Essays on Environmental and Health Economics

Muxin Zhai

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Robert Halvorsen, Chair

Judith Thornton

Jing Tao

Program Authorized to Offer Degree:
Economics

University of Washington

Abstract

Essays on Environmental and Health Economics

Muxin Zhai

Chair of the Supervisory Committee:
Professor Robert Halvorsen
Department of Economics

This dissertation examines the linkages between public policy, environment and human health. The first chapter evaluates the environmental impact of the world's largest Low Emission Zone (LEZ) in London. The second chapter reviews recent literature in assessing the health effects of air pollution, and at the same time, provides a national scale estimation on the impacts of four "criteria air pollutants" on fetal health using natality data in the United States from 1991 to 2008. The third chapter explores fetal sensitivity to different air pollution thresholds and analyzes the heterogeneous effects of air pollution across socioeconomic groups.

Chapter one concerns the environmental impact of the world's largest LEZ introduced in London in early 2008. The LEZ policy restricts heavily polluting diesel vehicles from entering Greater London by charging 100 to 200 euros per day. The empirical estimation relies on a difference-in-differences (DID) method that compares the air pollution concentrations before and after the introduction of the LEZ. Using daily level particulate matter (PM_{10}) and nitrogen dioxide (NO_2) data, I find that the initial phase of London's LEZ does not lead to substantially reduced air pollution inside London. The subsequent phases, on the other hand, generate an 8 to 13 percent reduction of air pollution. A further spatial analysis shows that a majority of this air pollution reduction occurs in industrial instead of roadside areas.

Combining data from traffic sectors, I provide evidence that the insignificant DID estimator in the initial phase is partially caused by the positive spillover effects nearby and far from Greater London. I supplement this result with evidence showing that drivers may substitute toward a non-subject vehicle to avoid paying the LEZ fee inside the zone.

Chapter two provides a thorough literature review on the health impact of direct and indirect exposures to air pollution. The reviewed articles span the fields of epidemiology, environmental health, medical science and economics. For each reviewed article, I discuss in detail the identification strategies, which include quasi-experimental design, instrumental variables approach and difference-in-differences method. The estimation results, however, show some degree of inconsistency. Motivated by the literature gap, the rest of Chapter 2 performs a comprehensive analysis on the health effect of maternal exposure to air pollution using data from over 20 million births in the United States from 1991 to 2008. For the four "criteria air pollutants" studied, carbon monoxide imposes the greatest health risk on fetuses during the first and third trimesters. Nitrogen dioxide, ozone and sulfur dioxide each shows some adverse effects on birth outcome, but their impacts are weaker than those of carbon monoxide. In addition, I find that maternal avoidance behaviors against air pollution may largely reduce the health risks of air pollution, especially for pre-term babies during the second trimester.

Chapter three constructs an innovative pollution index to empirically investigate whether there exists a "safety threshold" of air pollution. The econometric model is constructed based on the traditional approach but has an important revision on the key regressor. Specifically, a series of hypothetical "safety thresholds" are tested against the assumption that only pollution ranging above the threshold will convey explanatory power on health outcomes. The model is estimated using individual-level birth data in the U.S. from 1991 to 2008. Among the four pollutants analyzed, carbon monoxide imposes significant health risks even if its concentration approaches zero, whereas nitrogen dioxide, ozone and sulfur dioxide do not affect fetal health conditions unless their concentrations rise above certain thresholds. According to my estimation, the thresholds that generate the most sensitive changes of fetal birth weight are about 15.30 ppb for nitrogen dioxide, 0.029 ppm for ozone and 13.4 ppb for sulfur dioxide, all at the weekly average level. Meanwhile, the analysis also implies the existence of heterogeneity in the responsiveness to air pollution across different socioeconomic groups. In particular, smoking mothers giving subsequent births are found to be at high risk of environmental hazards.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my dissertation advisor Prof. Robert Halvorsen for his invaluable advice, encouragement and support. I would also like to thank the rest of my committee members Prof. Neil Bruce, Prof. Judith Thornton, Prof. Jing Tao and Prof. David Layton for their insightful comments and continuous support.

I am thankful to participants in the Department brownbags and PHEnOM seminars. I have received generous help from faculties and students in the Department of Economics, Department of Health Services and Evans School of Public Affairs. I am also thankful for the US Centers for Disease Control and Prevention for providing the non-public vital statistics files for research purpose. I appreciate staffs at CSDE for providing technical support to run large datasets on the simulation cluster.

I am grateful to Jorge Rojas-Vallejos who has offered me tremendous help during my most difficult time, and to Aseem Prakash, who guided me through my academic career.

My deepest gratefulness goes to my family and friends who provide unconditional love and understanding for me during my PhD life. This dissertation would never become possible without them.

TABLE OF CONTENTS

Chapter 1

PUBLIC POLICY AND AIR POLLUTION: EVIDENCE FROM THE WORLD LARGEST LOW EMISSION ZONE

1.1 Introduction

Air pollution is one of the leading health concerns in urban areas in Europe. According to the World Health Organization, outdoor exposures to air pollution cause nearly 600,000 premature deaths in Europe each year (WHO 2012). The pollution-related shortened life expectancy and loss of human capital are associated with 120 trillion pounds annual loss (equivalent to 137 trillion US dollars) based on the estimates by the European Environmental Agency (EEA) in 2013. Additionally, poor air quality may bring up issues in crop production, food safety and human fertility, leading to substantial social costs in the long run (Murphy et al. (1999); Zivin and Neidell (2011); Rosa et al. (2003)).

In Europe, one of the largest contributors to air pollution is transportation sector. According to EEA, transportation sector accounts for the emissions of about 58 percent total nitrogen oxides (NO_x), 30 percent carbon monoxide (CO), 21 percent sulfur dioxide (SO_2) and 21 percent particulate matter (PM_{10}) across 31 European countries in 2000 (EEA 2003). In response to the overwhelmingly increasing concern on traffic emissions, governments in European countries have established a series of policies and regulations over the recent decades, including taxes on gasoline, subsidies on green vehicles transactions, and command-and-control policies such as license plate programs. In this paper, I study a widely adopted incentive-based traffic regulation - Low Emission Zone (LEZ). A typical LEZ regulation restricts or disincentivizes heavily polluting vehicles from entering a targeted area. Advocates of LEZ argue that the policy helps alleviate government enforcement costs while inducing drivers to replace, upgrade or retrofit their non-compliant vehicles, creating both short-run and long-run social benefits. Nevertheless, there is rising concern that LEZ might cause behavioral adjustments in the form of detouring to avoid the zone or substituting toward non-subject vehicles. The potential existence of such unwanted consequences may substantially lower the policy's effectiveness, and the resulting environmental spillovers may drive down the average air quality improvements toward zero.

London has the world's largest LEZ. Launched in February 2008, London's LEZ covers almost the entire area of Greater London. The initial phase of the LEZ affects over 50 thousand most polluting heavy goods vehicles (HGVs) driving inside and across the border of Greater London. The extended phases additionally affect light goods vehicles (LGVs), buses and coaches, and impose more restrictive emission standards on the already affected vehicles. London's LEZ charges a daily entrance fee of £100 to £200 on vehicles violating the corresponding EU emission standards (e.g., EU III or EU IV) and a penalty fee of £250 to £1000 on each illegal entrance depending on vehicle type. The policy has generated a capital set-up cost exceeding 45 million pounds (about 65 million US dollars) and has an estimated annual maintenance fee of over 10 million pounds (about 15 million US dollars).

In this paper, I analyze the environmental effects of London's LEZ. I combine a unique dataset of daily average particulate matter (PM_{10}, airborne particulate matter with less than 10 micrometer in diameter) and nitrogen dioxide (NO_2) from nearly 100 air monitoring stations in the United Kingdom with daily weather conditions during 2005 and 2010. The empirical estimation adopts a difference-in-differences (DID) strategy which compares air pollution concentrations in traffic intensive areas versus those in background areas before and after the introduction of the LEZ. The background air quality is depicted by one of the two control groups: urban background areas (e.g. residential areas) inside the zone and traffic intensive areas far from the zone. To estimate the policy's spillover effect, I apply a similar DID model with the "treatment group" being the traffic intensive areas within 10 miles of the LEZ border. In both regressions, I incorporate a rich set of covariates, including year and month fixed effects, contemporaneous and lagged weather conditions as well as dummies for weekends and holidays.

The estimation results suggest that the initial phase of London's LEZ does not lead to a significant improvement of air quality along traffic intensive areas within the zone. Not until the operation of the most stringent phase of the LEZ, which affects a broader scope of vehicles, do the targeted areas show a statistically significant decline of air pollution. However, spatial analysis shows that the air quality improvements in London occur only in industrial areas, while the roadside air quality remains unchanged after the LEZ policy, controlling for other confounding factors. Nevertheless, the initial phase of the LEZ does not generate a substantial spillover effect nearby the zone. The most stringent phase, on the other hand, induces the roadside NO_2 nearby the zone to decline by over 10 percent compared with urban background areas outside the LEZ border.

I propose two possible explanations for the insignificance of the difference-in-differences estimates. First, a majority of the vehicles affected by London's LEZ are national commercial vehicles potentially traveling across different regions of England. The high daily charge of the LEZ provides economic incentives for these drivers to retrofit or replace their non-compliant vehicles, generating positive environmental externality outside London. These compliance behaviors make the air quality stations in distant locations incapable of showing background air pollution trends to help capture the true treatment effect of the LEZ. On the other hand, the policy may induce drivers to make undesirable adjustments such as detouring or substituting subject vehicles toward non-subject ones. These behavioral adjustments, if occurring aggregately, will result in increasing vehicle miles traveled and vehicle combustion. Using traffic volume data and traffic-related energy consumption data, I show evidence of these side-effects along major roads in London.

The rest of the paper proceeds as follows. Section 1.2 reviews the background information of London's LEZ. Section 1.3 provides a literature review of traffic regulation. Section 1.4 describes the primary dataset and provides summary statistics. Section 1.5 specifies the econometric model. Section 1.6 presents and discusses the estimation results. Section 1.7 provides further evidence for the estimation results. Section 1.8 concludes.

1.2 Policy Background

London's low emission zone (LEZ) is an economic incentive based traffic regulation targeting at regulating the air pollution emissions from heavily polluting commercial vehicles. The initial phase of the LEZ was officially launched in February 2008, about 9 months after the public announcement of the policy. In the subsequent years, the policy has been extended to incorporate more vehicles while tightening the emission standards for the already subject vehicles. Compliance with the policy may take various forms, including retrofitting, converting to gas and reorganizing fleet. Non-compliant vehicles driving inside the zone are subject to a daily fee ranging from £100 to £200 based on vehicle types.[1] A penalty fee of £250 to £1000 will be collected from vehicles illegally entering the zone. The LEZ covers almost the entire area of Greater London and operates 24 hours a day, 365 days a year, including weekends and national holidays. Table 1.1 summarizes the timeline and background information of different phases of the LEZ.

Drivers approaching the LEZ are notified by advance warning signs that allow them to detour to avoid entering the zone (Figure 1.2 Panel A). To monitor compliance behaviors inside the zone, Traffic for London (TfL) has installed over 300 fixed or mobile cameras on major roads in Greater London, accompanied with automatic number plate recognition (ANPR) technology (see Figure 1.2 Panel B). The ANPR reads the vehicle plate number, communicates with TfL database and reports automatically if the vehicle has complied with the emission standard or paid the entrance fee. The enforcement cost is estimated to exceed 45 million pounds for the initial installation with an additional maintenance fee of over 10 million pounds each year.

[1] The daily entrance fee incurs at midnight everyday. Vehicles on road over 23:30 pm and 12:30 am the next day are subject to two days' fees. Vehicles parking within LEZ do not have to pay the entrance fee.

1.3 Literature

Evaluating the effectiveness of traffic regulation on reducing urban air pollution is an area of active research in recent years. Davis (2008) analyzes the environmental effect of a license plate program *Hoy No Circula* (HNC) launched in Mexico City in 1989. Each day of week, the HNC banned a certain group of drivers from driving on roads based on the last digit of license place number. Using both ordinary least squares (OLS) and regression discontinuity (RD) design, Davis (2008) find that the HNC did not effectively reduce five major air pollutants (i.e. nitrogen dioxide, nitrogen oxides, ozone, carbon monoxide and sulfur dioxide) during its operating hours. Instead, behavioral adjustments in the form of acquiring a second, heavily polluting vehicle and substituting toward highly emitting taxis have led to an increase in air pollutants during the non-operating hours of the HNC. Chen et al. (2013) examine a slightly revised license plate program in Beijing during the 2008 Olympic Games. Compared with non-Olympic cities, Beijing has experienced a roughly 25 percent reduction of air pollution index (API) when the license place program was in operation. The magnitude of the air quality improvements, however, shortly faded away after the closing of the Olympic Games and the end of the policy. This result was confirmed in Viard and Fu (2015) by reevaluating the effectiveness of Beijing's driving restriction on reducing air pollution concentrations. Additionally, Viard and Fu (2015) show that the affected workers in Beijing have reduced their labor supply by 9 to 17 percent during the policy's period.

Several studies provide empirical estimations on the impact of LEZ established in various European cities. Boogaard et al. (2012) analyze the changes of air pollution concentrations following the implementation of LEZs in 5 Dutch cities. Using suburban locations as controls, Boogaard et al. (2012) show that the policy did not effectively reduce air pollution in all but 1 treated urban street where the traffic volume decreased sharply. Wolff (2013) investigates the impact of Germany's LEZs on the spatial distribution of air quality. The results indicate that the LEZs implemented in various areas in Germany have reduced PM_{10} by an average of 9 percent with larger LEZs yielding stronger effects.

London is among the most polluted cities in Europe and thus has received research focus in recent years. Jones et al. (2012) analyze the air quality data from 2 sites in London and 1 site in Birmingham during late 2007. Two air quality trends occurred shortly before this period. First, the United Kingdom introduced "sulfur free" diesel fuel on all highway vehicles. Second, the Traffic for London announced the launching of London's LEZ. These events jointly resulted in a one-third to two-thirds decline of NO_x and particulate matter. However, a significant portion of the reduction is expected to be caused by the sulfur regulation. Similarly, Ellison et al. (2013) use the air quality data from 3 sites within the LEZ and 1 site outside and show that the concentration of particulate matter has seen a slight reduction of about 3 percent while that of NO_x did not change significantly. Unfortunately, both papers have a relatively small sample size and the results are unlikely to represent the average treatment effect of London's LEZ.[2] Meanwhile, the articles lack a credibly econometric specification which controls for many of the confounding factors (e.g. weather conditions).

This analysis fills the literature gap by empirically estimating the environmental effect of London's LEZ using the most comprehensive data available in the United Kingdom. I carefully distinguish different environmental features of air monitoring stations to identify the spatially heterogeneous treatment effect of the LEZ. In addition, I combine data from traffic sectors and energy consumption to further support my findings. Understanding how the world's largest LEZ affects travel patterns and how effective it is in reducing vehicular emissions provides guidelines for policymakers to design, implement and enforce optimal traffic regulations to improve social wellbeing in the future.

1.4 Model and Identification Strategy

1.4.1 Difference-in-Difference Approach

The estimation goal is to identify the average treatment effect (ATE) of London's LEZ on air pollution concentrations. Since the UK has introduced "sulfur free" vehicles in late

[2] Less than 10 percent air monitoring stations in London are covered in these studies.

2007, it is essential for the identification strategy to credibly isolate any air quality changes caused exclusively by the LEZ implementation. To achieve this goal, I apply a difference-in-differences (DID) model controlling for both the air quality trends occurring all over the England and the unobservable across-region air quality differentials. Let i index air monitoring station and t index date. The baseline model has the following structure:

$$\log(AP_{it}) = \beta LEZ_t + \gamma LEZ_t Tr_i + \alpha Tr_i + Y_t + Q_t + \boldsymbol{X_{it}\theta} + \varepsilon_{it}. \tag{1.1}$$

The dependent variable is the logarithm of daily air pollution (AP), either PM_{10} or NO_2, at station i at time t. Variable LEZ_t equals 1 in the post-LEZ period and 0 otherwise. Tr_i is a time-invariant group indicator equalling 1 if station i is in the treatment group.[3] Y_t and Q_t capture the unobservable year fixed effects and quarter fixed effects, respectively. Including year fixed effects largely ensures that the long-run air quality trends, such as those resulted from business cycles, are eliminated. The quarter fixed effects account for seasonal variations of air quality that are in place even without the LEZ. $\boldsymbol{X_{it}}$ is a vector of control variables including current and one-day lagged weather conditions as well as a dummy variable indicating zero precipitation for 3 days. To account for the variations in travel flow during non-working days, I additionally include indicators for weekends and national holidays in $\boldsymbol{X_{it}}$. The idiosyncratic error term ε_{it} captures random shocks of air quality at station i at date t.

Equation (1.1) identifies the overall treatment effect of London's LEZ but is not able to isolate the additional change of air pollution in each phase of the LEZ. Furthermore, pre-LEZ compliance behaviors may start shortly after the policy's announcement, biasing the coefficient of interest γ toward zero. To deal with this issue, I revise the baseline DID model to allow for differential environmental effects when different groups of vehicles are targeted. The revised model has the following form:

$$\log(AP_{it}) = \sum_{\tau=0}^{\tau=3} \beta_\tau P\tau_t + \sum_{\tau=0}^{\tau=3} \gamma_\tau P\tau_t Tr_i + \alpha Tr_i + Y_t + Q_t + \boldsymbol{X_{it}\theta} + \varepsilon_{it} \tag{1.2}$$

[3] No sample station has been relocated three years before and after the implementation of the LEZ.

where $P\tau_t$ ($\tau = 0, 1, 2, 3$) is a phase indicator that equals 1 if date t falls in the post-phase-τ period (phase 0 refers to the post-announcement period) and 0 otherwise. The coefficient of interest, γ_τ, measures the ATE of the LEZ phase τ on air pollution concentrations. β_τ ($\tau = 0, 1, 2, 3$) depicts the generic trend of pollution concentrations experienced by both treated and control stations in phase τ of the LEZ. α captures the cross-sectional differentials of air quality between the treated stations and the controls. $\boldsymbol{\theta}$ is a vector of coefficients associated with the control variables.

Because stations monitor and report air quality data independently, ε_{it} is clustered at station level in the preferred specification. In some circumstances, the assumption underlying this clustering feature may be violated. For instance, random shocks of air pollution may occur aggregately, affecting all stations located within a certain region. To account for this possibility, I test alternative clusterings of the error term in robustness checks.

1.4.2 Test of Parallel Trends

The fundamental assumption of DID method is that pre-treatment trends across groups are parallel. In other words, all groups of stations should experience the same trends of air quality with the absence of the policy or regulation. To test if this assumption is valid, I estimate the following regressions[4]

$$\begin{aligned} \log(\overline{AP}_{Tr,ym}) &= \gamma_{Tr} YM_{ym} + \mu_{Tr,ym} \\ \log(\overline{AP}_{Con,ym}) &= \gamma_{Con} YM_{ym} + \mu_{Con,ym}. \end{aligned} \tag{1.3}$$

$\overline{AP}_{Tr,ym}$ and $\overline{AP}_{Con,ym}$ in equations (1.3) refer to the average air pollution in year-month ym across all treated (Tr) and control (Con) stations, respectively. YM_{ym} is a year-month indicator for ym. $\mu_{Tr,ym}$ and $\mu_{Con,ym}$ are standard errors. Equations (1.3) are estimated using data prior to releasing the information of LEZ to the general public (during the period that the LEZ was neither implemented nor publicly announced). I then test if the resulting

[4]The parallel trend assumption is typically tested using multiple years data for the treated and control groups. Since I only have 2 years data prior to the LEZ announcement, I instead consider using a year-month indicator to see if the trends are parallel over year-month combinations.

coefficients γ_{Tr} and γ_{Con} are identical. Failure to reject the null hypothesis implies that the parallel trends assumption is valid.

1.4.3 Spillover Effects

While the LEZ restricts non-compliant vehicles from entering Greater London, it may cause considerable spilloverts in surrounding areas. On one hand, drivers approaching the zone can choose to detour to avoid paying the fee. On the other hand, drivers already complying with the emission standards may also drive outside the zone and contribute fewer emissions to non-LEZ regions. These possibilities generate opposite spillovers and make the sign of the overall effect unclear. To empirically estimate the magnitude of spillover, I apply a similar DID model as follows:

$$\log(AP_{it}) = \sum_{\tau=0}^{\tau=3} \phi_\tau P\tau_t + \sum_{\tau=0}^{\tau=3} \eta_\tau P\tau_t Near_i + \lambda Near_i + Y_t + Q_t + \boldsymbol{X_{it}\theta} + \varepsilon_{it}. \tag{1.4}$$

where $Near_i$ equals 1 if station i, which locates nearby the LEZ border, is potentially affected by the policy. $Near_i$ equals 0 if the station is either located far from ($\geq$ 10 miles) the LEZ border or unlikely to be affected by the LEZ policy (e.g. urban residential stations).[5] All the other variables remain the same as those in equation (1.2). The coefficient of interest, η_τ, measures the spillover effect of the LEZ phase τ.

1.5 Data

1.5.1 Air Pollution Data

The study focuses on two highly correlated air pollutants, PM_{10} (particulate matter with less than 10 micrometer in diameter) and NO_2 (nitrogen dioxide). PM_{10} is the primary air pollutant under the regulation of London's LEZ. In particular, all the emission standards (i.e., EU III or IV) are based upon the vehicular emission of PM_{10}. Supplementary to PM_{10}, I consider NO_2, one component of nitrogen oxides, which is also largely attributed to traffic

[5]More details regarding how stations are clarified are provided in section 1.5.

emissions. Both PM_{10} and NO_2 are highly associated with respiratory symptoms and have caused thousands of premature deaths in London each year (Gowers et al. (2014)).

I collect and combine daily PM_{10} and NO_2 data from London Air Quality Network (LAQN) and Air Quality England (AQE) between 2005 and 2010. This period starts roughly two years before the initial phase of the LEZ and ends shortly after the implementation of phase III. The LAQN is maintained by King's College London. It currently contains over 100 air monitoring stations within and nearby Greater London. AQE, supported by Ricardo-AEA[6], collects and combines air quality data from Automatic Urban and Rural Network (AURN) and local authorities of England. I exclude stations located outside of England and stations with more than 25 percent missing data on daily PM_{10} and NO_2 during the sample period.[7] [8]

Air quality data after 2010 are dropped for two reasons. First, the local authorities of Greater London that provide support to LAQN faced deficits starting from late 2010. A large number of stations were shut down due to funding shortages, while others were subject to reshaping. Consequently, some pollutants were no longer reported and some others were reported less frequently. Dropping observations from 2011 and beyond helps maintain a balanced panel and retain a reasonably large sample size. Second, the City of London began considering further emission controls and regulations in 2010. In the following years, London boroughs and the City of London proposed supplementary transport policies for air quality management, including controlling vehicular emissions from taxis, usage of low emission buses and a central and inner low emission zone.[9] The adoption of multiple regulations would make it more difficult to isolate the treatment effect of any single policy.

---

[6]Ricardo-AEA is a worldwide association that provides data and advice on global energy and environmental challenges. For more information, visit www.ricardo-aea.com/cms/.

[7]Stations located outside of England are such that locate in Scotland, Wales, or Ireland.

[8]Stations with less than 25 percent missing data for at least one air pollutant are retained. Dropping stations with more than 20 percent missing data gives us similar results but substantially reduces our sample size.

[9]Visit www.cityoflondon.gov.uk for further details.

Depending on the relative distance to the LEZ border, each air monitoring station is classified as either inside, nearby, or far. Inside stations are located within the LEZ. An inside station in a traffic intensive area (e.g., roadside, curbside, etc.) is expected to experience the most significant treatment effect if the policy leads to a dramatic change in vehicular emission. Nearby stations are located within 10 miles of the LEZ border and are thus not directly exposed to the treatment.[10] However, it is likely that nearby stations in high traffic volume areas would experience a joint effect of negative externality created by aggregate behavioral adjustments (e.g. searching routes or detouring to avoid entering the zone) and positive externality resulted from vehicles taking actions to comply with the LEZ policy. Far stations are located at least 10 miles outside the LEZ border. Presumably, they receive neither the treatment nor the externality from the LEZ regulation and therefore can serve as a valid control group. The sample of control stations spans across South East, North West, East of England, and Yorkshire and The Humber. Figure 1.3 shows the geographical locations of the three groups of stations.

According to LAQN and AQE, each air quality station is labeled as either roadside, urban background, or industrial (hereafter Type R, U, I, respectively) based on its relative distance to major sources of pollution.[11] A station with a specific type, as noted by King's College London, can be considered as broadly representative of the environmental features of neighbouring locations.[12] Table 1.2 outlines the criteria of Type R, U and I stations. Because London's LEZ affects heavily polluting commercial vehicles that are allowed to drive only on truck roads, urban background stations which typically locate in residential areas isolated

[10]The 10 miles cutoff is calculated using the following method. First, researches have shown that PM_{10} can stay in the air for about 1 hour before settling down. Using the average wind speed of 4.2 meter per second in our sample, we obtain that the maximum distance PM_{10} could travel is approximately 8.9 miles. We then allow for possible variations in wind speed and re-suspension of particulate by increasing the distance to 10 miles.

[11]There are another two types of stations, curbside (C) and suburban (S). Due to missing data, we are unfortunately unable to include these two types in our study.

[12]For example, an industrial station record PM_{10} level in industrial areas that is similar and representative of those recorded by nearby industrial areas of similar size. Visit http://www.londonair.org.uk/london/asp/publicdetails.asp?region=0 for more information

from major roads and industrial plants are eligible to serve as another control group in the DID model. Figure 1.3 illustrates the geographical locations of each type of stations.

Figure 1.4 plots the probability density functions (PDF) for daily average PM_{10} and NO_2 below the station level 95 percentile.[13] With the 5 percent station-specific highest observations excluded, the daily distributions for PM_{10} and NO_2 are still left skewed. For PM_{10}, massive number of observations center around or below 20 $\mu g/m^3$ (40 percent of the EU 24-hour limit at 50 $\mu g/m^3$). However, roughly 6 percent roadside PM_{10}, 2.6 percent urban background PM_{10} and 20.3 percent industrial PM_{10} exceed the EU 24-hour limit during the period of 2005-2010.[14] For NO_2, all sample years have an annual average exceeding the EU annual limit at 40 $\mu g/m^3$, while about 0.5 percent daily observations in the sample reach above the EU 1-hour limit of NO_2 at 200 $\mu g/m^3$.[15]

1.5.2 Weather Data

I obtain comprehensive weather data from WeatherSpark operated by Cedar Lake Ventures Incorporation. WeatherSpark collects and combines data from the National Oceanic and Atmospheric Administration (NOAA), the Norwegian Meteorological Institute (NMI) and World Weather Online (WWO).[16] Compared with LAQN and AQE stations, WeatherSpark stations locate relatively sparsely, possibly due to the similarities of climate conditions in neighbouring areas. Based on the spatial distribution of air quality stations, I include 4 weather stations in my sample: Liverpool John Lennon Airport, London Heathrow Airport, Leeds Bradford International Airport and Benson Royal Air Force Base. To obtain a balanced panel, each air quality station is matched with the nearest weather station, regardless of station type. In the matched dataset, I consider a rich set of the weather conditions that may affect air pollution concentrations, including daily average humidity, hours of precipitation,

[13] Full-length PDFs for the entire sample have long and thin tails to the right.

[14] The EU 24-hour limit of PM_{10} cannot be exceeded more than 35 times each year.

[15] The EU 1-hour limit cannot be exceeded more than 18 times each year.

[16] Visit http://weatherspark.com/ for more information.

air pressure, average and maximum wind speeds and temperature. I allow differential impacts of temperature on air pollution concentrations in different seasons by interacting the daily temperature with quarter of the year. Finally, I create a dummy variable indicating zero precipitation for 3 days to account for rising un-suspended dusts if the weather is dry.

Table 1.3 provides summary statistics for the air pollution and weather conditions in my sample. The average daily humidity in London and other regions of England is about 77 percent and the average daily hours of precipitation is about 3.2 hours. Compared with many other European countries, England is relatively humid, with only 18.5 percent days experiencing no raining for at least 3 days. The climate conditions in England reveals the importance of including precipitation as one of the key control variables because the daily average concentrations of small particles may be substantially affected by the daily precipitation level. The year-round average daily temperature measured by the four weather stations in England is about 11.2 degrees centigrade, with a standard deviation of about 5.8 degrees. The daily average and maximum wind speeds are 4.2 and 6.7 meters per second, respectively. Both average and maximum wind speeds are controlled for in the regression because a moderate wind speed will help disperse the concentrated air pollution while a high wind speed may instead blow and distribute the already suspended particles. Besides pollution and weather variables, I generate additional indicators for non-working days, including approximately 29 percent weekends and 7 percent national and banking holidays.[17]

1.6 Estimation Results

1.6.1 Graphical Analysis

To begin with, I plot the time-series deseasonalized air pollution in the UK from 2005 to 2010. The deseasonalization takes the following two steps. First, I run the following ordinary

[17]The religious holidays are excluded here because they affect only a minority of population.

least squares (OLS) regression by station type[18]:

$$AP_{i,t} = \sum_{m=1}^{m=12} \rho_m MONTH_t + \mu_{i,t} \tag{1.5}$$

where $AP_{i,t}$ indicates daily average air pollution measured at station i at date t and $MONTH_t$ is a vector of dummy variables indicating whether date t falls in month m. Next, I obtain the deseasonalized air pollution $DesAP_{i,t}$ by subtracting the estimated month effects from the observed pollution level:

$$DesAP_{i,t} = AP_{i,t} - \rho_m[MONTH_t = 1]. \tag{1.6}$$

Figure 1.5 plots the locally weighted (lowess) smoothed curves for deseasonalized PM_{10} across station types and groups. In each graph, the bandwidth is chosen to reveal the long-run trends of air quality without over-smoothing so that local fluctuations are still reserved. The period of early 2005 and late 2010 may show some degree of noise because of the relatively sparse observations in the time window. From mid 2005 to mid 2010, a gradual decline of PM_{10} can be seen from all types of stations all over the England. The largest air quality improvement, however, occurs in industrial areas inside London where the daily average PM_{10} drops by more than 20 percent from above 50 $\mu g/m^3$ (i.e. the EU daily limit) to below 40 $\mu g/m^3$ (i.e. the EU annual average limit).

Figure 1.6 plots the lowess smoothed curves for deseasonalized NO_2 from 2005 to 2010. Instead of a downward trend, NO_2 in all areas of England stays relatively stable. Along roadside, the average daily PM_{10} inside London reaches above 65 $\mu g/m^3$, exceeding the EU annual limit by over 60 percent. Similarly, the average daily PM_{10} in urban areas inside London stays above 50 $\mu g/m^3$, exceeding the EU annual limit by roughly 25 percent.

1.6.2 Parallel Trends

The LEZ policy affects commercial vehicles attempting to drive inside Greater London and not compliant with the corresponding EU emission standards. Therefore, stations located

[18]Running the regression using pooled data generates very similar results.

in traffic intensive areas such as major truck roads inside the zone naturally serve as the treated group. As proposed earlier, the two control groups in the DID model are (a) urban background stations located inside the LEZ; and (b) traffic-intensive stations located far from the LEZ. For the estimation of spillover effect, I test whether stations located in heavy traffic areas within 10 miles of the LEZ border experience changes of air pollution compared with (a) nearby urban background stations; and (b) distant traffic stations far from the LEZ after the policy has been implemented.

For each treated-control combination, I first test if the parallel trends assumption holds using equation (1.3). Table 1.4 presents the test results. Both control groups are valid for the DID estimation of treatment and spillover effects.

1.6.3 Main Results

Next I proceed with the empirical estimation. Table 1.5 displays the estimation results based on the baseline specification (1.1) across control groups. For both PM_{10} and NO_2, the simplest model, shown by columns (1) and (4), accounts for only year FE and quarter FE. Columns (2) and (5) additionally include current and lagged weather conditions. Finally the richest model, represented by columns (3) and (6), additionally control for weekend and holiday dummies. In each column, robust errors clustered at station level are shown in parentheses.

The estimation results are very similar across control groups. According to Table 1.5 Panel A, the PM_{10} in traffic intensive areas inside London has seen no significant reduction after the LEZ implementation. Instead, a 3-4 percent insignificant increase in PM_{10} is observed in treated areas following the operation of the LEZ. This point estimate persists from the simplest model to the richest though the R^2 has improved when more confounders are incorporated. Similarly, Table 1.5 Panel B shows no statistically significant reduction of NO_2 during the post-LEZ period. The R^2 shows a similar stronger explanatory power of the model when accounting for weather conditions, weekends and holidays.

The point estimates in Table 1.5 could be biased down by pre-compliance behaviors.

Because the public notification regarding to Londont's LEZ was released by TfL roughly 9 months prior to the operation of the initial phase, drivers with vehicles affected by the policy may choose to retrofit, convert to gas or reorganize fleet to comply with the policy well in advance. Specification (1.2) takes care of this issue by disentangling the LEZ treatment effect into effects of different phases, including the announcement of the policy. Table 1.6 presents the estimation results using the richest covariates (i.e., year FE, quarter FE, contemporaneous and lagged weather conditions, weekends and holidays). From 2008 to 2010, the least stringent phase of the LEZ brings no desirable effect on air quality, while the more stringent phases jointly reduce PM_{10} by 5 to 8 percent and NO_2 by 12 to 13 percent. This implies that a higher level of policy effectiveness may largely result from a broader range of targeted vehicles and a more stringent restriction on vehicular emissions.

Unlike the results in Table 1.5, coefficients displayed in Table 1.6 using different control groups show some degree of inconsistency. After the LEZ announcement but before the policy operation, treated stations inside the zone show a 7 percent significant increase in NO_2 when the model uses urban background stations as controls. However, no similar effect is observed when the model uses traffic stations far from the zone as controls. Similarly, during the LEZ phase I the estimated increase in PM_{10} near high-traffic areas within the zone is 6 percent with controls being traffic areas in distant regions, with the point estimate roughly 2 percent higher and more significant than that obtained using controls as urban background stations in London. During later phases, using urban background stations as controls consistently generates larger and more significant coefficients for the reduction of PM_{10} and NO_2.

The existence of coefficient inconsistency can be caused by several reasons. First, random shocks of air pollution within a certain area could be correlated, thus it may be plausible to allow for potential correlations in standard errors across stations. This possibility is discussed in details in section 1.6.5. Second, the two control groups may each be very slightly affected by the LEZ regulation because of drivers' behavioral adjustments in different forms. Section 1.7 discusses and provides supportive evidence for this possibility.

1.6.4 Spatial Heterogeneity

The average treatment effect (ATE) of the LEZ regulation is sensitive to station locations. In particular, the point estimates obtained in Section 1.6.3 may change substantially if the proportion of roadside and industrial stations changes. In this section I explore the spatial heterogeneity of the treatment effect of the LEZ using specification (1.2). The two control groups are urban background stations inside London and distant traffic stations that have the same environmental classification as the treated stations. Because all the treated NO_2 stations locate along roadside of London, I provide the heterogeneity analysis for PM_{10} only.

Table 1.7 shows that roadside and industrial stations inside London experience different treatment effects of the LEZ. While roadside PM_{10} has increased by 5 to 9 percent following the initial phase of the LEZ, industrial PM_{10} do not change dramatically. In each of the extended phases, the more stringent emission standards on subject vehicles have led to a 7 to 12 percent decline of industrial PM_{10}, but their impacts on roadside PM_{10} are considerably lower and less significant. While this heterogeneity might be caused by the small sample size of industrial stations, it is more plausible to result from the differences in vehicle composition across areas. Specifically, a higher proportion of vehicles driving in industrial areas are heavily polluting diesel vehicles used for commercial purposes. Those vehicles are in general non-exempt from the LEZ regulation and the affected drivers are more likely to take actions to comply with the emission standards.

1.6.5 Robustness Check

Standard Errors

This section considers factors that would possibly affect the coefficient robustness. The first and most common concern is the quality of data and the reliability of pollution measurement. Let $AP^*_{i,t}$ be the true value of daily average pollution level. The observed $AP_{i,t}$ can be written as

$$AP_{i,t} = AP^*_{i,t}(1 + \nu_{i,t}) \tag{1.7}$$

where ν_{it} is the degree of bias from the true value. Plugging equation (1.7) into specification (1.2), we have

$$\log(AP_{it}) = \sum_{\tau=0}^{\tau=3} \beta_\tau P\tau_t + \sum_{\tau=0}^{\tau=3} \gamma_\tau P\tau_t Tr_i + \alpha Tr_i + Y_t + Q_t + \boldsymbol{X_{it}\theta} + \mu_{it} \tag{1.8}$$

where $\mu_{it} = \varepsilon_{it} + \log(1+\nu_{it})$ is the combined error. Depending on the source of measurement error, μ_{it} may or may not correlate with air pollution concentrations. When $\nu_{i,t}$ reflects pure random errors of the reported pollution values, it can be merged into ε_{it}, which the preferred specification can correctly handle. If the $\nu_{i,t}$ contains some level of aggregate shocks[19], the assumption that standard errors are station-clustered will no longer hold. This can be tested using an alternative clustering of standard errors at type-region level. Here, type refers to the station's environmental classification (i.e., R, I or U), and region refers to either London's borough (for inside stations) or city/district of England (for outside stations). Meanwhile, ν_{it} may be serially correlated within stations, reflecting some lags of random shocks over time. This possibility can be handled using Newey-West standard errors up to the maximum number of days that pollution shocks may persist.

Table 1.8 presents the results of robustness checks on the coefficients' significance levels. Columns (1) and (4) modify the preferred specification by assuming that the error is independent across stations and years. This assumption is appropriate if the annual maintenance on air monitoring networks leads to little correlation in the measurement errors over time. Using the two control groups, I find that this modification slightly increases the standard errors of the DID estimators. Nevertheless, I still find no evidence of a statistically significant decline of either PM_{10} or NO_2 following the LEZ announcement and phase I. Columns (2) and (5) display the estimation results using clustered standard errors by type-region. For both pollutants, the estimation coefficients are highly consistent with those presented in the preferred specification. This implies that the potential across-station random shocks do not alter the robustness of the point estimates. Lastly, columns (3) and (6) apply the heteroscedasticity-

[19]For instance, a heavily polluting truck traveling on ring road M25 is likely to increase the air pollution at multiple stations along the way.

consistent Newey-West standard errors. Following Stock and Watson's rule of thumb, I set $L = 20$ to incorporate relatively strong serial correlations. The Newey-West standard errors are very similar to those in the preferred specification. The results reconfirm the effectiveness of phase III in reducing air pollution and the insignificant treatment coefficients in the preceding, less strict phases of the LEZ (except for the announcement).

Extreme Outliers

Aside from standard errors, one may concern about the issue of endogeneity. In particular, if ν_{it} is correlated with one of the key regressors, the coefficient of interest γ_τ in equation (1.8) will suffer from omitted variable bias. For instance, certain weather conditions may cause both the true value of particulate matter and the measurement error to highly deviate from their normal levels.[20] The issue may exacerbate if this abnormality occurs aggregately.

While we can hardly determine the source of measurement error, we may detect possible outliers based on the shape of the station specific distribution of air pollution concentrations $AP_{i,t}$. According to Figure 1.4, the daily average PM_{10} and NO_2 depict highly left skewed distributions with thin and long right tails. To avoid bias from misreported high pollution concentrations, I re-estimate the model with exclusion of observations exceeding the station-specific 95th percentile.[21] This modification creates an unbalanced panel with gaps in date.

Table 1.9 estimates specification (1.2) using the modified sample. The displayed results are highly consistent with the one using the full sample, indicating that the heaviest polluting 5 percent observations do not bias the estimates. As columns (1) to (3) show, the most significant treatment effect still occurs after the implementation of phase III. Overall, using both urban background stations inside London and traffic stations far from London yield similar results that the targeted areas of the LEZ do not experience a statistically significant

[20]Specifically, the maximum record of daily PM_{10} reaches 383 $\mu g/m^3$, a value almost 8 times as high as the EU daily limit. Additionally, 0.4 percent PM_{10} observations range above 100 $\mu g/m^3$ and another 6 percent range above 50 $\mu g/m^3$.

[21]Excluding 90 percentile of station specific air pollution distributions lead to very similar results with the presented ones.

improvement of air quality until the broadest vehicles are targeted and the emission standards on subject vehicles are tightened.

1.6.6 Spillover Effects

The goal of the LEZ regulation is incentivizing heavily polluting vehicles to replace, upgrade or retrofit to generate long-run social benefits. It may, however, affect the surrounding areas of London if a substantial proportion of non-compliant vehicles choose to detour to avoid entering the zone. This unfavorable effect may offset the positive spillover caused by compliant vehicles driving outside the zone. To determine their joint impacts on high-traffic areas relatively close ($\leq$ 10 miles) to the LEZ border, I estimate equation (1.4) using (a) urban background stations nearby London; and (b) high-traffic stations far from London as controls.

Table 1.10 presents the estimated spillover effects. For both PM_{10} and NO_2, the secondary spillover effects have the same sign as the primary treatment effects. Across specifications, the PM_{10} spillover remains insignificant even with the most stringent emission standards. In contrast, the estimated NO_2 spillover during the LEZ phase III is as high as 14 percent using urban background stations nearby London as controls. This secondary effect rises roughly 6 percent above the estimated primary effect of the LEZ on NO_2 during phase III. However, using high-traffic stations far from London as controls yields a much smaller and insignificant coefficient for the same spillover effect. A possible reason could be that urban background locations nearby the zone are adversely affected by the detoured vehicles, while the nearby traffic areas are positively affected by the cleaning-up of subject vehicles. This potential sorting of vehicles close to the LEZ border could increase the estimated spillover effect of the LEZ captured by the DID estimator.

1.7 Discussion and Supportive Evidence

In earlier sections, I show that the introduction of London's LEZ does not effectively alleviate the concern of air pollution inside London until the most stringent phase. Moreover, only industrial areas close to power plants experience a significant treatment effect of the LEZ. This section provides two possible explanations for the t-statistics. These two hypotheses are then tested using traffic flow and fuel consumption data in London and other regions of England.

1.7.1 Behavioral adjustments

I first consider the possibility that the air quality measured by the second control group (i.e., high-traffic stations in distant locations) represents a lower level of treatment effect instead of pure background air quality when the policy is not in place. This possibility does not invalidate the parallel trends assumption which focuses on periods prior to the LEZ establishment. Because of the national flow of affected vehicles, other regions of England may also benefit from London's LEZ although they are not within the targeted area. To see this, consider the disparity of point estimates across control groups in Table 1.6 and 1.8. The treated stations near major roads and industrial sites show a larger treatment effect if compared with urban background stations inside London instead of the same type of stations far from London. Since the LEZ policy targets at commercial vehicles with high demand for road freight transport, drivers may have a strong incentive to comply with the emission standards instead of paying the fee to drive inside the zone. This strategic decision could potentially lead to a change of commercial vehicle fleet composition in other regions of England and thus reduce the traffic-related air pollution emissions all over the England.

To empirically test the existence of this positive externality in distant locations, I obtain the LEZ compliance data from Traffic for London (TfL). The compliance data consist of weekly compliance rate and the number of subject and non-compliant vehicles. To make use of the richest dataset, I linearly interpolate daily compliance rate and non-compliant statistics and merge that with air pollution concentrations in the South East, North West, East of England, Yorkshire and The Humber in which the control stations locate. Let $S\tau_t$, $NC\tau_t$ and $C\tau_t$ denote the total number of subject vehicles, non-compliant vehicles and compliant vehicles in phase τ, respectively, and let $CR\tau_t$ denote the compliance rate at date t in phase

τ. The interpolated daily number of non-compliant vehicles and compliance rate can be expressed by

$$NC\tau_t = \frac{7-w}{7} NC\tau_{s-1} + \frac{w}{7} NC\tau_s \tag{1.9}$$

$$CR\tau_t = \frac{C\tau_t}{S\tau_t} = \frac{S\tau_t - NC\tau_t}{S\tau_t} \tag{1.10}$$

where $s-1$ indexes the previous Sunday and s indexes the current Sunday when TfL releases the weekly compliance data. w indexes day of week at date t with Monday equalling 1.

The next step is to estimate whether the air quality in other England regions is affected by the compliance to London's LEZ. The econometric models applied to accomplish this goal are specified as follows:

$$\log(AP_{it}) = n_\tau NC\tau_t + \eta_i + Q_t + \boldsymbol{X_{it}\theta} + \varepsilon_{it}, \tau = 1, 2 \tag{1.11}$$

$$\log(AP_{it}) = c_\tau CR\tau_t + \eta_i + Q_t + \boldsymbol{X_{it}\theta} + \varepsilon_{it}, \tau = 1, 2. \tag{1.12}$$

In both specifications, η_i stands for unobservable station specific fixed effects, Q_t stands for quarter fixed effects. $\boldsymbol{X_{it}}$ contains contemporaneous and lagged weather conditions, day of week dummies and a holiday dummy. ε_{it} is heteroscedastic and clustered by station.

Specification (1.11) and (1.12) have three major distinctions compared with the main regression (1.2). First, the dichotomous treatment group indicator, the LEZ phase indicator and their interactions are replaced by continuous variable $NC\tau_t$ and $CR\tau_t$ which show drivers' compliance behaviors to the LEZ policy. Second, the year fixed effects in equation (1.2) has been dropped in equations (1.11) and (1.12) because the compliance data only span for two years from June 2007 to October 2009.[22] Third, in equations (1.11) and (1.12) the weekend dummy contained in $\boldsymbol{X_{it}}$ is replaced by a vector of day of week dummies because the compliance data from Monday to Saturday each week are interpolated from the reported ones on the neighboring Sundays. The day of week dummies will largely absorb any inaccuracy in the interpolation process given by equations (1.9) and (1.10).

[22]The time span only allows me to estimate the effects of compliance to phase I and II of the LEZ on air quality in other England regions.

Figure 1.7 plots the time series number of non-compliant vehicles in phase I and II and the corresponding compliance rate from mid 2007 to late 2009. The phase I compliance rate is about 75 percent shortly after the LEZ announcement and quickly reaches above 85 percent prior to the starting date of phase I. After the operation of phase I, the compliance rate continues rising to roughly 95 percent and stays stable thereafter. On the other hand, the compliance rate in phase II starts at roughly 60 percent in mid 2007 and gradually increase to about 80 percent after one year. The implementation of phase II further brings up the compliance rate to 90 to 95 percent from late 2008 to late 2009.

Table 1.11 presents the estimation results of specification (1.11) based on (a) all distant stations; and (b) stations in traffic intensive (i.e. R and I) areas only in distant locations. It is evident that a rising number of non-compliant vehicles to the LEZ regulation would also negatively contribute to the air quality in regions far from London. For every one thousand additional non-compliant vehicles entering London in phase I, distant locations in England experience a 0.5-0.6 percent increase in PM_{10} and a 0.4 percent increase in NO_2. Similarly, every one thousand increase in the number of non-compliant vehicles in phase II brings 0.4-0.5 percent increase in PM_{10} and 0.3 percent increase in NO_2 to the other England regions. Although some of these point estimates are not very significant, they play a role in driving down the DID estimators in the main specification toward zero.

Table 1.12 displays the results for specification (1.12) using all stations and traffic intensive stations only in other England regions. The point estimates are all negative, indicating that a higher LEZ compliance rate is associated with a lower level of air pollution far from London. This positive externality tends to be stronger in phase I than phase II, which could explain the insignificant t-statistics for the phase I coefficients. For every 10 percent increase in phase I compliance rate, distant locations experience a 2.7-3.4 percent decline of PM_{10} and a roughly 2.3 percent decline of NO_2. Similarly, every 10 percent increase in phase II compliance rate is expected to bring a 2.2-2.8 percent decrease of PM_{10} concentrations and a 1.3 percent decrease of NO_2 concentrations all over the England. This evident positive spillover effect serves as the external reason for the incapability of the DID estimator to capture a significant treatment effect of the LEZ regulation, especially in the early phases.

1.7.2 Traffic flow

The second plausible reason for the insignificant treatment effect is drivers' behavioral adjustments to avoid being charged or penalized. For instance, non-compliant commercial vehicle drivers may have an incentive to substitute toward a non-subject vehicle. I explore this potential behavioral adjustment by analyzing the changes of road usage and fuel consumption of different types of vehicles.

I first collect the traffic flow data from the United Kingdom Department of Transportation (DfT). The traffic data contain annual average daily traffic flow (AADF) by vehicle types on major roads in England. In London, traffic emissions from vehicles on these major roads take roughly 65 percent total emissions of PM_{10} and NO_x (a mixture of NO and NO_2). For each major road, I have data on the road location, type, length, and the traffic estimates based on the length between the major road conjunctions. I drop roads located in England regions which contain none of the air quality stations in the sample.

Figure 1.8 plots the percentage change of AADF by vehicle types across five major regions of England: London, East of England, South East of England, North West of England, Yorkshire and The Humber.[23] The base year is 2008 in which both phase I and phase II of the LEZ were launched. The traffic index is then obtained by taking the ratio of traffic flow for each type of vehicles in any given year to that in the base year. According to Figure 1.8 Panel A, London has seen a relative increase in the traffic volume of heavy goods vehicles (HGVs) in 2008 compared with other regions. A higher traffic volume of HGVs may emit considerable amounts of particles and other pollutants into atmosphere even if the compliance rate of HGVs continues rising. In contrast, the traffic volume of light goods vehicles (LGVs) travelling in London started to decline from 2007 (Figure 1.8 Panel

[23] I sum up traffic flows on the following types of road: M or Class A principal motorway, class A principal road in Rural area, class A principal road in urban area, M or class A trunk motorway, class A trunk road in rural area, and class A trunk road in urban area.

B). Although similar trends have also been found in other regions of England, London has experienced the highest percentage decline in LGVs since the operation of the LEZ phase II. In regards to the non-targeted vehicles from phase I to III, the traffic volume of buses and coaches (exempt from the LEZ regulation until the LEZ phase IV in early 2012) was very stable in London but gradually declining in other England regions during 2006 and 2009 (Figure 1.8 Panel C). Nevertheless, cars and taxis driving in London have experienced a dramatic drop of traffic volume prior to 2008 but a subsequent gradual increase since 2008. The turning point of 2008 clearly shows a substitution effect toward non-subject vehicles. In each phase of London's LEZ, private vehicles are generally not under regulation, but they keep contributing a considerable amount of traffic emissions. For instance, in 2010 the emissions from diesel cars contribute 12 percent of the total PM_{10} and 11 percent of the total NO_x. Petrol cars and taxis contributed 16 and 2 percent of the total PM_{10} and 7 and 2 percent of the total NO_x, respectively (TfL 2014).

1.7.3 Fuel consumption

Both diesel and petrol vehicles emit particulate matter. The former, however, emit substantially higher amounts of particulate than the latter. To investigate how the LEZ affects the fuel market, I collect annual fuel consumption data from the United Kingdom Department of Energy and Climate Change. Figure 1.9 plots the annual aggregate diesel and petrol consumptions in London. Although petrol consumption gradually declined over years, diesel consumption continuously rose from 2005 to 2008, even when both phase I and II were officially launched. Not until 2009 did the diesel consumption in London start to decline, but with a compensation of a lower speed of decrease of petrol consumption.

Figure 1.10 shows the annual diesel consumptions for four major types of diesel vehicles driving inside London from 2005 to 2010. Diesel cars, which consume over 30 percent of the total diesel from road traffic sectors, consistently increased their share by roughly 20 thousand tonnes each year. HGV and LGV each contributed to roughly 25 percent (equivalent to 300 thousand tonnes) of the total diesel consumption every year, making themselves the second largest diesel consumer in the traffic sector. The former, however, started to consume 20-25 thousand tonnes diesel more than the latter from 2008, and this gap persisted till the end of the sample period. Buses, the least diesel contributor among the four, consumed roughly 200 thousand tonnes of diesel each year.

Combining data from traffic flow and fuel consumption, we can see a clear trend of substitution from mid-sized petrol-powered vehicles to heavy-duty or privately used diesel-powered vehicles. In particular, the fact that diesel fuel consumption in London reached its peak in the year 2008 largely results from an increasing traffic flow of HGVs and diesel cars. Even after the establishment of the LEZ, diesel cars still experienced an upward trend of diesel consumption till the end of 2010. Unless an efficient regulation targets at heavily polluting private vehicles, rising emissions from diesel cars would be a big concern on reducing air pollution in London.

1.8 Conclusion

In this paper I analyze the environmental impacts of London's low emission zone. The preferred difference-in-differences estimation shows no evidence that the air quality in London has been successfully improved until the most stringent restriction starts to operate. Under the more stringent restrictions, however, a significant reduction of air pollution occurs only close to industrial sites instead of major truck roads, possibly due to the differences in vehicle composition across areas. Combining data from compliance statistics, traffic flow and fuel consumption, I propose that the insignificance of the DID estimators in the less stringent phases of the LEZ could be caused by a positive externality in other England regions. On the other hand, only a small percentage of HGVs are affected by the policy and those travelling inside London may have an intensive usage of roads. For LGV drivers, there is evidence showing a shift of vehicles toward heavily polluting, non-subject vehicles such as diesel cars. This substitution effect reinforces the increasing traffic flow of HGVs in the year 2008, serving as the internal force that drives down the DID estimators toward zero.

Reference

R.W. Atkinson, B. Barratt, B. Armstrong, H.R. Anderson, S.D. Beevers, I.S. Mudway, D. Green, R.G. Derwent, P. Wilkinson, C. Tonne, and F.J. Kelly. The Impact of the Congestion Charging Scheme on Ambient Air Pollution Concentrations in London. *Atmospheric Environment journal*, 43(34):5493–5500, 2009.

Hanna Boogaard, Nicole A.H. Janssen, Paul H. Fischer, Gerard P.A. Kos, Ernie P. Weijers, Flemming R. Cassee, Saskia C. van der Zee, Jeroen J. de Hartog, Kees Meliefste, Meng Wang, Bert Brunekreef, and Gerard Hoek. Impact of Low Emission Zones and Local Traffic Policies on Ambient Air Pollution Concentrations. *Science of the Total Environment*, 435–436:132–140, 2012.

David C. Carslaw and Sean D. Beevers. The Efficacy of Low Emission Zones in Central London as a Means of Reducing Nitrogen Dioxide Concentrations. *Transportation Research Part D*, 7:49–64, 2002.

Yuyu Chen, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. The promise of Beijing: Evaluating the impact of the 2008 olympic games on air quality. *Journal of Environmental Economics and Management*, 66:424–443, 2013.

City of London. City of London Air Quality Strategy 2011-2015, 2011.

Lucas Davis. The effect of driving restrictions on air quality in mexico city. *Journal of Political Economy*, 116(1):38–81, 2008.

Richard B. Ellison, Stephen P. Greaves, and David A. Hensher. Five years of Londons low Emission Zone: Effects on Vehicle Fleet Composition and Air Quality. *Transportation Research Part D*, 23:25–33, 2013.

European Environmental Agency (EEA). Air Quality in Europe - 2015 Report, 2015. European Environment Agency, Copenhagen.

European Environmental Agency (EEA). Explaining Road Transport Emissions: A Non-technical Guide, 2016. European Environmental Agency, Copenhagen.

European Topic Centre on Air Pollution and Climate Change Mitigation. Road traffics contribution to air quality in european cities, November 2012. European Topic Centre on Air Pollution and Climate Change Mitigation (ETC/ACM). Technical Paper 2012/14.

Freight Transport Association. FTA Compliance Guide to Greater London Low Emission Zone, 2010.

A M Gowers, B G Miller, and J R Stedman. Estimating Local Mortality Burdens associated with Particulate Air Pollution, 2014. Centre for Radiation, Chemical and Environment Hazard, Public Health England.

Colin P. Green, John S. Heywood, and Maria Navarro. Traffic Accidents and the London Congestion Charge. *Journal of Public Economics*, 133:11–22, 2016.

Christer Johansson, Lars Burman, and Bertil Forsberg. The Effects of Eongestions Tax on Air Quality and Health. *Atmospheric Environment*, 43:4843–4854, 2009.

Alan M. Jones, Roy M. Harrison, Benjamin Barratt, and Gary Fuller. A Large Reduction in Airborne Particle Number Concentrations at the Time of the Introduction of Sulphur Free Diesel and the London Low Emission zone. *Atmospheric Environment*, 50:129–138, 2012.

Jonathan Leape. The London Congestion Charge. *Journal of Economic Perspectives*, 20(4): 157–176, 2000.

J.J. Murphy, M.A. Delucchi, D.R. McCubbin, and H. J. Kim. The Cost of Crop Damage Caused by Ozone Air Pollution from Motor Vehicles. *Journal of Environmental Management*, 55:273–289, 1999.

Michele De Rosa, Stefano Zarrilli, Luigi Paesano, Umberto Carbone, Bartolomeo Boggia, Mario Petretta, Antonella Maisto, Francesca Cimmino, Giancarmelo Puca, Annamaria Colao, and Gaetano Lombardi. Traffic Pollutants Affect Fertility in Men. *Human Reproduction*, 18(5):1055–1061, 2003.

Cathryn Tonne, Sean Beevers, Ben Armstrong, Frank Kelly, and Paul Wilkinson. Air Pollution and Mortality Benefits of the London Congestion Charge: Spatial and Socioeconomic Inequalities. *Occupational and Environmental Medicine*, 65:620–627, 2008.

Traffic for London. Transport Emissions Roadmap: Cleaner Transport for a Cleaner London, September 2014.

V. Brian Viard and Shihe Fu. The Effect of Beijing's Driving Restrictions on Pollution and Economic Activity. *Journal of Public Economics*, 125:98–115, 2015.

Hendrik Wolff. Keep Your Clunker in the Suburb: Low Emission Zones and Adoption of Green Vehicles. *Economic Journal*, 124:F481–F512, 2013.

Hendrik Wolff and Lisa Perry. Trends in Clean Air Legislation in Europe: Particulate Matter and Low Emission Zones. *Review of Environmental Economics and Policy*, 4(2):293–308, 2010.

World Health Organization. Health Effects of Transport-Related Air Pollution (edited by michal krzyzanowski and birgit kuna-dibbert and schneider jurgen), 2005. Copenhagen: World Health Organization, Regional Office for Europe.

World Health Organization. Economic Cost of the Health Impact of Air Pollution in Europe: Clean Air, Health and Wealth, 2015. Copenhagen: World Health Organization, Regional Office for Europe.

Joshua S. Graff Zivin and Matthew J. Neidell. The Impact of Pollution on Worker Productivity. NBER Working Paper No. 17004, 2011.

LIST OF FIGURES IN CHAPTER 1

LIST OF TABLES IN CHAPTER 1

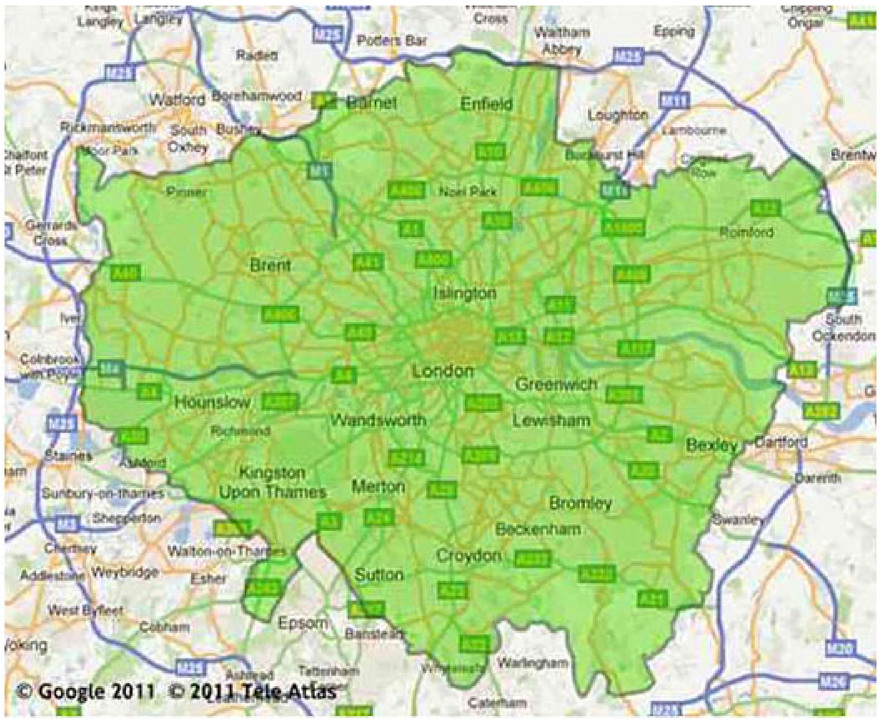


Figure 1.1: London Low Emission Zone Map (*Source: Traffic for London*)

Panel A: LEZ advance warning signs

Panel B: LEZ camera sign

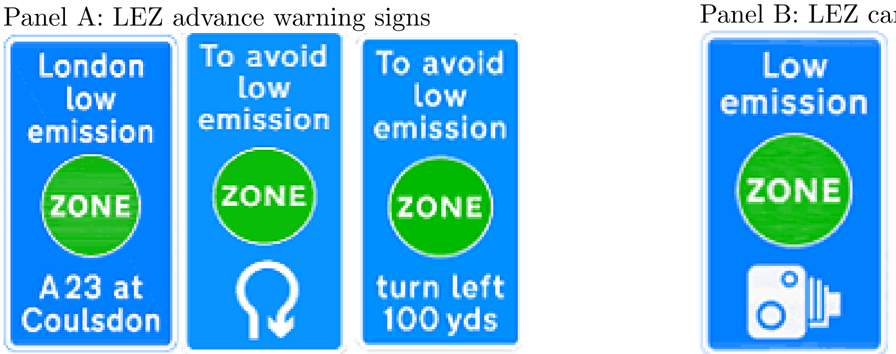


Figure 1.2: London Low Emission Zone Signs. (*Source: Traffic for London*)

Table 1.1: London's Low Emission Zone (LEZ) Timeline

LEZ announcement: May 3, 2007

LEZ scheme	Phase I	Phase II	Phase III	Phase IV
Date	Feb 4, 2008	Jul, 7, 2008	Oct 4, 2010	Jan 3, 2012
Subject vehicles	Heavy goods vehicles[b], Trucks[c]	Light goods vehicles[d]	Larger vans[e], Minibus[f]	Buses[g], Coaches[g], Heavy goods vehicles
Emission standard	Euro III	Euro III	Euro III	Euro IV
Daily fee	£200	£100	£100	£200
Penalty fee	£1000, or £500 if paid within 14 days	£500, or £250 if paid within 14 days	£500, or £250 if paid within 14 days	£1000, or £500 if paid within 14 days

Notes: Each phase specifies emission standards for a new group of vehicles while maintaining the emission requirements from all the previous phases.

a - Heavy goods vehicles, or HGVs, are good vehicles that are greater than 12 tonnes gross vehicle weight (GVW).

b - Affected trucks are diesel engine vehicles over 3.5 tonnes GVW.

c - Light good vehicles, or LGVs, are good vehicles between 3.5 and 12 tonnes GVW.

d - Affected larger vans are diesel engine vehicles between 1.205 and 3.5 tonnes GVW or motor caravans between 2.5 and 3.5 tonnes GVW.

e - Affected minibuses are diesel passenger vehicles with more than 8 seats (including driver's seat) and less than 5 tonnes GVW.

f - Affected busses and coaches are diesel passenger vehicles with more than 8 seats (including driver's seat) and more than 5 tonnes GVW.

Source: Transport for London

Table 1.2: Station classification criterion

Station Type	**Criterion**
Roadside (R)	Stations with sample inlets between 1 meter and 5 meters of the curb of a busy road. Sampling heights are within 2-3 meters of the ground.
Industrial (I)	Stations where industrial emissions make an significant contribution to pollution levels.
Urban Background (U)	Stations in urban locations away from major sources of pollution, broadly representative of town/city-wide background concentrations, e.g. urban residential areas.

Data source: King's College London

Table 1.3: Summary statistics

	PM10 sample			NO2 sample		
Variable	Mean	Std. Dev.	N	Mean	Std. Dev	N
Air pollutant						
PM_{10} ($\mu g/m^3$)	25.664	14.809	131783			
$NO2$ ($\mu g/m^3$)				46.668	24.684	148162
Station groups						
Treated (1=treated)	0.591	0.492	144606	0.649	0.477	162012
Nearby (1=nearby)	0.121	0.326	144606	0.162	0.369	162012
Control (1=control)	0.288	0.453	144606	0.189	0.391	162012
Station Types						
Roadside (1=roadside)	0.485	0.5	144606	0.608	0.488	162134
Industrial (1=industrial)	0.121	0.326	144606			
Urban background (1=urban)	0.394	0.489	144606	0.392	0.488	162134
Weather conditions						
Relative humidity (100%)	0.768	0.103	144523	0.768	0.103	161941
Hours of rain ($hours$)	3.249	4.392	144606	3.21	4.375	162012
Air pressure ($mBar$)	1014.651	10.78	144515	1014.718	10.726	161897
Temperature (^oC)	11.225	5.769	144571	11.256	5.793	161981
Average wind speed (m/s)	4.238	1.811	144567	4.203	1.788	161975
Maximum wind speed (m/s)	6.772	2.522	144567	6.715	2.486	161975
No rain for 3 days	0.185	0.388	144606	0.187	0.39	162134
Additional covariates						
Weekend (1=weekend)	0.286	0.452	144606	0.286	0.452	162012
Holiday (1=holiday)	0.068	0.252	144606	0.068	0.252	162012

Notes: Air pollution data sources are London Air Quality Network (LAQN) and Air Quality England (AQE). Weather data source is WeatherSpark.

Table 1.4: Test results of parallel trend assumption

Treated/Nearby	**Control**	χ^2	**Prob** $> \chi^2$
Panel A: PM_{10} sample			
For treatment effects			
Type R and I stations inside LEZ	Type U stations inside LEZ	0.11	0.7387
Type R and I stations inside LEZ	Type R and I stations far from LEZ	0.80	0.3713
For spillover effects			
Type R and I stations nearby LEZ	Type U stations nearby LEZ	0.03	0.8537
Type R and I stations nearby LEZ	Type R and I stations far from LEZ	0.25	0.6186
Panel A: NO_2 sample			
For treatment effects			
Type R and I stations inside LEZ	Type U stations inside LEZ	0.01	0.9351
Type R and I stations inside LEZ	Type R and I stations far from LEZ	0.17	0.6791
For spillover effects			
Type R and I stations nearby LEZ	Type U stations nearby LEZ	0.01	0.9431
Type R and I stations nearby LEZ	Type R and I stations far from LEZ	0.42	0.5161

Notes: Environmental classifications are based on LAQN and AQE.

Table 1.5: DID Estimation Results of LEZ Treatment Effects

	Control as U stations inside LEZ			Control as R and I stations far from LEZ		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: PM10						
Treatment effect	0.033	0.035	0.035	0.033	0.040	0.040
	(0.042)	(0.042)	(0.042)	(0.034)	(0.035)	(0.035)
Stations	39	39	39	37	37	37
Observations	77633	77599	77599	73944	73811	73811
R^2	0.107	0.356	0.375	0.049	0.272	0.296
Panel B: NO2						
Treatment effect	-0.000	-0.000	-0.000	-0.007	-0.004	-0.004
	(0.020)	(0.020)	(0.020)	(0.023)	(0.024)	(0.024)
Stations	48	48	48	40	40	40
Observations	96292	96248	96248	80160	80037	80037
R^2	0.158	0.406	0.445	0.107	0.320	0.361
controls						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Current weather	No	Yes	Yes	No	Yes	Yes
Lagged weather	No	Yes	Yes	No	Yes	Yes
Weekend	No	No	Yes	No	No	Yes
Holiday	No	No	Yes	No	No	Yes

Notes: This table shows the difference-in-differences estimation results based on specification 1.1 using urban stations inside LEZ and non-urban stations far from LEZ as controls, respectively. In each column, dependent variable is logarithm of daily average PM_{10} or NO_2 reported by each station. Current weather conditions include the current day's average relative humidity, hours of rain, air pressure, average and maximum wind speeds, and an interaction term between daily average temperature and quarter of year. Lagged weather conditions include previous day's relative humidity, hours of rain, air pressure, average and maximum wind speeds, an interaction term between previous day's temperature and current quarter, and a dummy variable for no raining for 3 days.

Robust standard errors clustered at station level are in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 1.6: DID Estimation Results of LEZ Treatment Effects from LEZ Announcement to Phase III

	Control as U stations inside LEZ			Control as R and I stations far from LEZ		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: PM10						
Treatment effect of announcement	0.022	0.022	0.022	-0.002	-0.002	-0.001
	(0.030)	(0.030)	(0.030)	(0.042)	(0.042)	(0.042)
Treatment effect of phase I	0.018	0.038	0.039	0.042	0.064*	0.063*
	(0.052)	(0.035)	(0.035)	(0.041)	(0.034)	(0.034)
Treatment effect of phase II		-0.023	-0.017		-0.025	-0.022
		(0.033)	(0.033)		(0.031)	(0.036)
Treatment effect of phase III			-0.062**			-0.030
			(0.030)			(0.073)
Stations	39	39	39	37	37	37
Observations	77599	77599	77599	73811	73811	73811
R^2	0.376	0.379	0.380	0.297	0.299	0.300
Panel B: NO2						
Treatment effect of announcement	0.071***	0.070***	0.070***	-0.018	-0.018	-0.018
	(0.025)	(0.025)	(0.025)	(0.033)	(0.033)	(0.033)
Treatment effect of phase I	-0.053**	-0.004	-0.004	0.009	0.051	0.051
	(0.024)	(0.023)	(0.023)	(0.022)	(0.035)	(0.035)
Treatment effect of phase II		-0.058**	-0.050**		-0.050	-0.042
		(0.023)	(0.023)		(0.038)	(0.041)
Treatment effect of phase III			-0.083**			-0.076
			(0.035)			(0.056)
Stations	48	48	48	40	40	40
Observations	96248	96248	96248	80037	80037	80037
R^2	0.445	0.446	0.446	0.361	0.361	0.362

Notes: This table shows the difference-in-differences estimation results based on specification 1.2. In each column, dependent variable is the logarithm of daily average PM_{10} or NO_2 reported by each station. Control variables include year FE, quarter FE, current and lagged weather conditions, and dummies for weekends and holidays.

Robust standard errors clustered at station level are in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 1.7: Heterogeneous Impact of LEZ Announcement to Phase III on PM_{10}

	Control as U stations inside LEZ		Control as R or I stations far from LEZ	
	Roadside	Industrial	Roadside	Industrial
	(1)	(2)	(3)	(4)
Treatment effect of LEZ announcement	0.016	0.004	-0.042	0.065
	(0.026)	(0.082)	(0.029)	(0.132)
Treatment effect of LEZ phase I	0.048	0.003	0.089^{**}	-0.003
	(0.035)	(0.055)	(0.039)	(0.077)
Treatment effect of LEZ phase II	0.001	-0.117^{**}	-0.025	-0.067
	(0.031)	(0.053)	(0.034)	(0.056)
Treatment effect of LEZ phase III	-0.046	-0.103^{**}	-0.001	-0.112
	(0.031)	(0.044)	(0.093)	(0.064)
Stations	34	17	29	8
Observations	67908	33654	58143	15668
R^2	0.427	0.484	0.335	0.468

Notes: This table shows the heterogeneous impacts of LEZ on PM_{10} concentrations. The difference-in-differences estimation is based on specification 1.2 with dependent variable being the logarithm of daily average PM_{10}. Control variables include year FE, quarter FE, current and lagged weather conditions, and dummies for weekends and holidays.

Robust standard errors clustered at station level are in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 1.8: Robustness Checks: Significance Level

	Control as U stations inside LEZ			Control as R or I stations far from LEZ		
	station-year clustered se	type-region clustered se	NW se	station-year clustered se	type-region clustered se	NW se
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: PM10						
Treatment effect of LEZ	0.022	0.022	0.022	-0.001	-0.001	0.000
announcement	(0.060)	(0.030)	(0.025)	(0.082)	(0.043)	(0.032)
Treatment effect of LEZ	0.039	0.039	0.039	0.063	0.063*	0.061
phase I	(0.077)	(0.035)	(0.038)	(0.102)	(0.034)	(0.045)
Treatment effect of LEZ	-0.017	-0.017	-0.017	-0.022	-0.022	-0.020
phase II	(0.058)	(0.035)	(0.033)	(0.075)	(0.037)	(0.039)
Treatment effect of LEZ	-0.062	-0.062**	-0.062*	-0.030	-0.030	-0.030
phase III	(0.047)	(0.030)	(0.035)	(0.094)	(0.078)	(0.055)
Observations	77599	77599	77599	73811	73811	73811
R^2	0.380	0.380		0.300	0.300	
Panel B: NO2						
Treatment effect of LEZ	0.070	0.070***	0.070***	-0.018	-0.018	-0.017
announcement	(0.066)	(0.025)	(0.025)	(0.088)	(0.024)	(0.033)
Treatment effect of LEZ	-0.004	-0.004	-0.004	0.051	0.051	0.050
phase I	(0.093)	(0.022)	(0.040)	(0.100)	(0.034)	(0.048)
Treatment effect of LEZ	-0.050	-0.050**	-0.050	-0.042	-0.042	-0.041
phase II	(0.075)	(0.021)	(0.036)	(0.080)	(0.044)	(0.044)
Treatment effect of LEZ	-0.083*	-0.083**	-0.083***	-0.076	-0.076	-0.076*
phase III	(0.048)	(0.035)	(0.032)	(0.069)	(0.045)	(0.046)
Observations	96248	96248	96248	80037	80037	80037
R^2	0.446	0.446		0.362	0.362	

Notes: This table provides robustness checks on the significance level based on specification 1.2. Column (1) and (3) displays results with clustered standard errors at station-year level. Column (2) and (4) cluster standard errors by station type (roadside, industrial and urban background) and region (London boroughs and districts of UK). Column (3) and (6) use Newey-West standard errors of up to 20 lags. All columns control for year FE, quarter FE, current and lagged weather conditions, and dummies for weekends and holidays.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 1.9: Robustness Checks: Extreme Outliers

	Control as U stations inside LEZ			Control as R and I stations far from LEZ		
	station-year clustered se	type-region clustered se	NW se	station-year clustered se	type-region clustered se	NW se
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: PM10						
Treatment effect of LEZ announcement	0.021	0.021	0.021	0.006	0.006	0.008
	(0.059)	(0.026)	(0.024)	(0.081)	(0.041)	(0.031)
Treatment effect of LEZ phase I	0.045	0.045	0.045	0.061	0.061*	0.058
	(0.077)	(0.033)	(0.036)	(0.105)	(0.033)	(0.045)
Treatment effect of LEZ phase II	-0.023	-0.023	-0.023	-0.003	-0.003	-0.002
	(0.059)	(0.035)	(0.031)	(0.078)	(0.040)	(0.038)
Treatment effect of LEZ phase III	-0.055	-0.055*	-0.055	-0.015	-0.015	-0.015
	(0.047)	(0.029)	(0.035)	(0.085)	(0.068)	(0.051)
Observations	73550	73550	73550	69962	69962	69962
R^2	0.362	0.362		0.272	0.272	
Panel B: NO2						
Treatment effect of LEZ announcement	0.075	0.075***	0.076***	-0.022	-0.022	-0.021
	(0.068)	(0.025)	(0.026)	(0.087)	(0.023)	(0.033)
Treatment effect of LEZ phase I	-0.007	-0.007	-0.007	0.048	0.048	0.046
	(0.093)	(0.021)	(0.040)	(0.099)	(0.030)	(0.048)
Treatment effect of LEZ phase II	-0.053	-0.053***	-0.053	-0.032	-0.032	-0.031
	(0.075)	(0.019)	(0.036)	(0.079)	(0.041)	(0.043)
Treatment effect of LEZ phase III	-0.077	-0.077**	-0.077**	-0.065	-0.065	-0.065
	(0.049)	(0.032)	(0.033)	(0.070)	(0.044)	(0.046)
Observations	91191	91191	91191	75830	75830	75830
R^2	0.410	0.410		0.329	0.329	

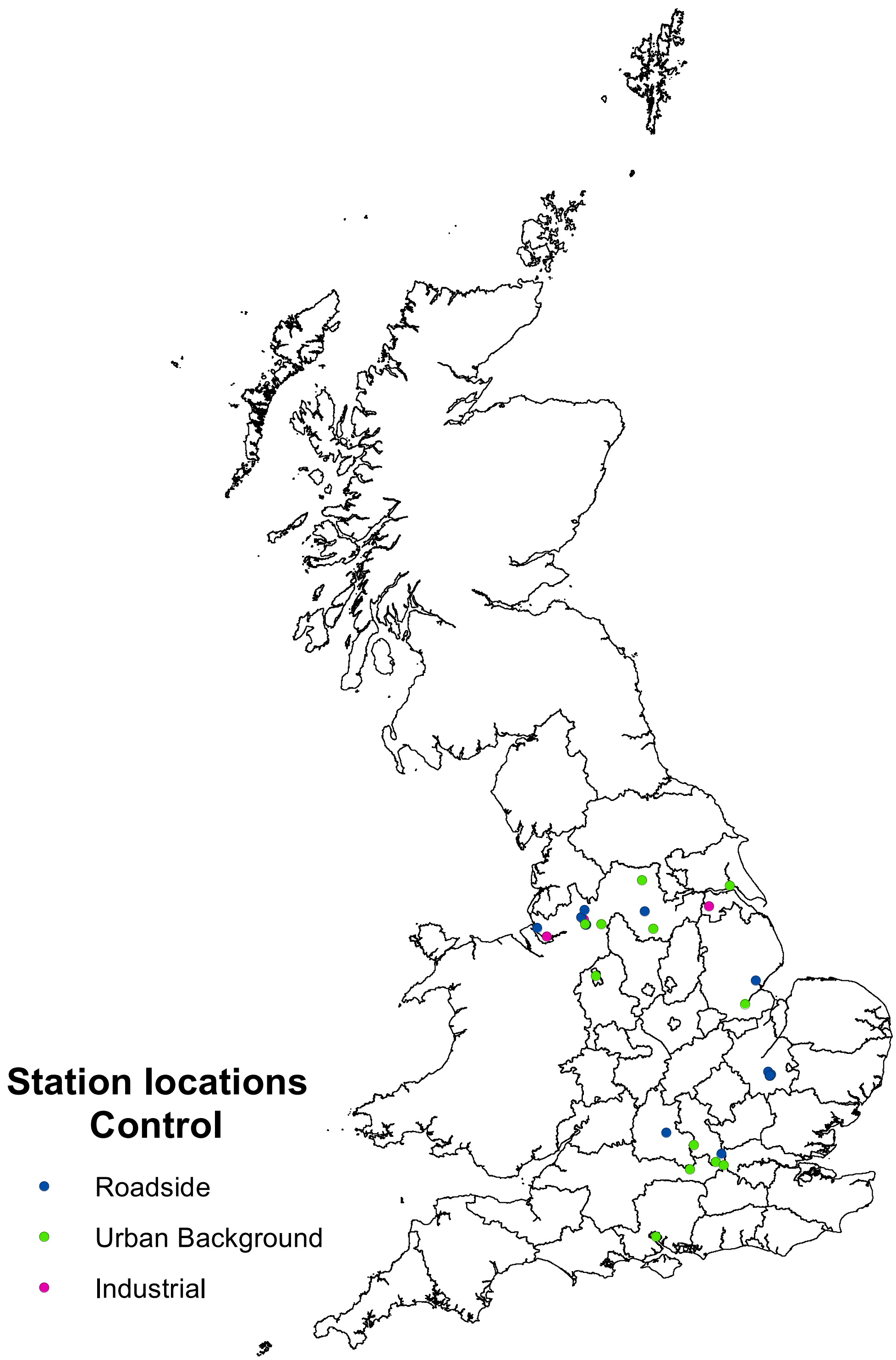
Notes: This table displays the difference-in-differences estimation results using observations below the 95 percentile of station level pollution distribution. All columns control for year FE, quarter FE, current and lagged weather conditions, and dummies for weekends and holidays.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Panel A: Control stations



(Continued on next page)

Panel B: Treated and outside stations

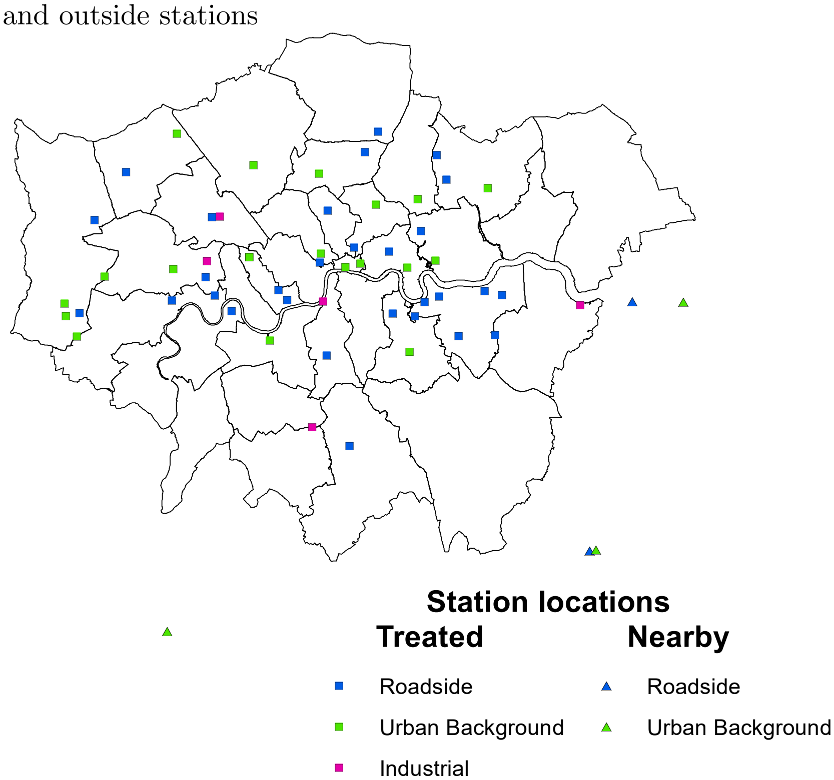


Figure 1.3: Geographical locations of sample stations

Notes: Panel A shows geographical location of 37 monitoring stations located in South East, North West, East of England, and Yorkshire and The Humber. Panel B shows geographical location of 39 treated stations inside Greater London. All the stations inside London locate within the low emission zone as well.

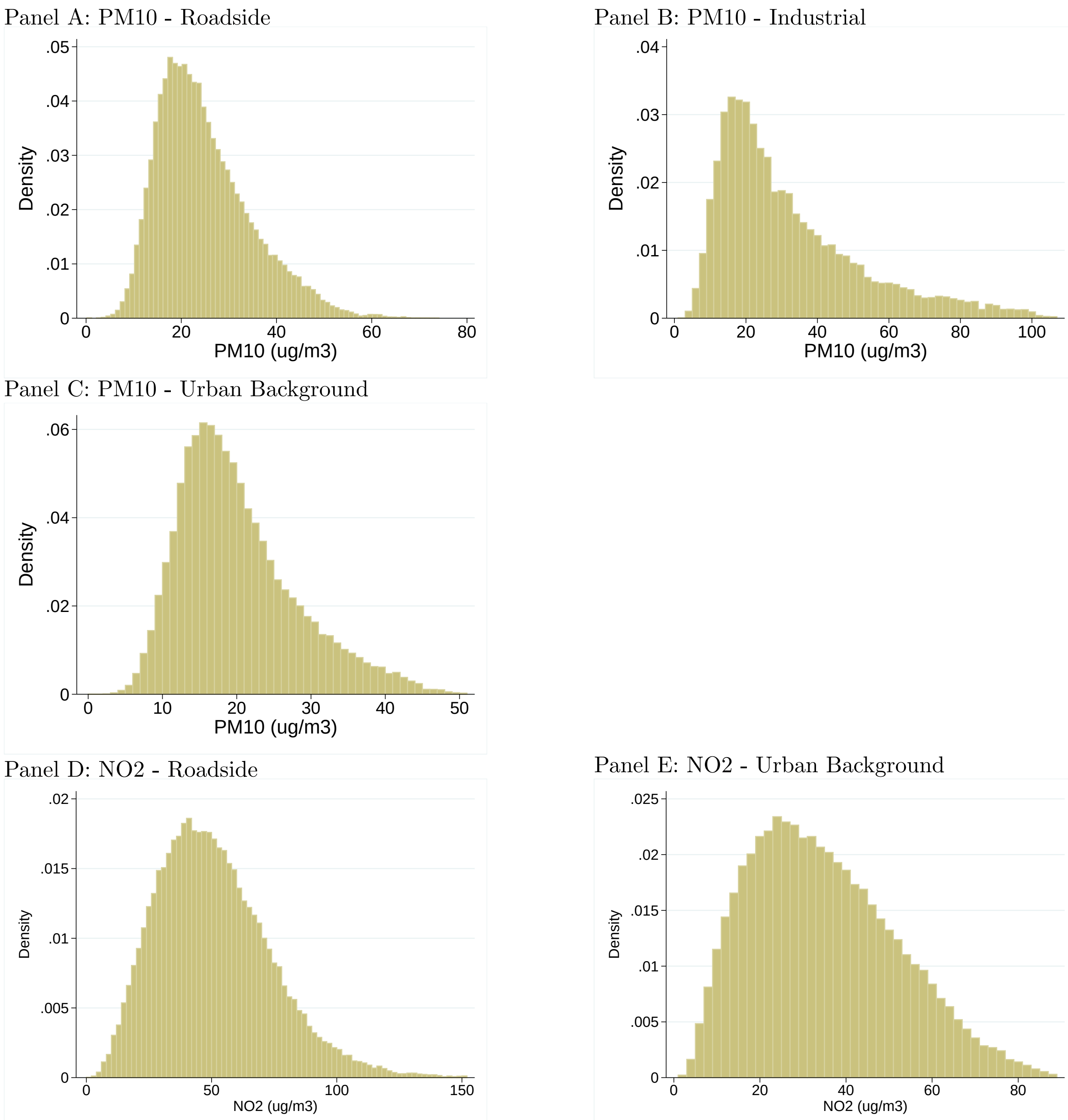
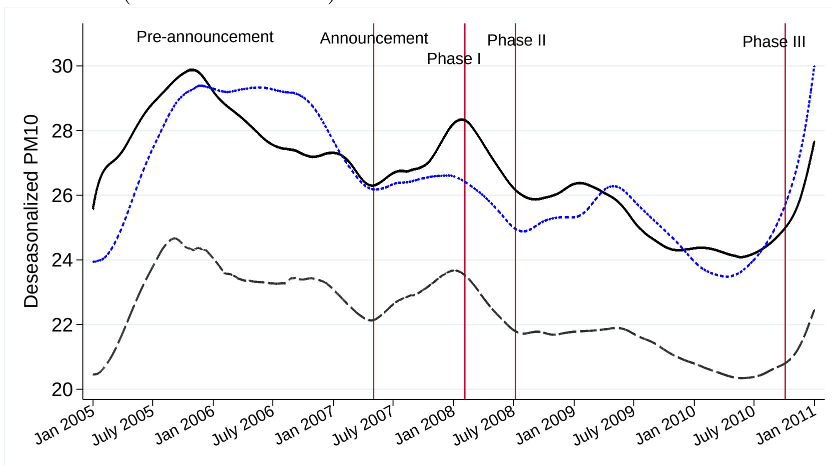


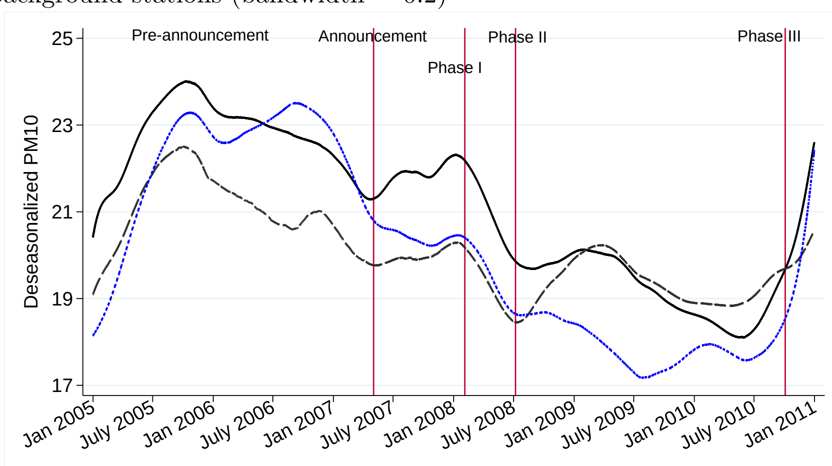
Figure 1.4: Probability density functions of daily average PM10

Notes: This figure shows the daily average PM_{10} and NO_2 below the station specific 95 percentiles from 2005 to 2010. *Data sources:* London Air Quality Network (LAQN) and Air Quality England (AQE).

Panel A: Roadside stations (bandwidth = 0.2)



Panel B: Urban background stations (bandwidth = 0.2)



Panel C: Industrial stations (bandwidth = 0.1)

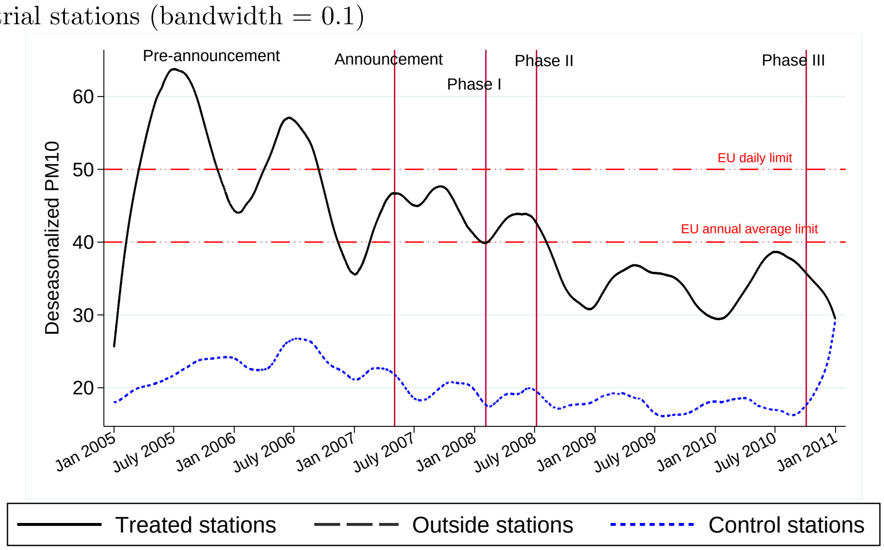
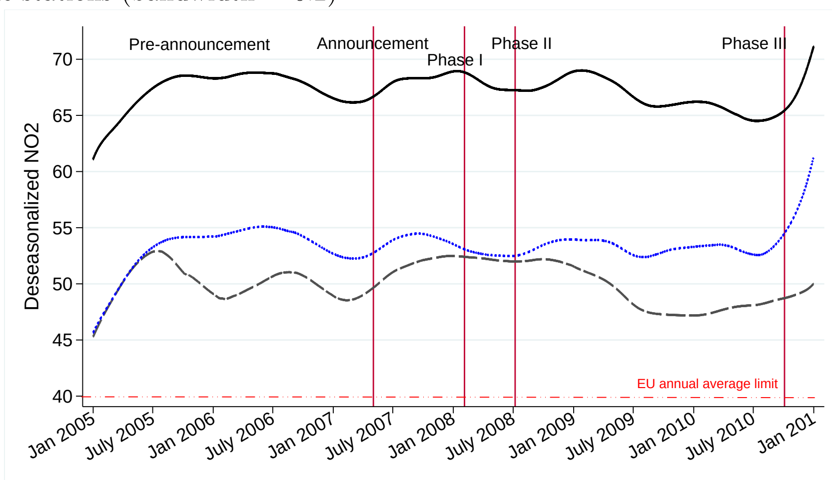


Figure 1.5: Lowess smoothed PM_{10} by type of stations

Notes: This figure shows the locally weighted (lowess) smoothed curves for deseasonalized PM_{10} in $\mu g/m^3$ generated by equation (3). In each panel, we choose a bandwidth such that the short-term fluctuations of particulate matter are visually presented without impairing long-run trend of air quality.

Panel A: Roadside stations (bandwidth = 0.2)



Panel B: Urban background stations (bandwidth = 0.2)

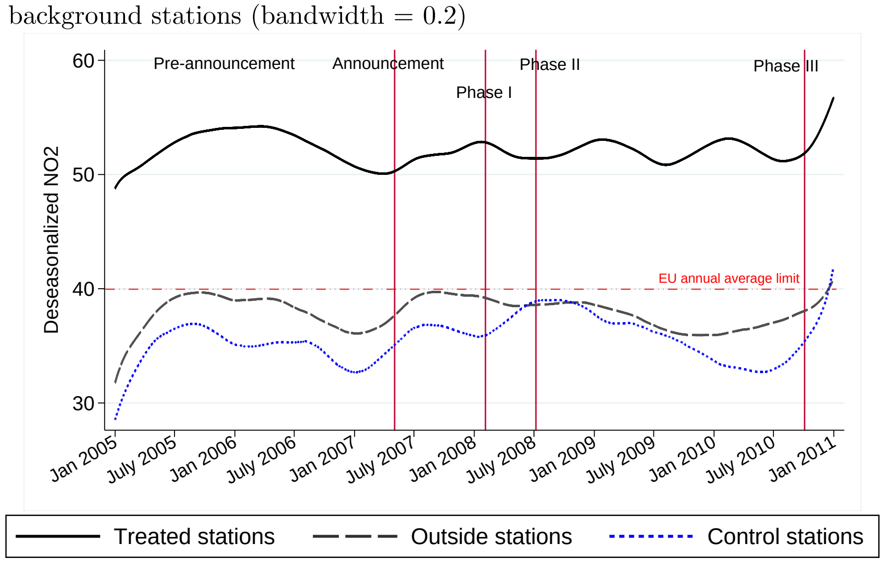


Figure 1.6: Lowess smoothed NO_2 by type of stations

Notes: This figure shows the locally weighted (lowess) smoothed curves for deseasonalized NO_2 in $\mu g/m^3$ generated by equation (3). In each panel, we choose a bandwidth such that the short-term fluctuations of particulate matter are visually presented without impairing long-run trend of air quality.

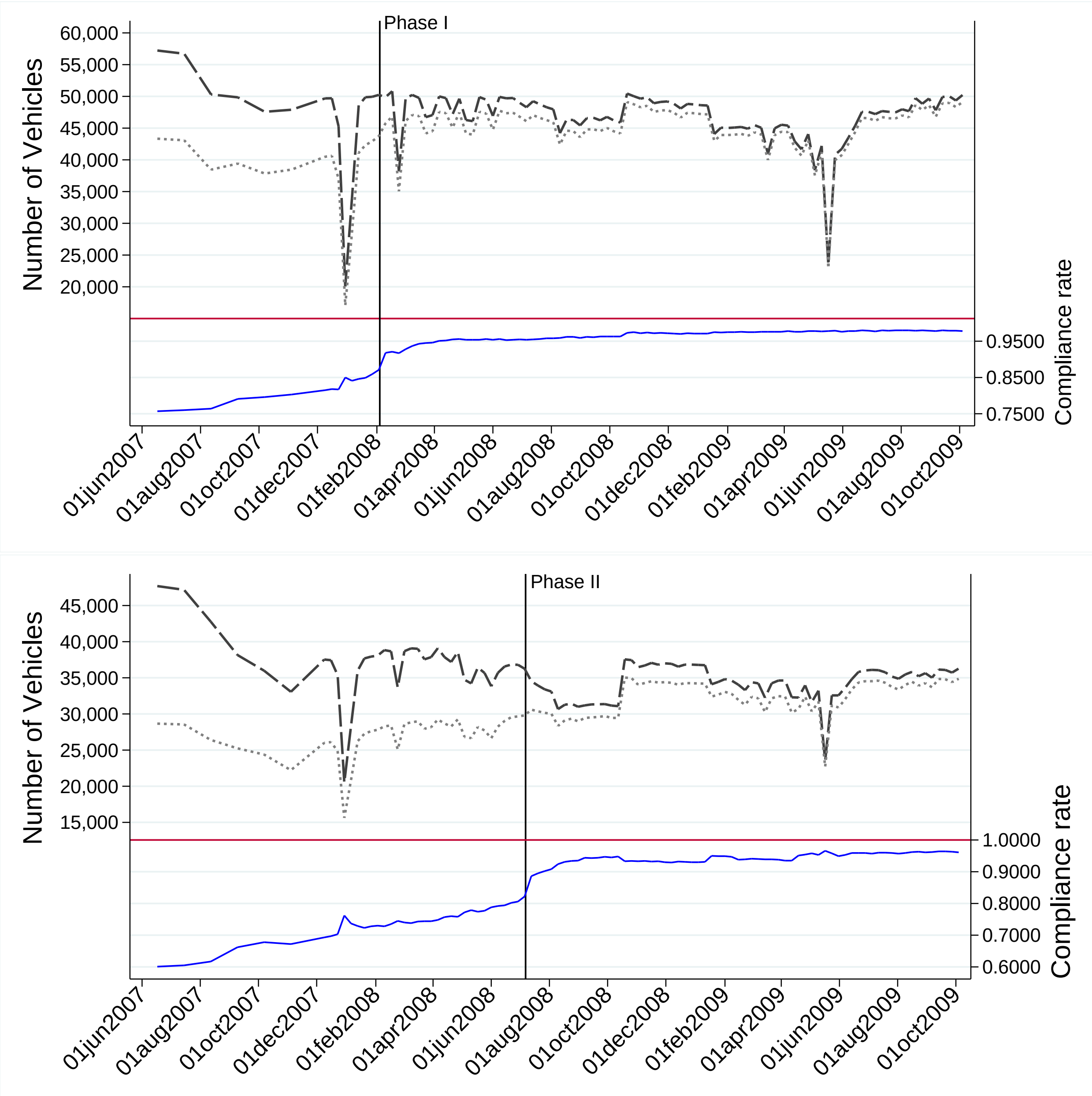


Figure 1.7: LEZ phase I and II non-compliant vehicles and lowess smoothed PM_{10}

Notes: This figure illustrates the time series non-compliant vehicles to phase I and phase II of the LEZ and the corresponding compliance rates.

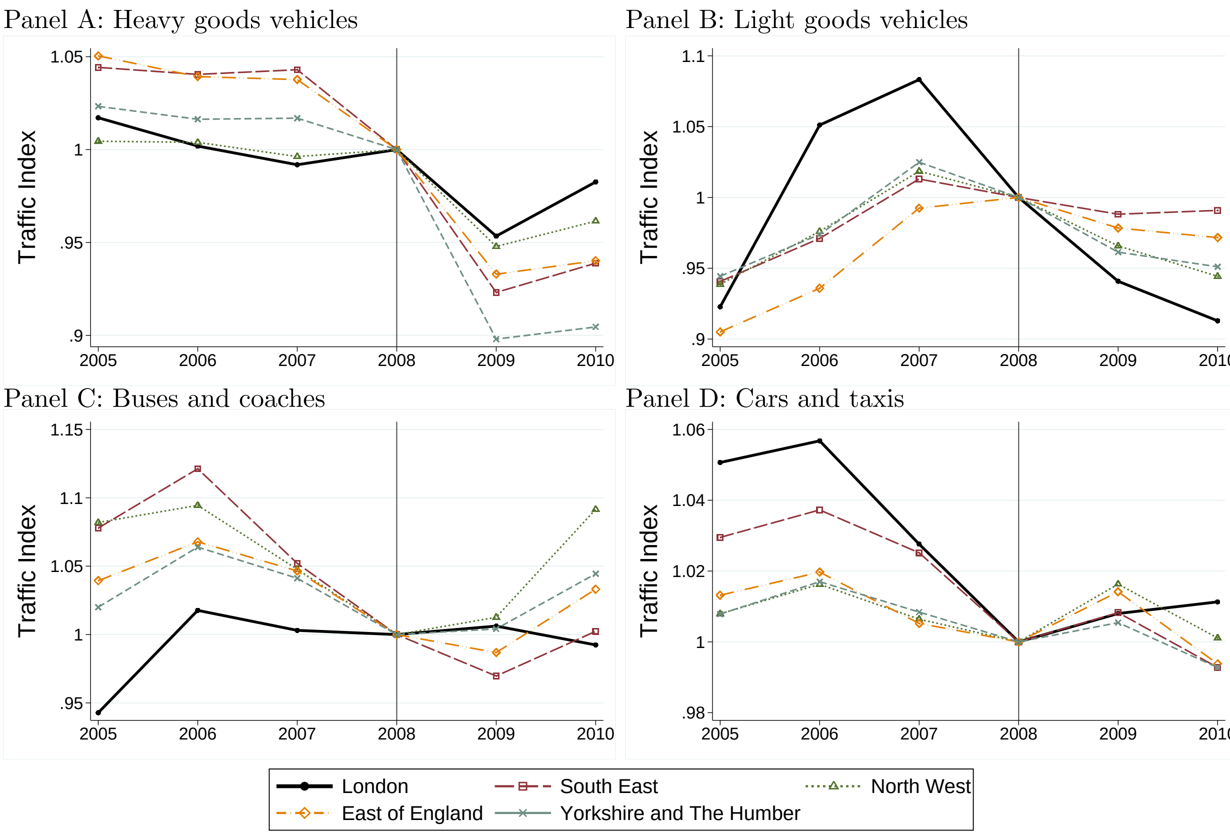


Figure 1.8: Annual traffic flow on major roads

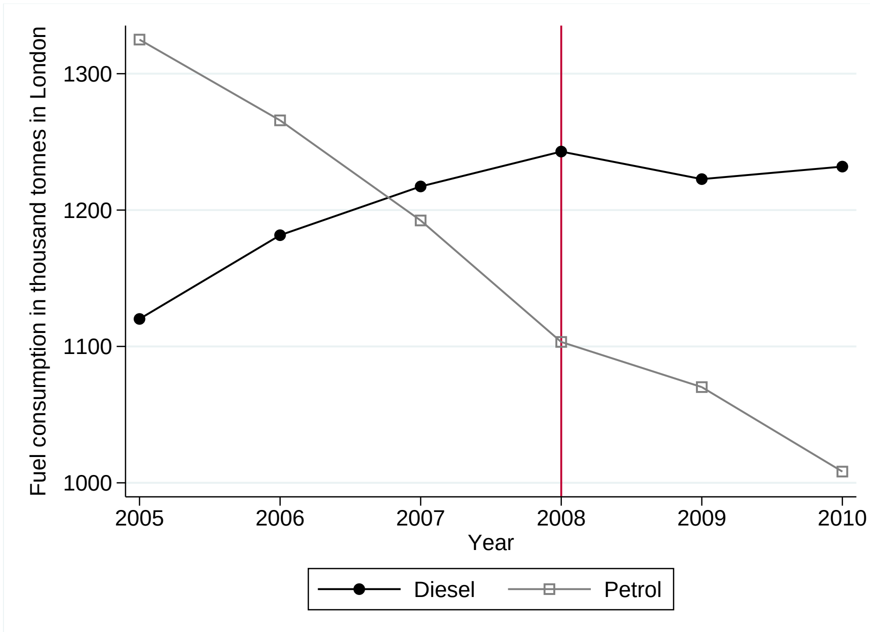


Figure 1.9: Annual aggregate fuel consumption in London

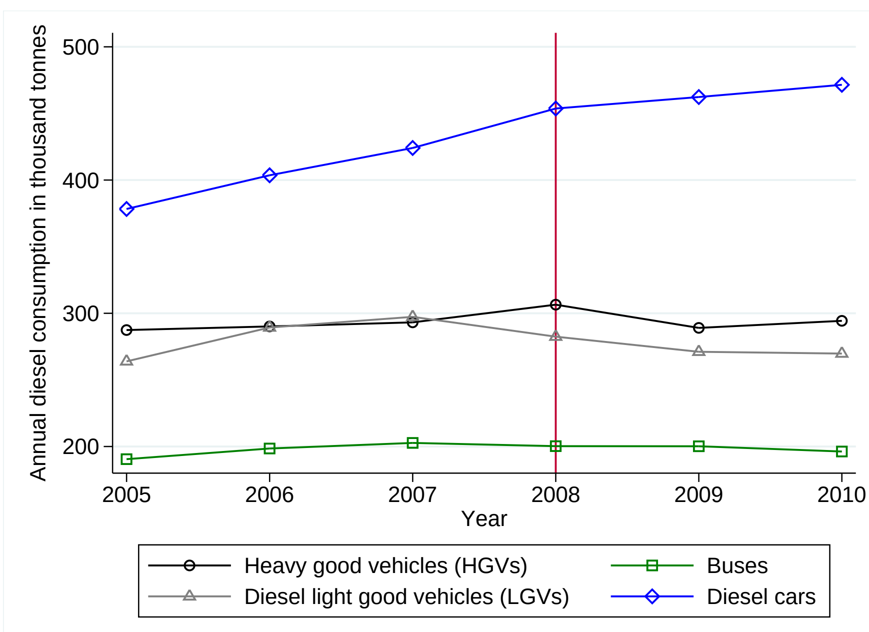


Figure 1.10: Annual diesel fuel consumption in thousand tonnes in London

Table 1.10: DID Estimation Results of LEZ Spillover Effects from LEZ Announcement to Phase III

	Control as U stations nearby LEZ			Control as R or I stations far from LEZ		
	station-year clustered se	type-region clustered se	NW se	station-year clustered se	type-region clustered se	NW se
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: PM10						
Spillover effect of LEZ	0.065	0.065	0.065	0.034	0.034	0.035
announcement	(0.089)	(0.063)	(0.041)	(0.085)	(0.058)	(0.039)
Spillover effect of LEZ	0.009	0.009	0.009	0.026	0.026	0.025
phase I	(0.105)	(0.054)	(0.058)	(0.101)	(0.051)	(0.056)
Spillover effect of LEZ	-0.068	-0.068	-0.068	0.017	0.017	0.018
phase II	(0.059)	(0.055)	(0.049)	(0.066)	(0.035)	(0.048)
Spillover effect of LEZ	-0.019	-0.019	-0.020	-0.049	-0.049	-0.049
phase III	(0.055)	(0.079)	(0.052)	(0.097)	(0.093)	(0.062)
Observations	15790	15790	15790	26150	26150	26150
R^2	0.384	0.384		0.297	0.297	
Panel B: NO2						
Spillover effect of LEZ	0.040	0.040	0.039	-0.018	-0.018	-0.018
announcement	(0.120)	(0.040)	(0.048)	(0.109)	(0.038)	(0.044)
Spillover effect of LEZ	0.031	0.031	0.033	0.081	0.081	0.081
phase I	(0.163)	(0.050)	(0.074)	(0.146)	(0.059)	(0.068)
Spillover effect of LEZ	-0.051	-0.051	-0.052	-0.097	-0.097	-0.097
phase II	(0.126)	(0.060)	(0.065)	(0.121)	(0.067)	(0.061)
Spillover effect of LEZ	-0.139	-0.139**	-0.139**	-0.047	-0.047	-0.047
phase III	(0.090)	(0.059)	(0.057)	(0.095)	(0.057)	(0.060)
Observations	24025	24025	24025	30241	30241	30241
R^2	0.506	0.506		0.304	0.304	

Notes: This table displays the difference-in-differences estimation results on the spillover effect of the LEZ. All columns control for year FE, quarter FE, current and lagged weather conditions, and dummies for weekends and holidays.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 1.11: Effect of LEZ Non-Compliant Vehicles on Air Pollution in Distant Regions of England

	PM10		NO2	
	All stations	Traffic stations only	All stations	Traffic stations only
	(1)	(2)	(3)	(4)
Phase I				
Non-compliant vehicles	0.006*	0.005	0.004	0.004**
	(0.003)	(0.004)	(0.003)	(0.002)
Observations	14673	7901	10852	7954
R^2	0.323	0.303	0.408	0.409
Phase II				
Non-compliant vehicles	0.005**	0.004	0.003	0.003
	(0.002)	(0.003)	(0.002)	(0.002)
Observations	14673	7901	10852	7954
R^2	0.324	0.304	0.407	0.409

Notes: This table shows the effect of LEZ non-compliant vehicles on air monitoring stations located far (i.e. over 10 miles away) from the border of the LEZ. Number of non-compliant vehicles are in thousand units. All columns control for quarter FE, current and lagged weather conditions, day of week, and a holiday dummy. Robust standard errors clustered by station in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 1.12: Effect of LEZ Compliance Rate on Air Pollution in Distant Regions of England

	PM10		NO2	
	All stations	Traffic stations only	All stations	Traffic stations only
	(1)	(2)	(3)	(4)
Phase I				
Compliance Rate	-0.337*	-0.272	-0.228	-0.227**
	(0.163)	(0.213)	(0.139)	(0.083)
Observations	14673	7901	10852	7954
R^2	0.323	0.303	0.408	0.409
Phase II				
Compliance Rate	-0.282***	-0.215	-0.125	-0.128*
	(0.097)	(0.122)	(0.097)	(0.068)
Observations	14673	7901	10852	7954
R^2	0.326	0.305	0.408	0.409

Notes: This table shows the effect of LEZ compliance rate on air pollution in distant locations. All columns control for quarter FE, current and lagged weather conditions, day of week, and a holiday dummy. Robust standard errors clustered by station in parentheses

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Appendix of Chapter 1

Table 1.13: Euro emission standards of PM for heavy-duty diesel engines

Stage	**Date**	**PM (g/kWh)**
Euro I	1992 ($\leq$ 85 kW)	0.612
	1992 ($>$ 85 kW)	0.36
Euro II	Oct 1996	0.25
	Oct 1998	0.15
Euro III	Oct 1999 (EEV[a] only)	0.02
	Oct 2000	0.10[b]
Euro IV	Oct 2005	0.02
Euro V	Oct 2008	0.02
Euro VI	Jan 2013	0.01

Notes: a - EEV refers to enhanced environmentally-friendly vehicles.

b - The standard is 0.13 for engines of less than 0.75 dm^3 swept volume per cylinder and a rated power speed of more than 3000 min^{-1}

Chapter 2

AIR QUALITY AND HUMAN HEALTH

2.1 Introduction

Air pollution and its health effect have been a longtime concern of the globe. Although the mechanisms of pollution caused diseases have not been fully explored, many epidemiological and socioeconomic studies nevertheless show a strong linkage between the exposure to air pollution and serious health issues (Samet et al. (1987); Ostro and Rothschild (1989); Schwartz (1995); Knittel et al. (2011); Janke et al. (2014); Kampa and Castanas (2015)).

Among all age groups, fetuses and infants may probably be the most vulnerable population to environmental hazards. Evidence has shown that the potential harm of air pollution through mother's channel can start as early as the first week of pregnancy and last till the birth delivery. In different gestational ages, air pollution may generate different health risks, including growth restriction, heart defects, and pre-term birth (births of less than 37 weeks) (Southern California Environmental Report Card, 2008). Infants born with these defects may have a higher mortality rate in early childhood and at higher risks of many chronic diseases in their adulthood. Since the birth delivery, direct exposure to air pollution will hurt infants whose immune system is not fully developed. Later in adolescent and adulthood, high levels of air pollution may still affect the respiratory system and cardiovascular system, causing cough, asthma, lung cancer and many other diseases.

This chapter begins with a thorough review of recent literature on air pollution and human health. I start with an array of epidemiological papers in the early 1990s and head toward the most recent published papers in health economics literature. I discuss the econometric issues arising in the earliest papers and show the strategies that later works employ to solve the identification issues. Using similar techniques, I estimate the impact of four "criteria air pollutants" on fetal health using individual-level natality data in the United States from 1991 to 2008. I compare my results with previous studies and discuss the advantages and disadvantages of my estimation. Lastly, I provide supportive evidence that even though my estimated results are robust across several dimensions, it is likely biased by maternal behavioral adjustments.

2.2 *Literature*

2.2.1 Health Effect of Air Pollution

A leading factor of adverse birth outcomes (e.g. fetal death, low birth weight, infant mortality, etc.) is environmental threat.[1] In particular, the relation between air pollution and birth outcome is gaining interests to researchers in the most recent decades. Wang et al. (1997) followed pregnant women in selected residential areas of Beijing from 1988 to 1991. They collect and merge individual birth data with daily sulfur dioxide (SO2) and total suspended particles (TSP) over the pregnancy periods. The authors' econometric models include both linear and logistic regressions, controlling for gestational age, infant sex, birth year, and mother's age. According to Wang et al. (1997), high levels of SO2 and TSP during the third trimester of pregnancy could significantly reduce infant birth weight. Bobak (2000) and Maisonet et al. (2001), respectively, replicate the analysis using birth data in selected districts in Czech during 1990-1991 and Northeastern United States during 1994-1996. In these two analyses, the researchers incorporate a richer set of control variables (e.g., mother's marital status, education level) to further reduce bias from confounding factors. Besides, each of the latter two papers studies a slightly different group of air pollutants.[2] According to Bobak (2000), there exists a strong association between low birth weight and SO2 exposure, with the greater impact being reported from the first trimester of pregnancy. Maisonet el al. (2001) nevertheless estimate that both CO and SO2 adversely affect infant birth weight,

[1] Other causes may include genetic issues, maternal characteristics, misusage of drugs and many others.

[2] Bobak analyzes the impact of SO2, TSP and nitrogen oxides (NOx). Maisonet et al. focus on SO2, carbon monoxide (CO), and particulate matter with diameter less than 10 micrometers (PM10).

especially during the third trimester.

Several issues arise in these epidemiological studies. First, all of the studies include gestational age as an independent variable. Gestational age is defined as the length between the mother's last menstrual date and the birth delivery, usually measured in total number of weeks. The gestational age could, however, largely co-vary with the outcome variable of interest, which in most studies is the infant birth weight. In other words, air pollution in utero affects birth weight by reducing the length of gestational period, therefore the coefficient for gestation age already absorbs a large portion of the health effect of air pollution. Including gestation as a key regressor, one will end up measuring the local variation of birth weight conditional on fetal age. Strictly speaking, this estimated effect should be referred to as the impact of air pollution on fetal growth retardation which is broadly understood as the intrauterine growth restriction (IUGR). The IUGR provides some useful information about the overall health condition of a fetus. However, it usually occurs when an infant or a mother has inherent health issues.[3] This means that the inclusion of gestational age as an independent variable will very likely lead to selection bias in the estimation results.

Second, the studies are each conducted in selected urban areas over a relatively short period of time. The cross-sectional structure of dataset omits issues with long-run time trends but encounters unobservable confounding factors of fetal health across regions. For instance, the differential birth outcome across residential areas in Beijing may be caused by water quality differences. Unfortunately, none of these studies control for area (e.g. city, district or residential areas) fixed effects. Failure to do so may lead to spurious results, or more specifically, what the estimated coefficients capture are random correlations instead of causal relations between air pollution and birth outcome.

The third concern in the estimation is the endogeneity issue. Evidence shows that weather conditions during pregnancy may significantly affect women's fertility and overall health conditions (Pitt and Sigle (1997); Buckles and Hungerman (2013)). For instance, seasonal

[3]For instance, pregnant women with high blood pressure, anemia, chronic diseases, substance abuse are at high risk of IUGR. Fetal factors such as infection and chromosome abnormality may also lead to IUGR.

diseases such as influenza can usually last for several weeks, increasing the risk of hospitalization for pregnant women. Since weather conditions and air pollution can be strongly correlated, omitting the former will lead to substantial bias in estimates.

The above concerns are largely attenuated in Chay and Greenstone's (2003a, 2003b) works which apply a first-differencing method with instrumental variables (IV) to assess the impact of total suspended particles (TSP) on infant mortality. Both papers employ a quasi-experimental research design by considering dramatic improvements of air quality induced by unexpected events. The two papers start with a similar cross-sectional specification

$$y_{ct} = T_{ct}\beta + {x_{ct}}'\theta + \varepsilon_{ct}, \qquad \varepsilon_{ct} = \alpha_c + u_{ct} \tag{2.1}$$

where y_{ct} is the infant mortality rate in county c in year t. T_{ct} is the average concentration of TSPs measured by all monitors in county c in year t. x_{ct} is a vector of county-level observed variables that may affect infant mortality. ε_{ct} is the error term that contains a county-specific component α_c and a random shock u_{ct}. The un-observability of α_c may cause omitted variable bias (OVB) since it is likely that $E(\alpha_c T_{ct}) \neq 0$.

When the counties are observed for at least two periods, the OVB can be removed using a first-differencing method as follows.

$$\Delta y_{ct} = \Delta T_{ct}\beta + \Delta {x_{ct}}'\theta + \Delta\varepsilon_{ct} \tag{2.2}$$

where $w_{ct} = w_{ct} - w_{ct-1}, (w_{ct} = y_{ct}, T_{ct}, x_{ct}$ or $\varepsilon_{ct})$. Equation (2.2) differences out the unobservable component α_c that does not vary over time. Since ΔT_{ct} and $\Delta\varepsilon_{ct}$ are less likely to be correlated, the model can be identified using a standard ordinary least squares (OLS) method.

Sensible estimation of β requires a substantial variation of T_{ct}. Chay and Greenstone (2003a) use the 1970 Clean Air Act Amendments (CAAA) as a regulatory intervention that significantly reduces TSPs in selected counties. Specifically, they estimate the following equation as a first step to equation (2.2).

$$\Delta T_{ct} = \Delta x_{ct}\delta + Z_{ct}\gamma + \Delta v_{ct}. \tag{2.3}$$

Here Z_{ct} equals 1 if county c has seen a TSPs level above the regulatory threshold before the implementation of CAAA. Since the legislation is enforced based on the non-attainment status of each county, the initial regulatory status of the county may serve as an instrumental variable that affects infant mortality only by sharply reducing the TSPs.

Similarly, Chay and Greenstone (2003b) use the 1981-1982 recession as an exogenous shock that dramatically drives down the income per capita and TSPs in some counties. Although the occurrence of recession does not depend upon the pre-existing level of TSPs, they instead find a strong correlation between changes of TSPs and lagged levels of TSPs and income per capita. This relationship, though weakens the exogeneity of treatment assignment, gives the possibility for lagged TSPs and income per capita to serve as instrument variables, conditional on that they are uncorrelated with changes in infant mortality shocks.

The estimation results of the two papers show that a 1 percent reduction of TSP is associated with 0.35 to 0.5 percent decline in infant mortality rate. This point estimate is both significant and robust along various dimensions. Converting to the number of life saved, the 1970 CAAA regulation has reduced roughly 1300 infant deaths, while the 1981-1982 recession has reduced roughly 2500 infant deaths.

A big advantage of Chay and Greenstone's (2003a, 2003b) papers is that they both well handle the omitted variable bias that commonly arises in cross-sectional studies. In addition, they include a wide variety of fixed effects (FE) (e.g. county FE, year FE, or state-year FE) to help further control for the unobservable factors contributing to infant mortality. One limitation, however, is that both papers ignore the harvesting effect prior to birth.[4] For example, infants with higher risk of mortality may have been exposed to severe air pollution in utero, and the effect of postnatal exposure to air pollution is to shorten the life expectancy. Meanwhile, the studies are both conducted at county-year level so substantial variations and heterogeneity of individual characteristics are omitted.

[4]Instead, the papers consider the harvesting effect after birth. Specifically, they exclude the possibility that previous (postnatal) exposures to air pollution lead to harvesting in current period, but they ignore the effect of maternal exposure to air pollution.

The data limitation gap was shortly filled by an array of studies that obtain individual birth data on a regional scale. Currie and Neidell (2005) examine the impact of three “criteria” air pollutants carbon monoxide (CO), particulate matter (PM10) and ozone (O3) on infant health in California over the 1990s. In this study, a spatial distribution of weekly air pollution concentrations over zip codes is generated based on the relative distance between each zip code and the nearest air quality monitor. Comparing with previous literature that generally uses annual average pollution across all stations within a county, the spatial assignment of weekly pollution largely improves the precision of pollution measurement. In regards to the research design, Currie and Neidell (2005) test the following two hypotheses. First, the postnatal exposure to air pollution has an impact on the risk of infant mortality. Second, the maternal exposure to air pollution has an impact on the probabilities of fetal death and low birth weight. Their econometric models are as follows.

$$y_{izt} = \alpha(t) + p_{zt}\beta + \boldsymbol{X_{izt}\theta} + \varphi_{zt} + Y_t \tag{2.4}$$

$$y_{izt} = p_{zt}\beta + \boldsymbol{X_{izt}\theta} + \varphi_{zt} + Y_t \tag{2.5}$$

Equation 2.4 estimates the health effect of postnatal exposure to air pollution. The dependent variable y_{izt} is the probability of death which equals 1 if the infant i living in zip-code c died in week t and 0 otherwise. $\alpha(t)$ is the duration spline measured in weeks since the baby was born.[5] p_{zt} is the postnatal pollution exposure. $\boldsymbol{X_{izt}}$ is a matrix of control variables including the infant's health condition at birth, maternal characteristics, and weather conditions after birth. Similarly, in equation (2.5) the dependent variable y_{izt} is either the probability of fetal death or the incidence of low birth weight. p_{zt} is the pollution exposure during three trimesters of pregnancy, and $\boldsymbol{X_{izt}}$ controls for maternal characteristics and weather conditions before birth. Both regressions account for zip-month FE φ_{zt} and year FE Y_t.

The results of Currie and Neidell's (2005) paper show a strong and significant impact

[5]To account for the fact that older infant has a lower risk of death, the spline has breaks after the 1st, 2nd, 4th, 8th, 12th, 20th and 32nd weeks.

of postnatal exposure to CO on infant mortality. For every 1 ppm decrease in CO, it is expected that 18 fewer infant deaths will incur among every 100,000 births. On the other hand, the exposure to other pollutants have a much smaller impact on infant mortality rate. Furthermore, the estimated coefficients in equation (2.5) for all three air pollutants in all trimesters are either insignificant or significant but with wrong sign, suggesting that the health effect of prenatal exposure to air pollutants may not be acute.[6]

To further explore the issue, Currie et al. (2009) use individual level birth data in New Jersey in the 1990s to reexamine the health impact of prenatal exposure to air pollution. This research further improves the precision of air pollution measurement by assigning mothers' exact residential addresses to the nearest air quality monitor. Meanwhile, the paper controls for unobservable maternal characteristics by including mother's fixed effect in the regression.[7] Their regression model has the following structure.

$$y_{ijmt} = \sum_{s=1}^{s=3} (p_{mt}^{s}\beta^{s} + w_{mt}^{s}\gamma^{s}) + \boldsymbol{X_{ijmt}\theta} + Y_t + \varphi_{mt} \times Q_t + \delta_j + \varepsilon_{ijmt} \tag{2.6}$$

The dependent variable y_{ijmt} indicates the birth outcome (i.e., birth weight in grams, indicator of low birth weight, gestational age, and infant death) for infant i of mother j near monitor m in time t. p_{mt}^{s} and w_{mt}^{s} are the average pollution and weather conditions in trimester $s(s = 1, 2, 3)$, respectively. $\boldsymbol{X_{ijmt}}$ is a vector of co-variates including mother's characteristics (i.e. age, educational attainment, race, marital status, smoker vs. non-smoker) and infant characteristics (i.e. multiple birth, birth order, and sex). Additionally, the model contains a rich set of fixed effects (i.e. year FE Y_t, monitor-quarter FE $\varphi_{mt} \times Q_t$, and mother FE δ_j) to eliminate the endogeneity issue.

According Currie et al. (2009), there exists a strong and adverse impact of CO exposure on newborns, especially during the third trimester. This result extends Currie and Neidell's

[6]The results for birth weight and fetal death are not explicitly reported in the paper, but the authors point out and discuss the insignificance of the coefficients.

[7]Omitting maternal fixed effects will cause bias to the estimates only if the omitted attributes of mother is correlated with air pollution exposure during pregnancy. For instance, if a mother's pollution-protective behaviors changes substantially during heavy polluting days, the main coefficient of interest will likely be biased.

(2005) findings in California by documenting that maternal exposure to CO consistently weakens fetal and infant health conditions. In addition, they explore the heterogenous impact of air pollution and find infants with smoking or older mothers are more negatively affected by air pollution.

A very closely related group of research focuses on the shocks to air pollution sources. According to the estimates of National Emissions Inventory (NEI 2011), approximately 52 percent CO emissions come from automobiles. Therefore, it is expected that an effective regulation on traffic emissions will lead to a substantial improvement of infant health. Motivated by this logic, Currie and Walker (2011) examine an electronic toll collection system called E-Zpass on state toll-ways in New Jersey and Pennsylvania. The E-Zpass allows equipped vehicles to drive through the toll lane without making a stop and thus attenuate both congestion and vehicular emissions. Because E-Zpass was implemented in different toll plazas at different time, the authors apply the following difference-in-differences strategy to evaluate the average health benefit brought by the installation of E-Zpass

$$y_{it} = \alpha + \eta T_{it} + \varphi C_{it} + \beta T_{it} \times C_{it} + \boldsymbol{\theta X_{it}} + \boldsymbol{\lambda Z_{it}} + Y(t) + \varepsilon_{it} \tag{2.7}$$

with i indicating mother and t indicating time period. The dependent variable is the probability of premature birth or low birth weight, both are key indicators for infant health at birth. Dummy variables T_{it} equals 1 if mother i lives with the closest plaza implementing the E-Zpass, regardless of the distance between the mother's residential address and the plaza. Dummy variable C_{it} equals 1 if mother i lives within 2 km of a toll plaza either implementing or not implementing the E-Zpass. The difference-in-differences estimator β picks up any post-treatment health benefit experienced only by treated mothers (i.e. mothers who have a toll plaza located within 2 km and implemented E-Zpass) as compared with untreated mothers (i.e. mothers with the closed plaza implemented E-Zpass but located more than 2 km away, and mothers who live within 2 km of a plaza without implementing E-Zpass). Additional controls in $\boldsymbol{X_{it}}$ include mother and infant characteristics. Meanwhile, $\boldsymbol{Z_{it}}$ includes indicators for the closest toll plaza and a continuous variable measuring the linear distance of a busy road. Year FE and month FE are captured by $Y(t)$ to absorb any unobservable long-run trends and seasonality. The robust estimates indicate that traffic congestion is a key contributing factor of prematurity and low birth weight.

Besides road traffic, another leading contributor to air pollution is industrial emissions. Lavaine and Neidell (2014) study the effect of energy production on newborns. The paper uses France strike as a natural experiment that induces a sharp decline of sulfur dioxide (SO2) emissions from oil refineries. Similarly with Currie and Walker (2011), Lavaine and Neidell (2014) apply a difference-in-difference estimation with areas nearby refineries being the treatment group. Specifically,

$$y_{ct} = T_t + C_c + \beta T_t \times C_c + \boldsymbol{\theta X_{ct}} + \varepsilon_{ct} \tag{2.8}$$

where y_{ct} is the infant health measured either as birth weight or gestational age in census tract c at time t. T_t is an indicator for strike period, and C_c is an indicator for the census tract containing an oil refinery. The difference-in-differences coefficient β captures variations of birth outcome occurring only in areas close to a refinery during the strike period. Control variables in $\boldsymbol{X_{ct}}$ include weather conditions and unemployment rate.

Lavaine and Neidell's (2014) paper builds a close connection between infant health and SO2 emissions from oil refineries, especially during the third trimester of pregnancy. One limitation, however, is that the data only provide birth outcome by month. Therefore, all births in the same month have to be reassigned to the first day of the month with the pollution and meteorology data aggregated accordingly. Bias will vary depending on the degree of mismatching between birth outcomes and the fetuses' exposed air pollution during the last several weeks before birth.

2.2.2 Air Quality and Human Capital

While many studies explore the health impact of air pollution exposure, a few papers focus instead on how the adverse health outcome is related to the loss of human capital. Currie et al. (2009) examine the impact of three "criteria pollutants", namely CO, O3 and PM10, on school attendance using data from over 1500 elementary and middle schools located in Texas from 1996 to 2001. To identify the causal relationship, they employ a triple-difference estimation taking into account the systematic differences across schools, years and periods of attendance. Contrary to pervious studies, this paper utilizes a non-linear pollution measure consisting of a group of indicators for the percentage of days with pollution concentrations lying in-between the thresholds of 25%, 50%, 75% and 100% of the air quality standards suggested by the US Environmental Protection Agency (EPA). This method allows researchers to analyze both the severity of air pollution and the length of period that an air pollution level has lasted. Meanwhile, it fully respects the increasing feature of marginal damage of air pollution by assigning different coefficients to pollution concentrations within different ranges. Consistent with Currie and Neidell's (2005) results, Currie et al. (2009) find that higher levels of CO is significantly associated with higher school absence rates.

A closely related study is conduced by Lavy et al. (2014) on students' cognitive performance after short-term exposures to CO and PM2.5. The study uses data on the standardized test scores for college entrance in Israel. The identification of air pollution's impact replies on that the test was held on multiple days at the same location for every student, therefore including both school and student fixed effects will control for the unobservable, time-invariant school and student specific factors contributing to cognitive performance. The research finds that students achieve significantly lower scores on standard tests during days with either high CO or high PM2.5 concentrations. This evidence indicates that an effective air quality strategy will not only improve human health, but also increase human capital and boost productivity.

2.2.3 Fetal Origins Hypothesis and Long-Run Effects of Air Pollution

Scholars have long suspected that a human's health condition in early childhood and later adulthood may originate from the period of gestation. Barker (1986) proposed the "fetal origins" hypothesis that relates many chronic diseases (e.g. heart disease and type II diabetes) with prenatal nutrition. The hypothesis triggers many epidemiologic studies later on focusing on long-run impacts of pool health in infancy. Bharadwaj et al. (2015) link birth weight data from all twins in Sweden between 1926 and 1958 with later-life educational achievement, earning, insurance status and adult mortality. For each twin i in twin pair j, Bharadwaj et al. (2015) express the outcome variable y_{ijt} which indicates the persons' lifetime performance (e.g. logarithm of income) in period t as

$$y_{ijt} = \beta_t BW_{ij} + \boldsymbol{\theta_t X_{ij}} + \lambda_j + \varepsilon_{ijt}. \tag{2.9}$$

The coefficient of interest β_t captures the differential impact of birth weight BW_{ij} on the outcome variable during different stages of life. Control variable $\boldsymbol{X_{ij}}$ includes individual specific characteristics such as educational attainment. The twin fixed effects λ_j accounts for any unobservable twin specific characteristics such as family background.

Bharadwaj et al.'s (2015) paper documents a remarkably strong correlation between birth weight and lifetime performance in adulthood. For example, heavier infants are found to more likely be employed, earn a higher wage and be at lower risk of adult mortality up to their 50s. This finding provides a direct evidence for the fetal origins theory and indicates that inequality may arise from periods prior to birth.

Based on the above discussions, one may expect that there exists a long-run relationship between maternal exposure to air pollution and lifetime performance. Although little has been done to demonstrate this relation so far, future research is needed to provide further support or critique to this hypothesis. One of the leading challenges, however, is that as the time goes a person's behavior and health condition will be affected by many other environmental and social factors, and therefore more and more confounders should be adjusted to provide statistically unbiased results. To help identify the impact of air pollution exposure in utero, researchers will need to search for a strong and exogenous event that dramatically changes air quality in some areas over a short period of time. Taking into account human's mobility in later life stages, the impact of fetal exposure to air pollution can be identified by comparing people living in the same area but experiencing different air qualities in utero (i.e. either people who were born in the same places but during slightly different periods or people who were born in different places) after controlling for a rich set of covariates.

The reviewed literature sheds light on the exposure-response relationship between air pollution and human health. They fill the gap left by experimental studies and contribute to the interdisciplinary research on epidemiology, environment economics, public policy and health studies. With well-built econometric models, they largely reduce omitted variable bias and attenuate issues related to measurement errors. Nevertheless, most existing literature estimates the health effect of air pollution at regional scale, while studies using national level data always encounter problems in the aggregation process. Up to now, there has been no study providing a national scale estimation on the health impact of prenatal exposure to air pollution in the United States. In this chapter, I fill the literature gap by combining over 20 million birth data in the US from 1991 to 2008 with county-week level air pollution and demographic features. For the four "criteria air pollutants" that I analyze, carbon monoxide (CO) and nitrogen dioxide are found to adversely affect both gestation and birth weight. The magnitude of the point estimates for CO, however, is much smaller than what has been estimated in Currie et al.'s (2009) paper.[8] Besides, ozone (O3) is found to negatively affect gestation but not birth weight. I find little effect of sulfur dioxide (SO2) on fetal health.

2.3 Model and Identification Strategy

2.3.1 Baseline

The baseline econometrics model has a similar form with previous literature. For fetus i whose mother j resides in county c, the birth outcome y_{ijct} (e.g. gestation length, birth weight) at time t can be expressed as

$$y_{ijct} = \sum_{\tau=1}^{\tau=3} (\beta_\tau P_{ijct}^\tau + \boldsymbol{\delta_\tau W_{ijct}^\tau}) + \boldsymbol{\theta X_{ijct}} + \mu_{ijct}. \tag{2.10}$$

[8]The smaller coefficients could be caused by averaging the point estimates across the entire nation, or the imprecise assignment of air pollution exposure, or other measurement issues. Chapter 3 discusses one of the possible reasons and provides a modification to the measurement strategy to overcome the issue.

P_{ijct}^{τ} represents the average pollution that fetus i experiences in trimester τ, which depends upon both maternal residential county c and the newborn's gestational age. $\boldsymbol{W_{ct}^{\tau}}$ is a vector of weather conditions (i.e., daily average precipitation, maximum and minimum temperatures) in county c during trimester τ. $\boldsymbol{X_{ijct}}$ includes fetal characteristics such as gender and birth order and maternal characteristics such as age, race, Hispanic origin, education level and marital status. μ_{ijct} is an idiosyncratic term that may include any measurement errors of birth outcome as well as random shocks to fetal health.

2.3.2 Omitted Variable Bias and Fixed Effects

Estimation based on equation (2.10) likely suffers from omitted variable bias (OVB). Specifically, μ_{ijct} may contain unobservable components co-varying with both air quality and fetal health. For instance, an outbreak of influenza may strongly affect maternal health conditions during pregnancy, whereas the period of influenza is usually highly correlated with seasonal air quality. Even after controlling weather conditions in all trimesters, substantial bias in the the estimated coefficients would still arise because of co-variance between the still-omitted variables in the error term and the key regressors. To eliminate this concern, I first decompose μ_{ijct} into a time specific component λ_t, a county-time specific component η_{ct}, a mother-time specific component ϕ_{jct}, and a pure random error ε_{ijct}:

$$\mu_{ijct} = \lambda_t + \eta_{ct} + \phi_{jct} + \varepsilon_{ijct}. \tag{2.11}$$

For instance, λ_t may include an overall improvement of standard of living and η_{ct} may include the health effect of business downturns that is likely to be differential across regions. ϕ_{jct} could be any unobservable family backgrounds (e.g. genetic attributes, mother's protective behaviors against environmental risks) that may strongly affect fetal health.

Ideally, coefficient β_τ in equation (2.10) can be fully identified by using corresponding fixed effects to capture λ_t, η_{ct} and ϕ_{jct}.[9] This method is valid if multiple incidents are

[9]The identification relies on the assumption that λ_t, η_{ct} and ϕ_{jct} do not change within groups considered to pertain that fixed effect. For instance, mothers giving multiple births over a time period may adjust

observed in each time, county-time and mother-time group. When the group identifier is not easily detected, researchers have to rely on other control variables to approximate the fixed effects.

The data used in this study allow me to include most, but not all of the fixed effects. Specifically, the data neither link each newborn with his or her mother, nor provide any information on whether the birth is singleton or twin. Because mother's identifier is unobservable, I cannot control for maternal fixed effects ϕ_{jct}. Fortunately, for each birth I observe how many times the mother has visited the health care provider during the pregnancy period. Including this variable could at least partially compensate the health effect of unobservable maternal protective behaviors. Additionally, I incorporate county level medium income to help absorbing the sorting bias.

The revised model takes the following form. Let s index the state of county c, the birth outcome can be expressed as

$$y_{ijcst} = \sum_{\tau=1}^{\tau=3} (\beta_\tau P_{cst}^\tau + \boldsymbol{\delta_\tau W_{cst}^\tau}) + \boldsymbol{\theta X_{ijcst}} + \lambda_t + \eta_{ct} + \gamma_{st} + \phi_{ijt} + \varepsilon_{ijcst}. \tag{2.12}$$

In specification (2.12), λ_t is the year fixed effects capturing any trends of fetal health caused by an overall improvement of standard of living and long-run business cycles. η_{ct} is the county-year medium income representing the newborn's family background and mother's nutrition intake during pregnancy.[10] γ_{st} is the state-quarter fixed effects incorporating any seasonal fluctuations of birth outcome specific to each state. ϕ_{ijt} stands for the total number of mother's prenatal visits to hospital. The other control variables are the same as in 2.10. In the preferred specification, I cluster standard errors ε_{ijcst} by county to allow random health shocks occurring within each county.

The identification of coefficient β_τ relies on substantial variations of air pollution concentrations P_{cst}^τ. In a cross-sectional setting with only aggregated data available, a first-

behaviors from one birth to another. In this situation, controlling for maternal fixed effects would still lead to a bias in estimates.

[10]Since income starts to affect fetal health from the date of pregnancy, the year used in η_{ct} is the year of pregnancy instead of year of birth.

differencing method (as adopted by Chay and Greenstone (2003a, 2003b)) requires dramatic air quality changes being experienced by some “randomly selected” counties but not the others.[11] When higher frequency (e.g., daily, weekly) air pollution data are available in a longer time period, simply using the disaggregated data will generate substantial variations in P^{τ}_{cst} across counties and over time. With the omitted variable bias correctly handled using various fixed effects, the variations of birth outcome can be attributed to the changes of air quality over a short period of time for fetuses with the same control variables. Here the corresponding health effects captured by β_{τ} can be interpreted as any fetal health outcomes induced by local and short-run variations of air pollution.

2.4 Data

2.4.1 Birth Data

The data used in this study come from differen sources. The birth data are obtained from the U.S. Centers for Disease Control and Prevention (CDC). The original data contain 50,435,066 newborns between Jan 1, 1991 and Dec 31, 2008. This eighteen-year time period is the longest period I can consider because air quality data prior to 1990 are not available. After 2008, mothers no longer report the date of last menstrual period (LMP) so I cannot infer the infant birthday from the available information. Each observation contains the birth month, day of week and gestational age (in whole weeks).[12] I combine this information with the mother’s self-reported LMP to derive birth date. The data additionally provide information on fetal characteristics such as gender, total birth order and live birth order, and maternal characteristics such as age, race, Hispanic origin, educational attainment and marital status. For a majority of births, I also observe the number of prenatal visits to hospital during pregnancy. Dropping observations with missing data on any of these variables leaves me

[11] Of course, it is very hard to have a purely natural experiment, but with properly defined instrumental variable Chay and Greenstone obtain a quasi-natural experiment setting.

[12] Gestational age is considered as the number of weeks from mother’s last menstrual period to the birth delivery.

with 31,471,035 births (about 62.4% of the original dataset).

Figure 2.1 Panel A plots the time series monthly average birth weight in the US from 1991 to 2008. Surprisingly, the average birth weight declines from about 3340 grams in early 1990s to less than 3280 grams in late 2000s.[13] The downward trend persists when only considering the first order births. Figure 2.1 Panel B plots the time series monthly incidence of low birth weight infants (i.e., infants born less than 2500 grams). Consistent with Panel A, I find an upward-sloping trend in low birth weight infants starting from less than 7 percent in 1991 to nearly 8 percent in late 2000s. This long-run change of fetal health may largely be driven by the overall delay of the first pregnancy for US women. According to the data, less than 1.5 percent mothers choose to give birth when they are older than 40 in 1991 while this percentage nearly doubles in year 2005.[14] Starting from 2006 and on, the average birth weight increases gradually from the trough and the incidence of low birth weight infants declines slightly.

Besides the long-run trend, Figure 2.1 also shows a clear seasonality of birth weight in each calendar year. Infants born in spring months are roughly 20 grams heavier than those born in winter months. Currie and Schwandt (2013) examine the influenza prevalence during the critical period of pregnancy as one source of seasonality of birth weight. Buckles and Hungerman (2013) further supplement the evidence that women giving birth in different periods of year generally have different socioeconomic backgrounds such as age, race, educational attainment and marital status. They explain this differential as partially caused by the change in fertility pattern induced by climate conditions across mothers with different characteristics. These studies illustrate the importance of controlling seasonality in identifying the impact of air pollution on fetal health.

---

[13] A similar trend has been found in the average gestational age. In 1991, a typical length from conception to birth lasted about 39.1 weeks while during 2005 and 2007 the length shortened to less than 38.6 weeks. See Appendix for the time series figure of average gestation.

[14] Mothers older than 40 years deliver lighter infants than younger mothers. According to the full sample with all non-missing mother's age, I find the average birth weight of infants whose mother is older than 40 years is about 46 grams less than that of infants whose mother is between 25 and 30 years.

Figure 2.2 plots the average birth weight of newborns at different gestational ages. The full sample contains 70,731 (about 0.67 percent) infants delivered between 17 weeks and 27 weeks of gestation. These infants do not yet enter the third trimester of pregnancy and are at extremely high risk of early childhood death and lifetime disabilities.[15] After 28 weeks of gestation, male infants start to show slightly higher birth weights than female infants. This gender differential exacerbates as infants grow to up to 40 weeks and stays roughly constant thereafter. For both genders, 40-weeks is the ideal gestational age that generates the highest average birth weight of about 3350 grams. After 40 weeks, birth weight declines gradually with longer gestations.

Figure 2.3 Panel A and Panel B plot the average gestation length and average birth weight, respectively, for male versus female fetuses with mothers at different ages. In general, female fetuses experience a slightly longer gestation than male fetuses, although the latter can be up to 110 grams heavier than the former at birth delivery. On the other hand, both genders experience the longest gestation if the mother is between 20 and 30 years old. Nevertheless, mothers about 30 years old deliver substantially heavier babies than mothers in other age groups regardless of the baby's gender. This is probably because in general, mothers in their early 30s have a better health condition than those in their 40s. At the same time, the former are also likely to have a higher income level to support nutrition intakes than mothers in their early 20s.

Figure 2.3 also shows very noisy average gestation length and birth for mothers below 15 years old and above 45 years old. The sharp variations can either reflect the measurement errors in dataset or be caused by the very small sample size (only 0.25 percent observations fall in these age groups). To maintain a sample with high data quality, I focus on fetuses with mothers between 15 and 45 years old at the time of birth.

Figure 2.4 plots the average gestation and birth weight by race for fetuses with mothers at different ages. Within the same age group, fetuses with white mothers consistently have

---

[15]In fact, the lowest birth weight associated with these extremely pre-term infants is 227 grams. This implies that the infant is most likely born dead.

about 0.6 week longer gestations than those with black mothers. Meanwhile, the former weigh approximately 200 grams more than the latter. This racial disparity may originate from the differences in socioeconomic backgrounds such as income, educational attainment and accessibility to health care providers across white and black mothers.

2.4.2 Air Pollution Data

I combine the birth data with air pollution data obtained from the U.S. Environmental Protection Agency (EPA). I focus on four "criteria pollutants" carbon monoxide (CO), nitrogen dioxide (NO2), ozone (O3) and sulfur dioxide (SO2) because EPA provides high frequency data on these pollutants.[16] Some of these pollutants, such as CO and O3, have been studied in previous literature (Currie and Neidell (2005), Currie et al. (2009), Janke et al. (2009), Knittel et al. (2011) and Coneus and Piess (2012)). However, the results are not very consistent, and most studies in the US use regional instead of national data. Thus it is interesting to see whether the results can be reproduced using individual level natality data at US national scale. Meanwhile, the other pollutants such as NO2 and SO2 have rarely been analyzed before. Understanding their health impacts will provide guidelines for government policymakers to establish optimal environmental regulations to better protect fetal development.

I obtain 1-hour pollution concentration data for CO, NO2 and SO2 from over 500 stations located in almost all states in the United States. For O3, I obtain 8-hour block average pollution concentrations from over 1000 stations in all US states. These high-frequency data are first used to generate the county-day average pollution levels. To create a balanced panel with high data quality, I drop counties containing more than 5 percent missing daily observations from 1990 to 2008.[17] For the remaining counties, I linearly interpolate the gaps

[16] Another pollutant of interest is PM10. However, because PM10 is only reported every 6 days I do not include it in this paper.

[17] I start collecting pollution data 1 year before the starting year of birth data to gather full information on the fetal exposure to air pollution during whole period of pregnancy.

in dates using the nearest two observations. This strategy largely increases the number of counties included in the study without over-manipulating the data. From daily observations, I then generate weekly average pollution levels and drop counties with missing observation in any week.

Table 2.1 provides summary statistics for the county-week air pollution from 1990 to 2008. Under current air quality legislations, all the pollutants considered have a 95 percentile of the weekly average distribution below the EPA primary standards.[18] Nevertheless, each of the pollution distributions is highly left-skewed, and the maximum weekly average O3 and SO2 are both above the EPA standards.

2.4.3 Weather and Demographic Data

Following previous literature, control variables in the study include weather conditions and demographic features that are likely to influence fetal health. The weather data are obtained from the U.S. National Centers for Environmental Information (NCEI) in the National Oceanic and Atmospheric Administration (NOAA). I collect data on the average daily precipitation, maximum and minimum daily temperatures from all counties with at least one air pollution station. The daily observations are then converted to weekly means at county level. To handle missing observations and reserve the largest possible sample size, I first linearly interpolate the the missing values if the two neighboring observations are non-missing. The remaining missing values are then imputed using the state-week averages obtained from all stations with non-missing data in that state. These strategies ensures that the sample contains no missing data in maximum and minimum temperatures and only 0.04 percent missing data in average precipitation.[19]

Finally, I collect county-year medium income data from the Small Area Income and

---

[18] The EPA primary standards are established to protect both normal and sensitive populations against environmental harm. See Chapter 3 Table 3.2 for a summary of current air quality standards

[19] When the data are later converted to trimester level based on the fetus' gestational age, no missing value exists in the sample.

Poverty Estimates (SAIPE) provided by the U.S. Census Bureau. The SAIPE combines survey data and administrative records in estimating the annual income and poverty level for each county. Prior to 2005, the survey data used in SAIPE was the Annual Social and Economic Supplements of the Current Population Survey, while since 2005 the American Community Survey (ACS) has been used. The SAIPE reports county-year medium income in years 1989, 1993, 1995 and every year thereafter. I linearly impute the values of the missing years using the neighbouring two years' data for each county.

The matching process between the birth data and the pollution-weather-income data takes the following steps. First, based on the maternal residential county and the (imputed) infant date of birth, I assign each birth with past weekly air pollution concentrations and weather conditions up to the gestation length. Next I generate average pollution concentrations and average weather conditions during the first (i.e., 0-12 weeks), second (i.e., 13-27 weeks) and third trimesters (i.e., 28 weeks till birth delivery). I exclude fetuses shorter than 28 weeks of gestation since their zero exposure to air pollution during trimester III will likely bias our estimates. Lastly, I merge the resulting dataset with county medium income based on the mother's residential county and the year of pregnancy.[20]

Because an air monitoring station may report some but not all of the air pollutants, the matching process generates four data sets, each for a specific air pollutant. Figure 2.5 illustrates the county locations for the four matched data sets. The birth-CO dataset contains 130 counties located in 38 states. The densest stations locate along the coastal areas of California and New Jersey. The latter also sees a 4.8 percent higher weekly average CO than the national level. The birth-NO2 dataset contains 86 counties in 27 states. It has the smallers sample size among all the samples, however the stations reporting NO2 are relatively sparsely located, and less than 0.01 percent county-level hourly NO2 data exceeds the national 1-hour limit. The birth-O3 dataset contains 143 counties in 23 states,

[20]I use the county medium income in the year of pregnancy instead of the year of birth delivery because the former is more likely to indicate mother's behaviors (e.g. nutrition intake) during the entire pregnancy period.

with the densest counties concentrating in California, New Jersey and mid-Illinois. Roughly 0.072 percent (0.058 percent) observations in these three states (other states) see an 8-hour average O3 exceeding the US national limit. The birth-SO2 dataset contains 185 counties in 41 states, with the majority of counties locate in the eastern regions of US. Approximately 0.02 percent county-level observations for hourly SO2 are above the national 1-hour limit. Counties located very close to each other may share some similarities in air quality, climate conditions and demographic features. A variance comparison test, for instance, shows we can be 100 percent confident that the instate weekly CO in California has a smaller variance than the cross-state CO.

Table 2.2 provides summary statistics for fetal and maternal characteristics in each matched data set. For births delivered between 1991 and 2008, the average pregnancy length for women between 15 and 45 years old is approximately 38.8 weeks in counties with available air quality information. The average birth weight of infants is about 3315 grams, with male infants weighing roughly 110 grams heavier than female infants of the same gestation. The average age of mothers when giving birth is about 27 years old, although on average women may choose to have more than two babies (over 85 percent of these babies survive at birth) over their life time. Roughly 65 percent women are married when a baby is born, and the marriage rate is substantially higher among more educated women (e.g. about 83 percent women with some post-graduate education are married when giving birth). During the entire pregnancy period, mothers visit the prenatal health care provider about 11 times, and older and more educated women receive health care more than the younger and less educated women.

2.5 Estimation Results

2.5.1 Main Results

Table 2.3 presents the estimation results for CO based on regression 2.12. The outcome variable is either the length of gestation in whole weeks (in columns (1) to (3)) or birth weight (in columns (4) to (6)). With each outcome variable, the simplest regression includes only year fixed effects and state-quarter fixed effects. The middle column in each section additionally controls for fetal characteristics (sex, total birth order and live birth order) and maternal characteristics (age, race, Hispanic origin, education and marital status). The richest model includes all the previously-controlled covariates and an additional vector of county medium income, mother's total number of prenatal visits, and weather conditions (average daily precipitation, minimum and maximum temperature) over three trimesters of pregnancy.

As Table 2.3 shows, the estimates are substantially higher in the simplest model when only fixed effects are accounted. Incorporating fetal and maternal characteristics absorbs nearly half of the estimated coefficient on air pollution. Additionally controlling weather, income and prenatal visits further reduces the estimates, possibly due to maternal behavioral adjustments against air pollution. Under the richest model, fetuses are negatively affected by CO during the first and third trimesters. A 1 ppm increase of CO in the first trimester is expected to shorten the gestation by 0.034 week and decrease birth weight by 5.3 grams. Similarly, a 1 ppm increase of CO in the last trimester will (insignificantly) decrease gestation by 0.008 weeks and significantly lower birth weight by 8 grams. These point estimates are smaller in magnitude but more significant than proposed by previous studies (Currie et al. (2009)). One possible reason is that I use county average air pollution to approximate the air pollution measured at the exact residential address of mothers. Using the former will tend to smooth out any acute health effect induced by sharp fluctuations of air pollution. Meanwhile, since the model cannot fully control for maternal protective behaviors, the point estimates presented here will likely be the lower bound of the true impact.

Similarly with previous studies (Currie and Neidell (2005); Currie et al. (2009)), I got positive point estimates for the air pollution's impact during the second trimester. Moreover, both the significance level and the magnitude of coefficient are slightly higher in this study. It is hard to find out any biological explanation for this, but from the econometrics perspective the wrong-signed coefficient does not necessarily indicate that more pollution in the second trimester will benefit fetuses. It is possible that mothers take special care of fetuses and avoid outside activities during heavily polluting days. These behavioral adjustments may lead to a positive coefficient in the second trimester which should not be attributed to air pollution. Section 2.6 provides supportive evidence for this argument.

Table 2.4 displays the estimation results for NO2. Similarly with Table 2.3, accounting for fetal and maternal characteristics as well as weather, income and prenatal care absorbs most of the variations of health outcome. Using the preferred richest model, I find that every 10 ppb increase in NO2 in the third trimester will statistically significant decrease gestation by 0.018 week and birth weight by 7.6 grams. The point estimates over the first and second trimesters are insignificant for both outcome variables, although the negative coefficients during the first trimester are considerably larger than the positive coefficient during the second trimester. This study is among the first few literature that focuses on the health impact of NO2, and the result indicates that an effective regulation on NO2 is expected to bring health benefits to future generations.

Table 2.5 presents the estimation results for O3 based on regression 2.12. A 0.01 ppm increase in O3 over trimesters II and III is estimated to shorten the gestation length by 0.015 and 0.021 week, respectively. However, there is no evidence that O3 would cause a significant change of birth weight. This result is in conflict with Currie et al.'s (2009) finding which documents a slightly significant impact of O3 on birth weight but not on gestation. Besides the previously discussed behavioral adjustment issue, another possible reason for this disparity is that the ground level average O3 over each trimester is too low to capture the health effect.[21] In other words, taking the average of O3 over the three months period for each trimester will likely smooth out the peaks and troughs of air pollution that fetuses might respond to and drive down the estimates toward zero.

Table 2.6 displays the regression results for SO2. Under the richest model, I find that

[21] From 1990 to 2008, the national average O3 has been reduced by roughly 16 percent and the 90 percentile of O3 has been reduced by as much as 20 percent. See `https://www3.epa.gov/airtrends/ozone.html` for more details.

every 10 ppb increase in SO2 in the first trimester will reduce gestation length by roughly 0.03 week. Similar with that for CO, the coefficient for the impact of SO2 in the second trimester on gestation is wrong-signed and significant. In the later section I show that the positive coefficient for SO2 in Table 2.6 column (3) is caused by the same reason as that for CO in Table 2.3 columns (3) and (6), and even though both coefficients are highly significant we cannot safely argue that air pollution during the second trimester will help improving fetal health.

The estimated effects of fetal and maternal characteristics on birth outcomes all have expected signs. Specifically, female fetuses on average enjoy a roughly 0.12 week longer gestation but are in general 112-115 grams lighter than male fetuses, holding other factors constant. Higher ordered births, regardless of gender, are expected to have both a shorter gestation and a lower birth weight, although the fact that the previous births are live would increase the chance that the newly-born babies are of normal weight. On the other hand, older mothers in general have a shorter gestational length than younger mothers. However, the former deliver infants of roughly the same weight as the former. Fetuses with unmarried mothers are also more likely to experience a shorter gestation by about 0.07-0.08 week and a lower birth weight by 45-52 grams. In all samples, mother living in wealthier counties deliver substantially heavier babies, and mothers who visit prenatal health care more during the pregnancy period see significantly better birth outcomes even with the same level of county medium income.

2.5.2 Robustness Checks

The estimation results shown in section 2.5.1 are robust along several dimensions. First, I consider the possibility that omitted weather variable may bias the estimates. For instance, daily hours of sunshine and wind speed can be correlated with both air pollution concentrations and maternal health condition. To account for this issue, I collect data from NCEI at NOAA on the daily average sunshine and wind speed over three trimesters and combine them with the birth sample. Table 2.7 columns (1) and (4) displays the estimation results including these additional weather variables. The missing data from sunshine and wind speed reduce the sample size by 25 to 30 percent. With the remaining non-missing data, I find that most estimates stays relatively consistent with those presented in the preferred specification. However, there exist two exceptions. First, the impact of CO in the first and third trimesters are adjusted. Among the four analyzed air pollutants, CO is the only one lighter than air. Therefore the originally omitted wind speed may be strongly correlated with CO concentration. Nevertheless, the estimated overall impact of CO during the whole pregnancy maintains roughly the same when including both sunshine and wind speed in the regression. Second, the impact of O3 in the second trimester on gestation length nearly doubles. This can be explained by the fact that O3 is formed by the chemical reactions of different pollutants (e.g., nitrogen oxides (NOx), volatile organic compound (VOC)) with the help of sunshine. When omitting hours of sunshine, the negative impact of hazardous ozone on sunny days will likely be offset by the positive health impact produced by sunlight (e.g., more vitamin D produced), driving the estimate towards zero. This slight bias from omitted sunshine, nevertheless, does not persist with dependent variable being the birth weight.

Second, I am concerned with the health effect of omitted parental attributes and genetic factors. In particular, biological and medical studies have shown that paternal genes and certain characteristics may contribute to birth outcomes (Little and Sing 1987; Griffiths et al. 2006; Taiwo and Akinde 2012). Even though the paternal impact may not be as strong as maternal impact, there is a potential threat of bias if the omitted paternal attribute is correlated with air pollution. Unfortunately, the sample contain limited information on the paternal side, and none of the observed variables could provide any information on paternal genetic attributes (e.g. height, weight). Based on the available information, I include father's age group (i.e., under 15, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, and 55-98 years), race and Hispanic origin in the regressions shown by Table 2.7 columns (2) and (5). A naive estimation indicates that the fetal exposure to air pollution is significantly related to these paternal features, therefore including them in the regression serves as a robustness check for the omitted variable bias. Nevertheless, I find little change in the coefficients when paternal attributes are accounted.

Finally I consider possible correlation of random errors across counties. For instance, an influenza outbreak occurring in one county may also affect residents in surrounding counties. Table 2.7 columns (3) and (6) allow this possibility by clustering standard errors at state instead of county level. This modification causes a slight adjustment to the magnitude of standard errors, but almost all of the estimates stay consistent with those in the preferred specification.

2.6 Discussion and Supportive Evidence

The main results presented in Tables 2.3 and 2.6 show surprisingly wrong-signed and significant coefficients for the health impacts of CO and SO2 in trimester II. As noted in section 2.5.1, this positive sign does not necessarily indicate a beneficial impact of these pollutants. Specifically, unobservable maternal protective behaviors against air pollution may lead to the opposite sign of coefficient. While it is infeasible to include mother fixed effects to alleviate this bias, we may instead examine whether the positive sign persists when a more homogenous sample of mothers is considered.

Table 2.8 displays the estimation results for all full-term babies (≥ 37 weeks and ≤ 42 weeks of gestation). Surprisingly, all of the previously positive and significant coefficients become insignificant, and some of them even become negatively signed. I further perform similar analysis (not shown here) with pre-term (< 37 weeks and $>= 28$ weeks) and post-term (> 42 weeks) babies and find that the positive coefficients only occur among pre-term babies. Furthermore, the data show that mothers who have pre-term and extremely pre-term babies tend to visit prenatal care more frequently than those having full-term babies.[22] This supportive evidence implies that mothers with pre-term babies may have been informed by their prenatal care provider of the fetal health conditions (especially during the second

[22]Mothers with infants born between 28 and 36 weeks of gestation visit prenatal care 0.3 time per week (with a maximum being 1.75) while mothers with infants born at least 37 weeks visit prenatal care 0.29 time per week (with a maximum being 1.32).

trimester when fetal development can be easily detected). Therefore it is highly likely that mothers will take special care of their babies and avoid been exposed to environmental hazards when the perceived air pollution concentration is high.

2.7 Conclusion

In this chapter, I review the previous literature on the topic of air pollution's impact on human health and human capital. These interdisciplinary studies benefit from precise datasets, advanced measurement methods and well-built econometric models. They reduce measurement errors and omitted variable bias by adopting pollution distribution measurements, controlling for possible confounders and employing various fixed effects. Applying similar methods and combining data from over 20 million births in the US from 1991 to 2008 with county-week level air pollution, weather and demographic features, I estimate the impact of four "criteria air pollutants" on fetal health. According to the estimation results, gestation length is strongly affected by CO, NO2 and O3 especially during the third trimester. During the second trimester, maternal act of protecting likely pre-term babies against air pollution have led to a wrong signed coefficient for CO and SO2.

On the other hand, CO and NO2 are found to strongly affect infant birth weight especially during the first and last trimesters. Their impacts are even higher on full-term babies between 37 and 42 weeks of gestation. I do not find any adverse impact of O3 and SO2 on birth weight.

Reference

Anna Alberini, Maureen Croppeb, Tsu-Tan Fu, Alan Krupnick, Jin-Tan Liu, Daigee Shaw, and Winston Harrington. Valuing Health Effects of Air Pollution in Developing Countries: The Case of Taiwan. *Journal of Environmental Economics and Management*, 32(2):107–126, 1987.

Douglas Almond, Kenneth Y. Chay, and David S. Lee. The Costs of Low Birth Weight. *The Quarterly Journal of Economics*, 120(3):1031–1083, 2005.

Douglas Almond, Janet Currie, and Katherine Meckel. Fetal Origins of Lifetime Health. *Encyclopedia of Health Economics*, 1:309–314, 2014.

David Barker and Clive Osmond. Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales. *Lancet*, pages 1077–1081, 1986.

Michelle L. Bell, Keita Ebisu, and Kathleen Belanger. Ambient Air Pollution and Low Birth Weight in Connecticut and Massachusetts. *Environmental Health Perspectives*, 115: 1118–1124, 2007.

Prashant Bharadwaj, Petter Lundborg, and Dan-Olof Rooth. Birth Weight in the Long Run. NBER Working Paper No. 21354, 2015.

Martin Bobak. Outdoor Air Pollution, Low Birth Weight, and Prematurity. *Environmental Health Perspectives*, 108(2):173–176, 2000.

Kasey S. Buckles and Daniel M. Hungerman. Season of Birth and Later Outcomes: Old Questions, New Answers. *The Review of Economics and Statistics*, 95(3):711–724, 2013.

Kenneth Y. Chay and Michael Greenstone. Air quality, infant mortality, and the clean air act of 1970. NBER Working Paper No. 10053, 2003a.

Kenneth Y. Chay and Michael Greenstone. The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by A Recession. *Quarterly Journal of Economics*, 118(3):1121–1167, 2003b.

Katja Coneusa and C. Katharina Spiess. Pollution Exposure and Child Health: Evidence for Infants and Toddlers in Germany. *Journal of Health Economics*, 31(1):180–196, 2012.

Jane Currie and Hannes Schwandt. Within-Mother Analysis of Seasonal Patterns in Health at Birth. *Proceedings of the National Academy of Sciences*, 110(30):12265–12270, 2013.

Janet Currie. Pollution and Infant Health. *Child Development Perspectives*, 7(4):237–242, 2013.

Janet Currie and Matthew Neidell. Air Pollution and Infant Health: What Can We Learn From California's Recent Experience? *The Quarterly Journal of Economics*, 120(3):1003–1030, 2005.

Janet Currie and Reed Walker. Traffic Congestion and Infant Health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, 3:65–90, 2011.

Janet Currie, Matthew Neidell, and Johannes F. Schmieder. Air Pollution and Infant Health: Lessons from New Jersey. *Journal of Health Economics*, 28:688–703, 2009.

Lee Ellis and Ping He. Sex Differences in Fetal Activity and Childhood Hyperactivity. *Research Journal of Developmental Biology*, 2014. doi: 10.7243/2055-4796-1-1.

Emissions Inventory and Analysis Group. Profile of the 2011 National Air Emissions Inventory.

Environmental Protection Agency (EPA). Air Quality Index: A Guide to Air Quality and Your Health, 2014.

Gary W. Evans. Child Development and the Physical Environment. *Annual Review of Psychology*, 57:423–451, 2006.

Lucy J. Griffiths, Carol Dezateux, Tim J. Cole, and Millennium Cohort Study Child Health Group. Differential Parental Weight and Height Contributions to Offspring Birthweight and Weight Gain in Infancy. *International Journal of Epidemiology*, 36:104–107, 2007.

Katharina Janke, Carol Propper, and John Henderson. Do Current Levels of Air Pollution Kill? The Impact of Air Pollution on Population Mortality in England. *Health Economics*, 18(9):1031–1055, 2014.

Wieslaw Jedrychowski, Frederica Perera, Dorota Mrozek-Budzyn, Elzbieta Mroz, Elzbieta Flak, Jack D. Spengler, Susan Edwards, Ryszard Jacek, Irena Kaim, and Zbigniew Skolicki. Gender Differences in Fetal Growth of Newborns Exposed Prenatally to Airborne Fine Particulat Matter. *Environmental Research*, 109:447–456, 2009.

Marilena Kampa and Elias Castanas. Human Health Effects of Air Pollution. *Environmental Pollution*, 151:362–367, 2015.

Christopher R. Knittel, Douglas L. Miller, and Nicholas J. Sanders. Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health. NBER Working Paper No. 17222, 2011.

Emmanuelle Lavaine and Matthew Neidell. Energy production and health externalities: Evidence from Oil Refinery Strikes in France. NBER Working Paper No. 18974, 2014.

Victor Lavy, Avraham Ebenstein, and Sefi Roth. The Long Run Human Capital and Economic Consequences of High-Stakes Examinations. NBER Working Paper No. 20647, 2014.

Mildred Maisonet, Timothy J. Bush, Adolfo Correa, and Jouni J.K. Jaakkola. Relation Between Ambient Air Pollution and low Birth Weight in the Northeastern United States. *Environmental Health Perspectives*, 109(3):351–356, 2001.

Bart D. Ostro and Susy Rothschild. Air Pollution and Acute Respiratory Morbidity: An Observational Study of Multiple Pollutants. *Environmental Research*, 50(2):238–247, 1989.

Mark M. Pitt and Wendy Sigle. Seasonality, Weather Shocks and the Timing of Births and Child Mortality in Senegal. In *Anniversary Conference of the Centre for the Study of African Economies*, 1997.

Beate Ritz and Michelle Wilhelm. Air Pollution Impacts on Infants and Children (Southern California Environmental Report Card). Technical report, Institute of the Environment, University of California, Los Angeles, 2008.

Jonathan M. Samet, Marian C. Marbury, and John D. Spengler. Health Effects and Sources of Indoor Air Pollution. Part I. *American Review of Respiratory Disease*, 136(6):1486–1508, 1987.

Nicholas J. Sanders and Charles F. Stoecker. Where Have All the Young Men Gone? Using Gender Ratios to Measure Fetal Death Rates. NBER Working Paper No. 17434, 2011.

Joel Schwartz. Short Term Fluctuations in Air Pollution and Hospital Admissions of the Elderly for Respiratory Disease. *Thorax*, 50:531–538, 1995.

I.A Taiwo and O.R Akinde. Predictability of Offspring Birth Weight Using Simple Parental Anthropometrics in A Government Hospital in Lagos, Nigeria. *International Journal of Medicine and Biomedical Research*, 1(3):206–214, 2012.

Xiaobin Wang, Hui Ding, Louise Ryan, and Xiping Xu. Association between Air Pollution and Low Birth Weight: A Community-based Study. *Environmental Health Perspectives*, 105(5):514–520, 1997.

LIST OF FIGURES IN CHAPTER 2

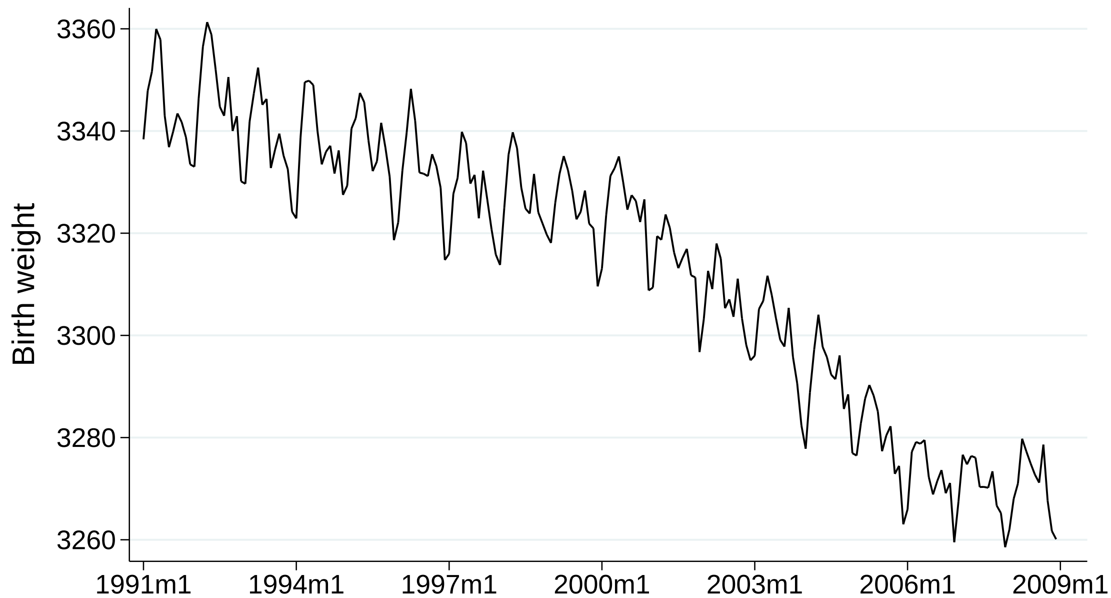
LIST OF TABLES IN CHAPTER 2

Table 2.1: Summary statistics of air pollution

Pollutants	Mean	Std. Dev.	Percentiles								N
			10	30	50	60	70	80	90	95	
CO (ppm)	0.74	0.44	0.30	0.46	0.64	0.74	0.86	1.02	1.30	1.59	128440
NO2 (ppb)	16.44	8.02	7.14	11.60	15.34	17.42	19.79	22.73	27.07	30.94	86944
O3 (ppm)	0.027	0.011	0.013	0.020	0.026	0.029	0.032	0.035	0.041	0.045	141284
SO2 (ppb)	5.25	4.71	1.01	2.53	4.12	5.11	6.33	8.00	10.73	13.40	182780

This table provides summary statistics for the county-week average CO, O3 and SO2 in the U.S. from 1991 to 2008. Number of counties are 130 for CO, 143 for O3 and 185 for SO2.

Panel A



Panel B

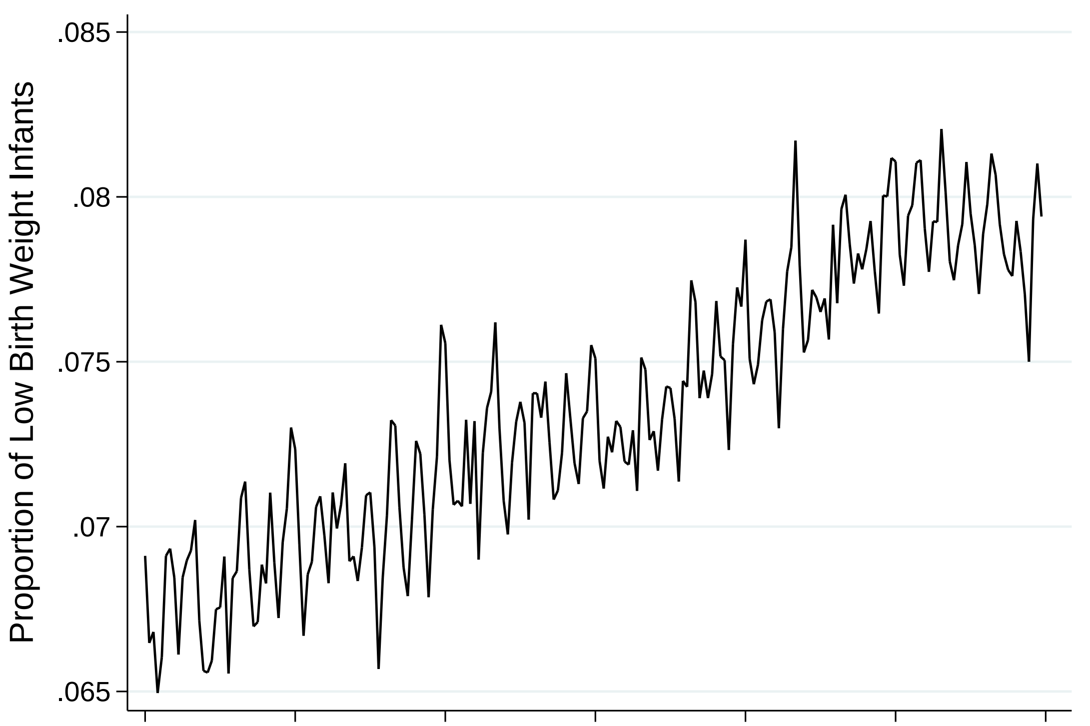


Figure 2.1: Time series monthly average birth weight and incidence of low birth weight infants

Notes: Panel A of this figure plots the monthly average birth weight in the United States from 1991 to 2008. Panel B of this figure plots the monthly average incidence of low birth weight infants (i.e., less than 2500 grams at birth) in the United States from 1991 to 2008.

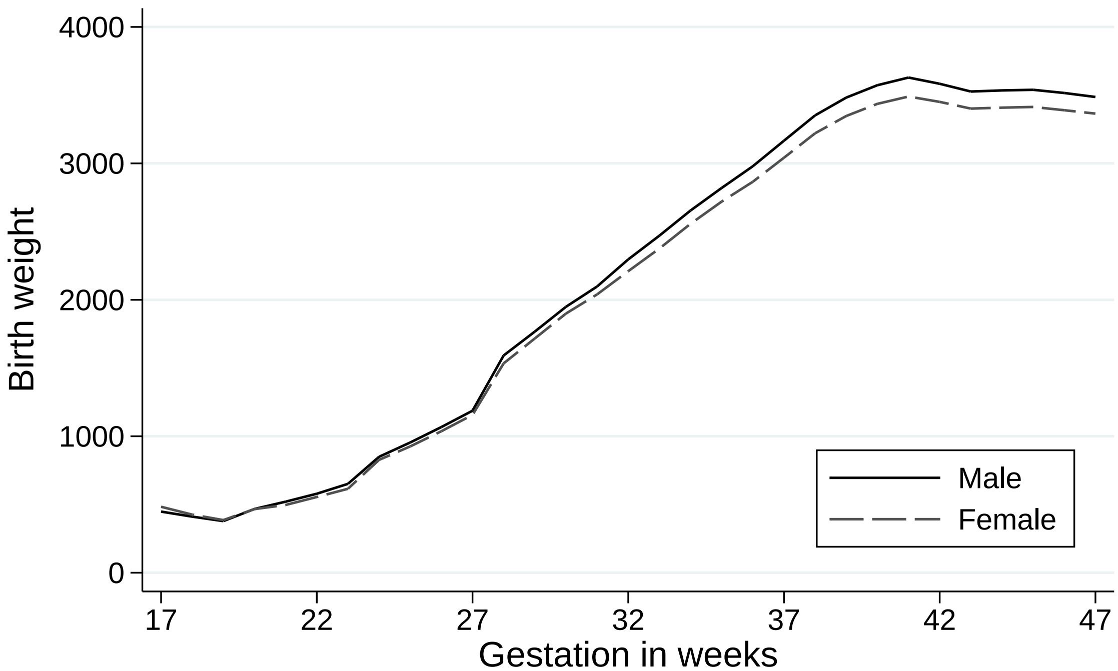


Figure 2.2: Average birth weight at each gestational age

Notes: This figure plots the average birth weight at each gestational age by sex in the U.S. from 1991 to 2008.

Panel A Panel B

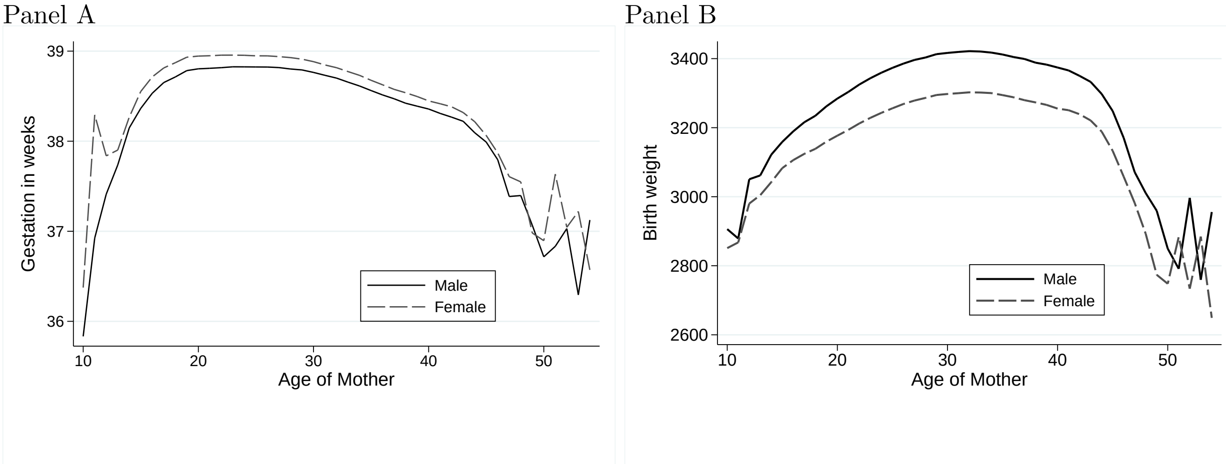


Figure 2.3: Average gestation length and birth weight at each age of mother (by fetus gender)

Notes: Panel A of this figure plots the average gestational age for male and female babies at different maternal ages between Jan 1, 1991 and Dec 31, 2008 in the U.S.. Panel B of this figure shows the average birth weight for male and female babies at different maternal ages between Jan 1, 1991 and Dec 31, 2008 in the U.S..

Panel A Panel B

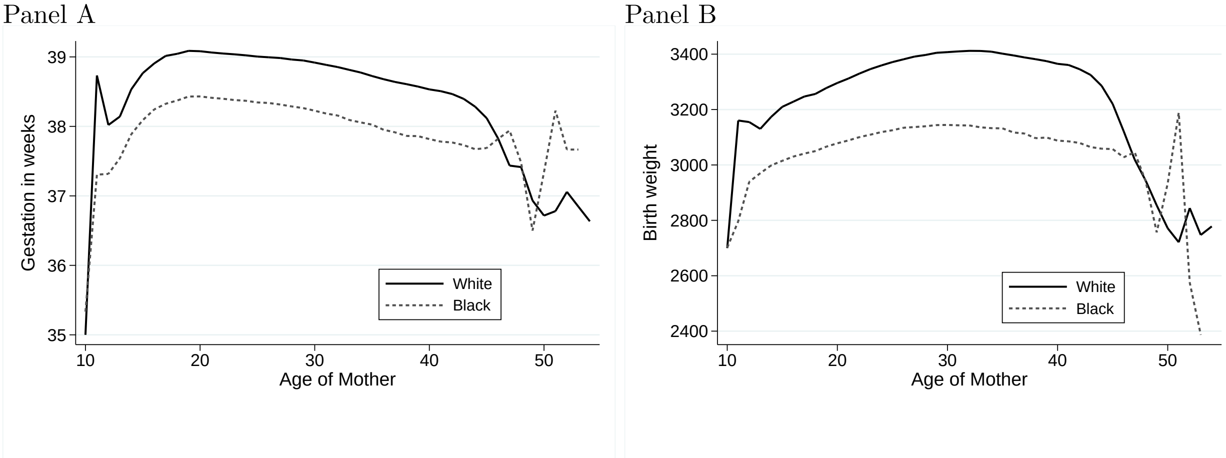


Figure 2.4: Average gestation length and birth weight at each age of mother (by mother's race)

Notes: Panel A of this figure plots the average gestational age for babies with white versus black mothers at different maternal ages between Jan 1, 1991 and Dec 31, 2008 in the U.S.. Panel B of this figure plots the average birth weight for babies with white versus black mothers at different maternal ages between Jan 1, 1991 and Dec 31, 2008 in the U.S..

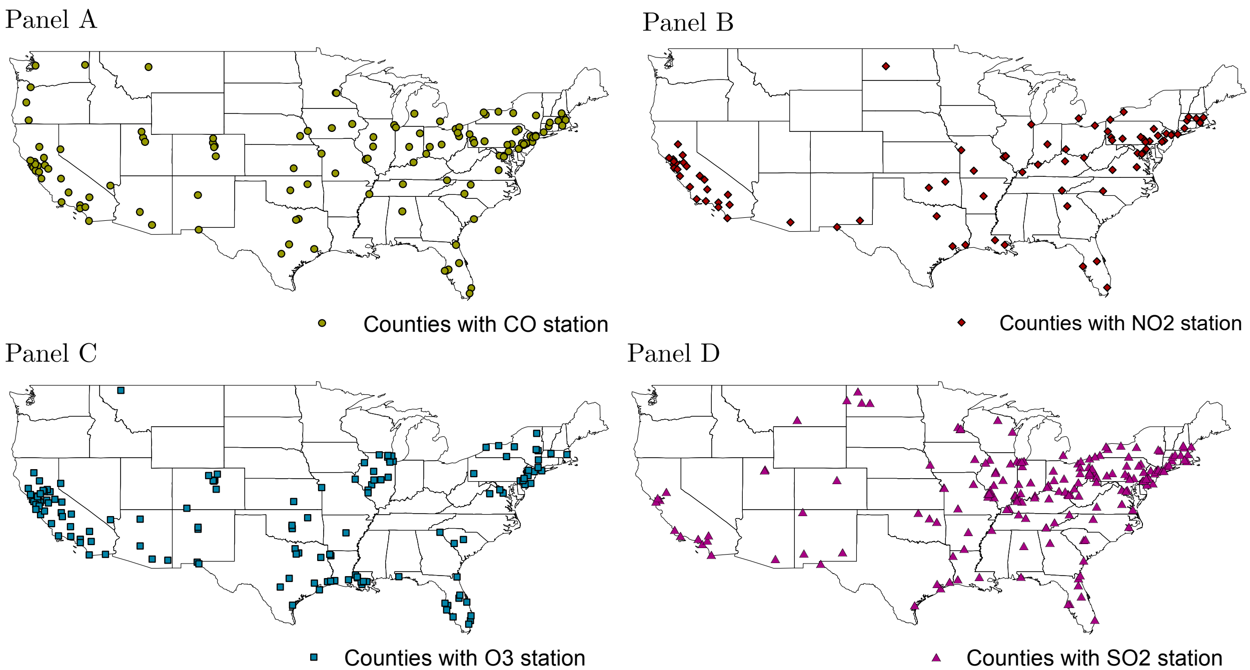


Figure 2.5: Location of counties with air pollution stations

Notes: This map shows the location of counties for each matched dataset between natality and air pollution.

Table 2.2: Summary Statistics for Matched Data

Variable	CO sample		NO2 sample		O3 sample		SO2 sample	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Outcome variables								
Birth weight	3314.484	590.724	3315.346	592.253	3315.292	585.967	3311.574	595.493
Gestational age	38.827	2.533	38.835	2.549	38.824	2.527	38.816	2.556
Infant characteristics								
Male	0.511	0.5	0.511	0.5	0.511	0.5	0.511	0.5
Live birth order	2.061	1.226	2.081	1.242	2.071	1.223	2.063	1.222
Total birth order	2.405	1.497	2.402	1.496	2.395	1.478	2.418	1.502
Mother characteristics								
Married	0.650	0.477	0.633	0.482	0.65	0.477	0.639	0.48
Age	27.421	6.125	27.365	6.157	27.318	6.141	27.408	6.106
Mexican	0.234	0.423	0.278	0.448	0.291	0.454	0.199	0.4
Puerto Rican	0.016	0.125	0.017	0.128	0.013	0.115	0.022	0.146
Cuban	0.003	0.059	0.004	0.06	0.004	0.064	0.002	0.048
Central or South American	0.041	0.198	0.048	0.214	0.042	0.201	0.043	0.202
Other/Unknown Hispanic	0.015	0.122	0.012	0.11	0.019	0.136	0.011	0.106
Non-Hispanic White	0.467	0.499	0.403	0.491	0.421	0.494	0.489	0.5
Non-Hispanic Black	0.150	0.357	0.165	0.371	0.134	0.340	0.173	0.378
Non-Hispanic other races	0.074	0.262	0.073	0.261	0.076	0.265	0.060	0.238
0-8 years education	0.081	0.272	0.096	0.295	0.093	0.291	0.074	0.262
9-11 years education	0.168	0.374	0.179	0.383	0.177	0.382	0.167	0.373
12 years education	0.301	0.459	0.303	0.459	0.303	0.460	0.308	0.462
13-15 years education	0.211	0.408	0.202	0.402	0.208	0.406	0.215	0.411
16+ years education	0.239	0.426	0.22	0.414	0.219	0.414	0.236	0.425
Other variables								
# prenatal visits	11.392	4.059	11.418	4.1	11.422	4.117	11.401	4.068
Precipitation (1st trim.)	23.441	33.176	22.24	17.277	21.834	33.589	25.004	32.939
Precipitation (2nd trim.)	23.471	32.131	22.135	16.497	21.87	32.324	25.056	31.228
Precipitation (3rd trim.)	23.257	32.874	21.941	17.697	21.66	33.416	24.807	32.365
Max temperature (1st trim.)	210.772	94.949	211.178	79.444	225.788	88.563	203.28	95.398
Max temperature (2nd trim.)	210.702	94.66	212.537	76.613	225.701	88.212	203.231	95.003
Max temperature (3rd trim.)	209.404	95.828	213.06	81.314	224.296	89.801	201.896	96.217
Min temperature (1st trim.)	93.259	85.313	94.066	70.281	105.741	81.288	89.203	86.155
Min temperature (2nd trim.)	93.180	85.057	95.201	67.805	105.643	80.972	89.144	85.795
Min temperature (3rd trim.)	92.593	85.048	96.331	70.807	104.941	81.19	88.535	85.836
County medium income	46603.209	9147.637	46148.631	9952.382	46920.531	9789.593	46041.6	9440.093
N. of counties	130		86		143		185	
N. of obs.	23543770		16248562		19958422		19123575	

This table provides summary statistics for the matched data sets between birth and air-pollution-weather-income from 1991 to 2008. I drop the observations falling in either of the following groups: 1. with missing data in birth weight, gestational age, (induced) birth date (using mother's date of last menstrual period (LMP)), fetal characteristics, maternal characteristics or number of prenatal visits; 2. with mothers younger than 15 years or older than 45 years. The total numbers of observations do not reveal missing data in weather or demographic variables. This table includes fetuses born prior to the 28th week of gestation, but when running regressions those fetuses are excluded.

Table 2.3: Effect of Average Carbon Monoxide in Three Trimesters on Fetal Health

	Dep. var = gestation			Dep. var = birth weight		
	(1)	(2)	(3)	(4)	(5)	(6)
Trimester I	-0.071***	-0.057***	-0.034***	-19.268***	-9.171***	-5.347*
	(0.013)	(0.012)	(0.012)	(3.867)	(2.720)	(2.803)
Trimester II	0.047***	0.051***	0.049***	9.675***	12.501***	6.053***
	(0.013)	(0.013)	(0.012)	(2.551)	(1.856)	(1.740)
Trimester III	-0.027**	-0.011	-0.008	-27.837***	-16.433***	-7.973***
	(0.013)	(0.012)	(0.017)	(3.977)	(2.710)	(2.951)
Controls						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Fetal characteristics	No	Yes	Yes	No	Yes	Yes
Maternal characteristics	No	Yes	Yes	No	Yes	Yes
Weather	No	No	Yes	No	No	Yes
County medium income	No	No	Yes	No	No	Yes
# prenatal visits	No	No	Yes	No	No	Yes
Observations	23396133	23396133	23379600	23396133	23396133	23379600
R^2	0.007	0.018	0.026	0.007	0.050	0.055

This table shows the estimation results for the effects of average CO during three trimesters on fetal health. The sample is restricted to births of at least 28 weeks in gestation. Fetal characteristics include sex, total birth order and live birth order. Maternal characteristics include age, race, Hispanic origin, education and marital status. Weather conditions include average precipitation, maximum and minimum temperatures over three trimesters.

Robust standard errors clustered by county in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.4: Effect of Average Nitrogen Dioxide in Three Trimesters on Fetal Health

	Dep. var = gestation			Dep. var = birth weight		
	(1)	(2)	(3)	(4)	(5)	(6)
Trimester I	-0.037***	-0.027***	-0.009	-9.599***	-4.145***	-2.418
	(0.006)	(0.006)	(0.008)	(2.137)	(1.522)	(1.690)
Trimester II	0.005	0.009***	0.004	2.633*	4.373***	0.674
	(0.004)	(0.003)	(0.005)	(1.528)	(1.218)	(1.259)
Trimester III	-0.026***	-0.015***	-0.018**	-17.805***	-11.203***	-7.580***
	(0.004)	(0.004)	(0.008)	(1.886)	(1.359)	(1.575)
Controls						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Fetal characteristics	No	Yes	Yes	No	Yes	Yes
Maternal characteristics	No	Yes	Yes	No	Yes	Yes
Weather	No	No	Yes	No	No	Yes
County medium income	No	No	Yes	No	No	Yes
# prenatal visits	No	No	Yes	No	No	Yes
Observations	16143617	16143617	16127084	16143617	16143617	16127084
R^2	0.006	0.017	0.025	0.007	0.050	0.055

This table shows the estimation results for the effects of average NO2 during three trimesters on fetal health. The sample is restricted to births of at least 28 weeks in gestation. Dependent variable is average NO2 in 10 ppb. Fetal characteristics include sex, total birth order and live birth order. Maternal characteristics include age, race, Hispanic origin, education and marital status. Weather conditions include average precipitation, maximum and minimum temperatures over three trimesters.

Robust standard errors clustered by county in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5: Effect of Average Ozone in Three Trimesters on Fetal Health

	Dep. var = gestation			Dep. var = birth weight		
	(1)	(2)	(3)	(4)	(5)	(6)
Trimester I	0.018**	0.006	0.004	5.387**	-1.224	-2.135
	(0.008)	(0.006)	(0.006)	(2.598)	(1.529)	(1.644)
Trimester II	-0.011**	-0.016***	-0.015*	-3.173***	-5.321***	1.032
	(0.005)	(0.005)	(0.009)	(1.053)	(0.689)	(1.564)
Trimester III	-0.010	-0.023***	-0.021***	7.435**	-0.129	-2.529
	(0.008)	(0.006)	(0.006)	(3.037)	(1.931)	(1.931)
Controls						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Fetal characteristics	No	Yes	Yes	No	Yes	Yes
Maternal characteristics	No	Yes	Yes	No	Yes	Yes
Weather	No	No	Yes	No	No	Yes
County medium income	No	No	Yes	No	No	Yes
# prenatal visits	No	No	Yes	No	No	Yes
Observations	19837801	19837801	19837801	19837801	19837801	19837801
R^2	0.008	0.018	0.025	0.007	0.048	0.053

This table shows the estimation results for the effects of average O3 during three trimesters on fetal health. The sample is restricted to births of at least 28 weeks in gestation. Dependent variable is average O3 in 0.01 ppm. Fetal characteristics include sex, total birth order and live birth order. Maternal characteristics include age, race, Hispanic origin, education and marital status. Weather conditions include average precipitation, maximum and minimum temperatures over three trimesters.

Robust standard errors clustered by county in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6: Effect of Average Sulfur Dioxide in Three Trimesters on Fetal Health

	Dep. var = gestation			Dep. var = birth weight		
	(1)	(2)	(3)	(4)	(5)	(6)
Trimester I	-0.062***	-0.050***	-0.031*	-12.918*	-5.424*	0.255
	(0.015)	(0.013)	(0.016)	(6.979)	(2.849)	(3.458)
Trimester II	0.052**	0.056***	0.056***	7.950	9.828***	2.543
	(0.020)	(0.015)	(0.017)	(5.841)	(2.638)	(3.108)
Trimester III	0.006	0.016	-0.014	-9.751	-2.696	1.208
	(0.016)	(0.012)	(0.015)	(7.821)	(2.960)	(3.049)
Controls						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Fetal characteristics	No	Yes	Yes	No	Yes	Yes
Maternal characteristics	No	Yes	Yes	No	Yes	Yes
Weather	No	No	Yes	No	No	Yes
County medium income	No	No	Yes	No	No	Yes
# prenatal visits	No	No	Yes	No	No	Yes
Observations	18997081	18997081	18997081	18997081	18997081	18997081
R^2	0.007	0.019	0.027	0.007	0.052	0.058

This table shows the estimation results for the effects of average SO2 during three trimesters on fetal health. The sample is restricted to births of at least 28 weeks in gestation. Dependent variable is average SO2 in 10 ppb. Fetal characteristics include sex, total birth order and live birth order. Maternal characteristics include age, race, Hispanic origin, education and marital status. Weather conditions include average precipitation, maximum and minimum temperatures over three trimesters.

Robust standard errors clustered by county in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.7: Robustness Checks

	Dep. var = gestation			Dep. var = birth weight		
	+sun, wind	+paternal char.	clust. by state	+sun, wind	+paternal char.	clust. by state
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: CO						
Trimester I	-0.011	-0.035***	-0.034***	-4.557	-4.917*	-5.347**
	(0.013)	(0.012)	(0.011)	(3.231)	(2.718)	(2.613)
Trimester II	0.043***	0.049***	0.049**	5.384***	6.193***	6.053***
	(0.014)	(0.013)	(0.020)	(1.835)	(1.698)	(1.500)
Trimester III	-0.043**	-0.008	-0.008	-10.347***	-7.568***	-7.973***
	(0.017)	(0.017)	(0.025)	(3.290)	(2.849)	(2.930)
Observations	17122758	23379600	23379600	17122758	23379600	23379600
R^2	0.026	0.026	0.026	0.057	0.057	0.055
Panel B: NO2 in 10 ppb						
Trimester I	0.011	-0.009	-0.009	-1.381	-2.210	-2.418
	(0.008)	(0.007)	(0.010)	(1.817)	(1.639)	(1.521)
Trimester II	-0.011	0.004	0.004	0.238	0.686	0.674
	(0.010)	(0.005)	(0.007)	(1.596)	(1.260)	(0.884)
Trimester III	-0.025***	-0.019**	-0.018	-7.560***	-7.425***	-7.580***
	(0.009)	(0.008)	(0.013)	(1.929)	(1.506)	(0.879)
Observations	11518549	16127084	16127084	11518549	16127084	16127084
R^2	0.025	0.025	0.025	0.057	0.057	0.055
Panel C: O3 in 0.01 ppm						
Trimester I	0.007	0.004	0.004	-1.041	-2.660	-2.135
	(0.006)	(0.006)	(0.010)	(1.687)	(1.607)	(1.713)
Trimester II	-0.029***	-0.015*	-0.015*	0.590	0.483	1.032
	(0.006)	(0.009)	(0.008)	(1.637)	(1.548)	(1.412)
Trimester III	-0.019***	-0.021***	-0.021***	-2.135	-3.226*	-2.529
	(0.007)	(0.006)	(0.007)	(1.865)	(1.904)	(2.073)
Observations	14755983	19837801	19837801	14761086	19837801	19837801
R^2	0.026	0.025	0.025	0.055	0.055	0.053
Panel D: SO2 in 10 ppb						
Trimester I	-0.018	-0.031*	-0.031**	1.209	0.568	0.255
	(0.016)	(0.016)	(0.014)	(3.171)	(3.431)	(2.733)
Trimester II	0.045**	0.056***	0.056***	0.357	2.568	2.543
	(0.020)	(0.017)	(0.019)	(3.563)	(3.089)	(3.288)
Trimester III	-0.013	-0.014	-0.014	0.299	1.217	1.208
	(0.018)	(0.015)	(0.016)	(2.959)	(3.102)	(2.854)
Observations	13918585	18997081	18997081	13918585	18997081	18997081
R^2	0.027	0.027	0.027	0.059	0.059	0.058

This table shows the robustness check results for the effects of air pollution on fetal health. The sample is restricted to births of at least 28 weeks in gestation. All regressions include fetal characteristics (sex, total birth order and live birth order), maternal characteristics (age, race, Hispanic origin, education and marital status) and weather conditions (average precipitation, maximum and minimum temperatures over three trimesters). Columns (1) and (4) additionally control for average hours of sunshine over each trimester. Columns (2) and (5) additionally control for paternal characteristics including age group (under 15, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, and 55-98 years), race and Hispanic origin. Columns (3) and (6) cluster standard errors by state.

Robust standard errors clustered by county in parentheses (except for columns (3) and (6).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.8: Effect of Air Pollution on Fetal Health - Full-Term Infants Only

	Dep. var = gestation				Dep. var = birth weight			
	preferred	+sun, wind	+paternal char.	clust. by state	preferred	+sun, wind	+paternal char.	clust. by state
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: CO								
Trimester I	-0.009	0.000	-0.009	-0.009	-5.942**	-5.879**	-5.363**	-5.942**
	(0.008)	(0.009)	(0.008)	(0.009)	(2.637)	(2.966)	(2.552)	(2.494)
Trimester II	0.002	0.007	0.002	0.002	-0.307	0.282	-0.186	-0.307
	(0.006)	(0.006)	(0.006)	(0.006)	(1.493)	(1.745)	(1.449)	(1.508)
Trimester III	-0.012	-0.028***	-0.011	-0.012	-8.514***	-9.374***	-8.003***	-8.514***
	(0.008)	(0.008)	(0.008)	(0.011)	(2.520)	(2.696)	(2.417)	(1.996)
Observations	20051920	14682650	20051920	20051920	20051920	14682650	20051920	20051920
R^2	0.026	0.025	0.026	0.026	0.064	0.065	0.066	0.064
Panel B: NO2 in 10 ppb								
Trimester I	-0.001	0.006	-0.001	-0.001	-3.184**	-2.697*	-2.879**	-3.184***
	(0.004)	(0.005)	(0.004)	(0.005)	(1.417)	(1.525)	(1.363)	(0.918)
Trimester II	-0.001	-0.002	-0.001	-0.001	-1.204	-0.393	-1.195	-1.204*
	(0.003)	(0.005)	(0.003)	(0.003)	(1.010)	(1.146)	(1.020)	(0.687)
Trimester III	-0.016***	-0.021***	-0.016***	-0.016***	-6.470***	-6.581***	-6.233***	-6.470***
	(0.005)	(0.004)	(0.005)	(0.004)	(1.623)	(1.775)	(1.543)	(1.027)
Observations	13821763	9867557	13821763	13821763	13821763	9867557	13821763	13821763
R^2	0.024	0.023	0.024	0.024	0.064	0.065	0.066	0.064
Panel C: O3 in 0.01 ppm								
Trimester I	-0.016***	-0.016***	-0.016***	-0.016***	-1.706	-1.272	-2.327	-1.706
	(0.004)	(0.003)	(0.004)	(0.004)	(1.556)	(1.573)	(1.511)	(1.487)
Trimester II	0.002	-0.004	0.002	0.002	1.104	1.312	0.543	1.104
	(0.004)	(0.003)	(0.004)	(0.002)	(1.429)	(1.517)	(1.397)	(1.623)
Trimester III	-0.019***	-0.016***	-0.019***	-0.019***	-2.085	-1.155	-2.864	-2.085
	(0.004)	(0.005)	(0.004)	(0.003)	(1.959)	(1.846)	(1.919)	(2.544)
Observations	16981884	12619709	16981884	16981884	16981884	12619709	16981884	16981884
R^2	0.026	0.026	0.026	0.026	0.062	0.063	0.064	0.062
Panel D: SO2 in 10 ppb								
Trimester I	-0.002	0.000	-0.002	-0.002	0.763	0.321	1.075	0.763
	(0.009)	(0.008)	(0.009)	(0.009)	(3.055)	(2.744)	(3.039)	(2.262)
Trimester II	0.003	0.003	0.003	0.003	-2.027	-2.883	-1.965	-2.027
	(0.009)	(0.010)	(0.009)	(0.008)	(2.634)	(2.779)	(2.633)	(2.729)
Trimester III	0.005	0.003	0.004	0.005	2.402	1.322	2.372	2.402
	(0.009)	(0.009)	(0.009)	(0.009)	(2.819)	(2.613)	(2.842)	(2.470)
Observations	16276001	11920840	16276001	16276001	16276001	11920840	16276001	16276001
R^2	0.026	0.026	0.026	0.026	0.066	0.067	0.068	0.066

This table shows the estimation results for the effects of air pollution on fetal health for all full-term babies between 37 and 42 weeks of gestation. All regressions include fetal characteristics (sex, total birth order and live birth order), maternal characteristics (age, race, Hispanic origin, education and marital status) and weather conditions (average precipitation, maximum and minimum temperatures over three trimesters). Columns (2) and (6) additionally control for average daily hours of sunshine and wind speed over each trimester. Columns (3) and (7) additionally control for paternal characteristics including age group (under 15, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, and 55-98 years), race and Hispanic origin. Columns (4) and (8) cluster standard errors by state.

Robust standard errors clustered by county in parentheses (except for columns (5) and (8).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3

AIR POLLUTION AND FETAL HEALTH: DOES A SAFETY THRESHOLD EXIST?

3.1 Introduction

Despite the global awareness of pollution-caused health issues, there is no consensus on which pollution threshold should be considered as "safe" for human beings. Taking fine particles (or commonly known as PM2.5, particulate with diameter less than 2.5 micrometers) as an example. In 1997, the United States Environmental Protection Agency (EPA) established an annual standard of PM2.5 at 15 $\mu g/m^3$ and a 24-hour standard at 65 $\mu g/m^3$.[1] In 2006, the 24-hour standard was revised to 35 $\mu g/m^3$ while the annual limit remained the same. Later in 2012, the EPA revised the annual standard of PM2.5 to 12 $\mu g/m^3$. At the same time, the European standard of PM2.5 has been maintained at 25 $\mu g/m^3$ for 1-year average and 20 $\mu g/m^3$ for 3-year average, each with an over 60 percent difference compared with the standard in the United States. Table 3.2 shows the current air quality standards for seven major air pollutants ("criteria air pollutants") in the US versus in the European Union. After converting to the same unit, the value of the standards as well as the time period during which the standards apply still vary substantially across regions.

This chapter applies an innovative measurement strategy to assess the health risk of air pollution, and at the same time, explore the empirical question of whether there exists a "safety threshold" of air pollution. To complete this goal, I establish a series of hypothetical "safety thresholds" for each air pollutant based on the observed distribution of weekly pollution concentrations. Given any safety threshold C, I then construct an index called "average

[1]The annual standard is based on a 3-year annual average of PM2.5 concentrations. The 24-hour standard is based on the 3-year average of the annual 98th percentiles of PM2.5 concentrations. More information can be obtained at `https://www3.epa.gov/airquality/particlepollution/designations/index.htm`.

pollution degree" (APD) that averages out all the high pollution levels ranging above C while ignoring those below C. This construction has two major advantages. First, it replicates the form of air quality index (AQI) established in many countries to inform the public of any possible health consequences related to outdoor air pollution exposure. In particular, human beings are expected to experience little or no health risk when exposed to the air with sufficiently low AQI.[2] In other words, adverse health consequences become measurable only if the air pollution exceeds a certain level. Second, with substantial governmental effort in reducing air pollution over the past decades, the air quality in the United States has been improved dramatically. Table 3.1 shows that from 1990 to 2008 the national average ambient CO, NO2, O3 and SO2 has been reduced by 73, 43, 16 and 51 percent, respectively (Environmental Protection Agency (EPA)).[3] The long-run decline of air pollution creates difficulties in the identification of health impact of air pollution because the very low concentration of pollutants may not cause severe health consequences. In addition, taking the average of air pollution over each trimester will further smooth out any acute fluctuations of air quality over a short period of time.[4] By weighting the high pollution concentrations heavily, the APD approach increases the explanatory power of pollution on birth outcomes.

This paper is the first empirical analysis on the sensitivity of fetuses to different air pollution thresholds. According to my estimation, even a nearly zero concentration of CO could generate severe health consequences. In contrast, NO2, O3 and SO2 do not cause significantly adverse health outcomes unless their weekly average concentrations rise above 15.3 ppb, 0.020 ppm and 13.4 ppb, respectively. This finding suggests that the ideal environment for fetal development should be the one that maintains roughly zero CO and relatively low NO2, O3 and SO2 concentrations. For the latter three pollutants, the estimated thresholds

[2]The current AQI for "good" air quality is 0 to 50, indicating no health concerns on human beings. "Moderate" air quality is associated with an AQI between 51 and 100, suggesting that some pollutants might impose moderate health concerns for the most sensitive group. "Unhealthy" air quality starts from an AQI above 101, where the generic population will likely experience pollution-caused symptoms.

[3]See `https://www3.epa.gov/airtrends/aqtrends.html` for more details.

[4]This can be seen from the estimated health impact of O3 and SO2 in Chapter 2. Neither pollutants are found to significantly affect infant birth weight.

that induce the largest and most significant change of birth outcomes are all below the current air quality standards established by the U.S. Environmental Protection Agency (EPA), implying that more restrictive legislations should be established.

The study also explores the heterogeneous impacts of pollution exposure across fetal and maternal characteristics. Specifically, I find that older mothers giving subsequent births are more likely to be affected by CO, especially during the first and third trimesters. This implies that an air quality legislation that effectively reduces CO concentrations will be expected to increase female fertility especially for older mothers. On the other hand, female fetuses are slightly more sensitive to NO2 while male fetuses tend to be more sensitive to O3 and SO2. The finding on gender disparity is related to biological studies that show male fetuses tend to be more active than female fetuses before birth delivery (Ellis and He (2014)), and social science research on the change of gender ratio led by the Clean Air Act Amendments of 1970 (CAAA) (Sanders and Stoeckel (2011)). While the estimation result on NO2 is not very consistent with the literature, in the presence of stronger protective behaviors for male fetuses against environmental hazards (among some socioeconomic groups), one would expect that the estimates on male fetuses would more likely be the lower bound. The gender disparity offer important policy implications. In particular, regions with increasing air pollution may experience a change of gender ratio in newly born infants, which could potentially affect marriage, productivity, violence, and crime rate in the long run (Angrist 2006; Hesketh and Xing 2006; Dyson 2012).

Besides the heterogeneity arising in the health effects of air pollution across socioeconomic groups, this paper also examines whether maternal cigarette and alcohol consumptions during pregnancy would exacerbate the health impact of air pollution. Using full-term infants as a homogenous sample, I find that a fetus whose mother smokes during pregnancy would weigh 169 to 197 grams lower than a fetus with a non-smoking mother, holding other fetal and maternal characteristics constant. Likewise, being a light drinker (i.e. drinking 1-2 times each week) during pregnancy would significantly reduce the infant birth weight by 58 to 134 grams. A heavy drinker (i.e. drinking more than 3 times per week), on the other hand, would further drive down the infant birth weight by 137 to 235 grams. Moreover, for most air pollutants (CO, NO2 and SO2), smoking and alcohol consumption would exacerbate the health effects of air pollution exposures by 1.4 to 17 times.

3.2 Model and Identification Strategy

3.2.1 Average Air Pollution Degree

Traditional approaches to quantifying the causal impact of pollution and health outcomes utilize the average pollution level over a time span as the key explanatory variable. This method relies on the assumption that any fluctuation in average pollution concentrations is considered to convey explanatory power on the health consequences. Under current air quality legislations, however, most regions in the United States maintain very low levels of major air pollutants. The lack of sufficient variation in pollution concentrations, especially when smoothing out the variables over a long period, makes the estimated coefficient unlikely to represent the adverse health outcomes caused by environmental hazards.

In this paper, I modify the key regressor to attenuate this concern. Consistent with previous literature (Currie et al. 2009; Maisonet et al. 2001; Bell et al. 2007), I distinguish the impact of fetal exposure to air pollution over three trimesters of pregnancy, assuming that the risk of pollution starts as early as the first week after conception and lasts till the birth occurrence. Let C be any hypothetical "safety threshold" of an air pollutant over a time segment t (e.g., day, week, etc.). The pollution degree (PD) at location c over a time period of length T is defined as

$$PD_c = \sum_{t=1}^{t=T} \max(P_{ct} - C, 0) \tag{3.1}$$

This construction resembles the format of cooling degree days (CDD) and heating degree days (HDD) which take into account only the temperatures ranging below or above a certain threshold (i.e. 65 degrees Fahrenheit for CDD and HDD) to indicate energy consumption. By setting the pollution degree to zero whenever P_{ct} ranges below C, equation (3.1) removes the very low pollution concentrations that may not cause adverse health outcomes. Meanwhile, it accounts for the accumulation of environmental risk before birth if the pregnancy period is longer.

While PD contains some favorable features comparing with the traditional approach, it is likely to cause an inverse sign of coefficient of interest. Since fetuses with a longer gestational age typically weigh heavier at birth, PD may generate the spurious result that more aggregated pollution leads to heavier infants. To solve this problem, I average out PD over the time period T. Specifically, I construct an index called average pollution degree (APD) that takes the following form

$$APD_c = \frac{PD_c}{T} = \frac{\sum_{t=1}^{t=T} \max(P_{ct} - C, 0)}{T}. \tag{3.2}$$

The APD can be viewed as the weighted average pollution concentration during a time of length T. It treats all the pollution exposures above threshold C as "determinants" of fetal health while ignoring any pollution below C. The flexible choice of different thresholds offers a direct way to examine the air quality standard that would most effectively increase the average birth weight of newborns if it were to be established.

3.2.2 Empirical Model

Let y_{ijct} be the birth weight of fetus i whose mother j resides in county c and gives birth at time t. I consider the baseline ordinary least squares (OLS) model taking the following form

$$y_{ijct} = \sum_{\tau=1}^{\tau=3} (\beta_\tau APD^{\tau}_{ijct} + \boldsymbol{\delta_\tau W^{\tau}_{ijct}}) + \boldsymbol{\theta X_{ijct}} + \mu_{ijct}. \tag{3.3}$$

The key regressor in specification (3.3) is APD^{τ}_{ijct} representing the average pollution degree (APD) that fetus i experiences in trimester τ. The value of APD depends upon both maternal residential county c and the newborn's gestational age. $\boldsymbol{W^{\tau}_{ct}}$ is a vector of weather conditions (i.e., precipitation, maximum and minimum temperature) in county c during trimester τ. $\boldsymbol{X_{ijct}}$ include infant's characteristics such as gender and birth order and mother's characteristics such as age, race, Hispanic origin, education level and marital status.

A naive estimation of equation (3.3) may suffer from omitted variable bias (OVB). Specifically, unobservable components in μ_{ijct} may correlate with both air quality and fetal health. For instance, an outbreak of influenza which is usually strongly correlated with seasonal air quality may affect maternal and fetal health conditions. To avoid mistakenly attributing this health consequence to air pollution, I revise model (3.3) by including various fixed effects to capture the unobservable components in μ_{ijct}. Let s index the state where county c locates in. The revised model takes the following form

$$y_{ijcst} = \sum_{\tau=1}^{\tau=3} (\beta_\tau APD_{cst}^{\tau} + \boldsymbol{\delta_\tau W_{cst}^{\tau}}) + \boldsymbol{\theta X_{ijcst}} + \lambda_t + \eta_{ct} + \gamma_{st} + \phi_{ijt} + \varepsilon_{ijcst}. \tag{3.4}$$

In specification (3.4), λ_t is the year fixed effects capturing any long-run trend of fetal health caused by business cycles and standard of living. η_{ct} is the county-year medium income that reflects the newborn's family background and mother's nutrition intake during pregnancy.[5] γ_{st} is the state-quarter fixed effects incorporating the seasonal fluctuations of birth weight specific to each state. ϕ_{ijt} stands for the total number of mother's prenatal visits. In my preferred specification, μ_{ijcst} is clustered at county level to incorporate any correlation among the random shocks within each county. I test alternative clustering and specifications in the robustness check.

It is worth noting that the traditional approach using the average pollution instead of APD can be viewed as a special case of specification (3.4) with the threshold setting to zero. A lower threshold weighs low pollution concentrations equally as high concentrations and assumes a unit change in pollution poses the same risk to fetuses regardless of the starting point of air pollution. As the threshold gradually increases, fluctuations in low levels of pollution are neglected, and more attention is put on higher pollution concentrations. The adjustment of the threshold shows a tradeoff between noise reduction and omission of

[5]Since income starts to affect fetal health from the date of pregnancy, the year used in η_{ct} is the year of pregnancy instead of year of birth.

information. When a threshold is set too high, very little information is extracted from each pollution observation, therefore the significance level of estimates may decline.

3.3 Data

The study uses data combined from U.S. Centers for Disease Control and Prevention (CDC), Environmental Protection Agency (EPA), National Centers for Environmental Information (NCEI) and Small Area Income and Poverty Estimates (SAIPE) from Census Bureau. I consider a time span of 1991 to 2008 that has the most comprehensive information on birth occurrences, air pollution, weather and demographic features.[6] The CDC provides individual birth data and related maternal characteristics (e.g., age, race, Hispanic origin, education, marital status). I merge these births with county-week air pollution data from EPA and county-week weather data from NCEI based on the (induced) date of birth and mother's residential county.[7] Finally these data are merged with county-year medium income data from SAIPE based on mother's residential county and year of pregnancy.[8] I drop observations with missing data in fetal characteristics (i.e., birth weight, gestation, sex, total and live birth orders), maternal characteristics (age, race, Hispanic origin, education and marital status) or other control variables (i.e. weather conditions, demographic features). I also drop mothers who are younger than 15 years or older than 45 years because the infant birth weights of these mothers are highly noisy. The resulting dataset has 23,379,600 observations from 130 counties in 38 states in the CO sample, 16,127,084 observations from 86 counties in 27 states in the NO2 sample, 19,837,801 observations from 143 counties in 23 states in the O3 sample and 18.997,081 observations from 185 counties in 41 counties in the SO2 sample.

---

[6]Before 1991, EPA does not report high-frequency air pollution data. After 2008, CDC no longer reports the date of mother's last menstrual period (LMP), therefore I cannot derive the exact birth day for each infant.

[7]The infants' date of birth are induced using the reported mother's last menstrual period (LMP), gestation in whole weeks, and birth month and day of week.

[8]I use mother's year of pregnancy instead of the year of infant birth to partially account for the family background during the pregnancy period.

3.4 *Estimation Results*

3.4.1 Main Results

As discussed above, one of the biggest advantages of applying the APD approach is the flexible choice of safety thresholds. The adjustment of different thresholds can be viewed as a noise-reducing process. Assuming that the pollution varying below the threshold C do not convey explanatory power on the changes of birth outcome, I extract the most useful information from the key regressor by focusing on the pollution concentration ranging above C and omitting the noisy random fluctuations of pollution below C. The empirical estimation process starts from establishing a series of hypothetical thresholds. Based on the observed weekly pollution distributions, I use the 10th, 30th, 50th, 60th, 70th, 80th, 90th and 95th percentile as the suggested threshold. The point estimates obtained under the APD approach should be interpreted as the change of birth weight associated with 1 unit change of weekly average pollution above C during each trimester.

Table 3.3 displays the estimation results for all fetuses entering the third trimester. All regressions control for year FE, state-quarter FE, fetal and maternal characteristics, weather conditions, prenatal visits and county medium income. A cross-column comparison shows a tradeoff between noise reduction and information truncation. As Panel A suggests, fetuses are vulnerable even to very low concentrations of CO (the wrong-signed coefficient in trimester II is caused by maternal protective behaviors as discussed in Chapter 2, more details follow.). Every 1 ppm increase in CO above the 10th percentile (or 0.30 ppm as weekly average) during the first and third trimesters would significantly decrease birth weight by 5.4 and 7.7 grams, respectively. With the gradual increase of threshold, the coefficient for CO in the first trimester rises but becomes less significant due to the omission of useful information ranging between the 10th and 30th percentiles. The coefficient for CO in the second and third trimesters declines along with the significance level. The trends of both the magnitude and significance level of coefficient imply that a low level of CO does not increase noise in estimators, and even a nearly zero concentration of CO would still convey explanatory power on the birth outcome. Because fetuses are very sensitive to even extremely low concentrations of CO, any effective air quality legislation targeted at reducing CO would be expected to bring health benefits to future generations.

Table 3.3 Panel B indicates that birth weight would be expected to fall by 7.5 grams for every 10 ppb increase in weekly NO2 in the last trimester. In the preceding trimesters, fetuses do not have a strong reaction to NO2 unless its weekly concentration rises above the 10th percentile (or 7.1 ppm). Overall, the largest joint effect of NO2 during the entire period of pregnancy occurs at a threshold of about 50 percentile (15.3 ppm). Above 15.3 ppm, every 10 ppb increase in weekly NO2 during the first and third trimesters is expected to reduce infant birth weight by 4.3 and 6.5 grams, respectively.

Similarly, Table 3.3 Panel C shows that the strongest effect of O3 occurs in the first and third trimesters. When the exposed weekly O3 in these periods increases by 0.01 ppm above the 30th percentile (or 0.02 ppm), infant birth weight is expected to drop by 4.3 and 4.1 grams, respectively. The largest and most significant fluctuation of birth weight occurs at a threshold of about 60 percentile (or 0.029 ppm), above which every 0.01 ppm increase in O3 in trimesters I and III is expected to reduce birth weight by 6.6 and 6.1 grams, respectively. Above 0.029 ppm the estimated coefficients for the health impact of O3 continue to rise but with a declining significance level.

The estimation results for SO2 are shown in Table 3.3 Panel D. Not until the threshold reaches to about 95 percentile (or 13.4 ppb) do the estimated coefficients start to capture a significantly adverse health effect. Meanwhile, unlike the previous pollutants the health effect of SO2 is the strongest during the second trimester. For every 10 ppb increase in SO2 above 13.4 ppb in the second trimester, birth weight is expected to decline by roughly 8.7 grams. This result is slightly different from Lavaine and Neidell's (2014) findings on the health effect of reduced SO2 from oil refineries. In the latter study, SO2 is estimated to influence birth outcome especially during the third trimester. However, because the sample contains no information on the exact date of birth, Lavaine and Neidell (2014) reassign all the births occurring in the same month to the first day of that month. This approximation could lead to some bias especially among pre-term births delivered at the end of each month because the last several weeks for these babies prior to birth fall into the second instead of the third trimester. Admittedly, the underlying mechanism of how different air pollutants affect fetal health remains unclear to researchers. It is likely that fetuses may be more sensitive to the exposure of one pollutant but not the others during a specific period in utero.

3.4.2 Supportive Evidence

The main results presented in Tables 3.3 Panel A show surprisingly wrong-signed and significant coefficients for the health impacts of CO and SO2 in trimester II. As noted in Chapter 2, this positive sign does not necessarily indicate a beneficial impact of these pollutants. Specifically, unobservable maternal protective behaviors against air pollution may lead to the opposite sign of coefficient. While it is infeasible to include mother fixed effects to alleviate this bias, we may instead examine whether the positive sign persists when a more homogenous sample of mothers is considered.

Table 3.4 displays the estimation results for all full-term babies ($\geq$ 37 weeks and $\leq$ 42 weeks of gestation) controlling for year FE, state-quarter FE, fetal and maternal characteristics, weather conditions, prenatal visits and county medium income. The 0th percentile threshold is equivalent to the traditional approach using average pollution as the key regressor. Surprisingly, all of the previously positive and significant coefficients become insignificant, and most of them even become negatively signed. I further perform similar analysis (not shown here) with pre-term ($<$ 37 weeks and $>=$ 28 weeks) and post-term ($>$ 42 weeks) babies, respectively. Results show that the positive coefficients only occur among pre-term babies. Furthermore, the data indicate that mothers who have pre-term and extremely pre-term babies ($<$ 28 weeks) tend to visit prenatal care more frequently than those having full-term babies.[9] This supportive evidence implies that mothers with pre-term babies may

[9]Mothers with infants born between 28 and 36 weeks of gestation visit prenatal care 0.3 time per week (with a maximum being 1.75) while mothers with infants born at least 37 weeks visit prenatal care 0.29 time per week (with a maximum being 1.32).

have been informed by their prenatal care provider of the fetal health conditions, especially during the second trimester when fetal development can be easily detected. Therefore it is highly likely that mothers will take special care of their babies and avoid been exposed to environmental hazards when the perceived air pollution concentration is high.

The thresholds for CO and NO2 that induce the strongest fluctuation of birth weight are similar in Table 3.4 as in Table 3.3. As illustrated by Table 3.4, full-term infants are vulnerable to a nearly zero level of CO, indicating that the ideal environmental for healthy fetal development would be one with abstinence of CO. Compared with pre-term and post-term babies, full-term babies are slightly more sensitive to NO2 in the first trimester. Nevertheless, the largest joint effect of NO2 over three trimesters is still obtained when the 50th percentile (15.3 ppm) is used.

Unlike CO and NO2, the fetal sensitivity to different thresholds of O3 and SO2 varies across full-term and non-full-term infants. For O3, no significantly negative effect can be measured even if the threshold increases to the 95th percentile, although the magnitude of coefficients substantially exacerbates with higher thresholds. For SO2, however, full-term infants seem to respond to a lower threshold than the entire population. Above the 50th percentile (4.12 ppb), every 10 ppb increase in SO2 in trimester II will decrease birth weight of full-term babies by 4.5 grams. This estimated effect of SO2 becomes higher with the same significance level if we consider a threshold at the 80th percentile (8.00 ppb). The entire population, on the other hand, does not show a strong responsiveness to SO2 until the threshold increases to the 95th percentile (13.40 ppb).

3.4.3 Robustness Checks

This section provides robustness check for the estimation results obtained in Tables 3.3 and 3.4. Using the APD approach, we are able to find two thresholds for each air pollutant. The first is the one at which fetuses just start to experience a statistically significant reduction of birth weight. The second is the one inducing the greatest and most significant joint impact over three trimesters. Table 3.5 displays the values of these thresholds. In the following robustness check and the heterogeneous analysis, I will focus on the second threshold because it is the one associated with the highest fetal sensitivity level. For each of these thresholds, I test whether the estimates will substantially differ across alternative specifications using the entire sample and the full-term births, respectively.

I first consider the assumption underlying clustered standard errors. In states like California and New Jersey where a relatively high proportion of counties is included in the study, there could exist considerable cross-county correlations in the random shock of birth weight. For instance, an influenza outbreak occurring in one county may also affect residents in surrounding counties. Thus, broader clusters of standard errors allowing for possible dependency of health shocks within state will be more appropriate. Table 3.6 columns (1) and (5) present the estimation results for all infants and full-term infants only, respectively, using state-level clustered standard errors. All other control variables remain the same as those in the preferred specification. This modification causes a slight adjustment to the magnitude of standard errors, but the significance levels of estimators stay highly consistent with those in the preferred specification.

Second, Figure 2.1 illustrates a strong seasonality of infant birth weight. This seasonality is handled in the preferred specification by including the state-quarter fixed effects. Nevertheless, there might be concerns that different months within a quarter convey distinct seasonal patterns that are unobservable to researchers. For instance, the peak flu season in most states typically starts from December and lasts till February. Therefore, infants born in different months within same season may experience differential health risks due to unobservable shocks. To attenuate this issue, I replace the state-quarter fixed effects by state-month fixed effects. Table 3.6 columns (2) and (6) contain the estimated effects of air pollution with the modified model. There are moderate adjustments of coefficients because of the possible collinearity between month and air pollution. In other words, the fixed effect parameters now absorb some of the impact of air pollution that should be captured by the coefficients of interest. Nevertheless, most of our conclusions still hold.

Third, one may concern about the health effect of omitted parental attributes and genetic factors. In particular, biological and medical studies have shown that paternal genes and certain characteristics may contribute to birth outcomes (Little and Sing (1987); Griffiths et al. (2006); Taiwo and Akinde (2012)). Even though the paternal impact may not be as strong as maternal impact, there is a potential threat of bias if the omitted paternal attribute is correlated with air pollution. Unfortunately, the sample contain limited information on the paternal side, and I do not observe what have been commonly regarded as strong contributing factors of infant birth weight (e.g. father's height and weight). Based on the available information, I include father's age group (i.e., under 15, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, and 55-98 years), race and Hispanic origin as additional controls. Simply regressing the fetal exposure to air pollution on these paternal characteristics generates statistically significant estimates, indicating that including these paternal variables serves as a valid robustness check for the omitted variable bias. Table 3.6 columns (3) and (7) show the estimation results. For all pollutants, adding paternal attributes leads to very little change in the estimated coefficients.

Lastly, I consider the possibility that omitted weather variable may bias the estimates. For instance, sunshine and wind can be correlated with both air pollution concentrations and maternal health condition. To account for this issue, I collect data from NCEI at NOAA on the daily average hours of sunshine and wind speed over three trimesters and combine them with the birth sample. Table 3.6 columns (4) and (8) displays the estimation results including these additional weather variables. The missing data from sunshine and wind speed reduce the sample size by 25 to 30 percent. Using all births in the regression, I find that the estimated impacts of CO, NO2 and O3 during the first trimester are substantially less significant when including sunshine and wind speed. In contrast, the estimated impacts of all pollutants on full-term infants stay highly consistent with those in the preferred specification. This disparity suggests that the negative effects of air pollution on pre- and post-term infants are largely affected by daily sunshine and wind speed. Here I provide two possible explanations for the underlying mechanism. For CO and other gases lighter than air, a high wind speed will help dispersing their concentrations, and at the same time, induce more pregnant women (especially those with high probability of having pre-term babies) to avoid outdoor activities. Therefore further controlling for wind speed will identify the health effect of behavioral adjustments on windy days, which is originally estimated as the health effect of low CO concentrations. On the other hand, O3 is formed by the chemical reactions between nitrogen oxides (NOx) and other pollutants (e.g., organic compound (VOC)) with the help of sunshine. When omitting hours of sunshine, the negative impact of hazardous ozone on sunny days will likely be offset by the positive health impact produced by sunlight (e.g., more vitamin D produced), driving the estimated coefficients towards zero.

3.4.4 Heterogeneity

Previous literature (e.g. Currie et al. 2009) suggests that infants with heterogeneous mothers characteristics (e.g., smoker vs. non-smoker, younger vs. older) may experience substantially different risks of infant mortality with the same degree of postnatal exposure to air pollution. In this section, I examine and discuss the heterogeneous impacts of prenatal exposure to air pollution on fetal development. To avoid selection bias due to unobservable attributes that may correlate with heterogeneous socioeconomic, I am going to focus on full-term births only.

I consider two groups of heterogeneity. First, fetuses with distinct characteristics may have different sensitivity levels to air pollution. For example, Ellis and He (2014) analyze self-reported fetal activity data from over six thousands mothers in the United States and Canada. Their results suggest that male fetuses are about 10 percent more active in the womb than female fetuses. Jedrychowski et al. (2009) study a sample of homogeneous pregnant women[10] and confirm that males respond more strongly to prenatal PM2.5 concentrations. In addition, Sanders and Stoeckel's (2011) examine the shock of gender ratio followed by the 1970 Clean Air Act Amendment (CAAA) and conclude that the dramatic improvement of air quality following the CAAA benefit male fetuses more than female fetuses. This paper

[10]The women eligible for the study are restricted to non-smokers, with singleton pregnancy, without infections and aging between 18 and 35 years old.

shows supplementary evidence on the gender disparities in the sensitivity levels to CO, NO2, O3 and SO2.

Table 3.7 columns (1) and (2) compare the estimated coefficients for the impact of air pollution by gender using the most sensitive threshold indicated by Table 3.5. Overall, male fetuses are more strongly affected by O3 and SO2 but less strongly affected by NO2. This finding is partially consistent with previous research. However, with the presence of son preference among some social groups (e.g., Asian immigrants) even in developed countries, the estimated impact of air pollution on male fetuses would more likely be biased down by the intensive health care provided on male rather than female fetuses ((Pabilonia and Ward-Batts (2007); Almond and Edlund (2007); Almond et al. (2009); Abrevaya (2009); Lhila and Simon (2008)). This gender heterogeneity in the responsiveness to air pollution exposure offers important policy implications. On one hand, any effective air quality legislation may have an impact on the gender ratio, affecting marriage, productivity, violence and crime rate in the long run (Angrist 2006; Hesketh and Xing 2006; Dyson 2012). On the other, changes in air quality may bring asymmetric welfare gains between male and female, requesting further reallocations of resource.

Besides gender disparity, there is extensive evidence that investment on children is negatively correlated with family size (Becker (1981); Black et al. (2005); Black et al. (2010); Mogstad and Wiswall (2010)). If the quality of health care on additional babies declines, one would expect to find a stronger effect of air pollution on higher order births. Table 3.7 columns (3) and (4) shows the estimation results for first order births and subsequent births. In the latter group, I include all the non-first order births from the second to up to the eighth. Overall, subsequent births are more likely to experience a large fluctuation of birth weight when exposed to higher-than-threshold pollution concentrations. This result, however, cannot totally rule out the possibility that subsequent babies are biologically weaker in health condition than firstborn babies, making them more vulnerable to environmental harm. As suggested by Modin (2002), later-born infants, in general, have a higher mortality rate than firstborn infants.

In addition to fetal characteristics, I consider the heterogeneous impact of air pollution across maternal characteristics. Currie et al. (2009) study the effect of air pollution on health at birth between younger (less than 19 years) and older (more than 35 years) mothers. The result shows that older mothers are highly sensitive to CO during the first trimester. Table 3.7 columns (5) and (6) compare the estimated coefficients for fetuses with mothers younger than 30 years and those with mothers older than 30 years. For all pollutants except for CO, older mothers tend to show slightly stronger resistance than younger mothers. Since women choosing to give birth older than 30 years generally have higher levels of income and education, fetuses with older mothers are likely to receive more parental investment than those with younger mothers, conditional on birth order. Nevertheless, fetuses with older mothers are more strongly affected by CO than those with younger mothers. Because even extremely low levels of CO are hazardous, extensive health care provided to fetuses by older parents can hardly offset the health risk of CO.

The second group of heterogeneity arises from maternal behaviors over the pregnancy period. Currie et al. (2009) document that the health risk of air pollution can substantially exacerbate when the mother smokes during pregnancy. This paper provides supplementary evidence to the joint effect of smoking and air pollution. Specifically, I estimate how much additional impact of air pollution a fetuses may experience if the mother is a smoker by interacting maternal smoking behaviors with air pollution above the most sensitive threshold. Because the sample only contains information on whether a mother answers "yes" to the question of whether she smokes during pregnancy, but not on how many packs of cigarette she consumes during each trimester, I cannot examine the intensity of smoking behavior and the corresponding health impact over time. Moreover, if a mother is well informed about the poor health condition of fetus (e.g. likely pre-term), she may choose to stop consuming cigarette especially when the baby is about to born. As discussed earlier, these behavioral adjustments may lead to an opposite sign of coefficient of interest. To avoid this bias, I estimate the impact of smoking on fetuses using full-term births only whose mothers are less likely to concern about potential adverse birth outcomes. Column (1) of Tables 3.8

displays the estimated effect of CO on smoking mothers as compared with non-smoking mothers. Being a smoking mother is expected to lower the infant birth weight by roughly 169 to 198 grams, almost 5 percent of the total birth weight for an average baby. The key toxic chemicals produced by tobacco include CO, NOx (a generic term for NO and NO2), particulate matter, hydrogen cyanide, and benzene, all of which are highly dangerous for human health. For instance, Klepeis et al. (1999) estimate that the average concentration of CO during a cigar social event can be comparable or even higher than that from a busy traffic highway.[11] Jaffe and Chavasse (1999) also demonstrate the average CO concentrations produced by a cigarette can be 10 times as much as that produced by a clean vehicle. Because a cigarette contains multiple harmful chemicals, their joint impacts on fetal health will be considerably high.

In regards to the joint effects of smoking and air pollution, I find that CO is twice as harmful to smoking mothers as to non-smoking mothers, while NO2 is roughly 1.4 times more harmful to smoking mothers. These two pollutants are the key components in cigarette, therefore a regularly smoking mother would experiences a slightly higher effect of ambience air pollution concentrations than non-smoking mothers. In contrast, the impact of SO2 on smoking mothers can be 12 times as much as that on non-smoking mothers. Yun (1975) conducts an experiment on the effect of CO-SO2 mixed gases on rats. The study finds strong evidence that the mixed gases negatively affect both the gestation period and birth weight of rats.

One exemption is the impact of O3 between smoking and non-smoking mothers. As Table 3.8 columns (3) shows, although maternal smoking has an expected negative impact on birth weight, the joint effect of smoking and O3 concentrations has a positive sign in all trimesters. Clearly, more research needs to be done to clarify the biological mechanism. But from chemical perspective, it is possible that O3 as a powerful oxidant may react with some

[11]The authors conduct two research, both of which suggest that the average CO from cigar-smoking events can be multiple times higher than the ambient CO. The former is comparable with the CO concentration in urban arterial highway in San Francisco Bay Area.

compounds in cigarette when the temperature is high (e.g. when the mother is smoking).

Another behavioral factor that contributes to poor fetal health is alcohol consumption. Mills et al. (1984) provide evidence that regularly drinking mothers impose a substantially higher risk of growth retardation to fetuses. However, more recent evidence shows that even a small dose of prenatal alcohol consumption may lead to adverse alcoholic effect (e.g. growth deficits) on fetuses, although the magnitude of the effect would not be as large as fetal alcohol syndrome (FAS) caused by a large dose of exposure (Larkby and Day (1997); Gabriel et al. (1998)). Ornoy and Ergaz (2010) confirm that fetal growth disturbances can occur even if the maternal consumption of alcohol is low and suggest the abstinence of alcohol during pregnancy period. This analysis provides further evidence that air pollution could aggravate the health concerns of fetuses whose mother drinks during pregnancy.[12] My sample contains about 1 percent mothers who report consuming alcohol during pregnancy, among whom roughly 70 percent are light drinkers. Since the sample size for alcohol consumers is rather small, I again interact the drinking behavior dummy with air pollution above the most sensitive threshold to estimate the additional impact of air pollution experienced by drinking mothers. All the other covariates are kept unchanged. Table 3.9 displays the estimated effect of consuming alcohol and the joint effect of air pollution and alcohol consumption for full-term births. As expected, alcohol consumption strongly affects fetal health. Being an light alcohol consumer statistically significantly drives infant birth weight down by 58 to 134 grams. As compared to nondrinker and light drinkers, being a heavily drinking mother would further reduce infant birth weight by 137 to 235 grams. In terms of the joint effects of air pollution and maternal alcohol consumption, I estimate that the health risk of CO, NO2 and SO2 are two to twelve times larger for light drinking mothers than for non-drinkers, and four to seventeen times larger for heavy drinkers than for light and non-drinkers. Again the coefficients for O3 are all positive, indicating a complex reaction may occur between O3 and

[12] I cannot fully exclude the possibility, though, that there might exist unobserved features for drinking mothers and they may correlate with both birth weight and air pollution.

ethanol.[13]

3.5 Conclusion

This chapter supplements the growing literature analyzing the causal impact of maternal exposure to air pollution on fetal health measured in terms of birth weight. By setting and adjusting a hypothetical "safety threshold" on four criteria air pollutants CO, NO2, O3 and SO2, I find that CO consistently and negatively affects birth weight even when its concentration approaches zero. The impact of CO, however, is not uniform across each period of gestation. The exposure to CO over the first and third trimesters impose substantially higher risks on fetuses than that over the second trimester. On the other hand, NO2, O3 and SO2 influence fetal health only if their concentrations rise above certain levels. The estimated weekly average safety levels for these pollutants are roughly 15.3 ppb for NO2, 0.02 ppm for O3 and 13.4 ppb for SO2.

Furthermore, this study analyzes the heterogeneous impacts of pollution on birth weight across fetal and maternal characteristics. I find disparities across gender and birth order in the responsiveness to air pollution. Meanwhile, I find that maternal smoking and drinking behaviors during pregnancy would jointly act with air pollution in affecting fetal health. These heterogeneities provide important implications. First, high levels of air pollution may potentially lead to a change in the gender ratio in the long run. Second, air quality strategies may generate asymmetric welfare gains across different socioeconomic groups. Third, fetuses with mothers who smoke or drink during pregnancy may particularly suffer from low birth weight and are generally at higher risks of being affected by environmental hazards.

---

[13] Alcohol (or ethanol) is a chemical that can easily be oxidized. Although there is no evidence that alcohol and O3 may have chemical reaction in room temperature, it is unknown if they can react with the presence of certain enzyme. (For instance, Waters et al. (1976) shows that under careful control, there could be "rapid and complete conversion of alcohols to ketones upon ozonation.")

Reference

Jason Abrevaya. Are There Missing Girls in the United States? Evidence from Birth Data. *American Economic Journal: Applied Economics*, 1(2):1–34, 2009.

Douglas Almond and Lena Edlund. TriverscWillard at Birth and One Year: Evidence from US Natality Data 1983-2001. *Proceedings of the Royal Society B: Biological Sciences*, 274 (1624):2491–2496, 2007.

Douglas Almond, Lena Edlund, and Kevin Milligan. O Sister, Where art Thou? The Role of Son Preference and Sex Choice: Evidence from Immigrants to Canada. NBER Working Paper No. 15391, 2009.

Josh Angrist. Consequences of Imbalanced Sex Ratios: Evidence from American's Second Generation. NBER working paper No. 8042, 2000.

Gary S. Becker. *A Treatise on the Family*. Havard Univ. Press, Cambridge, MA, 1981.

Sandra E. Black, Paul J. Devereux, and Kjell G. Salvanes. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*, 120(2):669–700, 2005.

Sandra E. Black, Paul J. Devereux, and Kjell G. Salvanes. Small Family, Smart Family? Family Size and the IQ Scores of Young Men. *Journal of Human Resources*, 45(1):33–58, 2010.

Janet Currie, Matthew Neidell, and Johannes F. Schmieder. Air Pollution and Infant Health: Lessons from New Jersey. *Journal of Health Economics*, 28:688–703, 2009.

Tim Dyson. Causes and Consequences of Skewed Sex Ratios. *The Annual Review of Sociology*, 38:443–461, 2012.

Lee Ellis and Ping He. Sex Differences in Fetal Activity and Childhood Hyperactivity. *Research Journal of Developmental Biology*, 2014. doi: 10.7243/2055-4796-1-1.

Environmental Protection Agency (EPA). Air Quality Index: A Guide to Air Quality and Your Health, 2014.

Kara Gabriel, Candace Hofmann, Maria Glavas, and Joanne Weinberg. The Hormonal Effects of Alcohol Use on the Mother and Fetus. *Alcohol Health and Research World*, 22 (3):170–177, 1998.

Therese Hesketh and Zhu Wei Xing. Abnormal Sex Ratios in Human Populations: Causes and Consequences. *Proceedings of the National Academy of Sciences of the United States of America*, 103(56):13271–13275, 2006.

Dan Jaffe and Laurie Chavasse. Comparing the CO Content of Cigarette Smoke and Auto Exhaust Using Gas Chromatography. *Journal of College Science Teaching*, pages 172–176, 1999.

Wieslaw Jedrychowski, Frederica Perera, Dorota Mrozek-Budzyn, Elzbieta Mroz, Elzbieta Flak, Jack D. Spengler, Susan Edwards, Ryszard Jacek, Irena Kaim, and Zbigniew Skolicki. Gender Differences in Fetal Growth of Newborns Exposed Prenatally to Airborne Fine Particulat Matter. *Environmental Research*, 109:447–456, 2009.

Neil E. Klepeis, Wayne R. Ott, and James L. Repace. The Effect of Cigar Smoking on Indoor Levels of Carbon Monoxide and Particles. *Journal of Exposure Analysis and Environmental Epidemiology*, 9:622–635, 1999.

Cynthia Larkby and Nancy Day. The Effects of Prenatal Alcohol Exposure. *Alcohol Health and Research World*, 21(3):192–198, 1997.

Emmanuelle Lavaine and Matthew Neidell. Energy production and health externalities: Evidence from Oil Refinery Strikes in France. NBER Working Paper No. 18974, 2014.

Aparna Lhila and Kosali I. Simon. Prenatal Health Investment Decisions: Does the Childs Sex Matter? *Demography*, 45(4):885–905, 2008.

R E Little and C F Sing. Genetic and Environmental Influences on Human Birth Weight. *American Journal of Human Genetics*, 40:512–526, 1987.

James L. Mills, Barry I. Graubard, Ernest E. Harley, George G. Rhoads, and Heinz W. Berendes. Maternal Alcohol Consumption and Birth Weight: How Much Drinking During Pregnancy Is Safe? *The Journal of the American Medical Association*, 252(14):1875–1879, 1984.

Bitte Modin. Birth Order and Mortality: A Life-long Follow-up of 14,200 Boys and Girls Born in Early 20th Century Sweden. *Social Science and Medicine*, 54:1051–1064, 2002.

Magne Mogstad and Matthew Wiswall. Testing the QuantitycQuality Model of Fertility: Estimation using Unrestricted Family Size Models. *Quantitative Economics*, 7:157–192, 2016.

Asher Ornoy and Zivanit Ergaz. Alcohol Abuse in Pregnant Women: Effects on the Fetus and Newborn, Mode of Action and Maternal Treatment. *International Journal of Environmental Research and Public Health*, 7:364–379, 2010.

Sabrina Wulff Pabilonia and Jennifer Ward-Batts. The Effect of Child Gender on Parents' Labor Supply: An Examination of Natives, Immigrants, and their Children. *American Economic Review*, 97(2):402–406, 2007.

Nicholas J. Sanders and Charles F. Stoecker. Where Have All the Young Men Gone? Using Gender Ratios to Measure Fetal Death Rates. NBER working paper No. 17434, 2011.

William L. Waters, Anthony J. Rollin, Cindy M. Bardwell, Jeffrey A. Schneider, and Thomas W. Aanerud. Oxidation of Secondary Alcohols with Ozone. *Journal of Organic Chemistry*, 41:889–891, 1976.

DR Yun. The Effects of Chronic Exposure to CO-SO2 Mixed Gases on the Gestation Pattern of the Rats. *Seoul Journal of Medicine*, 16(14):242–246, 1975.

LIST OF FIGURES IN CHAPTER 3

LIST OF TABLES IN CHAPTER 3

Table 3.1: Air Quality Trends from 1990 to 2008

Pollutants	Reduction of Mean	Reduction of 10th percentile	Reduction of 90th percentile
Carbon monoxide (CO)	73.3 %	68.6 %	75 %
Nitrogen dioxide (NO2)	43.4 %	32.7 %	49.1 %
Ozone (O3)	16 %	7.1 %	19.1 %
Sulfur dioxide (SO2)	50.9 %	34.6 %	46.0 %

Data source: Environmental Protection Agency (EPA) versus the European Union.

Table 3.2: Air quality standards: US versus Europe

Pollutants	US		Europe		
	Timing	Standard	Timing	Standard	Standard (converted)
Carbon monoxide (CO)	8-hour	9 ppm [a]	8-hour	10 mg/m^3	8.7 ppm
	1-hour	35 ppm [a]			
Nitrogen dioxide (NO2)	1-hour	100 ppb [a]	1-hour	200 $\mu g/m^3$	106.4 ppb
	1-year	53 ppb [a b]	1-year	40 $\mu g/m^3$	21.3 ppb
Ozone (O3)	8-hour	0.075 ppm [a b]	8-hour	120 $\mu g/m^3$	0.06 ppm
Sulfur dioxide (SO2)	1-hour [a]	75 ppb	1-hour	350 $\mu g/m^3$	133.6 ppb
	3-hour	0.5 ppm [b]	24-hour	125 $\mu g/m^3$	0.047 ppm
Lead	3-month	0.15 $\mu g/m^3$ [a b]	1-year	0.5 $\mu g/m^3$	
Particulate matter (PM2.5)	1-year	12 $\mu g/m^3$ [a]	1-year	20 $\mu g/m^3$ [c]	
	1-year	15 $\mu g/m^3$ [b]	1-year	25 $\mu g/m^3$	
	24-hour	35 $\mu g/m^3$ [a b]			
Particulate matter (PM10)	24-hour	150 $\mu g/m^3$ [a b]	24-hour	50 $\mu g/m^3$	

This table shows the air quality standards established by the United States Environmental Protection Agency (EPA) versus the European Union.

a. - primary standard; b. - secondary standard; c. - average exposure indicator (AEI)

Table 3.3: Fetal Sensitivity to Different Air Pollution Thresholds

	(1) 10 %	(2) 30 %	(3) 50 %	(4) 60 %	(5) 70 %	(6) 80 %	(7) 90 %	(8) 95 %
Panel A: CO in ppm								
Trimester I	-5.410*	-5.710*	-6.198*	-6.504*	-6.879*	-7.174	-7.442	-7.835
	(2.825)	(2.967)	(3.293)	(3.597)	(3.984)	(4.559)	(5.861)	(8.113)
Trimester II	5.954***	5.398***	4.504***	3.947**	3.360**	2.755	1.893	1.656
	(1.711)	(1.586)	(1.541)	(1.558)	(1.648)	(1.896)	(2.855)	(4.795)
Trimester III	-7.658**	-6.238**	-4.838	-4.159	-3.625	-3.159	-1.843	0.752
	(2.932)	(2.996)	(3.258)	(3.510)	(3.854)	(4.414)	(5.768)	(7.867)
Observations	23379600	23379600	23379600	23379600	23379600	23379600	23379600	23379600
R^2	0.055	0.055	0.055	0.055	0.055	0.055	0.055	0.055
Panel B: NO2 in 10 ppb								
Trimester I	-2.604	-3.407*	-4.333**	-4.739**	-5.001**	-5.232*	-5.178	-4.785
	(1.708)	(1.782)	(1.944)	(2.102)	(2.353)	(2.792)	(3.717)	(4.775)
Trimester II	0.631	0.623	0.787	0.869	0.851	0.807	0.842	0.972
	(1.259)	(1.246)	(1.239)	(1.269)	(1.348)	(1.447)	(1.632)	(1.970)
Trimester III	-7.541***	-7.040***	-6.465***	-6.141***	-5.725**	-5.079*	-3.946	-2.658
	(1.581)	(1.609)	(1.749)	(1.907)	(2.221)	(2.753)	(3.667)	(4.742)
Observations	16127084	16127084	16127084	16127084	16127084	16127084	16127084	16127084
R^2	0.055	0.055	0.055	0.055	0.055	0.055	0.055	0.055
Panel C: O3 in 0.01 ppm								
Trimester I	-2.959*	-4.306**	-5.797**	-6.631**	-7.292*	-7.902*	-9.471	-11.402
	(1.762)	(2.073)	(2.630)	(3.065)	(3.696)	(4.573)	(7.277)	(9.941)
Trimester II	1.811	2.698	2.795	2.729	2.645	2.468	1.228	-0.180
	(1.655)	(1.870)	(2.142)	(2.411)	(2.811)	(3.406)	(5.632)	(8.017)
Trimester III	-3.127	-4.135*	-5.328*	-6.075*	-6.811	-7.749	-10.868	-15.125
	(2.038)	(2.361)	(3.001)	(3.505)	(4.235)	(5.178)	(7.974)	(10.820)
Observations	19837801	19837801	19837801	19837801	19837801	19837801	19837801	19837801
R^2	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.053
Panel D: SO2 in 10 ppb								
Trimester I	0.058	-0.824	-2.261	-2.934	-3.566	-4.178	-4.179	-3.647
	(3.447)	(3.345)	(3.557)	(3.839)	(4.340)	(5.206)	(6.917)	(8.502)
Trimester II	2.383	0.958	-0.854	-1.768	-2.823	-4.096	-6.398	-8.703*
	(3.050)	(2.873)	(2.869)	(2.924)	(3.048)	(3.316)	(4.044)	(5.127)
Trimester III	1.247	0.676	0.142	0.049	0.197	0.735	2.129	2.444
	(3.042)	(2.979)	(3.162)	(3.394)	(3.738)	(4.240)	(5.112)	(6.106)
Observations	18997081	18997081	18997081	18997081	18997081	18997081	18997081	18997081
R^2	0.058	0.058	0.058	0.058	0.058	0.058	0.058	0.058

This table shows the estimation results of specification (3.4) using the 10th, 30th, 50th, 60th, 70th, 80th, 90th, and 95th percentile of weekly pollution distribution as the hypothetical threshold C. Each column controls for year FE, state-quarter FE, fetal and maternal characteristics, county medium income, weather conditions and number of prenatal visits during pregnancy. Robust standard errors clustered by county in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Fetal Sensitivity to Different Air Pollution Thresholds - Full-Term Infants Only

	(1) 0 %	(2) 10 %	(3) 30 %	(4) 50 %	(5) 60 %	(6) 70 %	(7) 80 %	(8) 90 %	(9) 95 %
Panel A: CO in ppm									
Trimester I	-5.942**	-5.921**	-5.913**	-6.242**	-6.451*	-6.796*	-7.176*	-7.691	-8.621
	(2.637)	(2.650)	(2.755)	(3.039)	(3.296)	(3.604)	(4.046)	(5.077)	(7.048)
Trimester II	-0.307	-0.310	-0.224	-0.281	-0.491	-0.778	-1.158	-1.836	-2.403
	(1.493)	(1.477)	(1.401)	(1.380)	(1.425)	(1.546)	(1.855)	(2.842)	(4.737)
Trimester III	-8.514***	-8.346***	-7.716***	-7.207**	-7.011**	-7.009**	-7.199*	-7.265	-6.885
	(2.520)	(2.519)	(2.614)	(2.895)	(3.160)	(3.511)	(4.063)	(5.338)	(7.405)
Observations	20051920	20051920	20051920	20051920	20051920	20051920	20051920	20051920	20051920
R^2	0.064	0.064	0.064	0.064	0.064	0.064	0.064	0.064	0.064
Panel B: NO2 in 10 ppb									
Trimester I	-3.184**	-3.349**	-3.898**	-4.559***	-4.844***	-4.991**	-5.117**	-5.079	-4.929
	(1.417)	(1.432)	(1.506)	(1.652)	(1.809)	(2.059)	(2.498)	(3.408)	(4.401)
Trimester II	-1.204	-1.224	-1.138	-0.956	-0.907	-0.944	-0.957	-0.956	-0.975
	(1.010)	(1.010)	(0.984)	(0.963)	(0.982)	(1.035)	(1.114)	(1.288)	(1.667)
Trimester III	-6.470***	-6.480***	-6.334***	-6.109***	-5.947***	-5.745**	-5.509**	-5.092	-4.570
	(1.623)	(1.637)	(1.705)	(1.874)	(2.034)	(2.308)	(2.757)	(3.536)	(4.461)
Observations	13821763	13821763	13821763	13821763	13821763	13821763	13821763	13821763	13821763
R^2	0.064	0.064	0.064	0.064	0.064	0.063	0.063	0.063	0.063
Panel C: O3 in 0.01 ppm									
Trimester I	-1.706	-2.031	-2.514	-3.223	-3.734	-4.173	-4.734	-6.654	-8.936
	(1.556)	(1.683)	(2.002)	(2.585)	(3.040)	(3.677)	(4.514)	(6.959)	(9.318)
Trimester II	1.104	1.399	1.367	0.728	0.252	-0.244	-0.850	-3.211	-5.956
	(1.429)	(1.484)	(1.639)	(1.913)	(2.153)	(2.530)	(3.141)	(5.358)	(7.708)
Trimester III	-2.085	-2.541	-3.253	-4.207	-4.868	-5.617	-6.640	-10.242	-14.932
	(1.959)	(2.086)	(2.416)	(3.013)	(3.500)	(4.178)	(5.029)	(7.520)	(9.944)
Observations	16981884	16981884	16981884	16981884	16981884	16981884	16981884	16981884	16981884
R^2	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062
Panel D: SO2 in 10 ppb									
Trimester I	0.763	0.521	-0.390	-1.623	-2.275	-3.015	-3.794	-4.051	-3.774
	(3.055)	(3.045)	(2.947)	(3.104)	(3.321)	(3.721)	(4.447)	(5.961)	(7.448)
Trimester II	-2.027	-2.257	-3.356	-4.528*	-4.845*	-5.093*	-5.326*	-5.847	-6.976
	(2.634)	(2.597)	(2.441)	(2.471)	(2.577)	(2.769)	(3.080)	(3.763)	(4.671)
Trimester III	2.402	2.288	1.118	-0.099	-0.659	-1.099	-1.407	-1.761	-2.725
	(2.819)	(2.803)	(2.669)	(2.812)	(3.008)	(3.295)	(3.718)	(4.520)	(5.416)
Observations	16276001	16276001	16276001	16276001	16276001	16276001	16276001	16276001	16276001
R^2	0.066	0.066	0.066	0.066	0.066	0.066	0.066	0.066	0.066

Notes: This table shows the estimation results of specification (3.4) using the 0th (traditional approach), 10th, 30th, 50th, 60th, 70th, 80th, 90th, and 95th percentile of weekly pollution distribution as the hypothetical threshold C. Each column controls for year FE, state-quarter FE, fetal and maternal characteristics, county medium income, weather conditions and number of prenatal visits during pregnancy. Standard errors clustered by state in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Proposed air pollution thresholds

Pollutant	Averaging Time	All infants		Full-term infants	
		Starting level to affect health	Most sensitive level	Starting level to affect health	Most sensitive level
CO	Weekly average	0 ppm	0 ppm	0 ppm	0 ppm
NO2	Weekly average	0 ppb	15.34 ppb	0 ppb	15.34 ppb
O3	Weekly average	0.013 ppm	0.029 ppm	-	0.045 ppm
SO2	Weekly average	13.40 ppb	13.40 ppb	4.12 ppb	8.00 ppb

Notes: This table displays the two thresholds obtained in section 3.4.1 and 3.4.2.

Table 3.6: Robustness Checks

	All infants				Full-Term infants only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: CO in ppm								
Trimester I	-5.347**	-4.858*	-4.917*	-4.557	-5.942**	-5.554**	-5.363**	-5.879**
	(2.613)	(2.839)	(2.718)	(3.231)	(2.494)	(2.687)	(2.552)	(2.966)
Trimester II	6.053***	8.425***	6.193***	5.384***	-0.307	0.323	-0.186	0.282
	(1.500)	(1.846)	(1.698)	(1.835)	(1.508)	(1.505)	(1.449)	(1.745)
Trimester III	-7.973***	-9.983***	-7.568***	-10.347***	-8.514***	-9.103***	-8.003***	-9.374***
	(2.930)	(2.834)	(2.849)	(3.290)	(1.996)	(2.535)	(2.417)	(2.696)
Observations	23379600	23379600	23379600	17122758	20051920	20051920	20051920	14682650
R^2	0.055	0.055	0.057	0.057	0.064	0.064	0.066	0.065
Panel B: NO2 in 10 ppb								
Trimester I	-4.333***	-4.299**	-4.146**	-2.902	-4.559***	-4.524***	-4.260***	-4.045**
	(1.382)	(1.985)	(1.886)	(2.012)	(1.041)	(1.623)	(1.591)	(1.715)
Trimester II	0.787	3.711***	0.811	-0.140	-0.956	0.409	-0.925	-0.304
	(1.028)	(1.331)	(1.225)	(1.486)	(0.748)	(1.030)	(0.962)	(1.129)
Trimester III	-6.465***	-9.065***	-6.372***	-6.722***	-6.109***	-7.322***	-5.921***	-6.195***
	(0.819)	(1.701)	(1.672)	(2.116)	(1.064)	(1.874)	(1.795)	(1.961)
Observations	16127084	16127084	16127084	11518549	13821763	13821763	13821763	9867557
R^2	0.055	0.056	0.057	0.057	0.064	0.064	0.066	0.065
Panel C: O3 in 0.01 ppm								
Trimester I	-6.631*	-7.515**	-7.461**	-3.691	-8.936	-10.546	-10.433	-4.943
	(3.689)	(3.111)	(3.018)	(3.088)	(8.791)	(9.520)	(9.223)	(9.323)
Trimester II	2.729	3.022	2.078	0.888	-5.956	-3.614	-7.177	-3.860
	(2.555)	(2.298)	(2.388)	(2.490)	(6.561)	(7.476)	(7.818)	(7.352)
Trimester III	-6.075	-5.358	-6.965**	-4.703	-14.932	-16.124	-16.513*	-9.963
	(4.768)	(3.720)	(3.466)	(3.164)	(9.720)	(9.980)	(9.913)	(9.175)
Observations	19837801	19837801	19837801	14755983	16981884	16981884	16981884	12619709
R^2	0.053	0.053	0.055	0.055	0.062	0.062	0.064	0.063
Panel D: SO2 in 10 ppb								
Trimester I	-3.647	-6.202	-3.699	-1.574	-3.794	-5.481	-3.705	-3.627
	(8.358)	(8.818)	(8.689)	(8.492)	(4.109)	(4.736)	(4.494)	(4.497)
Trimester II	-8.703*	-7.330	-8.476*	-5.974	-5.326	-3.066	-5.179*	-4.877
	(5.032)	(5.044)	(5.093)	(5.449)	(3.429)	(2.841)	(3.080)	(3.287)
Trimester III	2.444	1.018	2.514	1.001	-1.407	-2.926	-1.484	-2.713
	(5.500)	(6.152)	(6.215)	(5.991)	(3.242)	(3.883)	(3.731)	(3.629)
Observations	18997081	18997081	18997081	13918585	16276001	16276001	16276001	11920840
R^2	0.058	0.058	0.059	0.059	0.066	0.066	0.068	0.067

Notes: This table shows the results of robustness checks. An APD approach is applied with the most sensitive threshold C as shown in Table 3.5. Columns (1) to (4) include all fetuses at least 28 weeks of gestation. Columns (5) to (8) include only full-term babies between 37 and 42 weeks of gestation. Columns (1) and (5) cluster standard errors at state level, while all the other columns cluster standard errors at county level. Columns (2) and (6) replace state-quarter FE in the preferred specification by state-month FE. Columns (3) and (7) additionally control for father's characteristics, including age group (i.e., under 15, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, and 55-98 years), race and Hispanic origin. Columns (4) and (8) include average daily hours of sunshine and wind speed in the regression.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Impact of Air Pollution across Fetal and Maternal Characteristics - Full-Term Only

	Gender		Birth order		Maternal age	
	(1)	(2)	(3)	(4)	(5)	(6)
	Male	Female	First	Subsequent	< 30 yrs	≥ 30 yrs
Panel A: CO in ppm						
Trimester I	-5.441**	-6.430**	-4.252*	-6.655**	-4.735*	-7.064**
	(2.680)	(2.709)	(2.261)	(2.919)	(2.440)	(2.961)
Trimester II	-1.216	0.571	-1.197	0.135	-0.586	0.486
	(1.735)	(1.580)	(1.562)	(1.686)	(1.628)	(1.805)
Trimester III	-8.755***	-8.214***	-8.827***	-8.219***	-7.919***	-9.059***
	(2.722)	(2.488)	(2.218)	(2.797)	(2.277)	(2.981)
Observations	10207261	9844659	6803656	13248264	12399618	7652302
R^2	0.051	0.046	0.056	0.063	0.063	0.055
Panel B: NO2 in 10 ppb						
Trimester I	-3.783**	-5.343***	-3.707***	-5.094***	-4.351**	-5.049***
	(1.747)	(1.685)	(1.372)	(1.840)	(1.900)	(1.593)
Trimester II	-2.256*	0.366	-0.739	-1.031	-1.094	-0.821
	(1.268)	(1.002)	(1.195)	(1.061)	(1.039)	(1.022)
Trimester III	-5.873***	-6.338***	-5.785***	-6.294***	-6.488***	-5.454***
	(1.974)	(1.910)	(1.698)	(2.024)	(1.923)	(1.994)
Observations	7033374	6788389	4710853	9110910	8592210	5229553
R^2	0.051	0.045	0.054	0.062	0.062	0.054
Panel C: O3 in 0.01 ppm						
Trimester I	-4.314	-3.134	-3.542	-4.009	-4.170	-2.905
	(3.007)	(3.123)	(2.458)	(3.495)	(3.043)	(3.172)
Trimester II	0.798	-0.310	0.028	0.272	0.192	0.075
	(2.362)	(2.030)	(1.875)	(2.469)	(2.288)	(2.123)
Trimester III	-5.341	-4.374	-4.858	-5.008	-4.914	-4.390
	(3.483)	(3.551)	(3.110)	(3.816)	(3.551)	(3.483)
Observations	8641826	8340058	5748637	11233247	10626088	6355796
R^2	0.049	0.044	0.053	0.060	0.061	0.053
Panel D: SO2 in 10 ppb						
Trimester I	-4.943	-2.639	-5.881	-3.117	-2.260	-6.777
	(8.299)	(6.784)	(7.515)	(7.889)	(7.765)	(7.688)
Trimester II	-9.021*	-4.917	-7.379	-6.764	-6.293	-8.627
	(5.109)	(4.718)	(5.533)	(5.393)	(4.748)	(5.788)
Trimester III	-1.279	-4.123	-2.066	-3.285	-4.756	-0.315
	(5.706)	(5.605)	(6.006)	(5.799)	(5.227)	(6.646)
Observations	8282276	7993725	5447091	10828910	10064889	6211112
R^2	0.053	0.048	0.057	0.066	0.065	0.056

This table shows the heterogeneous impact of air pollution on birth weight for full-term infants only. In each column, dependent variable is birth weight in grams. Each column controls for year FE, county-quarter FE, fetal characteristics (excluding, if any, the specified characteristic showing the heterogeneity), maternal characteristics (excluding, if any, the specified characteristic showing the heterogeneity), weather conditions, prenatal visits and county medium income. Robust standard errors clustered at county level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8: Individual and Joint Effects of Air Pollution and Smoking - Full-Term Only

	(1) CO	(2) NO2	(3) O3	(4) SO2
Smoking	-168.893***	-184.736***	-197.906***	-192.090***
	(5.303)	(4.247)	(4.564)	(2.552)
Baseline effect of air pollution				
Trimester I	-5.058*	-3.497	-0.802	-0.698
	(2.730)	(2.564)	(2.924)	(5.876)
Trimester II	-0.714	-0.510	2.846	-5.503
	(2.055)	(1.773)	(2.465)	(5.422)
Trimester III	-7.028***	-3.936*	1.559	1.762
	(2.658)	(2.240)	(3.223)	(5.908)
Joint effect (air pollution $\times$ smoking)				
Trimester I	-14.947***	-9.503*	20.849***	-37.186***
	(3.621)	(4.843)	(5.120)	(13.999)
Trimester II	-6.421**	-5.627	10.219***	1.245
	(3.169)	(4.177)	(2.574)	(9.549)
Trimester III	-2.574	4.091	15.295***	-18.883
	(3.753)	(4.306)	(4.752)	(15.810)
Observations	13655782	7645780	10108303	11043464
R^2	0.079	0.063	0.058	0.066

Notes: This table shows the individual and joint impacts of air pollution and smoking on birth weight. In each column, dependent variable is birth weight in grams. All columns control for year FE, state-quarter FE, fetal and maternal characteristics, county medium income, weather conditions, number of prenatal visits, an indicator for smoking during pregnancy, as well as interactions between air pollution and smoking. Robust standard errors clustered at county level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9: Individual and Joint Effects of Air Pollution and Alcohol Consumption - Full-Term Only

	Light drinking				Heavy drinking			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	CO	NO2	O3	SO2	CO	NO2	O3	SO2
Light drinking	-58.290***	-79.668***	-133.924***	-107.576***				
	(12.233)	(14.354)	(16.434)	(11.119)				
Heavy drinking					-137.069***	-149.244***	-234.934***	-199.715***
					(27.639)	(22.737)	(23.489)	(12.884)
Baseline effect of air pollution								
Trimester I	-5.931*	-0.622	0.044	-5.300	-5.932*	-0.791	0.260	-4.666
	(3.155)	(2.289)	(2.812)	(6.775)	(3.175)	(2.280)	(2.815)	(6.711)
Trimester II	-1.678	-1.611	3.011	-5.611	-1.588	-1.924	2.997	-5.459
	(2.124)	(2.147)	(2.375)	(4.786)	(2.125)	(2.126)	(2.354)	(4.790)
Trimester III	-7.532**	-1.728	3.084	0.126	-7.409**	-1.870	3.246	0.683
	(2.941)	(2.137)	(3.059)	(5.858)	(2.942)	(2.159)	(3.048)	(5.818)
Joint effect (air pollution × drinking alcohol)								
Trimester I	-14.643**	-31.161***	49.590***	-38.665	-29.770*	-45.040**	90.461***	-78.737
	(7.279)	(9.326)	(17.605)	(37.864)	(15.985)	(19.512)	(25.233)	(48.333)
Trimester II	12.471*	-17.484*	12.013	-1.391	-9.085	-12.930	49.431***	92.262*
	(6.833)	(10.513)	(9.358)	(26.100)	(18.292)	(27.594)	(17.944)	(47.609)
Trimester III	-10.855	-0.613	42.992***	-52.599	-7.203	-20.411	81.293***	-82.254
	(9.578)	(14.158)	(15.752)	(41.347)	(17.928)	(19.231)	(25.320)	(53.072)
Observations	11504790	6267689	8454005	9706107	11531248	6267689	8454005	9706107
R^2	0.069	0.055	0.050	0.055	0.070	0.055	0.050	0.055

Notes: This table shows the individual and joint impacts of air pollution and smoking on birth weight. In each column, dependent variable is birth weight in grams. All columns control for year FE, state-quarter FE, fetal and maternal characteristics, county medium income, weather conditions, number of prenatal visits, an indicator for alcohol consumption during pregnancy, as well as interactions between air pollution and alcohol consumption. Robust standard errors clustered at county level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$