IDEC DP² Series Vol. 2 No. 11

Developmnet Discussion Policy Paper

Environmental Regulations on Air Pollution in China and Their Impact on Infant Mortality

Shinsuke Tanaka

June, 2012

Department of Development Policy
Division of Development Science
Graduate School for International Development and Cooperation (IDEC)
Hiroshima University
1-5-1 Kagamiyama, Higashi-Hiroshima
739-8529 JAPAN

ENVIRONMENTAL REGULATIONS ON AIR POLLUTION IN CHINA AND THEIR IMPACT ON INFANT MORTALITY*

Shinsuke Tanaka†

Abstract

Developing countries rank highest in air pollution worldwide, yet regulations of such pollution are still rare in these countries, thereby whether, and to what extent, those regulations lead to health benefits remain an open question. Since 1995, the Chinese government has imposed stringent regulations on pollutant emissions from power plants, as one of the first regulatory attempts on a large scale in a developing country. Exploiting the variation in the regulatory status across time and space, we find that infant mortality fell by 21 percent in the treatment cities designated as the so-called "Two Control Zones." The greatest reduction of mortality occurred during the neonatal period, highlighting the importance of fetal exposure as a biological mechanism, and was largest among the households with low mother's educational attainment. On the other hand, the regulations are found to be uncorrelated with deaths from causes unrelated to air pollution. When the regulatory status is used as an instrumental variable for air pollution reductions, we estimate that the impact of a unit change in total suspended particulates on infant mortality is of similar magnitude to that found in the U.S., but the elasticity is substantially higher in China, suggesting the greater benefits associated with regulations when pollution is already quite high.

JEL classification: Q56, I18, Q53, J13, O13
Keywords: environmental regulation, infant mortality, air quality, China

*I am indebted to Daniele Paserman, Dilip Mookherjee, Tavneet Suri, and Wesley Yin for invaluable advice and feedback. I am also grateful to Michael Greenstone, Hsueh-Ling Huynh, Kelsey Jack, Ginger Zhe Jin, Hiroaki Kaido, Kevin Lang, Michael Manove, Kenneth A. Rahn, Leena Rudanko, Marc Rysman, Johannes Schmieder, Jeremy Smith, and seminar participants at Boston University, Loyola Marymount University, National University of Singapore, Tufts University, University of Washington, 2010 NEUDC, and the 2011 Royal Economic Society for their comments and suggestions. I also thank the Chinese Center for Disease Control and Prevention for sharing the data. Financial support from the Institute for Economic Development at Boston University, as well as Hewlett/IIE Dissertation Fellowship in Population, Reproductive Health and Economics, is gratefully acknowledged. Jieshuang He provided excellent research assistance. All errors are my own.

†Shinsuke Tanaka, Assistant Professor of Economics, Tufts University, the Fletcher School of Law and Diplomacy, 160 Packard Avenue, Medford, MA 02155. Shinsuke.Tanaka@tufts.edu.

I. INTRODUCTION

There is little disagreement that air pollution poses a major environmental risk to human health. Improved air quality worldwide is correlated with the amelioration of numerous health problems, including respiratory infections, cardiovascular diseases, and lung cancer. Elevated air pollution is generally beyond the scope of individual control and falls to the public sector, especially in developing countries, whose children under age five are considered to be the most vulnerable population.

However, environmental regulations on air pollution are extremely contentious. This was evident at the 2009 United Nations Climate Change Conference, also known as COP15, held in Copenhagen, where many countries refused to commit themselves to a legal framework for reducing pollution emissions. Their opposition is largely due to the prevailing concern that the economic costs associated with pollution abatement may outweigh the health benefits.[1] Accordingly, air pollution regulations are still rare in developing countries, and whether, and to what extent, environmental regulations on air pollution lead to health benefits remains an important question yet to be answered.

In this paper, I examine the effect of environmental regulations pertaining to air pollution on infant mortality in China. As China's economy continued to grow at unprecedented rates for the last several decades, ambient air quality deteriorated to one of the worst levels in the world due to its heavy reliance on coal-fired energy generation. Since 1995, the Chinese government has imposed stringent regulations on pollutant emissions from power plants, in one of the first attempts on such a large scale in developing countries. Among them, the most striking was the Two Control Zone (TCZ) policy implemented in 1998 that designated nearly 175 prefectures exceeding the nationally-mandated pollution standards as the TCZ. In these areas, the power industry, which contributed more than 90 percent of air pollution, was enforced to reduce emissions and install new pollution control technologies, while also shutting down a massive number of small, inefficient power plants. For our purpose, the TCZ policy provides a quasi-experimental setting, wherein the intensity of exposure to the regulations can be defined by the TCZ regulatory status, and we are able to compare changes in infant mortality rate (IMR) before and after the policy reform, between the

[1] See, for example, Henderson (1996), Becker and Henderson (2000), and Greenstone (2002), for economic costs associated with the Clean Air Act in the U.S.

cities assigned and not assigned as the TCZ. Further, the variation in air pollution reductions across time and space generated by the regulations enables us to quantify the effect of air pollution on infant mortality. Evaluation of a credible relationship between the two has been hampered by a concern that unobserved factors in local economies may give rise to spurious correlations. Hence, any estimates that treat air-quality levels or changes as exogenous would be severely biased.[2] To the extent that the TCZ status serves as a valid instrumental variable (IV) for pollution changes, the 2SLS estimates provide a causal link between air pollution and infant mortality.

A major methodological challenge, however, is that the designation rule may not be orthogonal to unobservable characteristics that contribute to reductions in air pollution and infant mortality. The present study conduct a number of robustness checks and a falsification test to address this issue. First, we confirm that the TCZ status has little association with *changes* in observable covariates that are associated with the TCZ status in *levels*, assuring that there is no systematic difference in trends in observable characteristics between the TCZ and non-TCZ cities. Although this is not a direct test of exclusion restriction, since it requires that TCZ status not be correlated with changes in *unobservable* factors, this result leads us to believe that the treatment effect is less likely to be confounded by differential trends in unobservable factors as well (Altonji, Elder, and Taber 2005). Second, the estimates are robust to various alternative specifications aiming to restrict the sample to similar characteristics, i.e., controlling for extensive characteristics, a comparison under common support, a comparison within province × year, a comparison within urban areas, and including city-specific time trends. Lastly, the policy had no impact on infant deaths due to external causes of injury or poisons. Because there is no causal mechanism linking

[2] A sizable literature documents health risks associated with air pollution exposure. See, for example, Schwartz, Dockery, and Neas (1996), Levy, Hammitt, and Spengler (2000), and Samet et al. (2000), Chay and Greenstone (2003a), Currie and Neidell (2005), Currie, Neidell, and Schmieder (2009). It is generally accepted that the smaller particulates, the more detrimental it is to health. For example, PM_{10} or $PM_{2.5}$, whose particles are less than 10 or 2.5 μm in diameter, respectively, or toxic gas, such as SO_2, are considered to be the most hazardous because, when inhaled, these particulate matters or gas can reach deep into the lungs and interfere with internal gas exchange. Further, Laden et al. (2000) find that fine particles emitted from combustion sources (motor vehicles or coal combustion) have a stronger association with mortality than those from non-combustion sources. Alternatively, SO_2 becomes sulfuric acid when it interacts with water, which is the main component of acid rain that may have a direct or indirect impact on health. Yet, epidemiological evidence of the impact of SO_2 on mortality in developed countries is somewhat mixed. While Mendelsohn and Orcutt (1979) show close associations between the two, SO_2 is also considered as a less important determinant of mortality (Schwartz and Marcus 1990; Nielsen and Ho 2007). Hedley et al. (2002) is one of the few intervention studies that investigates changes in SO_2 caused by an overnight restriction for all power plants and road vehicles in Hong Kong on using fuel oil with more than 0.5 percent of sulfur content. They find that the intervention resulted in an immediate reduction in ambient SO_2 concentrations and a reduction in death rates particularly due to respiratory and cardiovascular reasons.

air pollution to these causes of death, a lack of significant relationship in this analysis reassures that differences in access to or quality of medical services and technologies cannot be the sources of bias. Overall, there is little evidence that the estimates are driven by inappropriate identification assumptions, leading us to believe that the treatment effect based on the TCZ status is not spurious but causal.

This study makes three major contributions to the existing literature. First, by exploiting regulation-induced changes in air quality, it addresses a policy-relevant question: to what extent do environmental regulations in developing countries lead to reductions in infant mortality? Several prior studies, on the other hand, have focused on variation in air quality, induced by recession (Chay and Greenstone 2003a), weekly fluctuations (Currie and Neidell 2005), wildfires (Jayachandran 2009), or wind directions (Luechinger 2009). Chay and Greenstone (2003b) provide compelling evidence for the linkage between the Clean Air Act of 1970 and infant mortality in the U.S. It remains to be determined, however, whether, and how effectively, environmental regulations can improve human health in developing countries. The present study is the first I am aware of to examine the effect of air pollution regulations on a large scale in developing countries. Further, the regulations we focus on targeted the power industry, which was a driving force behind China's rapid economic growth. The findings in this study highlight the important tradeoffs among economic growth, environmental quality, and human health.[3]

Second, it contributes to fill the gap in our understanding of the relationship between air pollution and infant mortality at greater concentration levels. Previous evidence is predominantly stemmed from the United States or other developed countries, where pollution is relatively low. Estimating the quantitative relationship in places subject to greater pollution levels is necessary, because we know little about the shape of dose-response relationship; different shapes lead to different elasticities for a given amount of changes, leading to different magnitudes of policy benefits. Air pollution in China is one of the highest in the world. Its total suspended particulates (TSP) level in 1995 was four times higher than the WHO standards, and four times higher than the one in the United States in 1970, when the Clean Air Act was amended, as examined in Chay and Greenstone (2003b). Thus, estimates in China provide compelling evidence applicable to the

[3] A relatively smaller-scale air quality regulatory regime, targeting a different industry, in another developing country, can be found in the Indian transportation sector, which was mandated to use compressed national gas vehicles in Delhi during working hours. See, Kumar and Foster (2007) for its effect on respiratory health.

contexts in developing countries, where air pollution levels are relatively high.

Third, it is still an open question in the literature as to what extent the pollution effect on health varies across socioeconomic status (Currie and Hyson 1999; Case, Lubotsky, and Paxson 2002; Jayachandran 2009). Infants in poor countries are considered to be the most susceptible to the effect of pollution, not only because of high pollution levels but also because of limited resources or knowledge to avoid it. Thus, findings in developed countries cannot be simply extrapolated to developing countries.

The current research design also has several advantages over previous studies. First, this study focuses on infants, not only because they are particularly vulnerable to air pollution due to their weak respiratory system, but because focusing on infant mortality mitigates concerns associated with adult mortality. For example, an association between air pollution and adult mortality is often impeded by unknown lifetime exposure to pollution. Most studies use current air-quality levels to proxy long-run exposure. Yet, this not only leads to significant measurement errors but also fails to capture variations in air quality over time. In addition, adult deaths depend more on their chronic disease conditions over time, rather than on acute changes in air quality. Further, adults may migrate into less polluted areas. Addressing infant mortality circumvents these issues, if not completely, because levels of exposure to air pollution during the first year of life can be relatively more precisely estimated, and because migration rates are low for pregnant women and infants.

Second, the current research design helps avert omitted variable bias problems that plague the identification of air-pollution effects. The dearth of exogenous variations in air quality, under cross-sectional or some panel settings, has resulted in little consensus in magnitudes of pollution impact on health.[4] The 2SLS estimates in this study are found to be far less sensitive to various alternative specifications over ones based on cross-sectional or fixed effect models, suggesting the limited scope of omitted variable bias.

Lastly, China is not only one of the most polluted countries but also one of the first developing countries that regulated air pollution on such a large-scale. It is evident that China serves as a rare

[4] In China, the meta-analysis shows that the magnitude of per 1-$\mu g/m^3$ impact of SO_2 ranges from 0.02 percent to 0.19 percent for all-cause mortality, from 0.02 percent to 0.18 percent for mortality due to cardiovascular diseases, and from 0.07 percent to 0.74 percent for mortality due to respiratory diseases. The estimates of the per 1-$\mu g/m^3$ impact of TSP (PM_{10}) range from 0.03 percent to 0.05 percent for all-cause mortality, from 0.04 percent to 0.13 percent for mortality due to cardiovascular disease, and from 0.04 percent to 0.30 percent for mortality due to respiratory diseases (Aunan and Pan 2004).

research environment to assess the impact of environmental regulations at greater concentration levels.

To implement the analysis, we have compiled two major data sources. IMR data come from the Chinese Disease Surveillance Points (DSP) system, that collected the universe of birth and death registrations for 145 nationally representative sites from 1991 through 2000. IMR, as defined by the number of infant deaths under age one, per 1,000 live births in a given year, is computed for each DSP site by year cell, linked with detailed information on their birth characteristics and parental attributes. We match this dataset to the TCZ regulatory status assigned to individual cities, based on the governmental report, and estimate the treatment effect of the regulations. In addition, in an effort to quantify air pollution effect on infant mortality, we have digitized air pollution data from the China Environment Yearbooks, which report the annual average concentration levels of TSP and sulfur dioxide (SO_2) for major prefectures.

The findings show that the air pollution regulations led to significant reductions in infant mortality. The difference-in-differences estimates suggest that the regulations have led to 3.33 fewer infant deaths per 1,000 live births, than it would have been in the absence of the regulations. This corresponds to a 21-percent reduction in IMR, compared to the average IMR in the pre-reform period, implying that 25,400 infant lives per year were saved by the regulations. 61 percent of the reduction in infant mortality occurred during the neonatal period, highlighting the importance of maternal exposure on fetal development, and the the greatest reduction of mortality occurred among the households with low mother's educational levels.

Using the matched sample of air pollution and infant mortality, we find that the regulations resulted in 53 $\mu g/m^3$ declines (a 17 percent reduction) in TSP and 20 $\mu g/m^3$ reductions (a 21 percent reduction) in SO_2. In addition, the 2SLS estimates suggest that a 1-$\mu g/m^3$ reduction in TSP leads to 0.049 fewer infant deaths per 1,000 live births, and a 1-$\mu g/m^3$ reduction in SO_2 leads to 0.14 fewer infant deaths per 1,000 live births. The corresponding elasticities of IMR to pollution concentrations imply that a one percent reduction in TSP (or SO_2) leads to a 0.95 (or 0.82) percent reduction in IMR. The estimated impact of a unit change in TSP is of similar magnitude to that found in the U.S., but the elasticity is substantially higher in China. This finding sheds light on the greater benefits associated with regulations when pollution is already quite high.

The rest of the paper is organized as follows. Section 2 provides the historical background in air pollution and national air pollution regulations in China. Section 4 describes the data and the descriptive statistics. Section 5 presents the econometric framework in addressing the effect of environmental regulations on infant mortality, and shows its results. Section 6 presents the econometric framework in quantifying the air pollution effect on infant mortality, and shows its results. Section 7 discusses the interpretation of these results and their magnitudes. Section 8 concludes.

II. BACKGROUND ON ENVIRONMENTAL REGULATIONS IN CHINA

China is infamous for its air pollution, primarily caused by pollutant emissions from the power sector due to its heavy reliance on coal for generating electric power.[5] As the world's largest coal producer, coal is abundant and relatively cheap in China, and it is the main energy resource endowment, accounting for 75.5 percent of total energy production in 1995 (National Bureau of Statistics of China 2006). However, coal generally emits more pollutants than other fossil fuels. As China underwent rapid economic growth, total SO_2 emissions increased from 18.4 million tons in 1990 to 23.7 million tons in 1995, and the ambient air pollution became detrimental (State Environmental Protection Agency [SEPA] 1996).

Figure 1 illustrates the world distribution of TSP and SO_2 concentration levels (in $\mu g/m^3$) in 1995. The TSP level in Beijing, the capital city of China, was 377 $\mu g/m^3$, almost four times higher than the WHO guideline of 90 $\mu g/m^3$, and its SO_2 concentration level was 90 $\mu g/m^3$, almost double of the WHO guideline of 50 $\mu g/m^3$ (WHO 2002). SO_2 is also an important precursor of acid rain. From the 1980s to the mid-1990s, the area experiencing acid rain expanded by more than 1 million km^2 (Yang and Schreifels 2003). Elevated levels of air pollution are concerned to have substantial adverse impacts on human health. The World Bank (1997) estimates that air pollution in China in 1995 resulted in 178,000 premature deaths, 346,000 registered hospital admissions, more than 6 million emergency room visits, and more than 75 million asthma incidences - the aggregate loss accounted for 2.9 percent up to 7.1 percent of the Chinese GDP.[6]

[5]It is common worldwide both in developed and developing countries that coal-fired power plants are one of the largest contributors to domestic air pollution.

[6]See also Nielsen and Ho (2007) for various other impacts and Matus et al. (2012) for substantially larger welfare

An increasing public concern toward these direct health impacts drove the Chinese government to formulate a series of regulatory policies. The first version of the succeeding China's environmental policies was enacted in 1987, known as the Air Pollution Prevention and Control Law (APPCL). This original law, however, failed to reduce air pollution, mainly because it excluded the power industry, the major contributor of SO_2 emissions (Qian and Zhang 1998). Even worse, SO_2 emissions continued to surge and areas affected by acid rain expanded. APPCL was, thereby, amended in 1995. The major part of the amendment was to include a section to regulate pollutant emissions and coal combustion, particularly the usage of high sulfur-content coal, at the power plants (Hao et al. 2007).[7] Although the 1995 APPCL had a weak enforcement mechanism and limited efficacy, a prominent feature of the amendment was to propose a future regional strategy, where priorities to improve air quality and to prevent the spread of acid rain would be focused upon.[8,9]

This was officially approved and implemented as the Two Control Zones (TCZ) policy, in January 1998 (State Council 1998). This legislation designated prefectures exceeding nationally mandated thresholds as either the acid rain control zone or the SO_2 pollution control zone. Based on the records in preceding years, prefectures were designated as an SO_2 pollution control zone if

- average annual ambient SO_2 concentrations exceed the Class II standard[10],
- daily average concentrations exceed the Class III standard, or
- high SO_2 emissions are recorded.

Alternatively, prefectures were designated as an acid rain control zone if

- average annual pH values for precipitation are less than or equal to 4.5,

losses, when taking into account of the cumulative impact.

[7] This helped lay the foundation for a series of regulations that would follow. First, the Chinese National Ambient Air Quality Standards (CNAAQS) were formulated in 1996, and then the Law on Coal was enacted in August 1996.

[8] Article 27 of the 1995 APPCL stipulates that:

> The environmental protection department under the State Council together with relevant departments under the State Council may, in light of the meteorological, topographical, soil and other natural conditions, delimit the areas where acid rain has occurred or will probably occur and areas that are seriously polluted by sulfur dioxide as acid rain control areas and sulfur dioxide pollution control areas, subject to approval by the State Council.

[9] It is a standard practice of policy experimentation in China to implement strategies in a particular region or for a set of time period, attempting to demonstrate their effectiveness before expanding their implementation to the entire nation.

[10] According to CNAAQS for SO_2, Class I standard means the annual average concentration level not exceeding 20 $\mu g/m^3$, Class II ranges 20 $\mu g/m^3 < SO_2 < 60$ $\mu g/m^3$, and Class III ranges 60 $\mu g/m^3 < SO_2 < 100$ $\mu g/m^3$. Cities should meet Class II, which is considered to be less harmful.

- sulfate deposition is greater than the critical load, or

- high SO_2 emissions are recorded.

In total, 175 prefectures across 27 provinces were assigned as TCZ (Figure 2), which covered 11.4 percent of the territory, shared 40.6 percent of national population, produced 62.4 percent of GDP, and emitted 58.9 percent of the total SO_2emissions in 1995 (Hao et al. 2001). The SO_2 pollution control zone concentrates on the north due to high SO_2 emissions for heating,[11] whereas the acid rain control zone is primarily in the south, where heat, humidity, and solar radiation are suitable in creating atmospheric acidity for a given amount of SO_2.[12]

In TCZ, more stringent regulations were enforced to accelerate a shift away from high-sulfur coal and developing clean coal technology at power plants. For example;

- Any new coal mines, using coal with more than 3-percent sulfur content, cannot be established, and the existing ones are gradually shut down or required to reduce output.

- Construction of any new coal-burning thermal power plants inside large and medium-sized prefectures is prohibited.

- Any new and renovated old power plants are required to use coal with less than 1-percent sulfur content.

- Existing power plants using coal with above 1-percent sulfur content are required to install flue gas desulfurization (FGD) equipment.

Various studies have documented the effectiveness of these regulatory actions in reducing pollutant emissions and improving air quality within the TCZ. For example, National SO_2 emissions fell from 23.67 million tons in 1995 to 19.95 million tons in 2000, and the percentage of prefectures exceeding the Class II standard fell from 54 percent in 1995 to 20.7 percent in 2000. Among TCZ, SO_2 emissions fell by about 3 million tons, and about 71 percent of all factories over 100 tons of emissions per year reduced their SO_2 emissions to the standard between 1998 and 2000 (He, Huo,

[11] See Almond et al. (2009) for the impact of heating policy, which created a discontinuity in air quality across north and south of the Huai River.

[12] In this sense, acid rain in the south cannot necessarily be attributed to SO_2 emissions that travel from the north, but rather due to its local emissions. This is even more plausible because acid deposition is the greatest in the summer, when winds direct from south to north.

and Zhang 2002). By the end of 1999, collieries, producing more than 50 million tons of high-sulfur coal, had been closed (Hao et al. 2001). By the end of 2000, the total power capacity with FGD equipment exceeded 10,000MW, and small thermal power plants, sized below 50MW, were actively shut down because they were relatively less efficient, had high coal consumption rates, and emitted massive pollutants. This reduced 10 million-ton of raw coal consumption and 0.4 million-ton of SO_2 emissions (Yang et al. 2002). Between 1998 and 2005, prefectures in the SO_2 pollution control zone meeting Class II rose by 12.3 percent, meeting Class III increased by 4.2 percent, and not meeting Class III fell by 16.5 percent. In the acid rain control zone, prefectures meeting Class II rose by 3.3 percent, meeting Class III increased by 7.9 percent, and not meeting Class III decreased by 11.2 percent (United Nations Environment Programme 2009).

III. DATA SOURCES AND DESCRIPTIVE STATISTICS

Our main empirical analysis traces the evolution of infant mortality over time across the TCZ and non-TCZ prefectures. One of the advantages in our study is a comprehensive dataset on infant mortality for nationally representative sample over a decade. This adds external validity to the results. We first merge this dataset to the TCZ regulatory status. Subsequently, we also try to quantify the effect of air pollution on infant mortality, which additionally requires to link infant mortality data with, relatively limited, air quality data. This section describes each data source and provides their descriptive statistics. Web Appendix 1 provides a more detailed description of data construction.

III.A. Data Sources

Infant mortality. The micro-level data on infant mortality come from the Chinese Disease Surveillance Points (DSP) system. The DSP covers 145 sites, established on the representative sample of the national population (Figure 3). The primary sampling units are at the county-level.[13] It reports the censuses of death and birth registrations for the sample population of 10 million residences (approximately 1 percent of the national population) in different geographic areas

[13]The administrative divisions of China primarily consist of four levels: the province, prefecture, county, and township. The county level refers to either districts, county-level cities, or counties. 17 out of 145 sites are at the prefecture level, while the rests are at one of the county-level divisions.

across 31 provinces, autonomous regions, and municipalities in China.

The birth record reports whether or not infants died within the calendar year, and if they did, what the cause was, using the International Classification of Disease, 9th Revision (ICD-9) codes. In addition, it contains variables relevant for the analysis, i.e., infant's characteristics such as gender, the date of birth, birth weight, length of gestation period, and birth order as well as maternal demographics such as age, education, and race.

Overall, the original data record approximately 500,000 deaths and 1,000,000 births from 1991 through 2000, from which our observations are aggregated to the DSP site by year level. The number of population in the DSP site × year cell is retained so that all regressions are population-weighted. Details on how these 145 sites were selected, and how the data were processed at each site can be found in Web Appendix 1.

TCZ designation. The TCZ regulatory status is reported in the document "Official Reply to the State Council Concerning Acid Rain Control Areas and Sulfur Dioxide Pollution Control Areas," published by the State Council in 1998. The assignments are primarily made at the prefecture level, while those in municipalities are at the county level. The document lists the names of all prefectures that are assigned as the acid rain control zone and as the SO_2 control zone. Hence, we can merge this information to the DSP sites, and those sites that are in the TCZ prefectures comprise the treatment group in the analysis.[14] In total, 61 of 145 DSP sites are the TCZ prefectures, and 84 sites are the non-TCZ prefectures. More detailed designation rules are provided in Web Appendix 1.

Air pollution. The data on air pollution, as measured by the concentrations of TSP or SO_2, primarily come from the China Environment Yearbooks (CEY) published by the SEPA. The Chinese government established a national air pollution monitoring system to pay a close attention to the deterioration in urban air quality. Currently more than 350 prefectures are equipped with monitoring stations to monitor pollutants of TSP and SO_2. CEY reports annual daily average concentrations of TSP and SO_2 for major prefectures. The data between 1991 and 1995 are compiled from CEY as part of the "Economics of Pollution Control Research" project by the World Bank

[14]Because the regulations are the same across the SO_2 and acid rain control zone, we consider both areas as comprising one treatment group.

and are available on its website. The data between 1996 and 2000 are digitized directly from CEY by the author. We also extracted air quality data from the City Statistic Yearbooks whenever they are available.

The resulting matched sample has 586 observations for 74 DSP sites with at least a one-year observation of both TSP and infant deaths, and 613 observations for 75 DSP sites with at least a one-year observation of both SO_2 and infant deaths.

Economic and weather variables. We also draw several economic variables from the National Bureau of Statistics of China. Because changes in local air pollution conditions are often driven by local economic activities, a systematic association between the regulatory status and local economies may indicate that infant mortality reductions are driven by unobservable economic conditions other than air quality. The key variables include GDP per capita (in yuan), the number of hospitals (in units), annual supply of tap water (in 100 million tons), and per capita electricity consumption for living (in kilowatt hour).

The data on rainfall come from the Global Historical Climatology Network (GHCN) project. The pollution impacts may be biased without controlling for rainfall because rainfall is negatively correlated with air pollution, as rain washes away pollutants in the air. In turn, rain may itself be correlated with infant mortality through water pollution or formation of acid rain.

III.B. Descriptive statistics

Table 1 presents the descriptive statistics of the baseline sample in the pre-reform period (prior to 1995). In Panel A and B, the weighted average means are calculated using population as the weight, from the observations at the DSP site × year level. The table shows that the mean IMR is 16.16, using the death record. As expected, the birth record yields a much lower level of IMR, 7.61, because it reports the occurrence of infant deaths only within a *calendar* year.[15]

The main analysis uses the birth record because it also reports infants and households characteristics. Further, the birth record contains information on birth weight and length of gestation

[15] Note that the definition of IMR is therefore different when using the death record or the birth record. In death record, it measures the number of infants who did not make it to the age one, per 1,000 live births, whereas in the birth record, it measures the number of infants who died within the same calendar year they were born. For example, if an infant is born on January 1st, there is almost a year to observe her death during the first year after birth, but if an infant is born on December 31st, there is only one day to observe her death.

period, while the death record reports none of them. It is worth noting that since the main analysis focuses on the *changes* in the number of infant deaths, the underestimated *levels* of IMR in the birth record does not threaten the validity of using the birth record. A concern, if any, is that the impact may be understated if reductions in infant deaths are truncated at zero. As the table shows, several DSP sites report no occurrence of infant deaths in both death and birth records. This is not surprising, given that they have a small number of births. To address this issue, we repeat the main analysis using the death record as a robustness check.

The average TSP concentration level before 1995 was 314.4 $\mu g/m^3$. This is more than three times higher than the record in the U.S. in 1970, the year the Clean Air Act was amended. The average SO_2 concentration level was 94.7 $\mu g/m^3$, twice as high as the WHO standards. It is clear that the Chinese people were exposed to extremely high levels of air pollution in years preceding the policy reform.

IV. EFFECT OF ENVIRONMENTAL REGULATIONS ON INFANT MORTALITY

The main objective of this study is to assess the effect of air pollution regulations on infant mortality. In an ideal research setting, the TCZ status is randomly assigned across cities, whereby creates variation uncorrelated with baseline characteristics. In the absence of a randomized controlled trial, the next subsection describes an econometric framework that addresses a potential concern about the correlation between the TCZ status and unobserved characteristics. Then, we present the main results, evidence on the biological mechanism, and heterogeneity in the treatment effect. Lastly, we present evidence from robustness checks.

IV.A. Empirical Framework and Its Validity

The main analysis uses a difference-in-differences approach;

$$Y_{jt} = (T_j * Post_t)\pi_1 + X_{jt}^{'}\delta_1 + \kappa_t + \mu_j + \varepsilon_{jt}, \tag{1}$$

where Y_{jt} is IMR in city j in year t, T_j is an indicator variable that takes on the value one if city j was assigned as a TCZ in 1998,[16] and $Post_t$ is an indicator variable that takes on the value one for being born in the post-reform period. The city fixed effects, μ_j, control for the permanent heterogeneities across cities, whereas the year fixed effects, κ_t, control for year-specific shocks that are common to both TCZ and non-TCZ cities. X_{jt} controls for an additional set of covariates that capture birth, parental, and city characteristics at the city $\times$ year level. The city-level time-varying disturbances, ε_{jt}, are clustered at the city level, allowing for an arbitrary correlation within cities over time in computing the standard errors.

The main analysis uses 1998 as a cut-off year for the post-reform period, rather than 1995 when the APPCL was amended, based on theoretical and empirical rationales. For theoretical motivations, it is plausible to take two to three years before the regulations are carried out to the full extent.[17] This is relevant in the current context because the regulations required power plants to alter the energy sources and install heavily costly technology (such as FGD).[18] Time-lags are also likely in China, where it is common for the government to set targets or outlines of policies, often very ambitious ones, without specifying the critical details until later, thereby extensively leaving implementations up to the local governments or individual firms. As mentioned earlier, the 1995 amendment had a weak implementation mechanism, and more drastic actions (such as shutting down numerous inefficient power plants and enforcement of stringent air pollution regulations) were enforced only after the TCZ policy in 1998. To the extent that the regulations were somewhat in effect between 1995 and 1997, using 1998 as a cut-off year only results in understating the true effect. On the other hand, there are also empirical rationales. The interaction term between

[16] Note that the non-TCZ cities are not equivalent to say "non-affected" cities. It is more appropriate to say that the regulations were more stringent in the TCZ cities, relative to the regulations in the non-TCZ cities. In such a context, it may be ideal for T_j to measure a continuous intensity of exposure to the regulations, measured by either per-capita amount of high-sulfur coal used or by per-capita amount of SO_2 emissions in the baseline years. A similar strategy is used in Qian (2008), where she uses the amount of tea planted in each county in China to measure the impacts of relative female income on sex ratio, where women have comparative advantage in picking tea, and the price of tea was dramatically increased by the post-Mao reforms. Also, Bleakley (2007) uses the pre-treatment hookworm infection rate to measure the benefits of hookworm eradication on school enrollment. Unfortunately, data limitation makes this infeasible, and we decided to use a discrete choice of being in the TCZ cities or not. This is still valid because the large amount of high-sulfur coal and SO_2 emissions were produced in the TCZ cities. To the extent that regulations affected the non-TCZ cities, the treatment effect is understated.

[17] Chay and Greenstone (2005) are based on a similar argument when they use 1975 nonattainment status as an IV in estimating the effect of 1970 Clean Air Act on housing prices between 1970 and 1980. In their context, the nonattainment status changes every year. By using the mid-decade regulation, they also take into account a two- to three-year lag before the policy was fully executed.

[18] Informal conversations with officials at local power plants provide anecdotal evidence to support this assertion.

the TCZ status and Post-1998 period better balances important determinants of infant mortality, compared with the one using 1995 as a cut-off year, suggesting that the former one serves as a better IV that meets the exclusion restriction[19] Nonetheless, robustness checks shown later using 1995 as the cut-off year or controlling for the "intermediate" period between 1995-1997 does not alter the main results.

The DID estimate of π_1 measures the reduced-form impact of air pollution regulations on infant mortality, capturing the difference in the changes in IMR before and after the regulations, between the cities that were and were not designated as a TCZ in 1998. If the air pollution regulations contributed to significant reductions in IMR in the TCZ cities relative to non-TCZ cities, π_1 is expected to be negative. The key identification assumption for a causal inference is that the non-TCZ cities provide valid counterfactual changes in infant mortality for the TCZ cities, had they not been treated. Two potential hypotheses may violate this assumption: (1) there is a systematic difference in pre-exiting trends in mortality reductions, and/or (2) the TCZ status is not orthogonal to factors explaining the reductions in infant mortality in the post-treatment period.

To address the pre-existing trend, I plot the evolution of IMR over time between the TCZ and the non-TCZ cities in Figure 4. The dotted vertical line indicates 1995, when the APPCL was amended, and the solid vertical line indicates the timing of TCZ policy implementation in January 1998.[20] The figure provides graphical support that IMR trends were similar in the pre-intervention period; infant mortality fell between 1991 and 1993, stayed somewhat constant until 1996, and fell again in 1997 in both sets of the cities. A trend break appears around 1998, when IMR continued to drop only among the TCZ cities, while IMR were somewhat constant within the non-TCZ cities. Notably, although the levels of IMR were continuously higher in the TCZ cities than those in the non-TCZ cities before 1998, this phenomenon is reversed after 1998, which is

[19] A systematic association between the interaction term and observable characteristics may indicate that the estimates are confounded by other factors. Chay and Greenstone (2005) also argue that the mid-decade nonattainment status is less correlated with observable variables and thus is a better candidate for an IV, compared with the early 1970s nonattainment status. In our setting, the interaction term is correlated with birth order and percentage of Han at the ten percent significant level (Table2), when using 1998, but when using 1995, maternal age, in addition to these two variables, is also correlated. Further, using a 1998 cut-off year enables us to restrict the sample to observations between 1996 and 2000, where the interaction term is not correlated with any observable variables. Therefore, using 1998 better averts omitted variable bias.

[20] Because each observation represents the annual average value, the dotted vertical line is located at 1995 (since the APPCL was amended in August, nearly in the middle of the year), while the solid vertical line is located between 1997 and 1998 (since the TCZ policy was implemented in January, early in the year). This is to clarify the timing of the TCZ policy implementation.

commensurate with the timing of the TCZ policy. A similar pre-trend between the two sets of the cities is likely to suggest that the post-trend would have been also similar in the absence of the regulations. More rigorous regression-adjusted evidence in robustness checks below also show that there is little systematic difference in trends in infant mortality up to 1998, and that the treatment effect is also robust to city-specific time trends. See Web Appendix 3 for further evidence on a trend break.

Another concern is that the regulations effect may be confounded by other concurrent changes in policies or factors affecting infant deaths. If the central government assigned the TCZ status solely based on the nationally-mandated standards, the designation should be less likely to be correlated with demands from local governments. However, because the TCZ status is based on air pollution level, and air pollution level is characterized by various local conditions, it may still be possible that the TCZ status simply proxies these other factors. For example, air pollution may be high in urban places where the number of births is decreasing due to high women's labor force participation or relatively more stringent one child policy. In this case, the negative association between the TCZ status and IMR would be confounded by the quantity-quality tradeoff. To address this concern, I investigate whether the TCZ status has any association with changes in observable characteristics. Although this is not a formal test of exclusion restrictions, as the assumption states that the treatment status should not covary with *unobservable* characteristics, a lack of significant correlation with observable characteristics suggests that there should not be significant correlations with unobservable variables either (Altonji, Elder, and Taber 2005).

Columns (1)-(2) of Table 2 examine potential associations between the TCZ status and levels of characteristics in the pre-intervention period. Each entry reflects the coefficients of TCZ status from separate regressions, when using respective characteristics as the dependent variable. Hence, each entry indicates the (average) difference in the weighted means of respective variable between the two sets of the cities. Significant correlations between the TCZ status and observable characteristics indicate that the two types of the cities are distinct in many dimensions. In particular, it shows that the TCZ cities tend to have a lower male proportion and a lower probability to be born in January. The latter may potentially bias the estimated regulations effect in the current research context, because their deaths are more likely to be recorded. Other differences in characteristics are typical features of urban areas, such as older mothers, greater mother's educational attainment, a larger share of Han, and a lower number of infants. Most differences remain significant even during a shorter time period: 1996-1997, which corresponds to the pre-intervention years during the 9th Five-Year Plan period (1996-2000). A possible explanation for few discrepancies among economic characteristics is that these variables are available only for urban areas and after 1996. Taken together, these findings suggest that there are significant differences in levels of characteristics between the TCZ and the non-TCZ cities, and thus the cross-sectional relationship is likely to suffer from severe omitted variable bias.

Column (3) addresses potential differences in trends of characteristics between the TCZ and non-TCZ cities, before and after the policy change. Each entry reflects the coefficient of the interaction term, π_1, in equation (1) from separate regressions, when each covariate is used as the dependent variable. Any significant estimate indicates a systematic difference in trend patterns and may confound the regulation effects. The table sheds light on little systematic difference in trends. Further, most of the point estimates become substantially small and virtually indistinguishable from zero. Importantly, the TCZ status balances the trends in important predictors of infant deaths, such as mother's age or educational level as well as the percentage of male and share of January birth. A little association with local economic conditions' trends suggests that underlying economic shocks are less likely to be a source of bias. In column (4), the sample is restricted to a shorter period, 1996-2000, which serves as a robustness check, because a number of policies are fixed under the 9th Five-Year Plan. In this case, none of the variables presents significant differences in trends.

Overall, these findings provide compelling support that the TCZ regulatory status is orthogonal to trends in observable determinants of infant deaths, reassuring that the current research design is less likely to be biased by changes in unobservable variables. A falsification test and robustness checks below provide further evidence.

IV.B. Empirical Results

1. Effect on infant mortality

We present the DID estimates of the regulations effect on infant mortality, by causes of death, in Table 3. Panel A reports the effect on infant mortality from all causes of death. Column

(1) provides the result without any controls except city fixed effects and year effects, indicating that the TCZ status is associated with 2.87 fewer infant deaths per 1,000 live births. Column 2 controls for a number of birth, parental, and city characteristics, directly addressing a concern that changes in infant mortality may be explained by changes in observable characteristics that vary over time and that are correlated with the TCZ status. Namely, it controls for the percentage of male, share of births in every month, birth order, mother's age, mother's education, percentage of Han, total number of births, total population, and precipitation rates. Because infant deaths for those who were born in earlier of the year are more likely to be in the data, the share of births across months controls for the possibility that the impact of air pollution regulations is driven by correlated changes in birth months. With these controls, the estimated coefficients become larger and are statistically significant at the one percent level. Column (3) additionally controls for birth weight and gestation periods, in an effort to control for conditions at birth. The estimated effects become even larger and remain highly significant.

A major concern with any non-experimental studies is that there may be omitted heterogeneity that gives rise to a spurious relationship. Controlling for the attributes, as in Column (2) and (3), may not solve this issue in case when there are no comparable non-TCZ cities to TCZ cities over the distribution of the attributes. We address this issue by focusing on the observations on common support of the propensity score, which potentially restricts the sample to cities that have similar observed characteristics. The estimated effect of the regulations on infant mortality, in Column (4), slightly increases and stays significant at the one percent level.

A related concern is that the regulations are confounded by other concurrent changes in factors contributing to the reductions in infant mortality. For example, in response to high exposure to pollution in the TCZ cities, the local governments may have increased healthcare spending, leading to an improvement in quality and/or quantity of healthcare services. Then, without directly controlling for a local health policy, a simple DID estimate would erroneously pick up effect through a healthcare policy, causing an upward bias. In order to rule out such a possibility, we examine the regulations effect on infant mortality by cause of death in Panel B and Panel C in Table 3. Because identifying exact diseases caused by air pollution has been difficult, we compute IMR based on internal deaths, the ones that are potentially associated with air pollution, aside infant deaths due to external causes of deaths, most obvious ones that are not pertaining to air pollution: injury and poisoning. Table A1 lists all causes; slightly more than half of infant deaths are due to diseases of the circulatory system, 19.4 percent comes from nervous systems and sense organs, and 19 percent is due to external causes. Low rates of respiratory diseases indicate that these are not typical causes of deaths for infants who tend to spend most of their time at home, while these diseases have been found to be a major cause of deaths for children.

We investigate whether the associations between the regulatory status and infant mortality reductions come from internal causes of deaths in Panel B. The point estimates are essentially unchanged and constantly statistically significantly different from zero. This suggests that the regulations are associated with reductions in infant mortality due to internal causes. The preferred estimate in Column (3) suggests that 3.33 fewer infants died per 1,000 live births in the TCZ cities than would have died in the absence of the regulations.

On the other hand, Panel C shows that the regulations have little association with infant deaths due to external causes of deaths. If households had had improved access to healthcare services, it is more plausible to believe that they would have helped reduce infant deaths by external causes as well. The point estimates are substantially lower and are not statistically different from zero, even under observations on commons support. Although this does not directly rule out a possibility that the regulations spuriously pick up the effect through unobserved changes that are correlated with internal causes of deaths but not with external ones, the finding suggests that the main results are not driven by any other channels, leading us to believe that the relationship is causal. We provide a number of further robustness checks in the next subsection.

2. Identifying the biological mechanism

In an effort to identify the biological mechanism through which air pollution affects infant mortality, we examine whether the regulations are associated with birth outcomes, namely birth weight and the length of gestation period. Maternal exposure to pollution during pregnancy is considered to retard fetal development, as absorbed pollutants into her blood limit nutrition and oxygen flows to fetuses. This sometimes results in low birth weight or shorter length of gestation, although exact channels or outcomes are not well-known (Dejmek et al. 1999; Perera et al. 1999).[21]

[21] For adult, air pollution is often linked to respiratory and cardiovascular disease, aggravation of asthma, heart disease, lung malfunction or cancer, stroke, and possibly carcinogenesis. For infants, high exposure to pollution in the

In Column (1), the dependent variable is the birth weight in grams. It shows that those who were born in the TCZ cities weighed more by 20.8 grams, yet it is not statistically significant. This result should not be interpreted that affected babies experienced no better in-utero environment. As shown in Table 1, the baseline birth weight was already high in China, and birth weights can be often controlled by parental preferences. Rather, a bigger issue is children born at low birth weight, which is often considered to have lasting impact on later health and socioeconomic status.[22] Hence in Column (2), we evaluate the effect on the share of births at low birth weight; the dependent variable is the number of births below 2,400 grams, the lowest one percent of birth weight, per 1,000 live births. The result shows that affected babies are less likely to be born at low birth weight by 3 infants per 1,000 live births. Columns (3) and (4) examine the corresponding impact on the length of gestation period and the share of gestation period below 32 weeks, the lowest one percent level. Neither estimates are significant, and this provides additional evidence that the impact on low birth weight is not driven by changes in length of gestation period.

The results above suggest that fetal exposure to pollution affects fetal development, particularly birth weight at the low tail. However, birth weight and length of gestation period may not reflect the entire effects on fetal development. Hence, we now explore the effect on deaths occurring different time periods. Infant deaths during the neonatal period (usually within 28 days after birth) are considered to be associated with poor fetal development.

Column (5) presents the impacts on infant deaths occurring within one day. The point estimate is small and indistinguishable from zero and implies that 16 percent of overall reduction in IMR occurred within one day.[23] The magnitude is in line with Chay and Greenstone (2003b), which estimates that roughly 22 percent of overall infant deaths occurs within one day. Yet, this contrasts to the finding in Chay and Greenstone (2003a), which attributes roughly 60 percent of overall impact to infant deaths within one day. Column (6) reveals that the regulations are disproportionately more associated with the probability of death within one month after birth. The estimates imply that 61 percent of the effect of the regulations on infant mortality is due to reductions in

neonatal period is likely to result in deaths by acute respiratory infections. Some recent studies show strong impacts coming from maternal exposure.

[22]See Behrman and Rosenzweig (2004), Case, Fertig, and Paxson (2005), and Almond, Chay and Lee (2005) for costs and returns of low birth weight.

[23]The point estimates are not exactly comparable to Table 3 because we eliminate the outliers in Table 4 (as a robustness check to these values). The comparable estimate is -3.00 (0.85) for the coefficient (standard error).

neonatal period. This corresponds to the findings in Chay and Greenstone (2003a) and Chay and Greenstone (2003b), whose 73-82 percent and 80 percent of infant mortality reductions occur in the neonatal periods, respectively. Overall, these findings emphasize weak fetal development via maternal exposure as an important biological mechanism.

3. Heterogeneity in the regulations effect

This part tests the hypothesis that the regulations on air pollution reductions may have a heterogeneous impact on infant mortality across groups, and the results are reported in Table 5. We consider whether the effect varies on infant mortality in Column 1, on the probability to be born at less than 2,400 grams in Column 2, and on the probability to be born in less than 32 weeks in Column 3. The first row presents estimated effect in the entire sample for the comparison purpose.

We first search for heterogeneity in the treatment effects between boys and girls. This may be the case, when there are biological differences by gender. In the literature, male fetuses are considered to be more physiologically sensitive than female fetuses to environmental changes. On the other hand, it may also reflect gender discrimination, particularly when infant deaths occur after birth. For example, if boys were initially more likely to be protected from pollution exposure or to receive medical treatments for health problems caused by pollution, then the effect of pollution reductions would be pronounced for girls. The second and third row report the estimated effects for boys and girls samples, respectively. The coefficients suggest that the effect on infant mortality is larger among girls than boys, yet the difference is not statistically significant. This indicates that pollution reduction effect on infant mortality was balanced between boys and girls. In contrast, the effect on the probability to be born at low birth weight is substantially higher for girls than for boys, and is statistically significant only for girls. This is consistent with the literature that female fetuses are more robust to changes in fetal environment, and thus they are affected more when pollution decreases from the high initial levels, compared to male fetuses. On the other hand, we do not find any significant effects on the length of gestation period.

Next, we study whether the effect differs by mother's educational level. The regulations effect may be amplified among households with low mother's education for two reasons. First, children of low mother's education tend to have lower initial health endowment, making them more susceptive to air pollution reductions. Second, mothers with greater educational attainment are more likely to have knowledge to protect their children from being exposed to pollution outside and/ or have more access to health services to treat their children. The fourth row reports the estimated effects for the sample of households with mother's education being less than a high school degree, and the fifth row for the sample of households whose mothers attained at least high school education. We find that the regulations effect is substantially higher among households with low mother's education, whereas the estimate drops close to zero and is statistically indistinguishable from zero for households with high mother's education. The finding suggests that the regulations effect on infant mortality should be stronger for the low-socioeconomic families that are more vulnerable to the effect of air pollution. On the other hand, there is little evidence that the regulations effect is different for birth weight and gestation period by mother's educational levels.

IV.C. Robustness Checks

The findings above leave little room for the scope of confounding factors. First, little association between TCZ status and trends in observable characteristics limits a possibility that the main results erroneously reflect time trends that vary systematically between the TCZ and the non-TCZ cities. Second, the consistency of the regulations effect in both magnitude and statistical significance when controlling for the set of key determinants of infant mortality or when restricting the sample on common support suggests that the estimated effects are robust to comparisons with similar characteristics. Third, the absence of treatment effect in the falsification exercise on external infant mortality provides strong evidence that health care system reform or medical technology advancement cannot be a source of bias.

In this subsection, we explore further robustness checks to rule out five other alternative stories. First, we attest that the finding is robust when using a different dataset. As discussed above, using the birth record, which reports the occurrence of deaths only within the calendar year, may understate the effect, if the number of infant deaths is truncated at zero. The death record allows us to compute IMR for all deaths occurring before age of one. As expected, the size of the estimate becomes larger, though not substantially different from the main result (Column (1) of Table 6). This shows that the effect is not sensitive to using the death record, and the main analysis may understate the overall impact, if any.

Second, we confirm that the treatment effect is not driven by other national or local policy changes. In Column (2), we restrict the sample to years between 1996 and 2000, which corresponds to the period of the 9th Five-Year Plan. The shorter time and being under the same policy plan help reduce a set of potential confounders in pre- and post-natal health environment other than pollution. Further in Column (3), I control for province $\times$ year effects, purging impacts through time-varying shocks at the provincial level. The estimated effects are essentially unchanged over these specifications, suggesting that various other national or local policy changes should not confound the effect.

Third, despite the fact that the treatment effect is significant for the DSP sites with similar characteristics, as in Table 3, there is still a concern that the TCZ status may be correlated with administrative divisions. For example, urban districts and county-level cities may be more likely to be treated, whereas poor counties may be less likely to be assigned as TCZ. To address this issue, we restrict the sample to only districts and county-level cities, mostly urban areas, in Column (4). The estimate is larger and remains significant at the five percent level, reassuring that the treatment effect is not driven by simple comparisons between urban and rural areas, or it does not simply pick up effects across administrative divisions.

An interesting disparity is found when the sample is restricted to counties, where we compare counties in the TCZ prefectures and counties in the non-TCZ prefectures, yet neither one is treated.[24] Given the proximity of these counties to the districts/county-level cities within the same prefecture, and given the scientific evidence that pollutants travel for a long distance within a short period[25], one would expect that air pollution reductions should have positive externalities to neighboring areas, as less pollutants spread out. However, the finding indicates that infant mortality increased in counties neighboring the TCZ cities, compared to the one in counties neighboring the non-TCZ cities. This may show that enforcing air pollution regulations may have negative externalities to surrounding areas, by providing polluting industries to increase production in areas with less stringent regulations.[26] However, these interpretations should require caution, because

[24]The TCZ status was assigned at the prefecture level, yet the regulation exempted impoverished counties even within the TCZ prefectures. See the Web Appendix 1 for details.

[25]The lifetime of TSP and SO_2 are considered to be up to one week. They can travel from 100 km to 1,000 km within these life spans.

[26]Empirical evidence to this so-called pollution haven hypothesis is rather limited. In the field of international

the estimate is not statistically significant. Micro-level data at the plant level would be able to answer some of these questions, which I leave as a future direction of research.

Fourth, the main analysis uses the 1998 as the cut-off year, focusing on the TCZ policy, yet a series of policies implemented after 1995 APPCL amendment may have been effective. This may not be a concern, since in this case the main analysis simply produces a lower bound. Nonetheless, we repeat the analysis that take accounts of these policies. In Column (6), we directly control for the period between 1995 and 1997, the intermediate period between the 1995 APPCL amendment and the implementation of the TCZ policy in 1998, finding the treatment effect after 1998 to stay highly significant. In Column (7), we use 1995 to re-define the post-reform period, instead of 1998, to account for a trend break around 1995, which coincides with the APPCL amendment. The treatment effect continues to be significant at the conventional level. Note that these pieces of evidence do not rule out a possibility that the amendment became effective with a few years lag. They still indicate that the trend break between the TCZ and non-TCZ cities becomes evident after 1998, in accord with the start of the TCZ policy.

The collection of robustness checks above substantially limits the scope of omitted variables, yet one may be still uncertain whether infant mortality itself may be trended systematically. Although this is not apparent from Figure 4, we now provide formal regression-adjusted evidence. In Column (7), we restrict the sample to pre-treatment period of 1991-1997. The estimated difference in trends is substantially small and insignificant, highlighting no systematic difference in trends up to 1998 (and see Web Appendix 3 for year-level analyses, which describe substantial changes in trends starting in 1998.). Lastly in Column (8), we directly control for the DSP site-specific linear time trends. Although the significance drops possibly due to a lower power of the statistical estimation, the point estimate is essentially unchanged. Taken all above together, there is little evidence to support that the main results are driven by inappropriate identification assumptions, leading us to

trade, Antweiler, Copeland, and Taylor (2001) show that international trade does not alter pollution concentration. Instead, trade helps disseminate technology and expand production scale, both of which contribute to a reduction in pollution. Eskeland and Harrison (2003) find similar results that foreign firms tend to pollute less using clean energy and energy efficient technology, compared to other domestic firms.

Some evidence from within a country shows that firms choose to locate or relocate based on regulations. For example, the Clean Air Act in the U.S., which imposed high costs for firms with severe regulations in nonattaiment counties that exceed the pollution standard, is found to have led to the relocation by polluting industries from nonattainment areas to attainment areas (Henderson 1996; Becker and Henderson 2000), and a reduction in births of new polluting industries in nonattainment areas (Becker and Henderson 2000). In a developing country, Kumar and Foster (2009) show that air pollution regulation in Delhi led to an increase in pollution in its neighborhoods.

believe that the relationship is causal.

V. QUANTIFYING THE AIR POLLUTION EFFECT ON INFANT MORTALITY

Another objective of this study is to quantify the effect of air pollution on infant mortality, using the matched sample where both IMR and air pollution data are available. Identifying the credible relationship between the two has been hampered by unobservable variables that correlate with air pollution and that also determines mortality. The TCZ regulatory status, which is shown to have little association with changes in various other determinants of infant mortality except via air pollution, naturally serves as a potential candidate for an IV. In the following subsection, we first describe cross-sectional and fixed-effects models, both of which are pervasive in the literature and serve as benchmark, yet we explain why their critical assumptions are less likely to hold. We then discuss how the current research design based on the TCZ status as an IV provides a rare research opportunity to circumvent identification problems. Then, the subsequent subsection provides empirical results.

V.A. Empirical Framework

Cross-sectional model. A large number of literature relies on cross-sectional models in estimating air pollution impacts;

$$Y_{jt} = \beta_0 P_{jt} + X_{jt}^{'}\gamma_0 + u_{jt}, \tag{2}$$

where Y_{jt} is an outcome variable, such as infant mortality, and P_{jt} is a pollution level in locality j at time t. Unobservable disturbances of Y, u, can be denoted as; $u_{jt} = \mu_j + \xi_{jt}$, where μ_j represents the permanent characteristics of locality and ξ_{jt} includes transitory shocks. The coefficient of interest, β_0, can be consistently estimated by the least-square estimation based on the assumption that $E[P_{jt}u_{jt}] = 0$.

However, cross-sectional estimates are likely to be biased because the differences in localities that are permanent (μ_j) or transitory (ξ_{jt}) are likely to explain much variation in IMR and air pollution.[27] The estimates are upwardly biased if high-air pollution localities are also characterized with other types of pollution (e.g., water pollution or hazardous waste). The estimates are downwardly biased if polluted localities concentrate on urban, industrial areas, where people are relatively wealthier. Thus, the cross-sectional models are less likely to provide unbiased estimates.

Fixed-effects model. A common approach in a panel setting is fixed-effects models, as a way to purge any time-invariant characteristics particular to a locality over time. Then, the equation (2) becomes:

$$\Delta Y_{jt} = \beta_0 \Delta P_{jt} + \Delta X'_{jt}\gamma_0 + \Delta\xi_{jt}, \tag{3}$$

where Δ indicates the demeaned values of respective variables. The consistency of β_0 requires $E[\Delta P_{jt}\Delta\xi_{jt}] = 0$, suggesting that there should be no unobservable transitory shocks that drive changes in both Y_{jt} and P_{jt}. This approach is powerful when time-invariant characteristics explain much variation in an outcome variable.

Unfortunately, this assumption is less likely to hold in the context of air pollution effect because changes in air pollution levels are largely driven by changes in underlying economic conditions, which are also associated with changes in mortality.[28] The bias could go in either direction. For example, Chay and Greenstone (2003a) show that the reductions in air quality between 1981 and 1982 in the United States are strongly associated with recession and a fall in income level. Hence, a lack of ability to adequately control for income variables understates the pollution impacts. The omitted variable bias is more critical in the studies in developing countries because many variables, notably income variables, tend to be noisier and are difficult to be controlled for. On the other hand, if air pollution changes are positively correlated with changes in other pollutions, the coefficient is overstated. Thus, fixed-effects models are still likely to suffer from severe omitted variable bias.

[27] Recall in Table 2, we show that the TCZ status, which can essentially proxy the initial pollution level, has strong correlations with various time-varying and time-invariant characteristics.

[28] Currie and Neidell (2005) and Currie, Neidell, and Schmieder (2009) rely on fixed effects models but seek to reduce this type of problem by using weekly variation. Yet, data on such frequent fluctuations in air quality are rarely available in developing countries.

Instrumental variable fixed effects (IV-FE) model. I seek to overcome the identification problems described above by using the TCZ regulatory status to instrument air pollution;

$$Y_{jt} = \beta_1 \widehat{P}_{jt} + X'_{jt}\gamma_1 + \kappa_t + \mu_j + \epsilon_{jt} \tag{4}$$

$$P_{jt} = (T_j * Post_t)\pi_2 + X'_{jt}\delta_2 + \kappa_t + \mu_j + \eta_{jt}. \tag{5}$$

The validity of the framework above rests on two identification assumptions. First, the regulations were effective in reducing air pollution. Various studies provide evidence that this was so. We can directly test this hypothesis by looking at whether π_2 is statistically significant or not. If the air pollution regulations contributed to significant reductions in air pollution, π_2 is expected to be negative. Second, the regulatory status should not be associated with any other determinants of infant mortality other than air pollution. Although this cannot be directly tested, the findings in the previous section support this. We also repeat the balancing test and some of the robustness checks for the matching samples in Web Appendix 3. Note that balancing observables does not require a specific functional form for consistency. This feature is particularly useful in the context of air pollution impacts because there is no conventional functional form in the literature. Otherwise, mis-specifications would also lead to biased estimates. To the extent that the IV is valid, β_1 can be interpreted as a causal impact of air pollution on infant mortality, purged of omitted variables bias.[29]

V.B. Empirical Results

We first present results based on the cross-sectional relationship between air pollution and infant mortality, using equation (2) in Panel A of Table 7. Air pollution is measured by the concentration levels of TSP in Columns (1)-(4), and of SO_2 in Columns (5)-(8). It shows that infant mortality is not associated with air pollution based on readings of either TSP or SO_2. The signs of the estimates alter, depending on the controlled variables, and none of the estimates is significant. A lack of significance and stability in cross-sectional estimates suggests that the findings in conventional literature using a similar approach may also suffer from severe omitted variable bias.

[29]Note that because the IV is exactly identified, the 2SLS estimate must be the ratio of the two reduced-form estimates ($\beta_1 = \frac{\pi_1}{\pi_2}$), where π_1 comes from equation (1). However, this is not the case when comparing the estimates in Table 3 and Table 7 because we can only assess the air pollution effect for restricted samples of the DSP sites.

In Panel B, we examine the within-city relationship between air pollution and IMR, using the fixed-effects model of equation (3). Compared to cross-sectional estimates, the fixed effects estimates based on the TSP matched sample show less sensitivity to different specifications; they have expected signs and more stable point estimates. The preferred estimate in Column (4) indicates that a 1-$\mu g/m^3$ reduction in TSP is associated with 0.01 fewer infant deaths per 1,000 live births, although the estimate is not significant. The findings based on SO_2 appear less intuitive. The estimate in Column (8) indicates that air pollution is *negatively* associated with IMR, and the relationship is statistically significant. This suggests that fixed-effects estimates may still be largely biased by substantial unobservable disturbances and may not provide causal inference.

The findings above highlight the important role of omitted variables in biasing conventional estimates. Our IV-FE model seeks to overcome this issue by exploiting variation in air pollution induced by the regulations. The second row of Panel C presents the first stage results, where the coefficients capture relative reductions in air pollution in the TCZ cities in the post-reform period, relative to the non-TCZ cities. We find that the TCZ cities achieved significant improvements in air quality relative to the non-TCZ cities. Using the weighted average of point estimates across specifications, with the inverse of sampling error as the weight, the regulations resulted in 53 $\mu g/m^3$ more reductions in the TSP concentration level in the TCZ cities, and the estimates are highly significant and robust. In addition, the TCZ cities experienced 20 $\mu g/m^3$ more reductions in the SO_2 concentration level, and the estimates are significant at conventional levels.

The first row presents the 2SLS impact of air pollution on infant mortality. The weighted average estimate suggests that a 1-$\mu g/m^3$ reduction in TSP level leads to 0.049 fewer infant deaths per 1,000 births. Importantly, the estimates are essentially unchanged to various specifications and stays highly significant. This adds robustness to the results. The findings also indicate that a 1-$\mu g/m^3$ reduction in SO_2 level leads to 0.14 fewer infant deaths per 1,000 births. However, the estimates in Column (7) and (8) cannot be overemphasized because the first-stage results are not strong enough. The difference in the extent of regulations effect on TSP and SO_2 may not be surprising because controlling TSP, which is relatively larger pollutants, is somewhat easier and less costly, than controlling gaseous SO_2, which takes longer time to develop or install costly technologies.

The comparison of estimates between fixed effects and IV-FE models reveals that the estimates are substantially and consistently higher in the IV-FE model than corresponding FE estimates, suggesting that fixed-effects estimates are negatively biased. This is consistent with Chay and Greenstone (2003b) and Luechinger (2009), both of which find IV estimates to be larger than the fixed effects estimates. The ratio of 2SLS estimates to the corresponding FE estimates is approximately 3.4, which is larger than 1.5 in Chay and Greenstone (2003a) and 1.9 in Luechinger (2009). This highlights a difficulty in controlling for appropriate variables in the studies in developing countries, most notably income levels. More importantly, the 2SLS estimates are far less sensitive across specifications than those in cross-sectional or fixed-effects models, leading us to believe that the relationship is not spurious but causal.

VI. MAGNITUDE OF THE OVERALL BENEFITS

This section aims to interpret the findings above, in an effort to estimate the overall magnitude of the benefits associated with the regulations and air pollution reductions in China. Although a formal test of cost-and-benefit analysis is far beyond the scope of this paper, we compare the findings in China with developed countries and discuss the effectiveness of environmental regulations in developing countries, where air pollution levels are already high.

In the medical literature, a dose-response relationship between air pollution and infant mortality is not known. Benefits of pollution reductions may be limited if there is a threshold above which pollution becomes detrimental, and if pollution moves in the range above the threshold. Thus, we compare our results to other studies based on lower initial pollution levels (Table A2). The comparison shows that the impact of a unit change in TSP in this study is close to the estimates in the literature, indicating that TSP has a similar marginal effect in China.[30]

This translates into substantially higher elasticities in this study, because initially more infants died from pollution (which drives the numerator high), or because a unit change accounts for a smaller percentage for a larger initial pollution level (which brings the denominator low), or both. This highlights substantially greater benefits of the regulations associated with reductions in air pollution when pollution levels are initially high. It is also likely that the benefits are understated

[30] The marginal effect of SO_2 appears to be larger in this study, compared to Luechinger (2009), yet the interpretation of SO_2 needs to be careful. It may indicate an upward sloping dose-response curve for SO_2, or it may be driven by relatively sensitive estimates in the analysis. See the discussions in the previous section.

because they do not take account of the forgone medical costs due to health problems caused by pollution or long-term impacts of early childhood health.

According to the estimates, the regulations led to 3.33 fewer infant deaths per 1,000 live births, which represents a 21 percent reduction in IMR. Multiplying this by 7.62 million infants, the average number of infants born per year among the TCZ cities between 1998 and 2000, suggests that the regulations saved approximately 25,400 infant lives every year. Further, the 2SLS estimates show that a 1-$\mu g/m^3$ decrease in TSP resulted in 0.049 fewer infant deaths per 1,000 live births. The back-of-the-envelope calculations based on the size of the reductions in TSP concentrations from its peak year to 2000 during the last decade indicate that TSP saved 9,439 infants per year.

The absolute size of benefits itself does not justify the policy implementation, without considering costs. Although the overall costs of the regulations are not known, it is presumably reasonable to believe that reducing a given amount of air pollution is relatively less costly at high levels than at low levels. Further, we find larger impacts on households with low socioeconomic status, which should be the target of the development policy. This study is informative by providing benchmark benefits of environmental regulations and pollution reductions in developing countries, where pollution levels are high.

VII. CONCLUSIONS

Air pollution in China is notorious, and its health effect has been an increasing public concern. Since 1995, China has implemented a series of environmental regulations. Among them, the TCZ policy, implemented in 1998, was one of the largest-scale air pollution regulations in developing countries, imposing stringent regulations on pollutant emissions from power plants in cities exceeding the nationally-mandated standards.

The major objective of this paper is to test the hypothesis that these regulations led to reductions in infant mortality within the TCZ cities subjected to particularly stringent regulations. Using the difference-in-differences approach, comparing changes in infant mortality between the cities assigned and not assigned as the TCZ, before and after the policy reform, we find substantial impacts; infant mortality reduced by 21 percent; a large fraction of the reduction occurred in the neonatal period, which can be attributed to fetal exposure; and infants from families with low mother's educational attainment benefited the most.

The set of falsification test and robustness checks limits the role of omitted variables in biasing these estimates, leading us to believe that the linkage between the air pollution regulations and infant mortality reduction is causal. First, we confirm that the treatment effect is absent for infant deaths caused by external causes, ruling out a potential mechanism through improved local healthcare system. Second, the regulatory status is found to be less associated with trends in observable characteristics, providing support that the regulatory status should also be less correlated with unobservable changes. Lastly, the estimates are robust to various alternative hypotheses; i.e., using the death record, controlling for other policies at national or local levels, restricted samples of similar characteristics, and addressing the local trends.

Another objective of this paper is to quantify air pollution effect on infant mortality, when air pollution levels are already high. Variation in air pollution induced by the regulations provides a quasi-experimental setting, where we can instrument the changes in air pollution by the TCZ regulatory status. The 2SLS estimates based on the TCZ status are found to be robust across specifications, relative to the ones from cross-sectional or fixed effects approaches. We find that air pollution is significantly associated with infant mortality. The estimates suggest that a one percent reduction in TSP leads to a 0.95 percent decline in infant mortality. The marginal effect of TCZ is close to those found in developed countries where pollution is lower, yet the elasticity is substantially higher in China, highlighting the greater benefits associated with air pollution regulations when pollution is already high. In total, we find that the regulations saved 25,400 infant lives per year.

The findings in this study shed light on a number of important policy debates. First, the question of whether, and to what extent, air pollution regulations in developing countries can lead to reducing infant mortality remains unanswered. This study highlights the significant reduction in infant mortality pertaining to air pollution reductions. Second, while there is increasing evidence on the link between air pollution and human health, their relationship at greater pollution level is not known. Our estimates illustrate that TSP has a similar marginal effect even at high concentration levels and substantially higher elasticities when pollution levels are high. Third, a biological mechanism and heterogeneity in the treatment effect found in this study identify pregnant women with low educational attainment as a targeting group of development policy.

Taken together, these results provide substantial evidence to support air pollution regulations in countries that suffer from high air pollution. Unfortunately, climate change currently may not appear to be a strong motivation for these countries to embark on more aggressive air pollution regulations, yet our findings cast little doubt that protecting environment must be an important part of development policies to improving domestic public health.

REFERENCES

Almond, Douglas, Kenneth Y. Chay, and David S. Lee, "The Costs of Low Birth Weight," *Quarterly Journal of Economics*, 120 (2005), 1031-1083.

Almond, Douglas, Yuyu Chen, Michael Greenstone, and Hongbin Li, "Winter Heating or Clean Air? Unintended Impacts of China's Huai River Policy," *American Economic Review Papers and Proceedings*, 99 (2009), 184-90.

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

Antweiler, Werner, Brian R. Copeland, and M. Scott Taylor, "Is Free Trade Good for the Environment?"*American Economic Review*, 91 (2001), 877-908.

Aunan, Kristin, and Xiao-Chuan Pan, "Exposure-response functions for health effects of ambient air pollution applicable for China: a meta-analysis," *Science of the Total Environment*, 329 (2004), 3-16.

Becker, Randy and Vernon Henderson, "Effects of Air Quality Regulations on Polluting Industries," *Journal of Political Economy*, 108 (2000), 379-421.

Behrman, Jere R. and Mark R. Rosenzweig, "Returns to Birthweight," *The Review of Economics and Statistics*, 86 (2004), 586-601.

Bleakley, Hoyt, "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics*, 122 (2007), 73-117.

Case, Anne, Angela Fertig, and Christina Paxson, "The Lasting Impact of Childhood Health and Circumstance," *Journal of Health Economics*, 24 (2005), 365-389.

Case, Anne, Darren Lubotsky, and Christina Paxson, "Economic Status and Health in Childhood: The Origins of the Gradient," *American Economic Review*, 92 (2002), 1308-1334.

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

Chay, Kenneth Y. and Michael Greenstone, "Air Quality, Infant Mortality, and the Clean Air Act of 1970," NBER Working Papers No. 10053, 2003b.

Chay, Kenneth Y. and Michael Greenstone, "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, 113 (2005), 76-424.

Currie, Janet and Rosemary Hyson, "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight," *American Economic Review*, 89 (1999), 245-250.

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

Currie, Janet, Matthew J. Neidell, and Johannes Schmieder, "Air Pollution and Infant Health: Lessons from New Jersey," *Journal of Health Economics*, 28 (2009), 688-703.

Dejmek, J., Sherry G. Selevan, Lvan Benes, Ivo Solansky, and Radim J. Sram, "Fetal growth and Maternal Exposure to Particulate Matter During Pregnancy," *Environmental Health Perspectives*, 107 (1999), 475-480.

Eskeland, Gunnar S. and Ann E. Harrison, "Moving to Greener Pastures? Multinationals and the Pollution Haven Hypothesis," *Journal of Development Economics*, 70 (2003), 1-23.

Greenstone, Michael, "The Impacts of Environmental Regulations in Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures," *Journal of Political Economy*, 110 (2002),1175-1219.

Hao, Jiming, Kebin He, Lei Duan, Junhua Li, and Litao Wang, "Air Pollution and Its Control in China," *Frontiers of Environmental Science and Engineering in China*, 1 (2007), 129-142.

Hao, Jiming, Shuxiao Wang, Bingjiang Liu, and Kebin He, "Plotting of Acid Rain and Sulfur Dioxide Pollution Control Zones and Integrated Control Planning in China," *Water, Air, and Soil Pollution*, 230 (2001), 259-264.

He, Kebin, Hong Huo, and Qiang Zhang, "Urban Air Pollution in China: Current Status, Characteristics, and Progress," *Annual Review of Energy and the Environment*, 27 (2002), 397-431.

Hedley, Anthony J., Chit-Ming Wong, Thuan Quoc Thach, Stefan Ma, Tai-Hing Lam, and Hugh Ross Anderson, "Cardiorespiratory and All-Cause Mortality After Restrictions on Sulphur Content Fuel in Hong Kong: An Intervention Study," *Lancet*, 360 (2002), 1646-1652.

Henderson, Vernon, "Effects of Air Quality Regulation," *American Economy Review*, 86 (1996), 789-813.

Jayachandran, Seema, "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires," *Journal of Human Resources*, 44 (2009), 916-954.

Kumar, Naresh and Andrew Foster, "Respiratory Health Effects of Air Pollution in Delhi and its Neighboring Areas, India," mimeo, 2007.

Kumar, Naresh and Andrew Foster, "Air Quality Interventions and Spatial Dynamics of Air Pollution in Delhi and its Surroundings," *International Journal of Environment and Waste Management*, 4 (2009), 85-111.

Laden, Francine, Lucas M. Neas, Douglas W. Dokery, and Joel Schwartz, "Association of Fine Particulate Matter from Different Sources with Daily Mortality in Six U.S. Cities," *Environmental Perspectives*, 108 (2000), 941-947.

Levy, Jonathan l., James K. Hammitt, and John D. Spengler, "Estimating the Mortality Impacts of Particulate Matter: What Can Be Learned from Between-Study Variability?" *Environmental Health Perspectives*, 108, (2000), 109-117.

Luechinger, Simon, "Air Pollution and Infant Mortality: A Natural Experiment from Power Plant Desulfurization," mimeo, 2009.

Matus, Kira, Kyung-Min Nam, Noelle E. Selin, Lok N. Lamsal, John M. Reilly, and Sergey Paltsev, "Health Damages from Air Pollution in China," *Global Environmental Change*, 22 (2012), 55-66.

Mendelsohn, Robert, and Guy Orcutt, "An Empirical Analysis of Air Pollution Dose-Response Curves," *Journal of Environmental Economics and Management*, 6 (1979), 85-106.

National Bureau of Statistics, 2006. (http://www.stats.gov.cn/)

Nielsen, Chris P. and Mun S. Ho, "Air Pollution and Health Damanges in China: An Introduction and Reviews," in *Clearing the Air: The Health and Economic Damages of Air Pollution in China*, ed. Ho, Mun S. and Chris P. Nielsen (Cambridge: the MIT Press, 2007)

Perera, Frederica P., Wieslaw Jedrychowski, Virginia Rauh, and Robin M. Whyatt, "Molecular Epidemiological Research on the Effects of Environmental Pollutants on the Fetus," *Environmental Health Perspectives Supplements*, 107 (1999), 451-460

Qian, J. and K. Zhang, "China's desulfurization potential," *Energy Policy,* 26 (1998), 345-351.

Qian, Nancy, "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance," *Quarterly Journal of Economics*, 123 (2008), 1251-1285.

Samet, Jonathan M, Scott L. Zeger, Francesca Dominici, Frank Curriero, Ivan Coursac, Douglas W. Dockery, Joel Schwartz, and Antonella Zanobetti, *The National Morbidity, Mortality, and Air Pollution Study. Part II. Morbidity, mortality, and air pollution in the United States*, (Boston: Health Effects institute, 2000)

Schwartz, Joel, Douglas W. Dockery, Lucas M. Neas, "Is Daily Mortality Associated Specifically with Fine Particles?" *Journal of the Air and Waste management Association*, 46 (1996), 927-939.

Schwartz, Joel, and Allan Marcus, "Mortality and Air Pollution in London: A Time Series Analysis," *American journal of Epidemiology*, 131 (1990),185-194.

State Council, "Official Reply to the State Council Concerning Acid Rain Control Areas and Sulfur Dioxide Pollution Control Areas," 1998.

State Environmental Protection Agency (SEPA), *The Report on Environmental Quality in 1991-1995*, SEPA, 1996.

United Nations Environment Programme, (2009). "Two Control Zone Plan and Program to Control Sulfur Pollution," available at http://www.ekh.unep.org/files/GP-2.pdf.

World Bank, *Clear Water, Blue Skies: China's Environment in the New Century*. (Washington, D.C.: World Bank, 1997).

World Bank, *World Development Indicators*. (Washington, D.C.: The World Bank, 1998).

World Health Organization (WHO), "Air Quality Guidelines for Europe," 2nd ed. (Copenhagen, 2002).

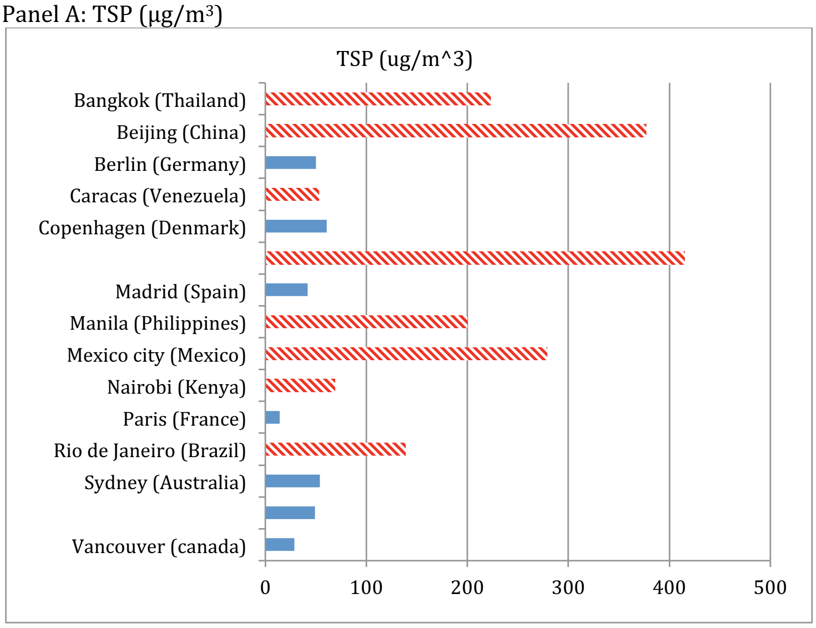
Yang, Gonghuan, Jianping Hu, Ke Quin Rao, Jeimin Ma, Chalapati Rao, and Alan D. Lopez. "Mortality Registration and Surveillance in China: History, Current Situation and Challenges," *Population Health Metrics*, 3 (2005).

Yang, Jintian, Dong Cao, Chazhong Ge, and Shuting Gao, "Air Pollution Control Strategy for China's Power Sector," *Chinese Academy for Environmental Planning*, (Beijing, China, 2002).

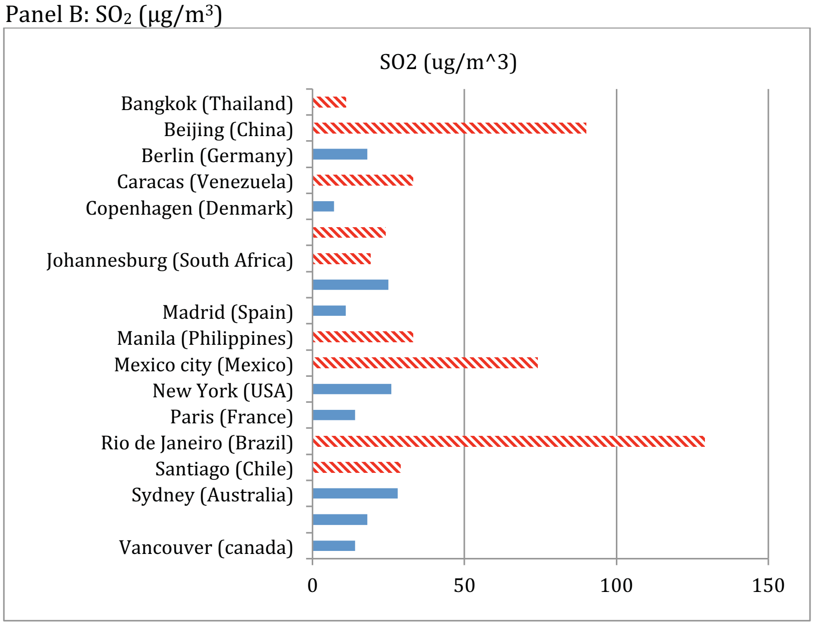
Yang, Jintian, and Jeremy Schreifels, "Implementing SO_2 Emissions in China," presented in OECD Global Forum on Sustainable Development: Emissions Trading, Paris, 2003.

Figure I: Air Pollution Across Countries

Panel A: TSP ($\mu g/m^3$)

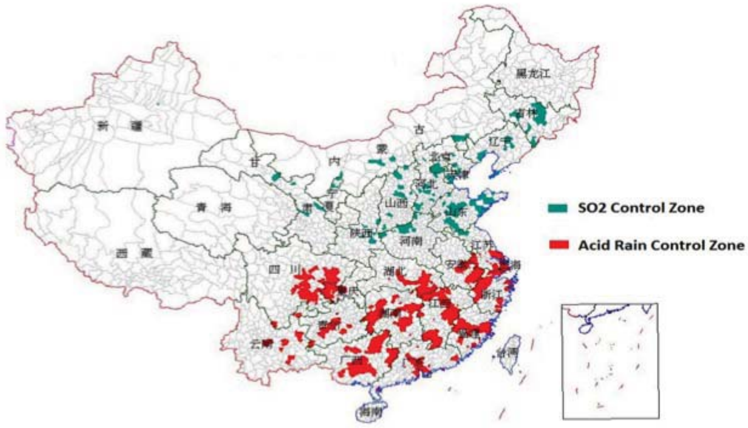


Panel B: SO_2 ($\mu g/m^3$)



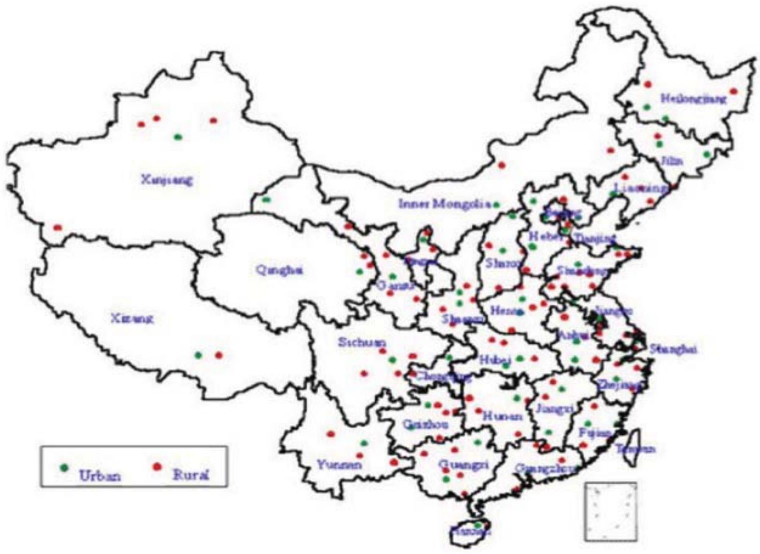
Note: These figures present the world distribution of the TPS in Panel A and SO_2 in Panel B in 1995. Red-striped bars indicate developing countries, whereas blue-solid bars indicate developed countries.
Source: World Bank (1998)

Figure II: Distribution of Two Control Zones



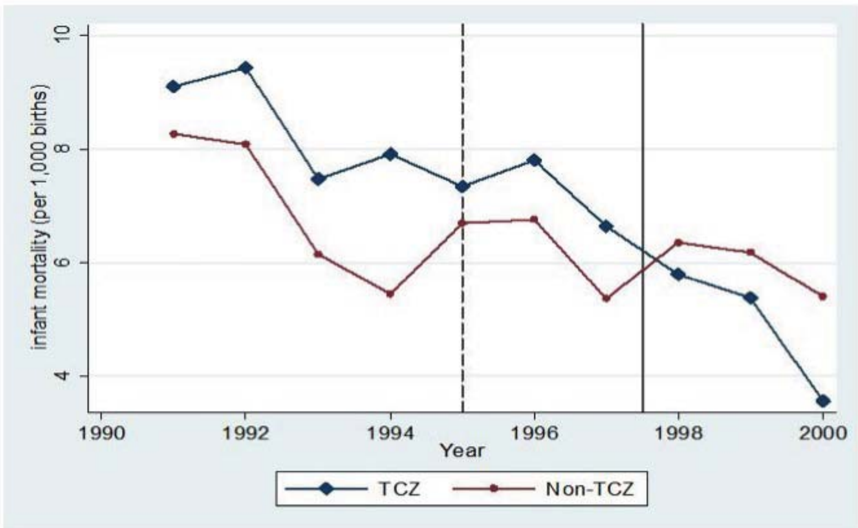
Notes: The green shaded prefectures represent SO_2 Control Zone, and the red shaded prefectures represent Acid Rain Control Zone, designated by the Two Control Zone policy in 1998. *Source*: China Atlas of Population and Environment (1990-1999).

Figure III: Distribution of DSP sites



Notes: Each dot represents one of the 145 DSP sites, that are nationally representative. Source: Yang et al. (2005).

Figure IV: Trends in Intant Mortality Rate



Notes: This figure plots the trend of infant mortality rate between the TCZ and non-TCZ cities. The annual mean is calculated using the population as the weight. The dotted vertical line indicates the timing of APPCL amendment in August 1995, and the solid vertical line indicates the timing of TCZ policy implementation in January 1998. Because each observation represents the annual average value, the dotted vertical line is located at 1995, while the solid vertical line is located between 1997 and 1998 to clarify the timing of their implementations.

TABLE I
DESCRIPTIVE STATISTICS OF BASELINE SAMPLE

	Mean	Std. dev.	Min.	Max.	Obs.
Panel A: Birth Characteristics					
IMR from death record	16.16	16.49	0	177.2	556
IMR from birth record	7.61	11.28	0	128.5	550
% of Male	0.55	0.04	0.333	0.708	550
Birth order	1.40	0.28	1	3.39	551
Born in January	0.11	0.07	0	1	548
Birth weight (gram)	3313	166.9	2500	3787	514
Gestation period (weeks)	39.09	1.80	30	42.31	46
Panel B: Parental Characteristics					
Mother's age	25.37	1.35	22.2	30.4	552
Mother's education >=H.S.	0.22	0.30	0	1	547
% Han	0.89	0.25	0	1	553
Panel C: City Characteristics					
% urban	0.35	0.48	0	1	562
Num. of Infant	1090	1006	43.5	6747	558
Total population	71972	49932	5267	251707	556
TSP concentration ($\mu g/m^3$)	314.44	116.95	98.09	634.04	198
SO_2 concentration ($\mu g/m^3$)	94.74	81.69	2	457.38	215
Precipitation rate (*mm*)	86.24	46.28	1.88	233.6	565

Notes: This table reports the means of the variables, the standard deviation, the minimum value, the maximum value, and the number of observations, based on the sample prior to 1995. For Panel A and B, the number of population is used as weight. See Web Appendix 1 for the definitions of variables.

TABLE II
BALANCING TESTS BY THE TCZ REGULATORY STATUS

	Level Difference		Trend Difference	
Years of sample:	Pre-1998	1996-1997	All Years	1996-2000
	(1)	(2)	(3)	(4)
Birth Characteristics				
% of Male	-0.012*	-0.008	0.004	-0.001
	(0.006)	(0.010)	(0.009)	(0.006)
Birth order	-0.056	-0.012	0.053*	-0.007
	(0.042)	(0.049)	(0.029)	(0.024)
Born in January	-0.013***	-0.018***	0.001	0.005
	(0.005)	(0.004)	(0.006)	(0.005)
Parental Characteristics				
Mother's age	0.977***	1.166***	0.138	0.025
	(0.237)	(0.241)	(0.113)	(0.100)
Mother's education >=H.S.	0.214***	0.231***	0.034	0.032
	(0.051)	(0.063)	(0.045)	(0.041)
% Han	0.094***	0.110***	0.038*	0.035
	(0.031)	(0.040)	(0.021)	(0.029)
City Characteristics				
% urban	0.389***	0.415***	-	-
	(0.081)	(0.083)		
# of Infant	-578.49**	-598.14**	131.67	108.47
	(234.65)	(253.92)	(136.28)	(98.54)
Total population	-11651.5	-15836.3	252.48	1647.85
	(11510.0)	(12869.7)	(3707.1)	(2306.80)
Economic Characteristics				
GDP per capita	2025.03		918.12	
(yuan)	(3144.60)		(740.88)	
Num. of hospitals	-20.22		-29.34	
(in units)	(63.90)		(32.70)	
Water supply	1.93		0.109	
(100 million tons)	(1.34)		(0.114)	
Electricity consumption	61.44		14.94	
(in kwh)	(52.65)		(26.87)	

Notes: Each entry reports the coefficient of the TCZ status (=1 if city is designated as a TCZ) in Columns (1) and (2) and of the interaction term of TCZ status and a post dummy (=1 if year is 1998 or later) in Columns (3) and (4), from a separate regression when using respective variable as the dependent variable. Regressions control for year effects in Columns (1) and (2) and additionally for city fixed effects in Columns (3) and (4). All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses. The sample includes the observations prior to 1998 in Column (1), between 1996 and 1997 in Column (2), all observations between 1991 and 2000 in Column (3), and observations between 1996 and 2000 in Column (4).
Significant at p<0.01***, p<0.05**, p<0.1* levels.

TABLE III
REGULATIONS EFFECT ON INFANT MORTALITY

	All Sample			Common Support
	(1)	(2)	(3)	(4)
Panel A: All causes of death				
TCZ × Post	-2.87***	-3.20***	-4.04***	-3.24***
	(1.10)	(1.09)	(1.08)	(1.12)
R^2	0.41	0.44	0.46	0.46
Panel B: Internal causes of death				
TCZ × Post	-2.58**	-2.82***	-3.33***	-2.85***
	(1.02)	(0.97)	(0.95)	(0.99)
R^2	0.44	0.46	0.48	0.48
Panel C: External causes of death				
TCZ × Post	-0.289	-0.387	-0.702	-0.394
	(0.432)	(0.479)	(0.492)	(0.483)
R^2	0.28	0.29	0.29	0.29
Observations	1340	1281	1235	1223
City FE	Y	Y	Y	Y
Year effects	Y	Y	Y	Y
Full controls	N	Y	Y	Y
Fetal development	N	N	Y	N

Notes: Each entry reports the coefficient of the interaction term in equation (1), when using all observations in Columns (1)-(3) and only observations on common support of propensity score in Column (4). All regressions include year and city fixed-effects. Columns (2)-(4) also include full controls (see Web Appendix 1 for the definitions), and Column (3) additionally controls for birth weight and length of gestation period. All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses. The dependent variable is the number of infant deaths per 1,000 live births for all causes in Panel A, for internal causes in Panel B, and for external causes in Panel C.
Significant at p<0.01***, p<0.05**, p<0.1* levels.

TABLE IV
IDENTIFYING THE BIOLOGICAL MECHANISM

	Dependent variable						
	Birth weight	Low birth weight	Gestation period	Short gestation period	Deaths w/in 1 day	Deaths w/in 1 mo.	Deaths w/in 6 mo.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
TCZ × Post	20.80	-3.18**	-0.304	-0.573	-0.473	-1.84**	-2.41***
	(15.49)	(1.56)	(0.277)	(3.99)	(0.387)	(0.76)	(0.86)
R^2	0.80	0.64	0.48	0.25	0.48	0.48	0.50
Observations	1187	1235	623	651	1225	1225	1226

Notes: Each column reports the coefficient of the interaction term in equation (1). The dependent variable is birth weight in grams in Column (1), the number of infants whose birth weight is less than 2,400 grams per 1,000 live births in Column (2), the length of gestation period in Column (3), the number of infants whose length of gestation period is less than 32 weeks per 1,000 live births in Column (4), the number of infants deaths by internal causes per 1,000 live births, that occur within 1 day after birth in Column (5), within 1 month in Column (6), and within 6 months in Column (7). All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses. The largest five percent of respective dependent variable is eliminated as the outliers.
Significant at p<0.01***, p<0.05**, p<0.1* levels.

TABLE V
HETEROGENEITY IN THE REGULATION EFFECT

	Dependent variable		
	Infant death	Low birth weight	Short gestation period
<u>*Subsample:*</u>			
Whole sample	-3.08***	-3.20**	-7.24
	(0.86)	(1.47)	(12.00)
Boys	-2.69***	-1.76	-2.61
	(0.96)	(1.41)	(12.52)
Girls	-3.57***	-4.89**	-10.39
	(1.15)	(2.08)	(11.70)
Mother low edu.	-4.40**	-3.69	8.14
	(1.81)	(2.60)	(7.21)
Mother high edu.	-0.34	-3.17	-0.099
	(0.70)	(1.94)	(0.42)

Notes: Each cell reports the coefficient of the interaction in equation (1) from a separate regression. All regressions include year and city fixed effects and full controls. Birth weight is also included in Columns (1) and (3), and the length of gestation period is also included in Column (1) and (2). The dependent variable is the number of infant deaths by internal causes per 1,000 live births in Column (1), the number of infants whose birth weight is less than 2,400 grams per 1,000 live births in Column (2), and the number of infants whose length of gestation period is less than 32 weeks per 1,000 live births in Column (3). The first row uses all observations as a benchmark, the second uses only boys, the third uses only girls, the fourth uses infants whose mother's educational attainment is middle school or less, and the fifth uses infants whose mother's educational attainment is high school or more. All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses.
Significant at p<0.01***, p<0.05**, p<0.1* levels.

TABLE VI

ROBUSTNESS CHECKS

	Change to specification								
	Use death record	Use 1996-2000	Include province × year effects	Only districts and cities	Only counties	Control for 1995-97	Use 1995 cut-off	Use 1991-1997	Add city trend
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
TCZ × Post	-3.86***	-3.24**	-3.23***	-3.95**	1.53	-3.85***	-2.35*	-0.63	-2.80
	(1.49)	(1.26)	(1.07)	(1.63)	(1.94)	(1.31)	(1.30)	(1.56)	(2.56)
R^2	0.69	0.60	0.65	0.44	0.53	0.48	0.47	0.49	0.57
Observations	1238	623	1235	699	503	1235	1235	919	1281

Notes: This table shows the results of robustness checks. Each column reports the estimates of the interaction term in equation (1). The change in specification is using the death record to compute infant mortality rate in Column (1), using observations only between 1996 and 2000 in Column (2), controlling for province × year effects in Column (3), using observations only in districts and county-level cities in Column (4), using observations only in counties in Column (5), including another interaction terms of the TCZ status and years between 1995 and 1997 in Column (6), using the interaction term of the TCZ status and a post dummy (=1 if after 1995) in Column (7), using observations only in the pre-reform period of 1991-1997 in Column (8), and including city-specific time trends in Column (9). All regressions include year and city fixed effects and full controls, and additionally birth weight and length of gestation period in all except Column (9) to increase the power of the estimation. All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses.

Significant at p<0.01***, p<0.05**, p<0.1* levels.

TABLE VII
EFFECTS OF AIR POLLUTION ON INFANT MORTALITY

Air pollution measured by:	TSP				SO_2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Cross-sectional model								
Air pollution	-0.0005	0.006	-0.002	0.004	0.012	0.007	0.009	0.005
	(0.005)	(0.004)	(0.005)	(0.004)	(0.008)	(0.008)	(0.009)	(0.008)
R^2	0.00	0.16	0.05	0.19	0.01	0.13	0.03	0.15
Panel B: Fixed-effect model								
Air Pollution	0.020***	0.015**	0.011	0.010	0.001	-0.010	-0.029**	-0.033**
	(0.007)	(0.007)	(0.008)	(0.008)	(0.013)	(0.012)	(0.014)	(0.013)
R^2	0.47	0.50	0.49	0.51	0.39	0.43	0.43	0.46
Panel C: IV-FE model								
Air Pollution	0.044***	0.036**	0.066**	0.075**	0.10***	0.11**	0.34*	0.42*
	(0.012)	(0.015)	(0.032)	(0.038)	(0.03)	(0.04)	(0.20)	(0.25)
1st-stage results								
TCZ × Post	-63.79***	-58.70***	-44.43***	-40.07***	-30.94***	-25.52***	-12.46**	-10.41*
	(8.81)	(9.33)	(12.62)	(12.14)	(5.28)	(6.24)	(5.97)	(5.52)
F-stat	52.41	39.55	12.39	10.90	34.31	16.74	4.36	3.55
Observations	581	565	581	565	608	593	608	593
Year effects	N	N	Y	Y	N	N	Y	Y
Full controls	N	Y	N	Y	N	Y	N	Y

Notes: Each entry reports the coefficient of air pollution, as measured by the concentration (in $\mu g/m^3$) of TSP in Columns (1)-(4) and SO_2 in Columns (5)-(8) from equation (2) in the first row, from equation (3) in the second row, and from equation (4) in the third row. The last row reports the coefficient of the interaction term in equation (5). The dependent variable is the number of infant deaths by internal causes per 1,000 live births. All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses.
Significant at p<0.01***, p<0.05**, p<0.1* levels.

Appendix

TABLE A.1
CAUSES OF INFANT DEATHS

Cause of death	percentage
Infectious and parasitic diseases	3.6%
Neoplasms	1.2%
Nutritional diseases and immunity disorders	0.0%
Mental disorders	0.2%
Diseases of the nervous system and sense organs	19.4%
Diseases of the circulatory system	51.0%
Diseases of the respiratory system	3.6%
Diseases of the genitourinary system	2.0%
Injury and poisoning	19.0%
Total num. of observations	9077

Note: The table lists all incidents of infant deaths by cause based on individual observations.

Appendix

TABLE A.2
COMPARISON OF THE MAGNITUDES

Study	**Study Area**	**Pollution**	**Unit effect**	**Elasticity**
Tanaka (2010)	China	TSP	0.049	0.95
Tanaka (2010)	China	SO_2	0.14	0.82
Chay and Greenstone (2003a)	United States	TSP	0.05	0.35
Chay and Greenstone (2003b)	United States	TSP	0.083	0.45
Currie and Neidell (2005)	California	CO	0.18	0.09
Currie, Neidell, Schmieder (2009)	New Jersey	CO	0.176	0.04
Luechinger (2009)	Germany	SO_2	0.051	0.14

Notes: The table compares unit effect of air pollution on infant mortality and its elasticity across studies. A unit is 1 $\mu g/m^3$ for TSP and SO_2, and is 1 ppm for CO.

NOT FOR PUBLICATION

I. WEB APPENDIX 1: DATA APPENDIX

This appendix provides the data description and variable definitions and construction used in the paper.

I.A. The Chinese Disease Surveillance Points (DSP) System.

The micro-level data on infant mortality come from the Chinese Disease Surveillance Points (DSP) system. The DSP covers 145 sites, established on the representative sample of the national population. It reports the universe of birth and death registrations for the sample population of 10 million residences in different geographic areas (approximately 1 percent of the national population) across 31 provinces, autonomous communities, and municipalities in China. Overall, the original data record approximately 500,000 deaths and 1,000,000 births between 1991 and 2000. We first describe the sampling process and quality management of the datasets, most of which are drawn from Yang et al. (2005). Then, we explain how the datasets we use in this study at the DSP site $\times$ year level were constructed from the original datasets.

Sampling processes. In an effort to form a sample population that represents variation in socioeconomic status and general dispersion of population across different geographical areas, the DSP's sampling is based on three levels of stratification and a multi-stage cluster probability. The first level of stratification is established over 7 different geographic communities (Northeast, North, East, South, Southwest, Northwest, and Central) and 3 municipalities (Beijing, Tianjin, and Shanghai). The second level is based on the comparison of primary sampling units (the city for urban areas, and the county for rural areas) between urban and rural areas. For rural areas, the third level is applied to break them down into 4 classifications based on socioeconomic status, such as literacy rates, GDP per capital, and dependency rates. Within urban areas, the third level is applied to classify them into 3 groups based on the population size, big cities with over 1 million population, middle sized cities with 0.5 - 1 million population, and small cities with 0.2 - 0.5 million population.

After the primary cluster units are reclassified based on the procedures above, sampling units are selected using sampling probability based on the population size given by 1982 Census data. Among the selected units, the unit of sampling (a 'neighborhood' in cities, and 'township' in counties) is selected using sampling probability based on the population size of the neighborhoods or townships. This results in 145 sites, representing the dispersion of population and socioeconomic status across the entire nation.

Data collection and quality control. Each DSP site has at least one township hospital, and the 'Disease Prevention Unit' in these hospitals is responsible for vital registration. In urban areas, almost half of the deaths occur at health facilities, while about 80 percent of adult deaths occur at home in rural areas. If someone dies in hospitals, the attending physicians completes a death certificate, while if someone dies at home, the Disease Prevention Unit sends a staff member to household and completes a death certificate. In both cases, the information from family members and available medical records or documents are used to identify the cause of the death. The death certificates are collected by the Center for Disease Control in Beijing, where the completeness and consistency of the data are evaluated and validated using an internal procedural check system and statistical measurements. See Yang et al. (2005) for more details.

Variable definitions and construction. The following variables are included as controls in specifications labeled "full control" in the tables. The word "indicator" refers to an indicator variable that is equal to one for meeting the respective category.

- Birth characteristics
  - Birth order
  - Indicator for male
  - Birth month dummies

- Parental characteristics
  - Mother's age
  - Indicator for mother having at least high school degree
  - Indicator for Han ethnicity
- City characteristics
  - Number of infants under age one
  - Number of total population
  - Weighted average precipitation rates (mm)

Additionally, the following variables are labeled as "fetal development" in the tables. Note that we use indicator variables for gestation period when they are used as a control variable to retain the number of observations given a large amount of missing values, whereas we use it as a discrete number of weeks when it is used as a dependent variable.

- Fetal development
  - Birth weight (grams)
  - Indicators for gestation period ¡=32, 33-37, 37-38, 38-39, 40, 41, 42, missing

Finally, individual observations are aggregated to the DSP level for each year.

I.B. Air quality.

The data on air quality, as measured by the concentrations of TSP or SO_2, primarily come from the China Environment Yearbooks (CEY) published by the SEPA. In response to an increasing public concern about urban air pollution, the Chinese government established a national air pollution monitoring system, and currently more than 350 cities are equipped with monitoring stations to monitor pollutants of TSP and SO_2 (recently, several other pollutants are also monitored, such as NO_x. CEY reports annual daily average concentrations of TSP and SO_2 for major cities. The data between 1991 and 1995 are compiled from CEY as part of the "Economics of Pollution Control

Research" project by the World Bank and are available on its website. The data between 1996 and 2000 are digitized directly from CEY by the author. We also extracted air quality data from the City Statistic Yearbooks for several cities whenever they are available. We believe that this constructs one of the most comprehensive dataset on air pollution in China.

If pollution data are available in a city exactly matched with a DSP site, we simply merge them. If pollution data are not available, we assign a pollution to a DSP site as follows:

1. We identify the geographical coordinates of the city centroid where air quality data are available, and we use this as the location of an air quality monitoring station.

2. We measure the distance between the centroid of each DSP sites and the monitoring stations, using the geographical coordinates.

3. We assign a pollution level to the DSP sites if it (i) is within the same province, (ii) locates within 150 kilometers, and (iii) shares the same TCZ status. (i) is adopted to avoid different trends across provinces, (ii) is known as a reasonable distance that pollutants can travel (they can travel up to 1,000 kilometers less than a week), and (iii) is crucial because different TCZ status is characterized as different pollution trends.

To capture variation in the individual level of exposure within the DSP site × year cell, we compute a weighted average pollution level during the first year of life, estimated as

$$P_{ijt} = P_{jt}W_{ijt} + P_{jt+1}(1 - W_{ijt}),$$

where $W_{ijt} = \frac{Daysleft}{365}$ and $Daysleft = (December\ 31st - date\ of\ birth)$.[31]

Finally, individual levels of pollution data are aggregated to the DSP × year level. This creates a panel dataset of air quality.

I.C. TCZ designation.

The TCZ regulatory status are reported in the document "Official Reply to the State Council Concerning Acid Rain Control Areas and Sulfur Dioxide Pollution Control Areas," published by the State Council in 1998. The assignments are mostly made at the prefecture level, while those

[31] This formula also helps mitigate measurement error in readings in a certain year.

in municipalities are at the district level. It lists the names of all prefectures that are assigned as the Acid Rain Control Zone and the SO_2 Control Zone. We follow the following rules to assign the TCZ status.

1. First, we assign the TCZ status to all districts/county-level cities/counties if they are listed as a TCZ or if they belong to the prefectures that are designated as a TCZ.

2. In case administrative divisions has changed, we use their divisional status (either district, county-level city, or county) as of 1998.

3. I modify the TCZ status to non-TCZ status if the DSP site is ranked as a county unless itself is listed as a TCZ. I apply this rule because the document explicitly states to exempt impoverished counties, even though it does not specify each individual country falling under this category.

I.D. Precipitation rates.

The data on rainfall are drawn from the Global Historical Climatology Network (GHCN) project. The GHCN provides monthly average precipitation rates in mm for given longitudes and latitudes with the minimum cell size of 0.5 degree $\times$ 0.5 degree, covering the entire area of China. For a given geographical coordinate, I first compute an annual average precipitation rate by taking an arithmetic average of all twelve months. Then, I calculate a weighted average precipitation rate for each DSP site by interpolating readings at the four nearest points, using the inverse squared distance as the weight, often known as the Shepard's method. For example, an interpolated precipitation rate of grid j using the four nearest points k is given as;

$$Rain_j = \sum_{k=1}^{4} \frac{Rain_k \times D_{jk}^{-2}}{\sum_{k=1}^{4} D_{jk}^{-2}}$$

where D_{jk} is the distance between j and k.

II. WEB APPENDIX II: BALANCING TESTS ON RESTRICTED SAMPLES

This appendix provides additional evidence from balancing tests on restricted samples that the TCZ regulatory status has little correlation with observable determinants of infant mortality. A major methodological challenge in any non-experimental research designs is that the treatment status may not be orthogonal to unobservable determinants of outcome. In this case, simple difference-in-differences estimates may suffer from two sources of bias.

The first bias arises from differential pre-existing trends, leading the DID term to find a spurious relationship, where there is no association. In the main analysis, we have conducted a number of approaches to show that there is no systematic difference in pre-trends between the two sets of the cities. This addresses the first concern.

The second bias concerns that the treatment is confounded by other concurrent changes in the post-treatment period, leading the DID term to erroneously pick up their effects. This is essentially more difficult to address, because we need to confirm that there is no *unobservable* factors that confound the effects, and by construction, we cannot directly test this hypothesis. Nonetheless, we have provided testable implications of the hypothesis in the main analysis. A major confounding factor may be a local health policy. If local governments were concerned about health impacts caused by air pollution, they may have improved the local health system, i.e., building more hospitals or improving quality of health services. We show evidence to rule out such a possibility; the treatment effect is absent on infant deaths due to injury or poison.

Another compelling evidence that significantly limits the scope of unobservable changes is a lack of treatment effects for other determinants of infant mortality. If there had been any other changes that affected infant mortality, these factors would have also affected these other determinants of mortality as well. Indeed, the absence of correlation between the treatment status and these other determinants suggests that it is less likely to be correlated with unobservable determinants as well. In the main analysis, we have shown that the TCZ status is less likely to be correlated with various births, households, and cities characteristics, except that a few variables show marginally significant differences. In the main analysis, we further restrict the observations with similar characteristics as a robustness check. The Web Appendix Table 1 presents evidence from their balancing tests.

In Column (1), we use observations only on common support. The estimates find no correlation between the TCZ status and any observable characteristics. In Columns (2) and (3), we present similar results using the TSP-matched sample and the SO_2 matched sample, respectively, reassuring that there is no systematic difference in trends in any observable characteristics.

Overall, the regulatory status well balances trends in characteristics, providing evidence to support the exclusion restriction.

WEB APPENDIX TABLE I

BALANCING TESTS UNDER RESTRICTED SAMPLES

	Trend Difference		
Sample	Common Support	TSP-matched	SO_2-matched
Birth Characteristics			
% of Male	0.004	-0.005	-0.008
	(0.009)	(0.008)	(0.009)
Birth order	0.046	0.011	0.070
	(0.029)	(0.036)	(0.045)
Born in January	0.003	0.005	0.002
	(0.006)	(0.010)	(0.012)
Parental Characteristics			
Mother's age	0.129	-0.197	-0.033
	(0.120)	(0.198)	(0.198)
Mother's education >=H.S.	0.029	-0.024	-0.062
	(0.047)	(0.074)	(0.084)
% Han	0.019	-0.025	-0.023
	(0.015)	(0.022)	(0.023)
City Characteristics			
% urban	-	-	-
# of Infant	146.25	91.31	165.01
	(122.05)	(195.30)	(243.30)
Total population	-1498.7	4061.73	5129.39
	(3264.34)	(6424.00)	(9027.24)
Economic Characteristics			
GDP per capita	970.91	875.31	602.88
(yuan)	'(1125.22)	(913.11)	(939.42)
Num. of hospitals	-20.61	-17.56	7.44
(in units)	(31.73)	(31.62)	(26.50)
Water supply	0.136	0.153	0.124
(100 million tons)	(0.122)	(0.138)	(0.127)
Electricity consumption	13.76	9.95	30.28
(in kwh)	(34.47)	(31.76)	(26.19)

Notes: Each entry reports the coefficient of the interaction term in equation (1), from a separate regression when using respective variable as the dependent variable. All regressions include year and city fixed effects. All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses. The sample includes observations on common support of propensity score in Column (1), observations in TSP-matched sample in Column (2), observations in SO_2-matched sample in Column (3).
Significant at p<0.01***, p<0.05**, p<0.1* levels.

III. WEB APPENDIX III: EXAMINATION OF TREND BREAKS

Whether there is a trend break in years other than 1998, or whether outcome variables are trended differentially over the entire years is an important identification assumption in assessing the treatment effect. So far, we have provided four testable implications. First, Figure 4 illustrates visual evidence that infant mortality reductions are accelerated when the TCZ policy was implemented in 1998. Second, the interaction term using 1998 is relatively more strongly associated with infant mortality reductions than the one using 1995. Third, the treatment effect is absent when using only observations prior to 1998, suggesting that there should be little systematic difference in trends before 1998. Lastly, the robustness to the inclusion of city-specific time trend provides evidence that mortality was not trended differentially.

In this Appendix, we provide regression-adjusted evidence on how differences in levels of mortality and air pollution evolved over years. Specifically, we regress an outcome variable on the interaction terms of the TCZ status in the birth city and birth-year dummy variables for all birth years:

$$Y_{jt} = \sum_{l=1992}^{2000} (T_j * d_l)\pi_l + X_{jt}\delta_3 + \kappa_t + \mu_j + \varepsilon_{jt}, \tag{6}$$

where d_l is a dummy variable for being born in year l. The comparison year is 1991, and thus its interactions are omitted. π_l captures the treatment effect on the outcome variable for being born in the TCZ city in year l, relative to 1991. π_l is expected to be constant until around a trend break, after which it should be negative.

The results are summarized in Web Appendix Table 2. Column (1) illustrates a clear change in trends in 1998; differences in IMR between the TCZ and the non-TCZ cities are essentially constant until 1997. Starting in 1998, the coefficients pick up substantial improvements in relative IMR in the TCZ cities. In Column (2), we use the TSP as the dependent variable. The point estimates are constant around 13 $\mu g/m^3$ up to 1995, indicating that the TCZ cities were on a relatively increasing trend in TSP, though this constitutes a very small difference compared to the average TSP concentration of 300 $\mu g/m^3$ prior to 1995. The point estimates become negative after 1997, indicating that the regulations after 1995 may have had some effect on reducing TSP, while somewhat significant differences in trends can be found only after 1998. SO_2 describes a clearer evidence. Although it is unknown why SO_2 makes some improvements in the TCZ cities in 1994, the difference in trends shows no further movement up to 1998. The TCZ cities underwent a significant improvement in 1999, and the difference in trends is significantly at the five percent level in 2000. Overall, these pieces of regression-adjusted evidence confirm the visual evidence that a statistically significant trend break occurs in or around 1998.

WEB APPENDIX TABLE II
EFFECTS OF BEING TREATED

	Dependent variable		
Birth year	IMR	TSP	SO_2
1992	-0.33	23.6	3.89
	(2.46)	(19.9)	(7.29)
1993	-1.09	19.6	2.52
	(2.79)	(20.5)	(11.1)
1994	0.353	15.3	-13.9
	(3.01)	(20.4)	(9.86)
1995	-2.24	15.0	-12.8
	(3.04)	(23.2)	(12.5)
1996	-0.767	5.81	-10.2
	(2.68)	(22.7)	(11.4)
1997	-0.973	-11.2	-13.3
	(2.62)	(21.7)	(12.7)
1998	-3.32	-22.4	-15.6
	(2.68)	(23.3)	(12.8)
1999	-4.10	-37.9	-20.2*
	(2.76)	(24.9)	(10.8)
2000	-4.73*	-44.9	-25.6**
	(2.60)	(28.5)	(11.2)

Notes: Entries in each column represents the coefficient of the interaction term in equation (6). All regressions include year and city fixed effects, full controls, birth weight, and length of gestation period. The dependent variable is the number of infant deaths by internal causes per 1,000 live births in Column (1), TSP concentration levels in Column (2), and SO_2 concentration levels in Column (3). All regressions are weighted by number of population, and the robust standard errors, clustered at the DSP site level, are reported in the parentheses.
Significant at p<0.01***, p<0.05**, p<0.1* levels.