 100 km (Column (3) and Column (4)). All variables included in the regression are measured as daily averages within a three-day period. Effect size indicates how much the web-search index for cough increases due to a one-standard-deviation of $PM_{2.5}$. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.8: Robustness Check — Alternative Wind Direction

Alternative Wind Direction Fires	Instrumented $PM_{2.5}\_3$					
	All Sides		**Upwind 90 Degree**		**Upwind 60 Degree**	
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage						
	0.310***	0.208***	0.514***	0.340***	0.659***	0.426***
	(0.065)	(0.047)	(0.097)	(0.070)	(0.117)	(0.091)
F-Statistic	22.492	19.752	28.232	23.409	31.569	21.921
Second Stage						
Ln(Cough_3)	0.095***	0.118**	0.106**	0.135*	0.119**	0.157*
	(0.037)	(0.058)	(0.047)	(0.074)	(0.053)	(0.085)
Effect Size	4.19%	5.20%	4.66%	5.94%	5.23%	6.90%
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Time FEs	Y	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y	N	Y
Observations	36149	36149	36149	36149	36149	36149

Notes: This table reports the robustness check by using alternative fire counts as instruments in the first stage. Fires are computed depending on different wind angles, the preferred all sides (Column (1) and Column (2)), upwind 90 degree (Column (3) and Column (4)), and upwind 60 degree (Column (5) and Column (6)). All variables included in the regression are measured as daily averages within a three-day period. Effect size indicates how much the web-search index for cough increases due to a one-standard-deviation of $PM_{2.5}$. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.9: Robustness Check — Alternative Instrument

Alternative Instrument	$PM_{2.5}$_3			
	Confidence-Weighted Fire_3		**Raw Fire_3**	
	(1)	(2)	(3)	(4)
First Stage				
	0.310***	0.208***	0.195***	0.132***
	(0.065)	(0.047)	(0.040)	(0.030)
F-Statistic	22.492	19.752	23.511	19.190
Second Stage				
Ln(Cough_3)	0.095***	0.118**	0.079**	0.094*
	(0.037)	(0.058)	(0.035)	(0.054)
Effect Size	4.19%	5.20%	3.46%	4.11%
Additional Controls				
City FEs	Y	Y	Y	Y
Time FEs	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y
Observations	36149	36149	36149	36149

Notes: This table reports the robustness check by using the raw number of fires as an instrument in the first stage (Column (3) and (4)). Column (1) and Column (2) replicate the preferred specification both without and with co-pollutants. All variables included in the regression are measured as daily averages within a three-day period. Effect size indicates how much the web-search index for cough increases due to a one-standard-deviation of $PM_{2.5}$. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.10: Robustness Check — Alternative Sub-sample

Alternative Sub-sample	**Full** (1)	(2)	**Fire-frequent months** (3)	(4)
		Panel A: OLS		
Ln(Cough_t)	0.019***	0.007	0.015**	0.012
	(0.004)	(0.005)	(0.007)	(0.009)
Effect Size	0.83%	-	0.67%	-
		Panel B: 2SLS		
First Stage				
	0.365***	0.260***	0.333***	0.248***
	(0.074)	(0.057)	(0.074)	(0.054)
F-Statistic	24.165	21.064	20.181	21.447
Second Stage				
Ln(Cough_t)	0.109***	0.140**	0.157***	0.201***
	(0.041)	(0.058)	(0.046)	(0.065)
Effect Size	5.34%	6.88%	6.93%	8.87%
Additional Controls				
City FEs	Y	Y	Y	Y
Region by year	Y	Y	Y	Y
Region by month	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y
Observation	36150	36150	14098	14098

Notes: Column (1) and Column (2) re-estimate the regression based on the full sample under both OLS and IV methods, without and with co-pollutants. Column (3) and Column (4) re-estimate the coefficients under the limited sample, which excludes the months with fewer fires (see Appendix Table A2.2 for the excluded months). All variables included in the regression are measured as daily averages within a three-day period. Effect size indicates how much the web-search index for cough increases due to a one-standard-deviation of $PM_{2.5}$. Time fixed effects include **region by year, region by month,** holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering the co-pollutants. Robust standard errors reported in parentheses are clustered at city level. (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.11: Robustness Check — Alternative Time Fixed Effects (OLS)

Alternative Fixed Effect (OLS)	(1)	(2)	(3) Preferred	(4)	(5)
Ln(Cough_3)	0.019***	0.014***	0.017***	0.015***	0.011***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Effect Size	0.83%	0.62%	0.75%	0.67%	0.47%
Additional Controls					
City FEs	Y	Y	Y	Y	Y
Region by year	**Y**	N	N	N	N
Region by month	**Y**	N	N	N	N
City by year	N	**Y**	**Y**	**Y**	**Y**
City by month	N	**Y**	**Y**	**Y**	**Y**
Year by month	N	N	**Y**	N	**Y**
Region by year by month	N	N	N	**Y**	N
Date FEs	N	N	N	N	**Y**
Holiday	Y	Y	Y	Y	Y
Day of week	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y
Co-pollutant	N	N	N	N	N
Observations	36150	36149	36149	36149	36149

Notes: This table reports the robustness check with OLS results under different temporal controls. From Column (1) to Column (5), the time fixed effects involved are more conservative. Column (3) replicates the preferred results shown in Column (3) of Table 2.2 with city by month, city by year and year by month fixed effects. All variables included in the regression are measured as daily averages within a three-day period. Effect size indicates how much the web-search index for cough increases due to a one-standard-deviation of $PM_{2.5}$. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.12: Robustness Check — Alternative Time Fixed Effects (2SLS)

Alternative Fixed Effect (2SLS)	**Without Co-pollutants**					**With Co-pollutants**				
	(1)	(2)	(3) Preferred	(4)	(5)	(6)	(7)	(8) Preferred	(9)	(10)
First Stage										
	0.365***	0.360***	0.310***	0.277***	0.310***	0.260***	0.253***	0.208***	0.186***	0.213***
	(0.074)	(0.079)	(0.065)	(0.067)	(0.069)	(0.057)	(0.059)	(0.047)	(0.048)	(0.050)
F-Statistic	24.165	20.902	22.492	17.271	20.197	21.064	18.333	19.752	14.761	18.039
Second Stage										
Ln(Cough_3)	0.109***	0.132***	0.095***	0.102**	0.100***	0.140**	0.166***	0.118**	0.129**	0.138**
	(0.041)	(0.034)	(0.037)	(0.043)	(0.038)	(0.058)	(0.051)	(0.058)	(0.066)	(0.056)
Effect Size	5.34%	5.81%	4.19%	4.47%	4.40%	6.88%	7.29%	5.20%	5.56%	6.08%
Additional Controls										
City FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Region by year	**Y**	N	N	N	N	**Y**	N	N	N	N
Region by month	**Y**	N	N	N	N	**Y**	N	N	N	N
City by year	N	**Y**	**Y**	**Y**	**Y**	N	**Y**	**Y**	**Y**	**Y**
City by month	N	**Y**	**Y**	**Y**	**Y**	N	**Y**	**Y**	**Y**	**Y**
Year by month	N	N	**Y**	N	**Y**	N	N	**Y**	N	**Y**
Region by year by month	N	N	N	**Y**	N	N	N	N	**Y**	N
Date FEs	N	N	N	N	**Y**	N	N	N	N	**Y**
Holiday	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Day of week	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Co-pollutant	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	36150	36149	36149	36149	36149	36150	36149	36149	36149	36149

Notes: This table reports the robustness check with IV results under different temporal controls. Column (1) through Column (5) display the outcomes without controlling the co-emissions. Column (6) through Column (10) take account of the co-pollutants. Column (3) and Column (8) repeat the preferred IV results shown in Column (1) and Column (2) of Table 2.5, respectively, with city by month, city by year and year by month fixed effects. All variables included in the regression are measured as daily averages within a three-day period. Effect size indicates how much the web-search index for cough increases due to a one-standard-deviation of $PM_{2.5}$. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.13: Placebo

Placebo	Preferred (1)	(2)	Half year before (3)	(4)	Half year after (5)	(6)
			Panel A: OLS			
Ln(Cough_3)	0.017***	-0.001	0.006	0.006	-0.004	0.005
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Effect Size	0.75%	-	-	-	-	-
Observation	36149	36149	31528	31528	31629	31629
			Panel B: 2SLS			
First Stage						
	0.310***	0.208***	-0.024	-0.024	0.071	0.060
	(0.065)	(0.047)	(0.023)	(0.023)	(0.069)	(0.069)
F-Statistic	22.492	19.752	-	-	-	-
Second Stage						
Ln(Cough_3)	0.095***	0.118**	-	-	-	-
	(0.037)	(0.058)	-	-	-	-
Effect Size	4.19%	5.20%	-	-	-	-
Observation	36149	36149	30472	30472	30215	30215
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Time FEs	Y	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y	N	Y

Notes: This table conducts the placebo test by replacing the pollutant in the regression with dis-matched date (half-year before and half-year later). All variables included in the regression are measured as daily averages within a three-day period. Effect size indicates how much the web-search index for cough increases due to a one-standard-deviation of $PM_{2.5}$. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 2.14: Falsification Test

Pollution_unrelated Keywords	OLS	IV		**Pollution_unrelated Keywords**	OLS	IV	
	(1)	(2)	(3)		(4)	(5)	(6)
Ln(Diarrhea_3)	-0.0016 (0.0014)	-0.0004 (0.0003)	-0.0006 (0.0005)	**Ln(Dislocation_3)**	-0.0001 (0.0001)	0.0005 (0.0015)	0.0010 (0.0028)
Observation	26250	25306	25306	Observation	5087	5087	5087
Ln(Food Poisoning_3)	0.0001 (0.0001)	0.0002 (0.0004)	0.0002 (0.0007)	**Ln(Stomachache_3)**	-0.00004 (0.00004)	-0.0005 (0.0005)	-0.0009 (0.0007)
Observation	21587	21587	21587	Observation	24832	24832	24832
Ln(Lamp_3)	0.00003 (0.00012)	-0.0004 (0.0006)	-0.0006 (0.0009)	**Ln(Desk_3)**	0.00003 (0.00005)	-0.0004 (0.0006)	-0.0008 (0.0011)
Observation	6335	6335	6335	Observation	15046	15046	15046
Additional Controls							
City FEs	Y	Y	Y		Y	Y	Y
Time FEs	Y	Y	Y		Y	Y	Y
Weather Controls	Y	Y	Y		Y	Y	Y
Co-Pollutant	N	N	Y		N	N	Y

Notes: This table shows the falsification test by replacing "cough" with pollution-unrelated keywords. Column (1) and Column (4) regress the Baidu Index for individual keywords on $PM_{2.5}$ using OLS. Other specifications are estimated under IV methods. Two non-disease keywords picked up randomly (lamp, and desk) are supplemented in the last row using the identical estimation. All variables included in the regression are measured as daily averages within a three-day period. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms of maximum, minimum temperature, relative humidity, wind speed, cloud coverage, and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Figure A2.1: Chinese Regions

Regions	Province (31)
Southern	Guangdong, Guangxi, Hainan
Southwestern	Chongqing, Sichuan, Guizhou, Yunnan, Tibet
Central	Henan, Hubei, Hunan
Eastern	Shanghai, Jiangsu, Zhejiang, Anhui, Jiangxi, Shangdong, Fujian
Northern	Beijing, Tianjin, Hebei, Shanxi, Inner Mongdia
Northeastern	Liaoning, Jilin, Heilongjiang
Northwestern	Shannxi, Gansu, Qinghai, Ningxia, Xinjiang

Notes: This table lists all the provinces (Mainland China) included in the sample and their individual geographical distribution, among which Inner Mongdia, Xinjiang, Guangxi, Ningxia, Tibet are 5 autonomous regions of China; and Beijing, Shanghai, Tianjin, Chongqing are 4 direct-controlled municipalities of China.

Table A2.1: Reduced Form

Reduced Form	Fire_3	
	(1)	(2)
Ln(cough_3)	0.030***	0.025**
	(0.011)	(0.012)
Additional Controls		
City FEs	Y	Y
Time FEs	Y	Y
Weather Controls	Y	Y
Co-Pollutant	N	Y
Observations	36149	36149

Notes: This table reports the reduced form outcomes by regressing city-level Baidu Index of cough on confidence-weighted number of fires within 150km to city center. All observations are involved based on daily averages within a three-day period. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms for maximum, minimum temperature, relative humidity, wind speed, cloud coverage and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level. (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A2.2: Daily Fire Counts in Each Month for Different Regions

	Southern (1)	Southwestern (2)	Central (3)	Eastern (4)	Northern (5)	Northeastern (6)	Northwestern (7)
January	**2.825**	**0.860**	2.128	1.024	0.708	1.177	0.392
	(7.293)	**(7.542)**	(12.024)	(4.371)	(2.162)	(4.353)	(2.263)
February	**1.642**	**0.770**	2.020	0.971	1.936	3.853	0.642
	(4.793)	**(3.216)**	(8.960)	(3.895)	(5.559)	(11.989)	(2.429)
March	1.113	0.561	1.193	0.903	**2.982**	**10.831**	0.785
	(4.017)	(2.702)	(3.339)	(2.616)	**(5.028)**	**(23.240)**	(1.850)
April	**2.289**	**0.997**	1.679	**2.047**	1.962	**9.397**	0.663
	(12.308)	**(3.280)**	(7.185)	**(8.066)**	(5.201)	**(24.486)**	(1.811)
May	0.641	**1.010**	**4.087**	**4.469**	**3.974**	1.536	**1.577**
	(1.757)	**(4.109)**	**(13.479)**	**(16.243)**	**(12.339)**	(3.725)	**(6.003)**
June	1.163	0.372	**6.722**	**2.270**	**6.768**	1.738	**1.443**
	(3.073)	(1.733)	**(33.359)**	**(11.158)**	**(20.743)**	(4.447)	**(3.806)**
July	1.092	**1.532**	**4.058**	**4.057**	**4.504**	1.770	**1.902**
	(2.916)	**(4.767)**	**(12.631)**	**(10.492)**	**(12.263)**	(4.093)	**(5.814)**
August	1.278	**1.131**	1.687	**3.557**	2.356	1.403	**0.890**
	(2.874)	**(3.616)**	(4.828)	**(10.633)**	(5.661)	(2.926)	**(2.776)**
September	1.192	0.147	0.642	0.419	1.444	2.624	0.561
	(2.948)	(0.894)	(2.043)	(1.332)	(3.793)	(6.267)	(1.459)
October	0.910	0.115	**3.029**	0.540	**2.949**	**21.417**	**1.074**
	(2.002)	(0.477)	**(15.385)**	(1.861)	**(7.975)**	**(53.124)**	**(2.604)**
November	0.623	0.098	0.398	0.458	1.496	**13.686**	0.431
	(1.673)	(0.382)	(1.284)	(1.417)	(4.252)	**(40.563)**	(1.389)
December	**1.365**	0.198	0.848	0.559	0.533	0.450	0.155
	(3.582)	(1.193)	(3.491)	(1.838)	(1.304)	(1.920)	(0.551)
Mean	1.342	0.651	2.377	1.784	2.636	5.827	0.879
	(5.031)	(3.502)	(13.060)	(7.951)	(9.053)	(23.036)	(3.225)

Notes: This table displays city-level daily average fire counts from January to December for each geographical region (Column (1) through Column (7)). Standard deviations are reported in parentheses. The last row shows daily mean counts in the corresponding region. Under each column, the values are bold if they are larger than the mean value reported in the bottom. Only the bold region-months will be kept for constructing fire-frequent sub-sample which is used for robustness check in Table 2.12.

Table A2.3: First Stage Instrumented by Fire Points — Alternative Wind Angles and Scopes

Fire_3	$PM_{2.5}$_3 (All Sides) (1)	(2)	$PM_{2.5}$_3 (Upwind 90 Degree) (3)	(4)	$PM_{2.5}$_3 (Upwind 60 Degree) (5)	(6)
First Stage	1.036***	0.656***	1.515***	0.979**	1.830***	1.152**
(within 50km)	(0.321)	(0.226)	(0.563)	(0.411)	(0.689)	(0.504)
F-Statistic	10.446	8.406	7.249	5.666	7.062	5.224
Strong Instrument	N	N	N	N	N	N
First Stage	0.544***	0.354***	0.922***	0.602***	1.151***	0.731***
(within 100km)	(0.118)	(0.078)	(0.208)	(0.140)	(0.260)	(0.177)
F-Statistic	21.327	20.722	19.717	18.576	19.671	17.136
Strong Instrument	Y	Y	Y	Y	Y	Y
First Stage	**0.310*** **	**0.208*** **	**0.514*** **	**0.340*** **	**0.659*** **	**0.426*** **
(within 150km)	**(0.065)**	**(0.047)**	**(0.097)**	**(0.070)**	**(0.117)**	**(0.091)**
F-Statistic	**22.492**	**19.752**	**28.232**	**23.409**	**31.569**	**21.921**
Strong Instrument	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
First Stage	0.190***	0.128***	0.311***	0.205***	0.382***	0.243***
(within 200km)	(0.039)	(0.032)	(0.062)	(0.053)	(0.081)	(0.073)
F-Statistic	23.206	15.805	24.853	14.819	22.351	21.921
Strong Instrument	Y	N	Y	N	Y	Y
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Time FEs	Y	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y	N	Y
Observation	36149	36149	36149	36149	36149	36149

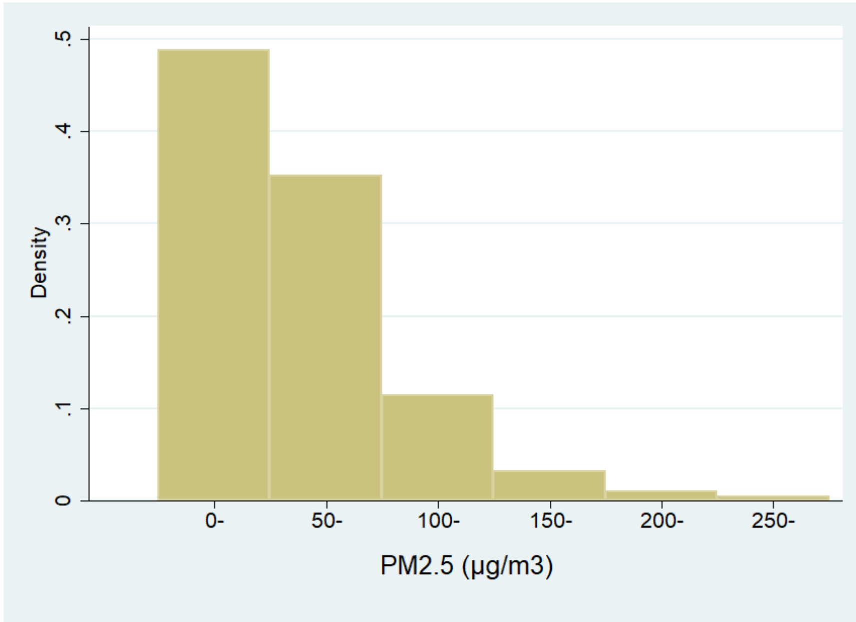
Notes: This table lists the first stage results regressing $PM_{2.5}$ on fires counted by alternative ways, varying from 50km to 200km scope, and from all sides to upwind 60 degree. Kleibergen-Paap rk Wald F statistic is used to check whether the instrument is weak or not by comparing with the critical value of Stock-Yogo weak ID test (10% maximal IV size) which is 16.38 in our case. All variables included in the regression are measured as daily averages within a three-day period. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms for maximum, minimum temperature, relative humidity, wind speed, cloud coverage and sea-level pressure. Co-pollutants include CO, NO_2, SO_2, and O_3. Robust standard errors reported in parentheses are clustered at city level. (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table A2.4: Distribution of $PM_{2.5\_}3$

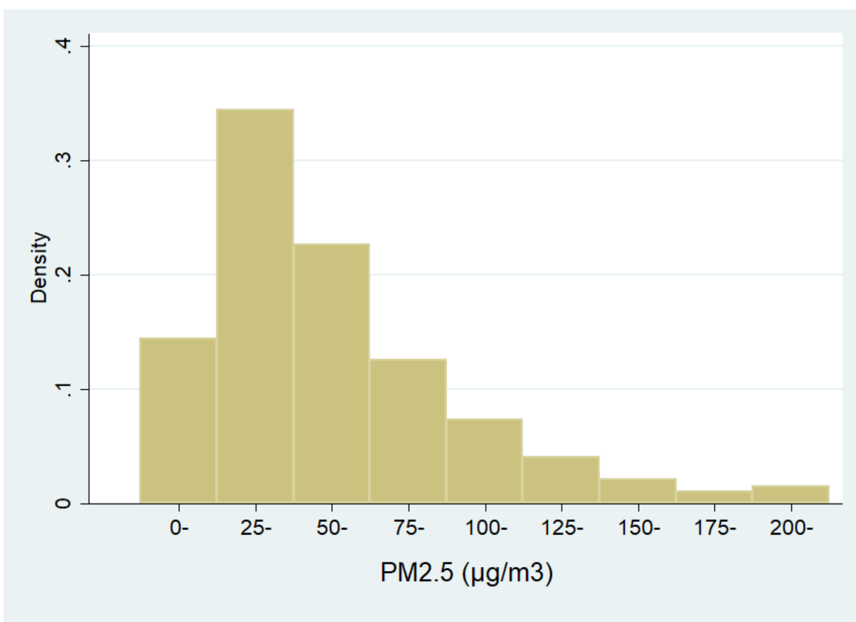
	Frequency	Percentage (%)	Cumulative Percentage (%)		Frequency	Percentage (%)	Cumulative Percentage (%)
	(1)	(2)	(3)		(4)	(5)	(6)
(0, 50]	22,641	48.83	48.83	(0, 25]	6,665	14.37	14.37
(50, 100]	16,308	35.17	84	(25, 50]	15,976	34.45	48.83
(100, 150]	5,287	11.4	95.4	(50, 75]	10,486	22.61	71.44
(150, 200]	1,469	3.17	98.57	(75, 100]	5,822	12.56	84
(200, 250]	462	1	99.56	(100, 125]	3,409	7.35	91.35
>250	202	0.44	100	(125, 150]	1,878	4.05	95.4
				(150, 175]	979	2.11	97.51
				(175, 200]	490	1.06	98.57
				>200	664	1.43	100
Total	46369	100		Total	46369	100	

Notes: This table displays the distribution of $PM_{2.5\_}3$ (the average level of $PM_{2.5}$ within three-day period). Column (1) through Column (3) classify $PM_{2.5\_}3$ with 50 $\mu g/m^3$ for each bin. Column (4) through Column (6) report $PM_{2.5\_}3$ in more detail with 25 $\mu g/m^3$ for each bin. The following histogram (Appendix Figure A2.2) plots the corresponding density distribution.

Figure A2.2: Distribution of $PM_{2.5}\_3$



(a)



(b)

Table A2.5: Alternative Time Frame — Daily, Two Days and Three Days

Comparisons	**Daily**		**Two days**		**Three days (Preferred)**	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: OLS			
Ln(Cough_t)	0.010***	-0.007**	0.013***	-0.006	0.017***	-0.001
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)
Effect Size	0.49%	-	0.59%	-	0.74%	-
			Panel B: 2SLS			
First Stage						
	0.221***	0.096***	0.222***	0.151***	0.310***	0.208***
	(0.063)	(0.025)	(0.055)	(0.037)	(0.065)	(0.047)
F-statistic	**14.416**	**14.146**	**16.446**	**16.797**	**22.492**	**19.752**
Strong instrument (compare with 16.38)	N	N	Y	Y	Y	Y
Second Stage						
Ln(Cough_t)	-	-	0.113***	0.145***	0.095***	0.118**
	-	-	(0.037)	(0.057)	(0.037)	(0.058)
Effect Size	-	-	5.19%	6.68%	4.19%	5.20%
Additional Controls						
City FEs	Y	Y	Y	Y	Y	Y
Time FEs	Y	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y	Y
Co-Pollutant	N	Y	N	Y	N	Y
Observation	124249	124249	58090	58090	36149	36149

Notes: This table re-conducts the estimation for both OLS and IV methods within alternative periods, varying from daily (Column (1) and (2)), two-day period (Column (3) and (4)) to the preferred three-day period (Column (5) and (6)). Effect size is measured relative to mean in term of one standard deviation of $PM_{2.5}$. Time controls include city by year, city by month, year by month, holiday and day of week fixed effects. Weather covariates contain squared terms for maximum, minimum temperature, relative humidity, wind speed, cloud coverage and sea-level pressure. CO, NO_2, SO_2, and O_3 are all included when considering co-pollutants. Robust standard errors reported in parentheses are clustered at city level.(* significant at 10%, ** significant at 5%, *** significant at 1%).

Chapter 3

Dreaming of Blue Skies: Air Pollution and Mobility Aspirations of Young People in Beijing using Online Search Behavior

3.1 Introduction

This paper investigates whether short-term air pollution causes young, highly-educated Chinese people in Beijing to express a desire to leave their home city to study elsewhere. We develop a proxy for aspiration that relies on the intensity of web searches for information about local versus out-of-city universities on the ubiquitous search engine Baidu. While we do not observe actual migration outcomes, (a) understanding intent can be interesting in its own right, and (b) internet search behavior has been shown to predict behavior in many settings (Goel et al., 2010; Choi and Varian, 2012). We bolster our causal claims by instrumenting for local pollution levels with plausible-exogenous variations in temperature inversion strength.

Those students who graduate from high school in China hope to progress to further education through attending the National Higher Education Entrance Examination for post-secondary education. This is a standardized national test, in which performance informs the sorting of students into universities. Conditional on achievement applicants have significant choice. In choosing a university they also choose a place to live, typically for at least the following four years. For many, the city in which they study subsequently becomes a long-term home. Our focus on students' migration has at least three advantages: First, as in other countries, movement for university is a key contributor to overall internal migration. Second, those moving belongs to the higher-skilled section of the labor force, the mobility of which is likely to be particularly influential for local economics outcomes. Third, since individuals are more 'footloose' at that stage in life, choice to move for university can more easily be sensitive to taste for local environment.

Beijing is well-endowed with highly-rated universities, including 21 in China's top 100 (see Table 3.2, the university ranking order in 2014). However, it is also one of the most heavily polluted cities in China. The elevated atmospheric pollution does significant damage to the physical and mental health of inhabitants (Neidell, 2004; Herrnstadt and Muehlegger, 2015), as well as to working life (Hanna and Oliva, 2015), in both the short-term and the long-term. Students in Beijing have been shown to be conscious about the ambient pollution (Wong, 2003). In response to these adverse impacts, students may take advantage of the opportunity to escape from the grey haze in local places.

In this paper, we provide the first study to investigate people's pollution-driven interest in out-of-city higher education. The emphasis will be placed on examining the changes in people's search behavior for universities both inside and outside Beijing, which is quantified by a search index.

There is by now a large and diverse body of research that exploits online search-behavior to proxy for current thinking and to predict future behavior (Rech, 2007; Goel et al., 2010). For example, searches for the symptoms diseases are correlated with actual disease prevalance

(Seifter et al., 2010; Anggraeni and Aristiani, 2016). Askitas and Zimmermann (2009) shows that Google search activity data can serve as a current and leading indicator for unemployment. Hand and Judge (2012) show that searches can predict cinema admissions and movie success. Wu and Brynjolfsson (2015) use Google search to forecast housing prices and sales. Choi and Varian (2012) have verified that the search trend data is strongly predictive of present economic activity through the comparisons between trend data and actual data in retail sales, automotive sales, home sales, and tourism, respectively.

Therefore, the indices of search intensity provide a useful window into the mind of internet users and a good predictor for the future actual behavior. We develop a metric that embeds how frequently people in Beijing are searching for information about Beijing-based universities compared to universities in other cities. We explore whether short-term bad air in Beijing causes the relative frequency of the former to decrease. We first use OLS to estimate the potential linear and non-linear association between air pollution in Beijing and the share of search in local universities. Then we supplement the analysis with the IV method by instrumenting for pollution with temperature inversion strength. The negative effect is substantial and robust after including different controls. In our preferred specification, when air pollution increases from 5% to 95% level within monthly mid-run, the search interest in local universities decreases by 3.6% relative to mean under OLS, and 11.2% under IV.

Our research complements two literatures. First, some works focus on the relationship between air quality at a location and the desirability of that location as a place to live. Such studies typically exploit hedonic methods based on house prices (Chay and Greenstone, 2005; Bayer et al., 2009; Greenstone and Gallagher, 2008). Second, there is a small emerging literature on the environmental determinants of migration (Chen et al., 2017; Qin and Zhu, 2018). For example, Chen et al. (2017) estimate the causal effect of air pollution on migrations across cities (or counties) in a mid to long-run setting. Perhaps closest to our paper in spirit, Qin and Zhu (2018) study how daily variation of city-level air pollution in China is correlated with next-day internet searches for the keyword “emigration”.

The rest of this paper is organized as follows. Section 3.2 outlines data sources. Section 3.3 explains the methodology strategy. In Section 3.4, we present our main results and robustness checks. Section 3.5 concludes.

3.2 Data

The purpose of our study is to estimate the causal effect of air pollution on potential decreased interest in local universities. The interest to leave local place (Beijing) for higher education is measured by the proxy of daily online search index. Air pollution and weather data are all integrated from hourly data into daily averages. In order to cope with the possible endogeneity problems, the IV method is supplemented based on the OLS specification. Due to the typical topographic conditions in Beijing, its frequent temperature inversions could serve as a strong instrument to indicate the variation of local air quality (Yang et al., 2015). All of the data involved will be discussed in the following sections.

3.2.1 Search Index

The main challenge for our work is to find a suitable measure for people's daily interest in each university. We introduce the use of search index on Baidu platform. Founded in 2000, Baidu is the largest Internet company in China specializing in search engine services. Baidu provides search index starting from January in 2011, computed by the weighted sum of search volume for each keyword people query on Baidu Web, providing the similar service as "Google Trend".[1] Although Baidu Company does not provide the exact formula under the search index, the highly linear correlation between Baidu Index and search volume has been verified in the work of Qin and Zhu (2018). Unlike Google Trend with the value ranging from 0 to 100, the Baidu Index is proportional to search volume without the maximum value. The larger the search frequency for one keyword is, the larger value of Baidu Index is.

Moreover, it could also locate the daily search in the specific city. The Baidu Index

[1] For the official explanation for Baidu Index, see https://zhishu.baidu.com/Helper/?tpl=help&word=#pdesc

analyses the hot trend of national concern based on the search behavior of hundreds of millions of netizens. It is an indispensable reference platform for big data analysis and plays an important predictive role in marketing decision-making.

We mainly focus on the top 100 universities in China.[2] A significant advantage of Beijing as our study setting is that the internet is very widely used there as the primary source of information, particularly among adolescents and young adults. The top-ranked universities are of higher quality and are more likely to attract the interest (The World University Rankings, 2018). In later robustness check, we will expand our range to include those universities in the top 200 to 500.

The universities in China at top 100 are divided into the ones in Beijing and the ones out of Beijing. Baidu search index for each university is sum up into those two parts. Now the main question is to see whether the share of search index for Beijing-based universities is related to the variations in air pollution.

It is worth mentioning that since Baidu Index has been scaled down the real search volume, mathematical operations based on Baidu Index, like addition and division, also make sense.

We define

$$U = U^1 + U^2 + U^3 + ... + U^{100} \tag{1}$$

as Beijing residents' total search activity for the top 100 universities, that is the sum of Baidu Index for each university within the range.[3] Among those 100 universities, 21 of which are in Beijing, 79 are out-of-Beijing. Accordingly, U is also composed of search index for Beijing universities (UB) and search index for non-Beijing universities (UNB).

The main dependent variable of interest in this study is the share of search index for universities in Beijing ($IB = UB/U$). Our principal interest is to check whether this ratio varies with daily changes in air pollution. Table 3.1 displays the summary statistics for

[2]The rank is based on the "China University Evaluation Research Report" released by Ai Ruishen Alumni Association (http://www.cuaa.net/), which is the third-party evaluation institution with good credibility.

[3]We count the top 100 universities according to the rank report in 2014.

this index ratio (IB) with different level groups (from top 100 to top 700) and various time frames (from past 10 days to 30 days) under Panel A.

Most of our results will be based on analysis conducted on the top 100 universities. Table 3.2 samples the ranking information and corresponding daily average search index. In general, the daily share of search index for local universities accounts for around 47% of the total search (also see Table 3.1 for IB_t under Panel A). When people weight and decide whether to leave for outside education, they are more likely to choose among the same ranking. Focusing on the universities in the top 100 is well targeted. Besides that, we also expand the top levels from the top 200 to the top 500 in the later discussions.

3.2.2 Air Pollution

The US Embassy built an air quality monitor that measures $PM2.5$ particles at the embassy in the Chaoyang District. It continues to update the historical air data ($PM2.5$ per hour), and the earliest one can be traced back to 2008.[4] Recall that the earliest Baidu Index can be obtained since January 2011. Therefore, compared with the data from MEP, the US Embassy source could enable us to carry out the regression analysis over a longer time span.[5] Although it only provides data for one monitoring station and does not cover the entire Baidu search area in Beijing, it has also been used as an alternative dataset due to its accuracy and openness (Chang et al., 2018; Qin and Zhu, 2018). Thus, we use the $PM2.5$ data from US Embassy spanning the period from January in 2011 to June in 2017 in the study.[6]

We aim to explore the causal effect of $PM2.5$ in a single city, Beijing, on the mobility aspirations of young people. $PM2.5$ is the primary pollutant in Beijing (Peng et al., 2017),

[4] Air pollution data reported from US embassy: https://aqicn.org/city/beijing/us-embassy/

[5] The BJMEP has published annual air reports since 2009, and the national MEP has released monthly air quality reports for 74 cities, including Beijing since 2013. The earliest daily historical data can be retrieved from the third party source *aqistudy.cn* since December 2013, which integrates hourly pollution data from all the monitoring stations and published daily measures, including AQI, $PM2.5$, $PM10$, $SO2$, CO, $NO2$, $O3_{8h}$. This platform processes data collection using the Application Programming Interface (API) provided by *PM25.in*, which is the direct mirror of national MEP data (Rohde and Muller, 2015).

[6] US Embassy does not release the historical air pollution after June 30 in 2017.

the variation of which could well indicate the local air quality.[7] As shown in Table 3.1, $PM2.5$ is reported under different time frames, including a ten-day mean, a fifteen-day mean, a twenty-day mean, and a thirty-day mean, with less deviation in the longer time frame. Their mean values are all around 90 $\mu g/m^3$ in magnitude, suggesting that most of days is lightly polluted.[8] Rolling-monthly mean of $PM2.5$ are studied in the main regression, with fewer than 4% of clean days under 50 μg/m^3, and most of observations in the group between 50–75 μg/m^3 and 75–100 μg/m^3.

3.2.3 Meteorological Data

Climate change and severe weather could trigger people's intention to study in other places, which we posit are closely related to the variation in ambient quality. We collect the weather data in Beijing from the National Oceanic and Atmospheric Administration (NOAA). Hourly data are all converted into daily levels and then averaged over various time frames. Weather covariates include the squared term in maximum temperature ($°C$), minimum temperature ($°C$), relative humidity (%), wind speed (km/h), cloud coverage (-/8), sea-level pressure (hPa), and precipitation (mm). The summary statistics over the past 30 days are displayed in Panel D of Table 3.1.

3.2.4 Temperature Inversion

In order to address the potential endogeneity issues (discussed in detail later), we instrument for local air pollution with temperature inversion. Normally air temperature becomes lower (air colder) as you move further from the ground. Sometimes air becomes inverted, which is to say colder close to the ground. Under such a temperature inversion, a layer of air is 'pinned' at the surface. In those conditions, pollution, most of which is generated

[7] $PM2.5$ is also the only pollutant released by US Embassy.

[8] In China, air quality is classified into six levels according to AQI, including excellent (0–50), good (51–100), lightly polluted (101–150), moderately polluted (151–200), heavily polluted (201–300) and severely polluted (>300). Referring to the formula published by the MEP in "Technical Regulation on Ambient Air Quality Index (http://210.72.1.216:8080/gzaqi/Document/aqijsgd.pdf)", $PM2.5$ is at 75 $\mu g/m^3$ when AQI is at 100.

by activity on the ground, rather than escaping upward and outward, is held at ground level (see Figure 1). The reversal of air temperature in the troposphere is well-understood to exacerbate local air pollution problems and has been used to instrument for short-term changes in air quality by a number of researchers (Jans et al., 2018; Sager, 2016; Chen et al., 2017).

We develop an instrument using both the existence and intensity of inversion in the vicintiy of Beijing. The air temperature data are obtained from MERRA-2, which provides outputs on the $0.625° \times 0.5°$ longitude-by-latitude grid (GMAO, 2015).[9] Beijing is spatially interpolated to twelve grids, with the Y-axis at 39.03°, 39.53°, 40.03°, 40.53° (degrees North latitude), and the X-axis at 115.0591°, 115.7257°, 116.3924° (degrees East longitude). Air temperature data are collected in the four dimensions of troposphere (1000hPa, 975hPa, 950hPa and 925hPa) and a frequency of four times per day (0:00, 6:00, 12:00 and 18:00).

Normally, the 1000hPa layer has the highest temperature, close to the surface level. The upper layers, like 975hPa and 950hPa, and 925hPa, have colder air. When temperature inversion occurs, the air temperature at higher layers, either in 975hPa, 950hPa or 925hPa, is warmer than that at the 1000hPa layer. The day in Beijing with the occurrence of reversal temperature in any of twelve grids at any of four time periods is recorded as an inversion day.

Beijing has very frequent temperature inversions due to its special topography, with the occurrence almost every one or two days. The upper panel in Figure 3.2 plots the monthly distribution of inversion days from January in 2011 to the end of 2017. The frequency of occurrence has a seasonal pattern, with higher frequencies in winter. The average number of inversion days within a month is about 20.

The lower panel in Figure 3.2 depicts the monthly $PM2.5$ under the study period. Intuitively, when temperature inversion in the upper panel happens frequently, the corresponding

[9] MERRA-2: the Modern-Era Retrospective analysis for Research and Applications version 2. (Data access from https://disc.gsfc.nasa.gov/datasets/M2I6NPANA_V5.12.4/summary?keywords=%22MERRA-2%22%20M2I6NPANA&start=1920-01-01&end=2017-01-16)

average $PM2.5$ of that month is also high.[10]

Inversion Strength

There are two ways to construct an instrument based on the temperature inversion: one is a binary variable that indicates the occurrence of inversion with a value of 1, otherwise being 0, which is used in Sager (2016); the other one is inversion strength - the difference between the air temperature at the ground (1000hPa) and at the higher layer (975hPa, 950hPa, and 925hPa), which is more widely used as the instrument for air pollution as in, like the work in Jans et al. (2018) and Chen et al. (2017). Since the reversal of any one of twelve grids at any time would be recorded in the binary variable with value 1, which is too rough to identify the overall severity of inversion in Beijing, inversion strength works better to quantify the average level of the whole city. Therefore, we utilize the inversion strength to construct the instrument in the study.

We follow the previous research in focusing on the temperature strength between the base layer 1000hPa and the second layer 925hPa.[11] The other two higher layers are added later for robust checks. The difference is recorded only when the temperature at the higher layer is higher than that at base layer ($Temp^{925hPa}_{grid_{it\_h}} > Temp^{1000hPa}_{grid_{it\_h}}$), otherwise the value is recorded as 0. Then we calculate average strength for those twelve grids within four 6-hour periods to construct daily average inversion strength in Beijing ($Inverstrenghth_t$).

$$Inverstrenghth_t = \sum_{h=0:00}^{18:00} \sum_{i=1}^{12} Temp^{925hPa}_{grid_{it\_h}} - Temp^{1000hPa}_{grid_{it\_h}} \tag{2}$$

Figure 3.3 displays the frequency of temperature inversion at four periods: 0:00, 6:00, 12:00 and 18:00, respectively. The horizontal axis represents the number of days temperature inversion has occurred in the past 30 days. The histogram plots the frequency of the corresponding situation indicated on the horizontal axis, with the value listed on the left

[10]The $PM2.5$ data after June in 2017 are supplemented by the data from BJMEP.

[11]The count of temperature inversion based on 925hPa is often used in previous study when instrumenting temperature inversion for air pollution, like in Sager (2016) and Jans et al. (2018).

axis. The error bar indicates the variability (both mean level and standard error) of $PM2.5$ according to each bar, with the value shown at the right axis. From panel (a) at 0:00 to panel (c) 12:00, $PM2.5$ shows a growing trend along with the increasing number of inversion episodes. At 18:00 in panel (d), $PM2.5$ fluctuates a lot without a clear trend.

3.3 Methodology

The objective of this study is to investigate the causal link between air pollution in Beijing and people's search interest in local universities via the proxy of the Baidu Index. We start from the most direct way using ordinary least squares (OLS) method for discovering the possible relationship in a linear regression model. We observe the changes in the coefficient of interest while gradually increasing temporal controls, time trend and weather covariates. Furthermore, the non-linear effect is studied by including the categorical bins for $PM2.5$, as well as using the spline function with corresponding knots. After identifying the linear model, the IV method is further supplemented so as to address the potential endogeneity problem.

3.3.1 OLS estimation

Recall that people's interest in whether to stay or leave for higher education is proxied by the proportion of search index for local universities, that is IB.[12] Moreover, search behavior could be seen as prepartion before action, which is also a good predictor of people's future behavior. If this ratio changes with the elevated $PM2.5$, then there could be the causal link from air pollution to future out-migration for education. The relation is examined by the following model:

$$IB_t = \alpha_0 + \beta P_t + \gamma W_t + \sum_{l=1}^{2} \kappa_l trend^l + \theta_t + \epsilon_t \tag{3}$$

The dependent variable IB_t is average share of the index for Beijing universities from date t to $t-29$. In the later specification, interval t varies from short-term to mid-term ranging

[12] See section 3.2.1, $IB = UB/U$.

from 10, 15, 20 to 30 days. In particular, the monthly mean under 30 days' time frame is mainly discussed in the regression.

P_t is the mean level of air pollution $PM2.5$ within the corresponding time frame. β is the primary coefficient of interest, which estimates the effect of $PM2.5$ on people's interest in Beijing-based universities. The sign is expected to be negative, which means that people reduce their attention on local universities when $PM2.5$ exposure increases. Conversely, their interest in the universities outside of Beijing is also motivated by local air pollution, with effect indicated by $1 - \beta$.

Climate change and extreme weather could also affect people's interest in migration. Weather covariates, W_t, are constructed in the same time frame, including in a quadratic form of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. Among them, all the weather variables are calculated as the mean over the past 30 days.

The time trend is included in the form of squared term, which helps to capture the increasing trend in search index due to the improved technology.[13] θ_t is a vector of temporal controls: Year by month dummies are used to control for seasonal patterns of people's search behavior for universities, especially before and after the National Higher Education Entrance Examination in June; The inclusion of categorical number of holiday days during the latest t days helps to identify the possible increase in search index due to the availability of more leisure time; ϵ_t is the error term.

Non-linear Effect — Categorical Levels

Given that people's response to air pollution may change with its degree of deterioration, non-linear effect is further analysed by converting the continuous P_t into seven groups: $25 - 50\ \mu g/m^3$, $50 - 75\ \mu g/m^3$, $75 - 100\ \mu g/m^3$, $100 - 125\ \mu g/m^3$, $125 - 150\ \mu g/m^3$, $150 -$

[13] In the unreported result, time trends in the form of higher-order are also tested and found to have little effect on the main estimate β.

175 $\mu g/m^3$ and $\geq$ 175 $\mu g/m^3$.[14] P_t bins are included in the regression with the cleanest group as the omitted base:

$$IB_t = \alpha_0 + \beta_1 P_{t(50-75)} + \beta_2 P_{t(75-100)} + ... + \beta_6 P_{t(>175)} + \gamma W_t + \sum_{l=1}^{2} \kappa_l trend^l + \theta_t + \epsilon_t \quad (4)$$

where the reported coefficients (β_1, β_2 ... β_6) measure the differences in the share of search index for Beijing universities compared with the base level. The negative relationship between P_t and IB_t could again be identified by the increasing values from β_1 to β_6 in the negative direction.

Non-linear Effect — Spline Regression

We also introduce the linear splines to evaluate the relationship between air pollution (P_t) and people's search behavior (IB_t) within different severity groups. Six knots of $50\mu g/m^3$, $75\mu g/m^3$, $100\mu g/m^3$, $125\mu g/m^3$, $150\mu g/m^3$, and $175\mu g/m^3$ are interpolated based on the same group division listed above.

$$IB_t = \alpha_0 + \beta_0 P_{t(<50)} + \beta_1 P_{t(50-75)} + \beta_2 P_{t(75-100)} + ... + \beta_6 P_{t(>175)} + \gamma W_t + \sum_{l=1}^{2} \kappa_l trend^l + \theta_t + \epsilon_t \quad (5)$$

The estimate before each group of pollution P_t (β_0, β_1, β_2 ... β_6) measures the slope of the corresponding linear segment. The linearity of the specification could be examined by comparing these β values to see whether they have the same slope throughout the entire range.

It is worth noting that the nonlinearity check is very sensitive to the sample size and outliers. To address this issue, we incorporate the frequency of individual group while plotting the non-linear effect estimates (see Panel B in Table 3.1). Besides that, the augmented component-plus-residual scatterplot that is proposed by Mallows (1986) also visually helps to detect the nonlinearity. We add both lowess smooth and median splines (with the increasing

[14]No observations are under 25 $\mu g/m^3$.

number of bands or knots from five, ten to fifteen) to diagnose the fitted regression line.

3.3.2 IV estimation

There could be an inevitable measurement error in air pollution. The big data provided by Baidu Index could only be accurate at the city level without releasing any searcher's location information, which makes it impossible for us to know the exact air pollution exposure of each relevant user. On the other hand, people could take the avoidance behavior to alleviate the adverse impact of ambient pollution, which is also unobserved to us. Therefore, neither the pollution data from one monitor in U.S. embassy nor the average from ground stations established by Chinese MEP could precisely indicate the air pollution exposure of users involved in the index. The presence of such measurement error in our explanatory variable could lead to the attenuation bias of the coefficient estimate.

We construct the search index for universities (UB and U) by the mathematical process of addition. The aggregated index is included in the regression equation without considering any individual information. It is likely that there exist some omitted variables that drive the air pollution and index together, which leads our independent variable to be endogenous. For example, the new infrastructure in the university could not only increase the air pollution in the built environment but also increase people's attention. Besides that, the economic related factor is also omitted because of its difficulty to control during the short to mid-term, which is closely associated with local air quality as well as the attention to local universities.

Therefore, in order to address the bias generated from both measurement errors and omitted variables, we supplement the OLS analysis by the IV method. As mentioned before, air pollution is instrumented by temperature inversion strength, which is computed by the difference of atmosphere temperature between base layer (1000 hPa) and higher layer (925 hPa, 950hPa and 975 hPa).

The validity of our instrument relies on its correlation with air pollution in the first stage and its independence with our search pattern in the main regression. It requires that the inversion episode is randomly assigned as a weather phenomenon, which is highly correlated with the variation in air pollution but does not directly affect people's search behaviors. However, the internal link between inversion episodes and cognitive behaviors cannot be testable. The major concern about the violation is their interaction with weather factors. Temperature inversion could affect the surrounding weather conditions, which in turn have an impact on people's cogitation. To address this issue, we have controlled the meteorological variables as many as possible. The underlying assumption is that the only channel by which the increased temperature inversion strength could reduce people's search interest in local universities is through the local ambient pollution. This independence assumption is crucial to exploit temperature inversion as the instrument to our application.

3.4 Results

In this section, we report the main results generated by the OLS and IV methods. Graphs and figures are used to help present the estimates effect. Robustness checks are added to examine how the coefficients behave when we modify time frame, data source, calculation of standard errors, time period, and top levels. We also study people's attitudes towards the universities in clean cities (INB^{clean}) and verify their motivated preference to cleaner places by local pollution from the opposite side .

3.4.1 OLS result

Table 3.3 presents estimates for the effect of $PM2.5$ on the proportion of search index for universities in Beijing (IB). As outlined above, we mainly discuss those universities within the top 100. Alternative top levels are added in the robustness check. The daily ratio is integrated into the rolling monthly mean within the latest 30 days, that is $IB_{t=30}$. The specifications in Column (1) through Column (4) add the additional controls one by one, from temporal dummies, time trend to weather covariates. The usual robust standard errors are reported in parentheses.

After the base control of year by month dummies, values are quite stable in size before incorporating weather conditions in Column (4), whose coefficient is obviously increased in the negative direction. It is apparent that weather is closely related to both $PM2.5$ and people's migration intentions and thus is an important factor in determining the estimate.

All the estimates shown in the table are negative and significant at 1% level, which indicates that the search for universities in Beijing decreases due to the elevated air pollution, all other factors held constant. It also indicates that local education becomes less attractive to residents in Beijing because of the local poor ambient quality. Each estimate we list in the table represents the percentage change in the share of search index for local universities due to one unit increase in $PM2.5$ in the mid-term frame of 30 days. When $PM2.5$ increases by 100 units, for example when air quality moves from good days (51-100 $\mu g/m^3$) to moderately polluted days (151-200 $\mu g/m^3$), the estimate decreases by 1.8 point. Accordingly, the interest in local education will also decrease by 3.8% relative to the mean level.[15]

Identify the nonliear effect — Categorical Bins

In order to test that whether our original linear model fits the data well, we re-estimate the regression based on seven bins of different $PM2.5$ levels listed in Table 3.4. As shown in Column (2), compared to the cleanest group of $PM2.5 < 50$ $\mu g/m^3$ (no data less than 25 $\mu g/m^3$), people decrease their search for local universities at dirtier group $50-75$ $\mu g/m^3$, even less at $75-100$ $\mu g/m^3$, $100-125$ $\mu g/m^3$, and $125-150$ $\mu g/m^3$; rebounds a bit at $150-175$ $\mu g/m^3$; then decrease again when $PM2.5$ over 175 $\mu g/m^3$. We plot the points estimates and the 95% confidence interval in Panel B of Figure 3.5. Since the reported estimates are very sensitive to the sample size within group, we also add the percentage distribution of $PM2.5$ in upper Panel A.

The graph in Panel B displays a clear linear trend when the observations of $PM2.5$ is less than 150 $\mu g/m^3$. Although the trend line starts to deviate significantly in the group of

[15] The reduced portion due to air pollution accounts for 3.8% of the mean proportion of search for universities in Beijing, that is $1.8/47.57 = 3.8\%$.

$150 - 175\ \mu g/m^3$, it returns to linear decreases after $175\ \mu g/m^3$. The only deviation does not overturn the overall linear trend, because only a small number of samples are in that bin (2.7% of the sample).[16]

Identify the nonliear effect — Linear Splines

Column (3) in Table 3.4 conducts a spline regression whose estimates representing the slope within each group. All the coefficients are negative, except for the group of $150 - 175\ \mu g/m^3$, which deviates in the same pattern as shown in Column (2). The magnitude of the estimates fluctuates among different groups of $PM2.5$, and the significance is lost in two of them. We can not tell the linear trend by comparing the values with that in Column (1). A further investigation for splines could be studied with the assistance of graph.

Figure 3.6 demonstrates the augmented component-plus-residual versus our main independent variable $PM2.5$. A clearer downward trend could be visually identified. The red line plots the linear regression and indicates the slope equal to the coefficient under Column (1). The yellow lines attempt to fit the scatter plot in various ways: (a) adds a lowess smoothing of the plotted points; (b) (c) and (d) add median splines with increasing numbers of bands from five, ten to fifteen units. Although the yellow line has a large fluctuation around the red line at the tail of the heavily polluted days over $150\ \mu g/m^3$, they do not invalidate the entire linear trend because they only accounts for small number of the our sample (3.9%).

3.4.2 IV Results

To address the potential problem from measurement error and endogeneity under OLS method, we implement the regression analysis again in the way of IV. $PM2.5$ in Beijing is instrumented by local temperature inversion strength, which could be computed by the difference in air temperature between the higher layer and the base layer (1000 hPa). In Table 3.5, three higher layers (925 hPa, 950 hPa, and 975 hPa) are introduced to construct the

[16]See Panel B in Table 3.1, the number of the observations within $150 - 175\ \mu g/m^3$ group is 65, accounting for 2.7% of the entire sample.

instruments. All the estimates in the first stage are statistically significant with Kleibergen-Paap Wald rk F statistics at around 200 in value, much larger than the Stock-Yogo weak ID test critical value of 16.38, suggesting that inversion strength is a strong instrument for $PM2.5$. We select the one constructed via 925 hPa, since this is most common layer used in previous research.

The second stage regresses the proportion of search for universities in Beijing on the predicted $PM2.5$ obtained above. The preferred outcome shown in Column (1) is significant at the 1% level, with magnitude two times larger than that under OLS. Accordingly, people's interest in local universities will decrease by 11.8% when $PM2.5$ increases by 100 units.[17]

3.4.3 Reduced form

We re-estimate the OLS regression by replacing the air pollution variable on the right-hand side with our instrument—temperature inversion strength. In Table 3.6, all the inversion strength estimates display significantly negative relationships with the search for local education. From Figure 3.4, the bottom panel (b) depicts the reduced form without any controls, which also clearly shows a decreasing trend. Based on our assumption, the inversion only affects people's search interest through its exogenous indication for air pollution. Thus, the underlying channel is that temperature inversion episodes cause the elevated air pollution as shown on the upper panel (a), and then the dirty air in Beijing decreases people's interest in local higher education.

3.4.4 Robustness check

In this section, we modify the various conditions to examine how the coefficient estimates behave, like altering the time frame, replacing the data source, calculating the standard errors in alternative ways, considering the different time periods, and expanding the top levels.

[17] The value is computed by the increased pollution $(100 * 0.056)$ over the mean level (47.57).

Alternative time frame

Until now, all the variables involved above are converted to the mean within the latest 30 days. We re-construct the variables with alternative time frames from the short-term (like 10 days) to relatively longer periods like 15 days and 20 days. Both OLS and IV results are presented in Table 3.7. Column (4) replicates the preferred OLS outcome in Column (4) of Table 3.3 and the estimate under IV in Column (1) of Table 3.5.

Intuitively, when time spans are extended progressively from Column (1) to Column (4), the coefficients are shown to increase for longer terms, but the exact effect size should also take into consideration the individual distribution of $PM2.5$ under each interval, as well as the corresponding search index for that period.[18] As for the upper panel OLS estimates, all the estimates are negative, with two of them significant at the 1% level.

For the IV method, the instruments in the first stage are still strong enough with F-statistics all over 100 in value; they are higher in value for shorter periods. The estimates in the second stage are statistically significant at the 1% level except for the 10 day's one at the 5% level, displaying a plausibly increasing trend. When air quality in Beijing moves from good-day level to moderately-polluted level, that is when $PM2.5$ increases by 100 units, the search for local education decreases by 3.4%, 7.4%, 7.1%, and 11.8% respectively.[19] It seems that people lose more interest in local education when they are exposed to air pollution in the longer term.

Alternative $PM2.5$ data

Compared with the $PM2.5$ data from the Chinese MEP accessible since 2014, the one from U.S. embassy is mainly used in the study because of its longer time span starting from January in 2011 to June in 2017. We re-estimate the regressions based on the alternative data over various periods. Column (1) in Table 3.8 reports the preferred results from both

[18] The mean search index is quite stable during various terms (See Panel A in Table 3.1).

[19] The effect size for 15 days is a bit larger than that for that under 20 days. There is not much difference between them.

OLS and IV using the data from the U.S. embassy. Column (4) replaces the data source and is based on the same specification as Column (1) during using shorter period.

Column (2) and Column (3) present the estimates obtained under the intersection span (January in 2014 to June in 2017) from the U.S. embassy and Chinese MEP, respectively. Both of the outcomes are similar in magnitude with the same significance level of 1% under both OLS and IV methods. A 100 units increase in air pollution leads to a decrease in the search index for Beijing universities of 2.5% under OLS and of 26.7% under IV if we use U.S. embassy data. For data from the Chinese MEP, it is 3.6% under OLS and 28.2% under IV. There is not much difference between the estimation drawn from the two data sources. Therefore, the results remain robust with alternative $PM2.5$ data.

Comparing the effect sizes between Column (1) and Column (2), which both adopt the same data source but based on the different time spans, the latter one from 2014 has the larger effect than that from 2011, suggesting that the impact of air pollution on the migration intentions for education gets stronger, and people's tolerance for dirty air declines over time.

Alternative standard errors

In the main study, robust standard errors are reported to indicate the statistical inference of the estimates (see Column (1) in Table 3.9). There might be concerns about the internal correlation among errors. The search index residual is likely to be correlated within a short period since people's same-intentioned search behaviors usually last for a few days. Although such correlation is diluted after we convert the daily data into a 30-day mean, we still re-estimate the standard errors clustered at year by month and year by week level as robustness checks. After introducing the year by month in Column (2), the standard errors clearly increase, leading to the OLS estimates significant at 10%, while for IV at 5%. The significance levels shown in Column (3) are little disturbed after we account for the cluster-robust standard errors at year by week.

Another concern is autocorrelation that often happens in time series data, in which the errors might be correlated over time. We introduce the Newey-West standard errors to improve the statistical inference reported (lags are chosen to be the number of monthly observations minus one, that is 29). As shown in Column (4), the standard errors increase a bit compared to those in Column (1), but the estimated coefficients retain their significance levels at 1% under IV.

Different top of tiers universities

When we construct the dependent variable which reflects Beijing citizens' intention to pursue higher education at a local place, we select the top 100 universities as the choice set. According to the university ranking information displayed in Table 3.3, most of the universities in Beijing stay at the top levels. People pay more attention to both inside and outside education within the top 100 which have largest search index values in Column (2) and Column (5). Now we relax the ranking range from top 200 to 500 and re-estimate the work to study people's migration interests when they have more choice with respect to outside education.

In Table 3.10, all of the estimates are statistically significant at the 1% level. The magnitudes are visually shown as an increasing trend from Column (1) to Column (5). After considering the individual means under different levels listed in the bottom line, from left to right, when air pollution increases by 100 units, the search index in Beijing universities increases by 3.8%, 6.1%, 7.3%, 7.8%, and 8.4% under OLS, and 11.8%, 19%, 22.8%, 26%, and 28.4% under IV. All the estimates are substantial in magnitude and significantly negative.

3.4.5 Different time period

The causal link from air pollution to people's search behavior has been tested above in the same time frame, which is the past 30 days. It could also be interesting to test whether the air pollution lagged one period and led one period would affect people's choices. In Table 3.11, Column (1) replaces the independent variables (including both $PM2.5$ and weather covariates) with the ones lagged last period without overlap, which is the monthly average 30 days ago. The estimates are negative with smaller size than those in Column (2), significant at the 5% for OLS and at the 1% level for IV, indicating that the value of $PM2.5$ in the last period also has an effect on people's present search interest in local education. Column (3) replaces the independent variables (including both $PM2.5$ and weather covariates) with the ones in lead period without overlap that is monthly average 30 days ahead. Both of the estimates under OLS and IV are insignificant, suggesting that people are not clearly affected by $PM2.5$ in the forecast.

3.4.6 Heterogeneous effects among different clean levels

We have identified the causal relationship between air pollution and the decrease of interest in local education. It can be inferred that when local air quality deteriorates, people in Beijing are more likely to choose those substitutable universities in cleaner cities. Therefore, we analyse people's search behavior on the opposite side by looking at the search index for those universities outside of Beijing, especially the ones in cleaner cities. The regression is implemented in the same way of OLS and IV discussed above :

$$Ln(INB_t^{clean}) = \widetilde{\alpha}_0 + \widetilde{\beta} P_t + \widetilde{\gamma} W_t + \sum_{l=1}^{2} \widetilde{\kappa}_l trend^l + \widetilde{\theta}_t + \widetilde{\epsilon}_t \tag{6}$$

where dependent variable $Ln(INB_t^{clean})$ is log form of the proportion of search index for non-Beijing universities in clean cities. The parameter of interest, $\widetilde{\beta}$ (actually $\widetilde{\beta}$*100), represents the percentage change in INB_t^{clean} due to each unit increase in $PM2.5$, that helps to discover whether the elevated $PM2.5$ would stimulate Beijing citizens' interest in seeking higher education in cleaner cities.

Estimates are reported in Table 3.12 in the percentage form by multiplying 100 to each original value. Column (1) through Column (4) construct different version of INB_t^{clean} based on different clean levels where the non-Beijing universities are located. After comparing each city's mean $PM2.5$ within our time span, "<Beijing" includes all the searches for those universities located in cleaner cities than Beijing. "<Beijing-10" includes on those cities 10 $\mu g/m^3$ cleaner than Beijing. "<Beijing-20" and "<Beijing-30" expand the gap to 20 $\mu g/m^3$ and 30 $\mu g/m^3$, respectively. From both OLS and 2SLS panels, it is obvious that the estimates increase for those universities located in cleaner cities, suggesting that the decreased ambient quality drives local people's intention to study in cleaner places.

3.5 Conclusion

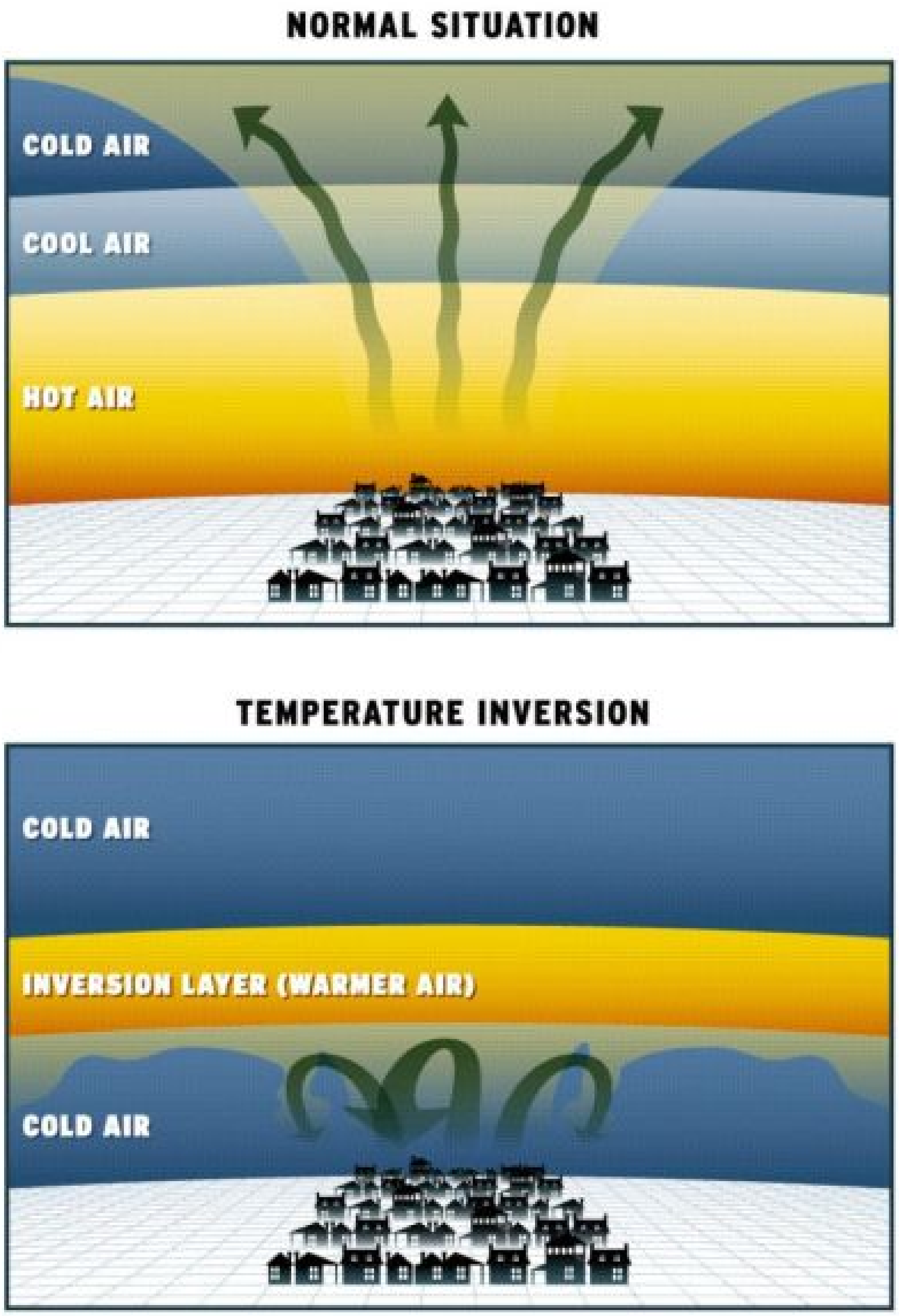
This paper explores the causal link between air pollution and people's migration intention for out-of-city education. It is challenging to analyse the hidden effect by directly considering the yearly migration ratio data, since the actual migration process is laborious, time-consuming and also costly, which leads to a slow response to air pollution exposure. However, students are freer to migrate when they choose the universities to continue their education after they complete the high school, and their choice can be more sensitive to the local environment. Our paper adds to the nascent literature on how environmental factors might motivate intentions to migrate and focuses on whether and how air quality affects people's interest in local education.

People's intention to study locally is proxied for by their search index for local universities. We first adopt the OLS method by regressing the proportion of search index for universities in Beijing on measure of local air pollution. Then we incorporate the temperature inversion strength to instrument for the exogenous variation in air pollution, supplementing the study by IV method so as to mitigate the threat from measurement errors in air pollution exposure and possible endogenous issues.

Both of the OLS and IV results demonstrate the clear negative effect of air pollution on people's search for local education. When air quality in Beijing moves from good-day level to moderately-polluted level, people's search volume in local universities declines by 3.8% under OLS and by 11.8% under IV. Moreover, there is also a corresponding increase in searches for universities outside of Beijing, with the effect particularly pronounced for universities in cleaner cities.

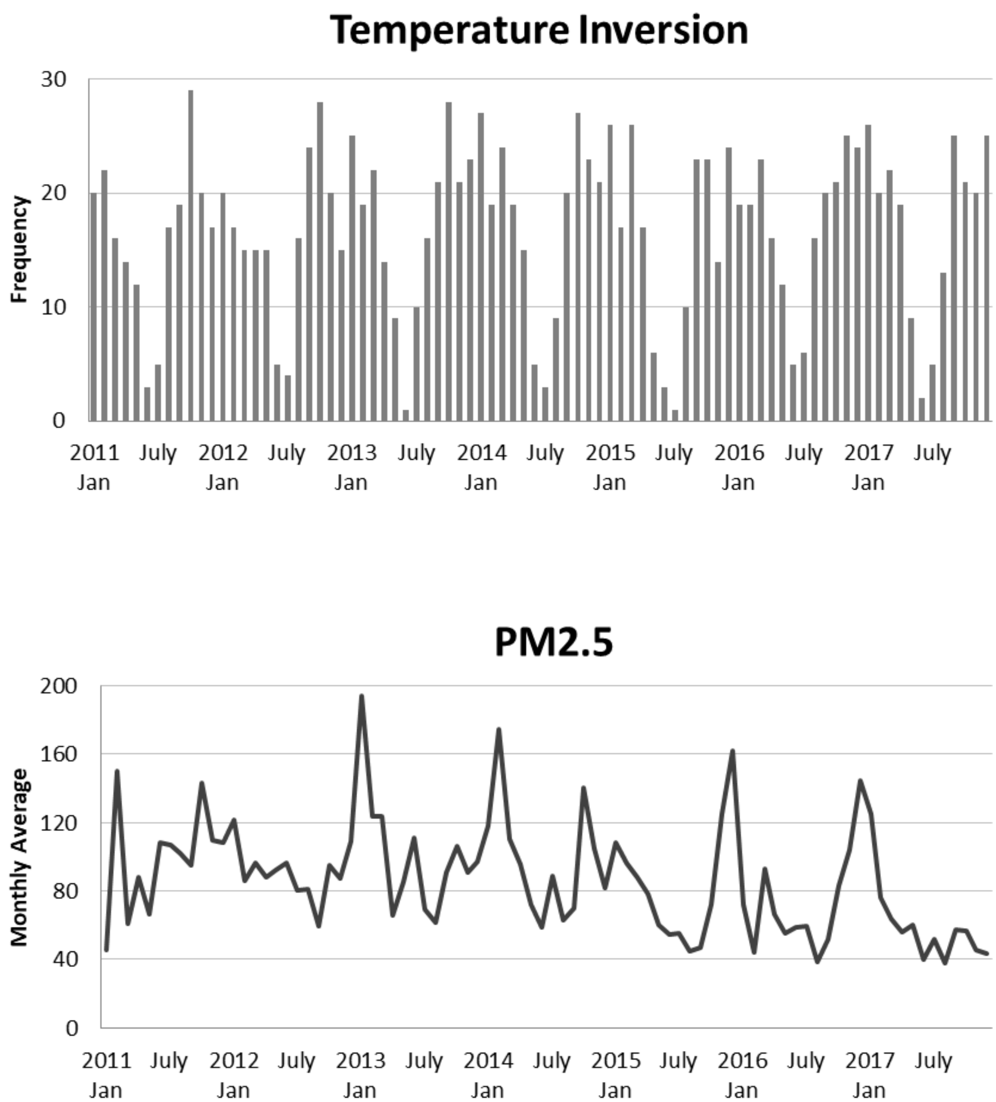
The decline in search activity for local education due to the elevated ambient pollution releases the signal that people are less interested in local education because of the dirty air in the city. Although Beijing gathers most of the best universities in China, it seems to become less attractive due to the serious air pollution problem. Local students are more likely to travel far away from their home place to cleaner cities for higher education.

Figure 3.1: Temperature Inversion



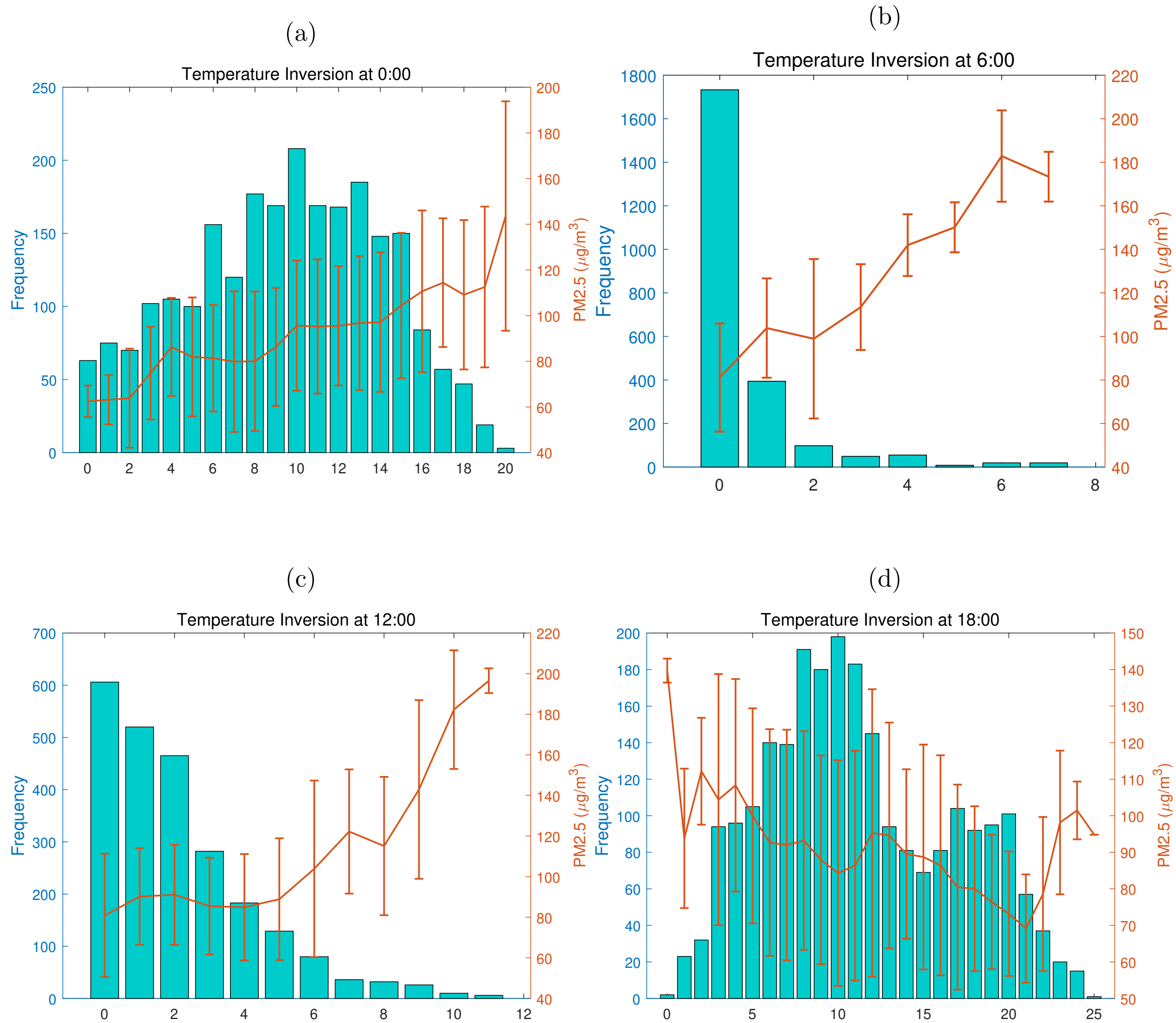
Source: Winter Smog in Canada.ca
(https://www.canada.ca/en/environment-climate-change/services/air-pollution/issues/smog-causes-effects/winter.html)

Figure 3.2: Monthly Distribution of Temperature Inversion and $PM2.5$



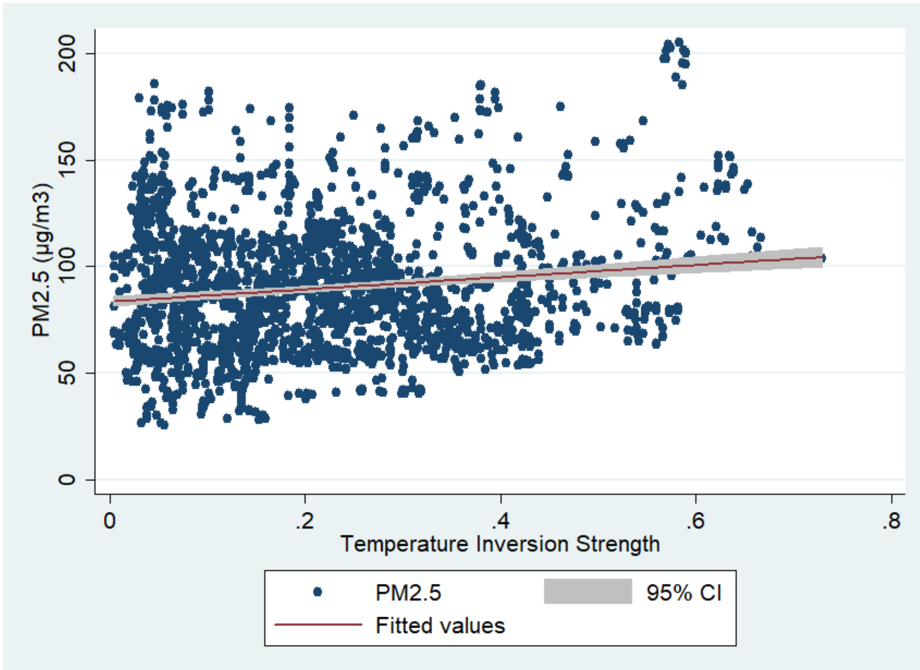
Note: The upper panel shows the frequency of monthly temperature inversions from 2011 to 2017. Correspondingly, the lower panel depicts the distribution of the monthly average $PM2.5$ ($\mu g/m^3$) within the same period (those missing historical $PM2.5$ data on U.S. embassy website after Jun 2017 are supplemented by those from Chinese MEP).

Figure 3.3: Frequency of Temperature Inversion and $PM_{2.5}$ at Four Time Periods

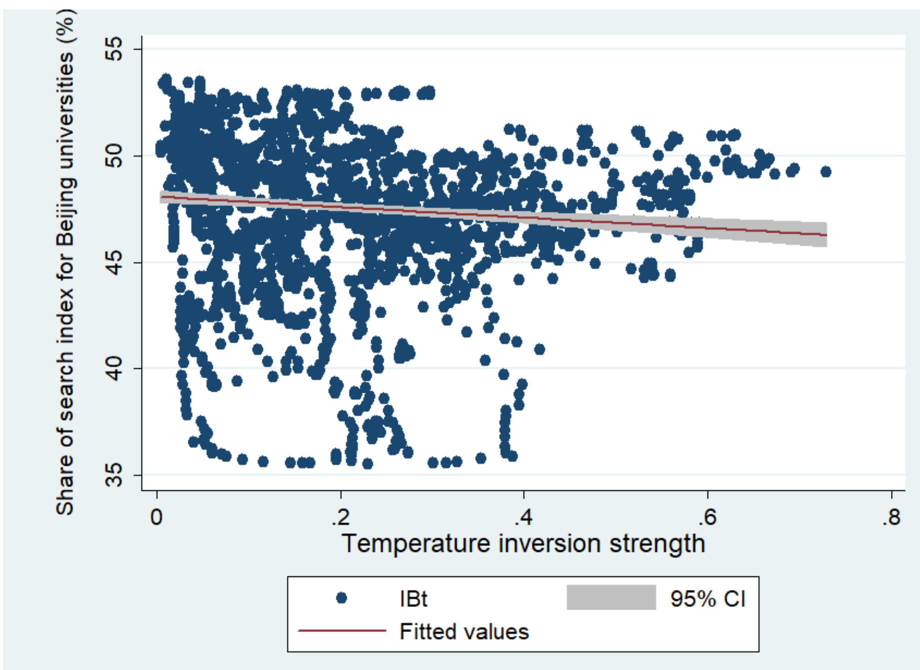


Notes: This diagram displays the whole frequency of temperature inversion at four time periods: 0:00, 6:00, 12:00 and 18:00, respectively. All the variables involved in this diagram are counted as the basis of the rolling 30 days. The horizontal axis represents the number of days temperature inversion has occurred in the past 30 days. The histogram plots the distribution of occurrence of the corresponding abscissa, with frequency listed at the left axis. The error bar indicates the variability (both mean level and standard error) of PM2.5 according to each bar, with the value shown at the right axis.

Figure 3.4: Instrumental Variable — Temperature Inversion Strength



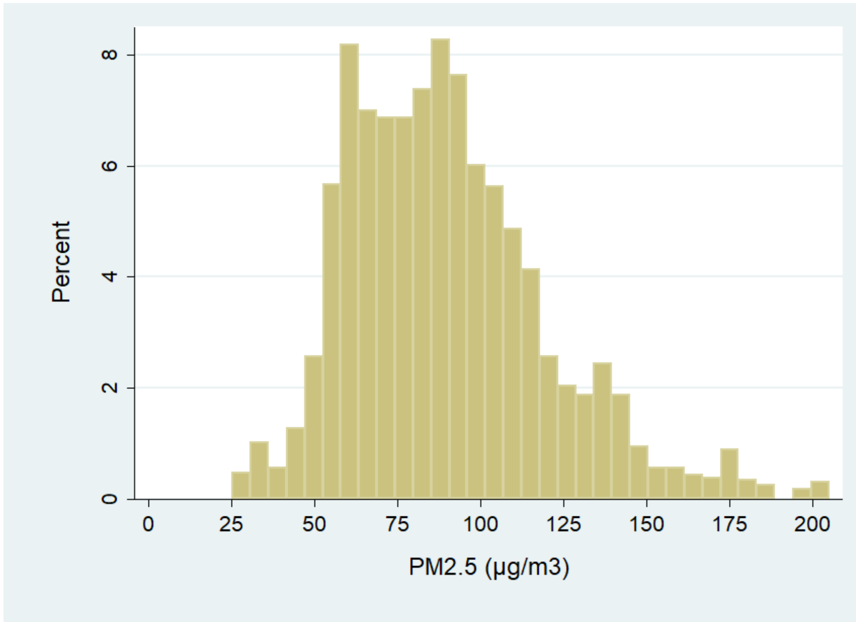
Panel (a)



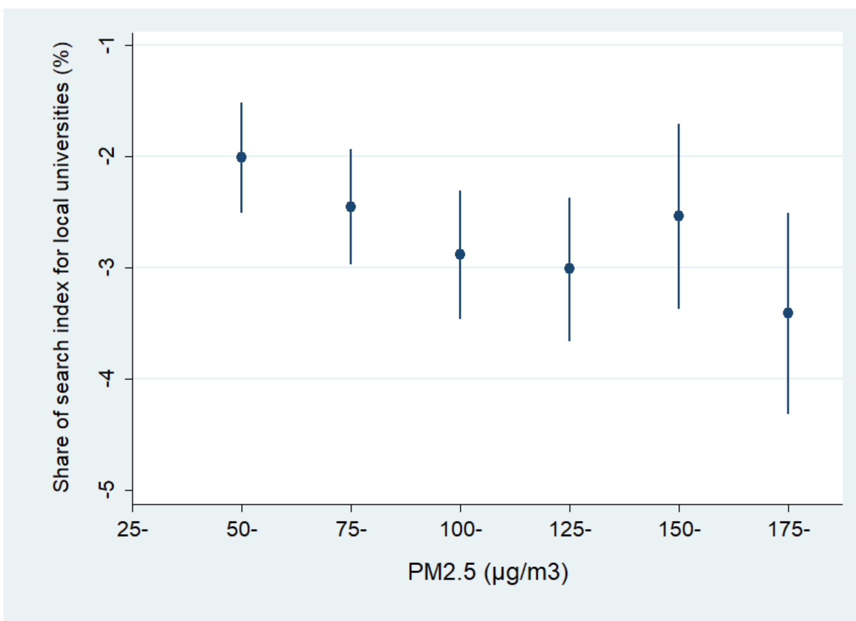
Panel (b)

Note: The upper scatter diagram (Panel (a)) graphs the relationship between inversion strength and $PM2.5$. The lower figure plots the reduced form of inversion strength and the share of the search index for Beijing Universities (IB_t).

Figure 3.5: Non-linear Effect of $PM2.5$



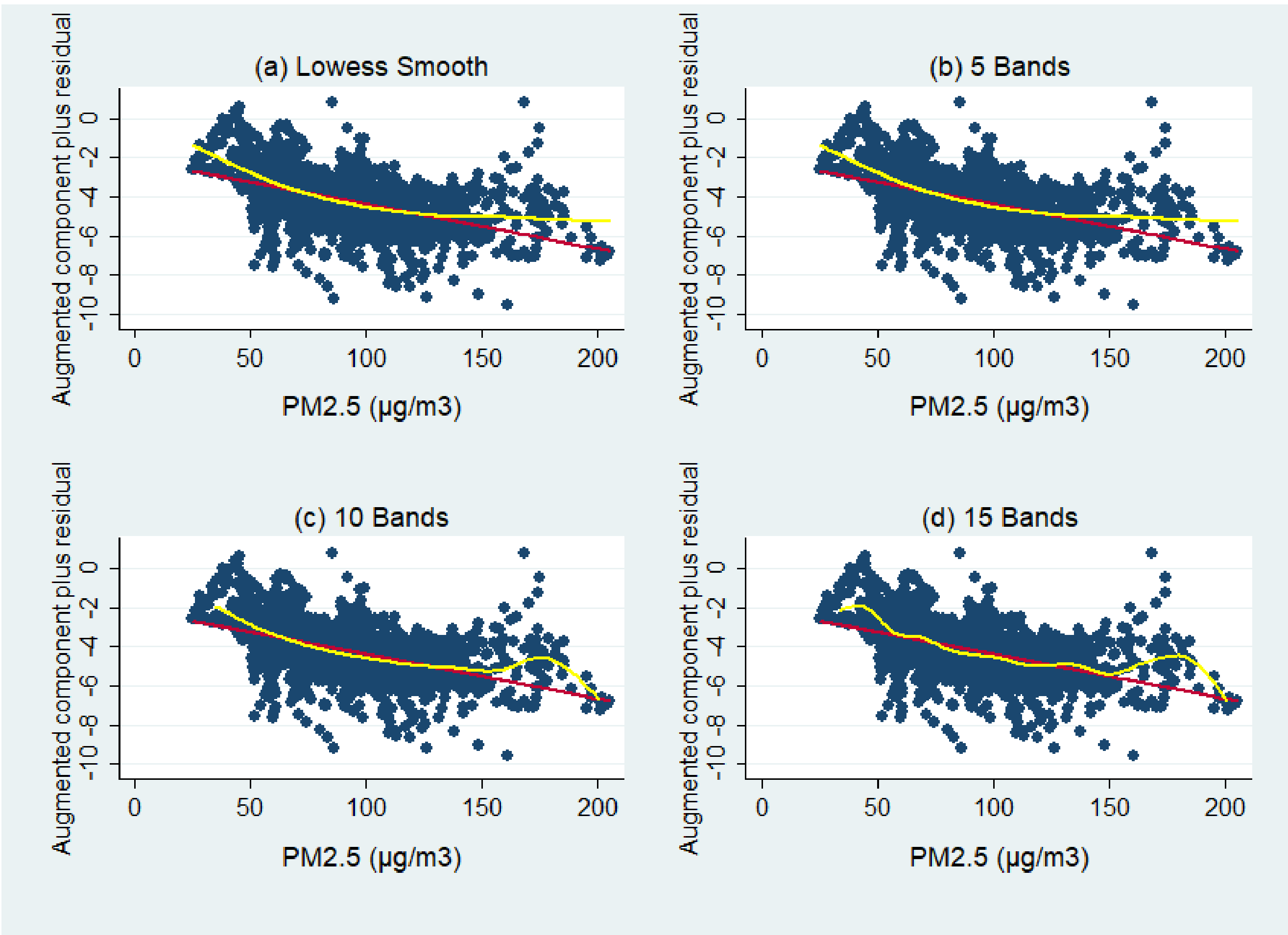
(a)



(b)

Note: Graph (a) displays the percentage distribution of $PM2.5$ (rolling monthly average). Graph (b) plots the point estimates and 95% confidence intervals shown in Column (2) of Table 3.4, indicating the non-linear effect of $PM2.5$. Each little bar represents the change in the share of search index for local universities in each bin compared to that in the base group 25 to 50 $\mu g/m^3$.

Figure 3.6: Nonlinearity — Augmented Component-plus-residual Plot



Note: This diagram displays four augmented component-plus-residual plots for identifying the nonlinearity between $PM2.5$ and the proportion change in IB_t. The red fitted line has the slope equal to the coefficient of that obtained from the original regression (shown in Column (1) of Table 3.4). The yellow line helps to assist in detecting the nonlinear relationship: Lowess smooth of the plotted points are added in (a); Median splines with different band options are added in (b), (c) and (d).

Table 3.1: Summary Statistics

Variable	Obs (1)	Mean (2)	Std. Dev (3)	Min. (4)	Max. (5)
Panel A: University Searches					
Beijing Univ/Total Uni - Past 30 Days' Mean					
IB_t (Top 100)	2344	47.566	3.577	35.530	53.584
IB_t (Top 200)	2344	36.342	3.564	24.961	42.725
IB_t (Top 300)	2344	30.223	3.329	19.800	36.309
IB_t (Top 400)	2344	29.633	3.244	19.340	35.427
IB_t (Top 500)	2344	27.433	2.838	17.938	32.483
IB_t (Top 100) - Various Time Frames					
Past 10 Days' Mean	2364	47.567	4.158	31.255	54.759
Past 15 Days' Mean	2359	47.568	3.981	32.479	54.260
Past 20 Days' Mean	2354	47.567	3.831	33.705	54.043
Panel B: Air Pollution $PM2.5$ ($\mu g/m^3$)					
Pollution ($PM2.5$)- Various Time Frames					
Past 10 Days' Mean	2364	89.906	41.416	13.233	286.554
Past 15 Days' Mean	2359	89.965	36.446	17.964	262.159
Past 20 Days' Mean	2354	90.033	33.562	21.158	246.420
Past 30 Days' Mean	2344	90.083	29.924	25.340	205.026
Pollution Bins ($PM2.5$)- Past 30 Days' Mean					
$PM2.5_{\_30} \in [25, 50)$	103	40.826	6.959	25.340	49.990
$PM2.5_{\_30} \in [50, 75)$	707	63.501	6.556	50.039	74.991
$PM2.5_{\_30} \in [75, 100)$	795	87.509	6.897	75.012	99.875
$PM2.5_{\_30} \in [100, 125)$	451	110.647	6.558	100.028	124.990
$PM2.5_{\_30} \in [125, 150)$	197	136.576	6.418	125.207	148.761
$PM2.5_{\_30} \in [150, 175)$	65	164.073	8.281	150.522	174.944
$PM2.5_{\_30} \geq 175$	27	189.010	10.019	175.332	205.026

Continued

Variable	Obs (1)	Mean (2)	Std. Dev. (3)	Min. (4)	Max. (5)
Panel B: Air Pollution $PM2.5$ ($\mu g/m^3$)					
$PM2.5_{\_30}$ - U.S. Embassy and Ground Data					
$PM2.5_{\_30}$ (U.S. Embassy 2014.01-2017.06)	1248	83.933	31.617	25.340	185.484
$PM2.5_{\_30}$ (Ground Data 2014.01-2017.06)	1248	77.743	26.532	24.8	172.767
$PM2.5_{\_30}$ (Ground Data 2014.01-2017.12)	1432	74.003	26.747	24.8	172.767
Panel C: Temperature Inversion (Base 1000 hPa)					
Past 30 Days' Mean Inversion Strength - Various Layers					
925 hPa_30	2344	0.209	0.147	0.005	0.730
950 hPa_30	2344	0.302	0.170	0.009	0.860
975 hPa_30	2344	0.351	0.155	0.030	0.780
Inversion Strength (925 hPa) - Various Time Frames					
Past 10 Days' Mean	2365	0.210	0.213	0	1.393
Past 15 Days' Mean	2359	0.209	0.185	0	1.084
Past 20 Days' Mean	2354	0.200	0.169	0	0.886
Panel D: Weather Covariates					
Past 30 Days' Mean Weather Variables					
Max Temperature ($^{\circ}C$)	2344	18.160	10.856	-1.840	32.927
Min Temperature ($^{\circ}C$)	2344	8.276	10.645	-9.797	24.363
Avg Humidity (%)	2344	50.778	13.104	23.510	77.079
Wind Speed (km/h)	2344	8.017	1.267	5.467	11.433
Cloud Coverage (-/8)	2344	4.796	1.197	1.3	7.2
Sea-level Pressure (hPa)	2344	1016.919	9.129	1001.193	1032.623
Precipitation (mm)	2344	1.9536	4.1186	0	32.1767

Notes: The table displays mean statistics for all variables included in the regressions which are grouped under each panel. The sample interval spans from January in 2011 to June in 2017.

Table 3.2: Comparison between Beijing Universities and Non-Beijing Universities under Different University Levels

University Level	Beijing Universities (1)	Search Index (2)	Cumulative Percentage (%) (3)	Non-Beijing Universities (4)	Search Index (5)	Cumulative Percentage (%) (6)	Mean Ratio (Beijing/ Non-Beijing) (7)
Top 100	21	22133.740	65.47%	79	24968.590	22.93%	46.99/53.01
100 to 200	3	2390.646	72.54%	98	19173.100	40.53%	11.09/88.91
200 to 300	1	1020.947	75.56%	98	16395.450	55.59%	5.86/94.14
300 to 400	7	4543.415	89.00%	95	12365.190	66.94%	26.87/73.13
400 to 500	5	3112.085	98.21%	98	16261.190	81.87%	16.06/83.94
500 to 600	1	288.893	99.06%	94	10326.810	91.35%	2.72/97.28
600 to 700	1	317.112	100.00%	98	9418.713	100.00%	3.26/96.74

Notes: The table presents the information of both Beijing and non-Beijing universities according to the rank report in 2014. Column (1) and Column (4) list the number of Beijing universities and non-Beijing universities under each level. Column (2) and Column (5) report the daily average search index for universities within each rank. Cumulative percentage for search index are displays in Column (3) and Column (6). Column (7) calculates the ratio of daily mean search index for Beijing universities to non-universities, that is Column (2) over Column (5).

Table 3.3: The Effect of $PM2.5$ on Search Interest in Local Education — OLS

OLS	Beijing Universities Relative to Top 100 (%) (IB_t)			
	(1)	(2)	(3)	(4)
$PM2.5_t$	-0.009***	-0.010***	-0.010***	-0.018***
	(0.003)	(0.002)	(0.002)	(0.003)
Additional Controls				
Year_month	Y	Y	Y	Y
Holiday	N	Y	Y	Y
Time Trend	N	N	Y	Y
Weather Controls	N	N	N	Y
Observations	2344	2344	2344	2344

Notes: This table lists the OLS results under additional controls from Column (1) to Column (4). Dependent variable is the share of search index for universities in Beijing. The daily index ratio is converted into monthly mean during the past 30 days (from t to t-29). In particular, only the top 100 universities are studied in the main regression. Independent variable is the linear form of $PM2.5$ computed in the same time frame ($PM2.5_{30}$). Temporal controls and the second order of the time trend are added one-by-one from Column (1) to Column (3). Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.4: Non-linear Effect of $PM2.5$ on Search Interest in Local Education

OLS	Beijing Universities Relative to Top 100 (%) (IB_t)		
	(1) Linear	(2) Bins	(3) Spline Function
$PM2.5_t$ ($\mu g/m^3$)	-0.018***		
	(0.003)		
$PM2.5_t \in [25, 50)$		-	-0.108***
		-	(0.019)
$PM2.5_t \in [50, 75)$		-2.007***	-0.060***
		(0.248)	(0.008)
$PM2.5_t \in [75, 100)$		-2.452***	-0.008
		(0.262)	(0.006)
$PM2.5_t \in [100, 125)$		-2.880***	-0.025***
		(0.291)	(0.007)
$PM2.5_t \in [125, 150)$		-3.011***	-0.002
		(0.328)	(0.009)
$PM2.5_t \in [150, 175)$		-2.535***	0.049***
		(0.422)	(0.015)
$PM2.5_t \geq 175$		-3.409***	-0.142***
		(0.459)	(0.019)
Additional Controls			
Year_month	Y	Y	Y
Holiday	Y	Y	Y
Time Trend	Y	Y	Y
Weather Controls	Y	Y	Y
Observations	2344	2344	2344

Notes: This table reports the linear and non-linear effects of $PM2.5$ on the share of the search index for local universities (IB_t). Column (1) replicates the preferred linear results with full controls as shown in Column (4) of Table 3.3. Non-linear regressions are estimated based on seven bins of $PM2.5$ with 25 $\mu g/m^3$ for each bin. Column (2) involves seven dummies with $PM2.5$ less than 50 $\mu g/m^3$ as the omitted category. Column (3) displays the spline function with seven bands to fit the variability of data. Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. All the variables involved take in the form of past 30 days' mean. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).
3

Table 3.5: The Effect of $PM2.5$ on the Interest in Local Education — IV

2SLS	Beijing Universities Relative to Top 100 (%) (IB_t) (1) Preferred	(2)	(3)
First Stage			
Inversion Strength	47.913***	46.239***	63.757***
	(3.653)	(3.488)	(4.145)
Kleibergen-Paap Wald rk F Statistic	173.114	175.704	234.561
Stock-Yogo weak ID test critical values: (10% maximal IV size)		16.38	
Second Stage			
$PM2.5_t$	-0.056***	-0.063***	-0.053***
	(0.009)	(0.009)	(0.009)
Additional Controls			
Year_month	Y	Y	Y
Holiday	Y	Y	Y
Time Trend	Y	Y	Y
Weather Controls	Y	Y	Y
Layer (base 1000 hPa)	925 hPa	950 hPa	975 hPa
Observations	2344	2344	2344

Notes: This table reports the effect of $PM2.5$ on Beijing citizens' search interest in their local post-secondary education under the IV method. $PM2.5$ is instrumented by temperature inversion strength, calculated by the difference in temperature between base layer at 1000 hPa and the higher layer. Column (1) through Column (3) present the IV results based on varying instruments with decreasing layers from 925hPa, 950hPa to 975hPa. Kleibergen-Paap Wald rk F statistics and critical values are displayed to identify the strong instruments. Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. All the variables involved take in the form of past 30 days' mean. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.6: Reduced Form

Reduced Form	Beijing Universities at Top 100 (%) (IB_t)		
	(1)	(2)	(3)
Inversion Strength	-2.593***	-2.927***	-3.397***
	(0.410)	(0.410)	(0.544)
Additional Controls			
Year_month	Y	Y	Y
Holiday	Y	Y	Y
Time Trend	Y	Y	Y
Weather Controls	Y	Y	Y
Layer (base 1000 hPa)	925 hPa	950 hPa	975 hPa
Observations	2344	2344	2344

Notes: This table presents the reduced form results by regressing the share of search index for local universities on temperature inversion strength. Column (1) through Column (3) display the estimates based on varying instruments with decreasing layers from 925hPa, 950hPa to 975hPa. Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. All the variables involved take in the form of past 30 days' mean. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.7: Robustness Check — Alternative Time Frame

Alternative Time Frame	Beijing Universities Relative to Top 100 (%) (IB_t)			
	(1) 10 Days' Mean	(2) 15 Days' Mean	(3) 20 Days' Mean	(4) 30 Days' Mean
OLS				
$PM2.5_t$	-0.003	-0.005***	-0.003	-0.018***
	(0.002)	(0.002)	(0.002)	(0.003)
2SLS				
First Stage				
Inversion Strength	45.351***	32.038***	40.118***	47.913***
	(3.071)	(2.329)	(3.699)	(3.653)
Kleibergen-Paap Wald rk F Statistic	208.631	193.128	119.029	117.114
Stock-Yogo weak ID test critical values: (10% maximal IV size)		16.38		
Second Stage				
$PM2.5_t$	-0.016**	-0.035***	-0.034***	-0.056***
	(0.007)	(0.007)	(0.008)	(0.009)
Additional Controls				
Temporal Controls	Y	Y	Y	Y
Time Trend	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y
Observations	2364	2359	2354	2344

Notes: This table presents both OLS results and IV results under different time frames, whose regression processes are estimated the same as those in Column (5) of Table 3.3 and Column (1) of Table 3.5, respectively. From Column (1) to Column (4), time frames vary from short-term to mid-term, ranging from 10 days, 15 days, 20 days to 30 days. Variables involved are converted into means under different intervals. Temporal controls include the dummies for month of year and holiday. Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.8: Robustness Check — Alternative $PM2.5$ Data

Alternative $PM2.5$ Data	Beijing Universities Relative to Top 100 (%) (IB_t)			
	(1)	(2)	(3)	(4)
OLS				
$PM2.5_t$	-0.018***	-0.012***	-0.017***	-0.018***
	(0.003)	(0.005)	(0.005)	(0.005)
2SLS				
First Stage				
Inversion Strength	47.913***	45.563***	42.182***	43.444***
	(3.653)	(5.380)	(4.992)	(4.974)
Kleibergen-Paap Wald rk F Statistic	117.114	68.781	68.828	74.068
Stock-Yogo weak ID test critical values: (10% maximal IV size)		16.38		
Second Stage				
$PM2.5_t$	-0.056***	-0.127***	-0.134***	-0.118***
	(0.009)	(0.018)	(0.020)	(0.018)
Additional Controls				
Temporal Controls	Y	Y	Y	Y
Time Trend	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y
PM2.5 Data	US Embassy (2011.01-2017.06)	US Embassy (2014.01-2017.06)	Chinese MEP (2014.01-2017.06)	Chinese MEP (2014.01-2017.12)
Observations	2344	1248	1248	1432

Notes: $PM2.5$ data from U.S. embassy are used in our main work because of its longer available time span, that is from January in 2011 to June in 2017. Column (1) presents the preferred results for both OLS and IV using data from U.S. embassy. Column (2) and Column (3) report the estimates obtained under the intersection span from the U.S. embassy and Chinese MEP, respectively. Column (4) presents the outcomes based on the full four years' data (2014–2017) from Chinese MEP. Temporal controls include the dummies for month of year and holiday. Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.9: Robustness Check — Alternative Standard Error

Alternative Standard Error	Beijing Universities Relative to Top 100 (%) (IB_t) (1)	(2)	(3)	(4)
OLS				
$PM2.5_t$	-0.018***	-0.018*	-0.018**	-0.018**
	(0.003)	(0.01)	(0.006)	(0.008)
2SLS				
First Stage				
Inversion Strength	47.913***	47.913***	47.913***	47.913***
	(3.653)	(11.747)	(7.566)	(10.724)
Kleibergen-Paap Wald rk F Statistic	117.114	16.797	40.381	19.981
Stock-Yogo weak ID test critical values: 16.38 (10% maximal IV size)				
Second Stage				
$PM2.5_t$	-0.056***	-0.056**	-0.056***	-0.056***
	(0.009)	(0.025)	(0.017)	(0.021)
Additional Controls				
Temporal Controls	Y	Y	Y	Y
Time Trend	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y
Standard Error	Robust Std. Err.	Cluster Year_month	Cluster Year_week	Newey-West Std. Err.
Observations	2344	2344	2344	2344

Notes: Column (1) replicates the preferred results for both OLS and IV methods with statistical inference based on robust standard errors. Column (2) and Column (3) cluster the standard errors at year by month and year by week level, respectively. Newey-West standard errors with a lag of 29 are reported in Column (4). Temporal controls include multiple dummies for year by month and holiday. Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.10: Robustness Check — Allow for Different Sets of Top Levels for Universities

Different Top Levels	Beijing Universities over Top n (%) (IB_t)				
	(1) Top 100	(2) Top 200	(3) Top 300	(4) Top 400	(5) Top 500
OLS					
$PM2.5_t$	-0.018***	-0.022***	-0.022***	-0.023***	-0.023***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
2SLS					
$PM2.5_t$	-0.056***	-0.069***	-0.069***	-0.077***	-0.078***
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Additional Controls					
Temporal Controls	Y	Y	Y	Y	Y
Time Trend	Y	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y	Y
Mean Ratio (UB_t/NUB_t)	47.57/52.43	36.34/63.66	30.22/69.78	29.63/70.37	27.43/72.57
Observations	2344	2344	2344	2344	2344

Notes: This table lists the effect of $PM2.5$ on search index for local universities over different top levels. The search for top 100 universities are mainly discussed in the our work. Column (1) reprints the preferred work for both OLS and IV. Column (2) through Column (5) expand overall group of universities based on various top levels indicated in the first row, from top 200 to top 500. Temporal controls include multiple dummies for year by month and holiday. Weather covariates contain quadratic forms of average temperature and relative humidity, as well as their interaction term, wind speed, cloud coverage, sea-level pressure, and precipitation. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.11: Comparison among the Effects of $PM2.5$ at Different Periods

Different Period	Lag One Period (1)	Present Period (2)	Lead one Period (3)
OLS			
$PM2.5_t$	-0.007**	-0.018***	-0.002
	(0.003)	(0.003)	(0.004)
2SLS			
$PM2.5_t$	-0.032***	-0.056***	-0.002
	(0.009)	(0.009)	(0.004)
Additional Controls			
Temporal controls	Y	Y	Y
Time trend	Y	Y	Y
Weather Controls	Y	Y	Y
Observations	2315	2344	2254

Notes: Column (2) replicates the preferred results on both OLS and IV methods based on dependent variable (IB_t) and independent variables ($PM2.5_t$ and W_t) at the same time period of past 30 days' mean. Column (1) replaces the independent variables with values lagged one period without overlap, which is monthly average 30 days ago. Column (3) replaces the independent variables with values led one period without overlap, that is monthly average 30 days ahead. Temporal controls include multiple dummies for year by month and holiday. Weather covariates contain quadratic forms of maximum and minimum temperature, relative humidity, wind speed, cloud coverage, sea-level pressure, and precipitation. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

Table 3.12: Heterogeneous Effects among Different Clean Levels

Heterogeneous Effects	Cleaner Universities over Total (%) ($Ln(INB_t^{clean})$)			
	(1)	(2)	(3)	(4)
	<Beijing	<Beijing-10	<Beijing-20	<Beijing-30
OLS				
$PM2.5_t$	0.031***	0.033***	0.035***	0.042***
	(0.003)	(0.004)	(0.004)	(0.004)
2SLS				
$PM2.5_t$	0.095***	0.103***	0.119***	0.130***
	(0.010)	(0.011)	(0.012)	(0.013)
Additional Controls				
Temporal Controls	Y	Y	Y	Y
Time Trend	Y	Y	Y	Y
Weather Controls	Y	Y	Y	Y
Observations	2344	2344	2344	2344

Notes: This table helps to discover that whether the elevated $PM2.5$ would stimulate Beijing citizens' interest in seeking higher education in cleaner cities. Dependent variable is the log-form of the share of search index for non-Beijing universities, $Ln(INB_t^{clean})$. Estimates are reported in percentage form by multiplying 100 to each original value. "<Beijing" (Column (1)) includes all the index for those universities located in cleaner cities than Beijing. "<Beijing-10" (Column (2)) focuses on those cities 10 $\mu g/m^3$ cleaner than Beijing. "<Beijing-20" (Column (3)) and "<Beijing-30" (Column (4)) expand the gap to 20 $\mu g/m^3$ and 30 $\mu g/m^3$, respectively. Temporal controls include multiple dummies for year by month and holiday. Weather covariates contain quadratic forms of average temperature and relative humidity, as well as their interaction term, wind speed, cloud coverage, sea-level pressure, and precipitation. Robust standard errors are reported in parentheses (* significant at 10%, ** significant at 5%, *** significant at 1%).

References

Alberini, A., Cropper, M., Fu, T. T., Krupnick, A., Liu, J. T., Shaw, D., Harrington, W., 1997. Valuing health effects of air pollution in developing countries: The case of Taiwan. Journal of Environmental Economics and Management, 34(2), 107-126.

Altevogt, B. M., Colten, H. R. (Eds.), 2006. Sleep disorders and sleep deprivation: An unmet public health problem. National Academies Press.

Anderson, M. L., 2015. As the wind blows: The effects of long-term exposure to air pollution on mortality. Working paper, National Bureau of Economic Research.

Andrew, M., 2016, April 21. Why companies are willing to pay to make sure you get a good night's sleep. Retrieved from http://www.executivestyle.com.au/why-companies-are-willing-to-pay-to-make-sure-you-get-a-good-nights-sleep-gobjfg

Anggraeni, W., Aristiani, L., 2016. Using Google Trend data in forecasting number of dengue fever cases with ARIMAX method case study: Surabaya, Indonesia. In 2016 International Conference on Information & Communication Technology and Systems (ICTS) (114-118). IEEE.

Angrist, J. D., Pischke, J. S., 2008. Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Arceo, E., Hanna, R., Oliva, P., 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. The Economic Journal, 126(591), 257-280.

Archsmith, J., Heyes, A., Saberian, S., 2018. Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. Journal of the Association of Environmental and Resource Economists, 5(4), 827-863.

Asgeirsdottir, T. L., Zoega, G., 2011. On the economics of sleeping. Mind & Society, 10(2), 149.

Askitas, N., Zimmermann, K. F., 2009. Google econometrics and unemployment forecasting. Applied Economics Quarterly (formerly: Konjunkturpolitik), 55(2), 107-120.

Baiardini, I., Braido, F., Fassio, O., Tarantini, F., Pasquali, M., Tarchino, F., Berlendis, A., Canonica, G. W., 2005. A new tool to assess and monitor the burden of chronic cough on quality of life: Chronic Cough Impact Questionnaire. Allergy, 60(4), 482-488.

Bakirov, N. K., Szekely, G. J., 2006. Student's t-test for Gaussian scale mixtures. Journal of Mathematical Sciences, 139(3), 6497-6505.

Barnes, C. M., Wagner, D. T., 2009. Changing to daylight saving time cuts into sleep and increases workplace injuries. Journal of Applied Psychology, 94(5), 1305-1317.

Barreca, A. I., Neidell, M., Sanders, N. J., 2017. Long-run pollution exposure and adult mortality: Evidence from the acid rain program. Working paper, National Bureau of Economic Research.

Barton, D. N., 1999. The quick, the cheap and the dirty: Benefit transfer approaches to the non-market valuation of coastal water quality in Costa Rica. Agricultural University of Norway, Department of Economics and Social Sciences.

Barwick, P. J., Li, S., Rao, D., Zahur, N. B., 2018. The morbidity cost of air pollution: Evidence from consumer spending in China. Working paper, National Bureau of Economic Research.

Bayer, P., Keohane, N., Timmins, C., 2009. Migration and hedonic valuation: The case of air quality. Journal of Environmental Economics and Management, 58(1), 1-14.

Baylis, P., 2015. Temperature and temperament: Evidence from a billion tweets. Energy Institute at HAAS working paper.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates?. The Quarterly Journal of Economics, 119(1), 249-275.

Biddle, J. E., Hamermesh, D. S., 1990. Sleep and the allocation of time. Journal of Political Economy, 98(5, Part 1), 922-943.

Billings, M. E., Gold, D. R., Leary, P. J., Szpiro, A., Aaron, C. P., Kaufman, J. D., Redline, S. S., 2017. Relationship of air pollution to sleep disruption: The Multi-Ethnic Study of Atherosclerosis (MESA) Sleep and MESA-Air Studies. Paper presented at American Thoracic Society Conference.

Blazquez, D., Domenech, J., 2018. Big data sources and methods for social and economic analyses. Technological Forecasting and Social Change, 130, 99-113.

Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8.

Brunekreef, B., Holgate, S. T., 2002. Air pollution and health. The Lancet, 360(9341), 1233-1242.

Bucks, R. S., Gidron, Y., Harris, P., Teeling, J., Wesnes, K. A., Perry, V. H., 2008. Selective effects of upper respiratory tract infection on cognition, mood and emotion processing: A prospective study. Brain, Behavior, and Immunity, 22(3), 399-407.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proceedings of the National Academy of Sciences, 115(38), 9592-9597.

Cain, K., 2015, September 21. An introduction to Sina Weibo: Background and status quo. Retrieved from http://www. whatsonweibo.com/sinaweibo/

Cameron, A. C., Gelbach, J. B., Miller, D. L., 2008. Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), 414-427.

Cameron, A. C., Miller, D. L., 2015. A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317-372.

Cappuccio, F., Miller, M. A., Lockley, S. W. (Eds.), 2010. Sleep, health, and society: From aetiology to public health. Oxford University Press, USA.

Chamberlain, S., Garrod, R., Birring, S. S., 2013. Cough suppression therapy: Does it work?. Pulmonary Pharmacology & Therapeutics, 26(5), 524-527.

Chan, M., Wu, X., Hao, Y., Xi, R., Jin, T., 2012. Microblogging, online expression, and political efficacy among young Chinese citizens: The moderating role of information and entertainment needs in the use of Weibo. Cyberpsychology, Behavior, and Social Networking, 15(7), 345-349.

Chang, T., Graff Zivin, J., Gross, T., Neidell, M., 2016. Particulate pollution and the productivity of pear packers. American Economic Journal: Economic Policy, 8(3), 141-169.

Chang, T. Y., Huang, W., Wang, Y., 2018. Something in the air: Pollution and the demand for health insurance. The Review of Economic Studies, 85(3), 1609-1634.

Chay, K. Y., Greenstone, M., 2005. Does air quality matter? Evidence from the housing market. Journal of Political Economy, 113(2), 376-424.

Chen, Y., Jin, G. Z., Kumar, N., Shi, G., 2012. Gaming in air pollution data? Lessons from China. The BE Journal of Economic Analysis & Policy, 12(3).

Chen, R. C., Lai, K. F., Liu, C. L., Luo, W., Zhong, N. S., 2006. An epidemiologic study of cough in young college students in Guangzhou. Zhonghua Liu Xing Bing Xue Za Zhi, 27(2), 123-126.

Chen, J., Li, C., Ristovski, Z., Milic, A., Gu, Y., Islam, M. S., Wang, S., Hao, J., Zhang, H., He, C., Guo, H., 2017. A review of biomass burning: Emissions and impacts on air quality, health and climate in China. Science of the Total Environment, 579, 1000-1034.

Chen, S., Oliva, P., Zhang, P., 2017. The effect of air pollution on migration: Evidence from China. Working paper, National Bureau of Economic Research.

Chen, Y. L., Yang, X. J., Shi, C. F., 2004. Elementary study on sleep quality and influence factors for common population in Xiamen, China (2004). Strait Journal of Preventive Medicine, 10, 4–6.

Chen, X., Ye, J., 2015. When the wind blows: Spatial spillover effects of urban air pollution in China. Journal of Environmental Planning and Management. Retrieved from https://doi.org/10.1080/09640568.2018.1496071

Chestnut, L. G., Ostro, B. D., Vichit-Vadakan, N., Smith, K. R., Tsai, F. C., 1998. Health effects of particulate matter air pollution in Bangkok. Final Report Prepared for the World Bank and the Royal Thai Government.

Chiu, C., Lin, D., Silverman, A., 2012, May. China's Social-media Boom. McKinsey & Company. Retrieved from https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/chinas-social-media-boom

Choi, H., Varian, H., 2012. Predicting the present with Google Trends. Economic Record, 88, 2-9.

Chung, K. F., Cheung, M. M., 2008. Sleep-wake patterns and sleep disturbance among Hong Kong Chinese adolescents. Sleep, 31(2), 185-194.

Currie, J., Greenstone, M., Moretti, E., 2011. Superfund cleanups and infant health. American Economic Review, 101(3), 435-41.

Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M., Rivkin, S. G., 2009. Does pollution increase school absences?. The Review of Economics and Statistics, 91(4), 682-694.

Currie, J., Neidell, M., Schmieder, J. F., 2009. Air pollution and infant health: Lessons from New Jersey. Journal of Health Economics, 28(3), 688-703.

Currie, J., Walker, R., 2011. Traffic congestion and infant health: Evidence from E-ZPass. American Economic Journal: Applied Economics, 3(1), 65-90.

Daley, M., Morin, C. M., LeBlanc, M., Grégoire, J. P., Savard, J., Baillargeon, L., 2009. Insomnia and its relationship to health-care utilization, work absenteeism, productivity and accidents. Sleep Medicine, 10(4), 427-438.

Delfino Jr, M., Holt, E. W., Taylor, C. R., Wittenberg, E., Qureshi, A. A., 2008. Willingness-to-pay stated preferences for 8 health-related quality-of-life domains in psoriasis: A pilot study. Journal of the American Academy of Dermatology, 59(3), 439-447.

DeLuca, K.M., Brunner, E., Sun, Y., 2016. Weibo, WeChat, and the transformative events of environmental activism on China's wild public screens. International Journal of Communication, 10(2016), 321-339.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., Reif, J., 2016. The mortality and medical costs of air pollution: Evidence from changes in wind direction. Working paper, National Bureau of Economic Research.

Dicpinigaitis, P. V., Eccles, R., Blaiss, M. S., Wingertzahn, M. A., 2015. Impact of cough and common cold on productivity, absenteeism, and daily life in the United States: ACHOO Survey. Current Medical Research and Opinion, 31(8), 1519-1525.

Diekelmann, S., Born, J., 2010. The memory function of sleep. Nature Reviews Neuroscience, 11 (1), 114-126.

Doornik, J.A., 2009. Improving the timeliness of data on influenza-like illnesses using Google search data. In 8th OxMetrics User Conference.

Driscoll, J. C., Kraay, A. C., 1998. Consistent covariance matrix estimation with spatially dependent panel data. The Review of Economics and Statistics, 80(4), 549-560.

Dubourg, R., 1998. The value of preventing respiratory illness in Kuala Lumpur and Petaling Jaya, Malaysia. Draft final report to the Economic Planning Unit, Government of Malaysia.

Ebenstein, A., Fan, M., Greenstone, M., He, G., Zhou, M., 2017. New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy. Proceedings of the National Academy of Sciences, 114(39), 10384-10389.

Ebenstein, A., Lavy, V., Roth, S., 2016. The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. American Economic Journal: Applied Economics, 8(4), 36-65.

EPA (U.S. Environmental Protection Agency), 1996. Review of the national ambient air quality standards for particulate matter: Policy assessment of scientific and technical information. OAQPS Staff Paper.

Esarey, J., Menger, A., 2019. Practical and effective approaches to dealing with clustered data. Political Science Research and Methods, 7(3), 541-559.

Ettredge, M., Gerdes, J., Karuga, G., 2005. Using web-based search data to predict macroeconomic statistics. Communications of the ACM, 48(11), 87-92.

Fang, L., Clausen, G., Fanger, P. O., 1998. Impact of temperature and humidity on the perception of indoor air quality. Indoor Air, 8(2), 80-90.

Finlayson-Pitts, B. J., Pitts, J. N., 1997. Tropospheric air pollution: Ozone, airborne toxics, polycyclic aromatic hydrocarbons, and particles. Science, 276(5315), 1045-1051.

Fu, S., Viard, B., Zhang, P., 2017. Air quality and manufacturing firm productivity: Comprehensive evidence from China. Unpublished.

Gao, Q., Abel, F., Houben, G. J., Yu, Y., 2012. A comparative study of users? Microblogging behavior on Sina Weibo and Twitter. Presented at International Conference on User Modelling, Adaptation, and Personalization.

Gayo-Avello, D., 2012. "I wanted to predict elections with Twitter and all I got was this lousy paper"– A balanced survey on election prediction using Twitter data. Preprint (arXiv:1204.6441).

Gayo-Avello, D., Metaxas, P. T., Mustafaraj, E., 2011. Limits of electoral predictions using Twitter. Presented at Proceedings of the Fifth International Association for the Advancement of Artificial Intelligence (AAAI) Conference on Weblogs and Social Media.

Gendron-Carrier, N., Gonzalez-Navarro, M., Polloni, S., Turner, M. A., 2018. Subways and urban air pollution. Working paper, National Bureau of Economic Research.

Gerber, M. S., 2014. Predicting crime using Twitter and kernel density estimation. Decision Support Systems, 61(2014), 115-125.

Giglio, L., 2015. MODIS collection 6 active fire product user's guide revision A. Department of Geographical Sciences, University of Maryland.

Gilliland, F. D., Berhane, K., Rappaport, E. B., Thomas, D. C., Avol, E., Gauderman, W. J., London, S. J., Margolis, H. G., McConnell, R., Islam, K. T., Peters, J. M., 2001. Effects of ambient air pollution on school absenteeism due to respiratory illnesses. Epidemiology, 12(1), 43-54.

Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., Watts, D. J., 2010. Predicting consumer behavior with Web search. Proceedings of the National Academy of Sciences, 107(41), 17486-17490.

Gordon, A. M., Chen, S., 2014. The role of sleep in interpersonal conflict: Do sleepless nights mean worse fights?. Social Psychological and Personality Science, 5(2), 168-175.

Graff Zivin, J., Neidell, M., 2012. The impact of pollution on worker productivity. American Economic Review, 102(7), 3652-73.

Greenstone, M., Gallagher, J., 2008. Does hazardous waste matter? Evidence from the housing market and the superfund program. The Quarterly Journal of Economics, 123(3), 951-1003.

Gupta, A., Spears, D., 2017. Health externalities of India's expansion of coal plants: Evidence from a national panel of 40,000 households. Journal of Environmental Economics and Management, 86(2017), 262-276.

Hamilton, N. A., Nelson, C. A., Stevens, N., Kitzman, H., 2007. Sleep and psychological well-being. Social Indicators Research, 82(1), 147-163.

Hand, C., Judge, G., 2012. Searching for the picture: Forecasting UK cinema admissions using Google Trends data. Applied Economics Letters, 19(11), 1051-1055.

Hanna, R., Oliva, P., 2015. The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City. Journal of Public Economics, 122, 68-79.

Harden, J. J., 2011. A bootstrap method for conducting statistical inference with clustered data. State Politics & Policy Quarterly, 11(2), 223-246.

He, L. M., Wang, W. J., Wang, Q., Wei, B., Li, Q., Wang, C. Z., Liu, X. M., 2007. Evaluation of the agricultural residues burning reduction in China using MODIS fire product. Environmental Monitoring in China, 23(1), 42-50.

Herrnstadt, E., Heyes, A., Muehlegger, E., Saberian, S., 2016. Air pollution as a cause of violent crime: Evidence from Los Angeles and Chicago. Unpublished.

Herrnstadt, E., Muehlegger, E., 2014. Weather, salience of climate change and congressional voting. Journal of Environmental Economics and Management, 68(3), 435-448.

Herrnstadt, E., Muehlegger, E., 2015. Air pollution and criminal activity: Evidence from Chicago microdata. Working paper, National Bureau of Economic Research.

Heyes, A., Zhu, M., 2018. Sleepless in Shanghai: Evidence of the effect of air pollution on sleep using social media metrics. Working Paper.

Hoechle, D., 2007. Robust standard errors for panel regressions with cross-sectional dependence. The Stata Journal, 7(3), 281-312.

Huber, R., Ghilardi, M. F., Massimini, M., Tononi, G., 2004. Local sleep and learning. Nature, 430(6995), 78-81.

Ibáñez, A. M., McConnell, K. E., 2011. Valuing morbidity: Acute respiratory illnesses in Bogotá, Colombia Draft. In AERE Workshop.

Ireland, J. L., Culpin, V., 2006. The relationship between sleeping problems and aggression, anger, and impulsivity in a population of juvenile and young offenders. Journal of Adolescent Health, 38(6), 649-655.

Jans, J., Johansson, P., Nilsson, J. P., 2018. Economic status, air quality, and child health: Evidence from inversion episodes. Journal of Health Economics, 61, 220-232.

Keech, M., Scott, A. J., Ryan, P. J. J., 1998. The impact of influenza and influenza-like illness on productivity and healthcare resource utilization in a working population. Occupational Medicine, 48(2), 85-90.

Killgore, W. D., Balkin, T. J., Wesensten, N. J., 2006. Impaired decision making following 49 h of sleep deprivation. Journal of Sleep Research, 15(1), 7-13.

Knittel, C. R., Miller, D. L., Sanders, N. J., 2016. Caution, drivers! Children present: Traffic, pollution, and infant health. Review of Economics and Statistics, 98(2), 350-366.

Lai, K., Pan, J., Chen, R., Liu, B., Luo, W., Zhong, N., 2013. Epidemiology of cough in relation to China. Cough, 9(1).

Lavaine, E., Neidell, M., 2017. Energy production and health externalities: Evidence from oil refinery strikes in France. Journal of the Association of Environmental and Resource Economists, 4(2), 447-477.

Lavy, V., Ebenstein, A., Roth, S., 2014. The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation. Working paper, National Bureau of Economic Research.

Li, M. S., Ren, X. X., Yu, Y., Zhou, L., 2016. Spatiotemporal pattern of ground-level fine particulate matter (PM2.5) pollution in mainland China. China Environmental Science, 36(3), 641-650.

Luo, J., Zhu, G., Zhao, Q., Guo, Q., Meng, H., Hong, Z., Ding, D., 2013. Prevalence and risk factors of poor sleep quality among Chinese elderly in an urban community: Results from the Shanghai aging study. PloS One, 8(11), e81261.

Mallows, C. L., 1986. Augmented partial residuals. Technometrics, 28(4), 313-319.

Maquet, P., 2001. The role of sleep in learning and memory. Science, 294(5544), 1048-1052.

Meegan, R.L., 1998. Contingent valuation and reduced morbidity. Master's thesis, California State University.

Meijer, A. M., Reitz, E., Deković, M., Van Den Wittenboer, G. L., Stoel, R. D., 2010. Longitudinal relations between sleep quality, time in bed and adolescent problem behavior. Journal of Child Psychology and Psychiatry, 51(11), 1278-1286.

Mendell, M. J., Heath, G. A., 2005. Do indoor pollutants and thermal conditions in schools influence student performance? A critical review of the literature. Indoor Air, 15(1), 27-52.

Moretti, E., Neidell, M., 2011. Pollution, health, and avoidance behavior evidence from the ports of Los Angeles. Journal of Human Resources, 46(1), 154-175.

National Research Council, 1993. Protecting visibility in national parks and wilderness areas. National Academies Press.

Neidell, M. J., 2004. Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma. Journal of Health Economics, 23(6), 1209-1236.

Nielsen, B., 2006. Order determination in general vector autoregressions. IMS Lecture Notes-Monograph Series, 52, 93–112.

Norman, J., 1988. Chinese. Cambridge University Press.

Obradovich, N., Migliorini, R., Mednick, S. C., Fowler, J. H., 2017. Nighttime temperature and human sleep loss in a changing climate. Science advances, 3(5), e1601555.

Okamoto-Mizuno, K., Mizuno, K., 2012. Effects of thermal environment on sleep and circadian rhythm. Journal of Physiological Anthropology, 31, Article 14.

Peng, H., Zhao, X. H., Bi, T. T., Yuan, X. Y., Guo, J. B., Peng, S. Q., 2017. PM2.5 obtained from urban areas in Beijing induces apoptosis by activating nuclear factor-kappa B. Military Medical Research, 4(1), Article 27.

Pollinger, M., 2014. Individuals willingness to pay for health and wellness in the built environment. Journal of Environmental and Resource Economics at Colby, 1(1), Article 5.

Qin, Y., Zhu, H, 2018. Run away? Air pollution and emigration interests in China. Journal of Population Economics, 31(1), 235-266.

Rangel, M. A., Vogl, T., 2016. Agricultural fires and infant health. Working paper, National Bureau of Economic Research.

Ransom, M. R., Pope III, C. A., 1992. Elementary school absences and PM10 pollution in Utah Valley. Environmental Research, 58(1-2), 204-219.

Rech, J., 2007. Discovering trends in software engineering with Google Trend. ACM SIGSOFT Software Engineering Notes, 32(2), 1-2.

Rechtschaffen, A., 1971. The control of sleep. In Hynt WA (ed): Human Behavior and Its Control. Cambridge, MA; Shenkman Publishing Company, Inc.

Rohde, R. A., Muller, R. A., 2015. Air pollution in China: Mapping of concentrations and sources. PloS One, 10(8), e0135749.

Rosekind, M. R., Gregory, K. B., Mallis, M. M., Brandt, S. L., Seal, B., Lerner, D., 2010. The cost of poor sleep: Workplace productivity loss and associated costs. Journal of Occupational and Environmental Medicine, 52(1), 91-98.

Rozan, A., 2004. Benefit transfer: A comparison of WTP for air quality between France and Germany. Environmental and Resource Economics, 29(3), 295-306.

Sager, L., 2016. Estimating the effect of air pollution on road safety using atmospheric temperature inversions. Grantham Research Institute on Climate Change and the Environment working paper, 251.

Schlenker, W., Walker, W. R., 2016. Airports, air pollution, and contemporaneous health. The Review of Economic Studies, 83(2), 768-809.

Seifter, A., Schwarzwalder, A., Geis, K., Aucott, J., 2010. The utility of "Google Trends" for epidemiological research: Lyme disease as an example. Geospatial Health, 4(2), 135-137.

Shao, M., Tang, X., Zhang, Y., Li, W., 2006. City clusters in China: Air and surface water pollution. Frontiers in Ecology and the Environment, 4(7), 353-361.

Steptoe, A., O'Donnell, K., Marmot, M., Wardle, J., 2008. Positive affect, psychological well-being, and good sleep. Journal of Psychosomatic Research, 64(4), 409-415.

Stoerk, T., 2016. Statistical corruption in Beijing's air quality data has likely ended in 2012. Atmospheric Environment, 127, 365-371.

Strøm-Tejsen, P., Zukowska, D., Wargocki, P., Wyon, D. P., 2016. The effects of bedroom air quality on sleep and next-day performance. Indoor Air, 26(5), 679-686.

Sullivan, J., 2012. A tale of two microblogs in China. Media, Culture & Society, 34(6), 773-783.

Sun, W., Yu, Y., Yuan, J., Li, C., Liu, T., Lin, D., Lau, A., Zhong, C., Xu, T., Shan, G., 2015. Sleep duration and quality among different occupations — China national study. PloS One, 10(3), e0117700.

The World University Rankings, 2018. Best universities in China 2019. Retrieved from https://www.timeshighereducation.com/student/best-universities/best-universities-china

Valent, F., Bartolomeo, S. D., Marchetti, R., Sbrojavacca, R., Barbone, F., 2010. A case-crossover study of sleep and work hours and the risk of road traffic accidents. Sleep, 33(3), 349-354.

Van Someren, E. J., 2006. Mechanisms and functions of coupling between sleep and temperature rhythms. Progress in Brain Research, 153, 309-324.

Vassanadumrongdee, S., Matsuoka, S., Shirakawa, H., 2004. Meta-analysis of contingent valuation studies on air pollution-related morbidity risks. Environmental Economics and Policy Studies, 6(1), 11-47.

Vaughan, L., Chen, Y., 2015. Data mining from web search queries: A comparison of Google Trends and Baidu Index. Journal of the Association for Information Science and Technology, 66(1), 13-22.

Wang, J., Ogawa, S., 2015. Effects of meteorological conditions on PM2.5 concentrations in Nagasaki, Japan. International Journal of Environmental Research and Public Health, 12(8), 9089-9101.

Wolkoff, P., 1998. Impact of air velocity, temperature, humidity, and air on long-term VOC emissions from building products. Atmospheric Environment, 32(14-15), 2659-2668.

Wooldridge, J. M., 2010. Econometric analysis of cross section and panel data. MIT press.

Wong, K. K., 2003. The environmental awareness of university students in Beijing, China. Journal of Contemporary China, 12(36), 519-536.

Wu, L., Brynjolfsson, E., 2015. The future of prediction: How Google searches foreshadow housing prices and sales. The Economics of Digitization: An Agenda, National Bureau of Economic Research.

Yang, Y., Liu, X., Qu, Y., Wang, J., An, J., Zhang, Y., Zhang, F., 2015. Formation mechanism of continuous extreme haze episodes in the megacity Beijing, China, in January 2013. Atmospheric Research, 155, 192-203.

Yu, Y., Lu, B. S., Wang, B., Wang, H., Yang, J., Li, Z., Wang, L., Liu, X., Tang, G., Xing, H., Xu, X., 2007. Short sleep duration and adiposity in Chinese adolescents. Sleep, 30(12), 1688-1697.

Yuan, Q., Nsoesie, E. O., Lv, B., Peng, G., Chunara, R., Brownstein, J. S., 2013. Monitoring influenza epidemics in china with search query from Baidu. PloS One, 8(5), e64323.

Zahran, S., Iverson, T., McElmurry, S. P., Weiler, S., 2017. The effect of leaded aviation gasoline on blood lead in children. Journal of the Association of Environmental and Resource Economists, 4(2), 575-610.

Zammit, G. K., Joish, V. N., Kong, M. C., Balkrishnan, R., Lerner, D., Rosekind, M., 2010. Impact of nighttime awakenings on worker productivity and performance. Journal of Occupational and Environmental Medicine, 52(5), 513-518.

Zanobetti, A., Redline, S., Schwartz, J., Rosen, D., Patel, S., O'connor, G. T., Lebowitz, M., Coull, B. A., Gold, D. R., 2010. Associations of PM10 with sleep and sleep-disordered breathing in adults from seven US urban areas. American Journal of Respiratory and Critical Care Medicine, 182(6), 819-825.

Zeidner, M., Shechter, M., 1988. Psychological responses to air pollution: Some personality and demographic correlates. Journal of Environmental Psychology, 8(3), 191-208.

Zhang, J., Lioy, P. J., 1994. Ozone in residential air: Concentrations, I/O ratios, indoor chemistry, and exposures. Indoor Air, 4(2), 95-105.

Zhang, L., Liu, Y., Hao, L., 2016. Contributions of open crop straw burning emissions to PM2. 5 concentrations in China. Environmental Research Letters, 11(1), 014014.

Zheng, S., Cao, J., Kahn, M. E., Sun, C., 2014. Real estate valuation and cross-boundary air pollution externalities: Evidence from Chinese cities. The Journal of Real Estate Finance and Economics, 48(3), 398-414.

Zheng, S., Wang, J., Sun, C., Zhang, X., Kahn, M. E. (2019). Air pollution lowers Chinese urbanites' expressed happiness on social media. Nature Human Behaviour, 3(3), 237–243.