The Asian Games, Air Pollution and Birth Outcomes in South China: an Instrumental Variable Approach*†

Xiaoying Liu‡ Huazhang Miao§ Jere R. Behrman¶ Emily Hannum||
Zhijiang Liang** Qingguo Zhao††

September 7, 2021

*Liu, Behrman and Hannum acknowledge funding for this project from the National Science Foundation (No:1756738) and the China Research and Engagement Fund and School of Arts and Sciences Dean's Integrative Global Inquiries Fund, both at the University of Pennsylvania. We thank Janet Currie, Guy Grossman, Heather Schofield, Herbert Smith, Harsha Thirumurthy, Fan Wang, and two anonymous referees and the editor for helpful suggestions. The authors all declare they have no actual or potential competing financial interests.

†This research has been reviewed by the Research Ethics Committee at Guangdong Women and Children Hospital [201601032] and the Institutional Review Board at the University of Pennsylvania [#825326].

‡*Corresponding author.* Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104 (email: xiaoyliu@sas.upenn.edu)

§Guangdong Women and Children Hospital, Guangzhou, Guangdong Province, China

¶Department of Economics and Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA

||Department of Sociology and Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA

**Guangdong Women and Children Hospital, Guangzhou, Guangdong Province, China

††*Corresponding author.* Epidemiological Research Office of Key Laboratory of Male Reproduction and Genetics (National Health and Family Planning Commission), Family Planning Research Institute of Guangdong Province, Guangzhou, Guangdong Province, China (email: zqgfrost@126.com).

Abstract

We estimate the effects of air-pollution exposure on low birthweight, birthweight, and prematurity risk in South China, for all expectant mothers and by maternal age group and child sex. We do so by exploiting exogenous improvement in air quality during the 2010 Guangzhou Asian Games, when strict regulations were mandated to assure better air quality. We use daily air-pollution levels collected from monitoring stations in Guangzhou, the Asian Games host city, and Shenzhen, a nearby control city, between 2009 and 2011. We first show that air quality during the Asian Games significantly improved in Guangzhou, relative to Shenzhen. Next, using birth-certificate data for both cities for 2009 to 2011 and using expected pregnancy overlap with the Asian Games as an instrumental variable, we study the effects of three pollutants (PM_{10}, SO_2, and NO_2) on birth outcomes. Four main conclusions emerge: 1) air pollutants significantly reduce average birthweight and increase preterm risk; 2) for birthweight, late pregnancy is most sensitive to PM_{10} exposure, but there is not consistent evidence of a sensitive period for other pollutants and outcomes; 3) for birthweight, babies of mothers who are at least 35 years old show more vulnerability to all three air pollutants; and 4) male babies show more vulnerability than female babies to PM_{10} and SO_2, but birthweights of female babies are more sensitive than those of male babies to NO_2.

Key words: ambient air pollution, birthweights, preterm births, instrumental variable, China

JEL codes: I1, I12, I14, I18, Q51, Q52, Q53, Q54

1 Introduction

In 2016, 91% of the world's population lived in areas with air pollution that exceeded the World Health Organization air-quality guidelines for ambient fine particulate matter $PM_{2.5}$ – airborne particles less than or equal to 2.5 micrometers in aerodynamic diameter.[1] However, nearly all (86%) of the most extreme concentrations – above 75 $\mu g/m^3$ – were experienced by populations in China, India, Pakistan, and Bangladesh (Health Effects Institute, 2019). The World Health Organization estimates that 4.2 million premature deaths every year are linked to ambient air pollution (World Health Organization, 2019). Pregnant women and children may be particularly vulnerable to the ill effects of air pollution. Despite natural barriers protecting infants in utero, prenatal air pollution exposures have been linked to adverse birth outcomes (Fu et al., 2019; Huang et al., 2019; Šrám et al., 2005; Tsoli, Ploubidis and Kalantzi, 2019; Currie et al., 2014), and effects may persist into later childhood (Almond, Edlund and Palme, 2009; Almond, Currie and Duque, 2018; Currie et al., 2014; Tang et al., 2014; Bharadwaj et al., 2017; Sanders, 2012).[2] However, reviews suggest that evidence from epidemiological studies is inconsistent and causal links are hard to identify (Šrám et al., 2005; Bosetti et al., 2010).

In a separate line of research, economists have capitalized on novel natural experiments to come closer to estimating effects of air pollution during the prenatal period on infant health. For example, Jayachandran (2009), Rangel and Vogl (2019), and Kim et al. (2017) exploit incidents of natural fire or biomass fire and wind direction to identify the effect of air-pollution exposure on birth outcomes or longer-term health in Indonesia and Brazil. Currie and Walker (2011) exploit the introduction of electronic toll collection in the US and show reduced traffic pollution brought by electronic toll collection reduced prematurity and low birthweight among mothers within two kilometers of a toll plaza relative to mothers two to ten kilometers away. Luechinger (2014) exploits mandated desulfurization at power plants and prevailing wind directions to study the effects on infant mortality in Germany.[3]

Our paper follows this strand of literature by exploiting an international sports event: the

[1] The standard is an annual mean concentration of 10 $\mu g/m^3$.

[2] Mechanisms of air pollution effects on fetal development have been studied in animal models. For example, prenatal exposure to SO_2 leads to developmental and functional toxicities; NO_2 suppresses antioxidant defense systems; CO interferes with oxygen delivery to fetuses by displacement of oxygen from hemoglobin; and early fetal exposure to PM can alter trophoblast formation and vascularization of the placenta (Shah, Balkhair and Knowledge Synthesis Group on Determinants of Preterm/LBW births, 2011).

[3] See Currie et al. (2014) for a review of earlier studies.

2010 Guangzhou Asian Games, during which air quality was exogenously improved. We estimate, for the first time to our knowledge, air-quality impact on birth outcomes in South China. Specifically, assuming 39 weeks of gestation, we use the time during pregnancy that overlapped with the Asian Games as an instrumental variable to identify exogenous variation in air-pollution exposure and estimate the causal impacts on birth outcomes in response to such an exogenous change. Our approach is also novel in assessing heterogeneity in the impact of air pollution on birth outcomes by child sex and by maternal age group.

This paper is structured as follows. In Section 2, we summarize several issues commonly seen in the literature on air pollution and birth outcomes. In Section 3, we describe the air-quality regulations implemented before and during the 2010 Guangzhou Asian Games. We introduce the data we use and measurements of air pollution and weather exposures in Section 4. The empirical methods and main results are presented in Sections 5 and 6, respectively. Section 7 examines heterogeneity of effects by child sex and maternal age. Section 8 provides discussion and conclusions.

2 Background and Research Objectives

Deterioration in China's air quality since the 1990s and increased public availability of air quality data since 2014 have prompted a proliferation of research on the health effects of air pollution (Dong et al., 2013; Guan et al., 2016), including studies on air pollution and fetal growth and pregnancy outcomes (Qian et al., 2016; Zhu et al., 2017; Fu et al., 2019; Lu et al., 2019). Much of this literature studied associations between air pollution and birth outcomes, including infant mortality, low birthweight, and prematurity risk. However, most existing studies using individual birth information from China have suffered from one or both of the following limitations.

First, most of these studies use individual-level birth data from a single city within a relatively short time period–two to three years.[4] But multi-city and multi-year studies have statistical power advantages (Ito, Thurston and Silverman, 2007). In the single-city design, moreover, it is difficult to disentangle air pollution from meteorological factors as they usually co-vary and both can have impacts on birth outcomes. Multi-city samples with bigger variation in both air pollution and weather therefore have advantages because different cities

[4] Two recent papers from China are exceptions: Sun et al. (2019) covers all of Zhejiang Province, though this work does not control for weather, and Liu et al. (2019) covers the whole of Guangdong Province.

usually have different compositions of air pollution and patterns of interactions between air pollution and weather.

Second, most epidemiological studies fail to consider pregnancy timing and self-selection of residence. Pregnancy timing, either conscious or unconscious, is a significant phenomenon in many parts of the world. For example, using China's Fifth National Population Census, Yang (2021) finds there is a salient peak of birth numbers in October, which is mainly driven by labor migrants who often get married and conceive in the Chinese New Year period (usually between late January and mid February). Since labor migrants, on average, have lower educational attainment than non-migrants, ceteris paribus, failing to consider changes over time in a year in the composition of the pregnant population by controlling for maternal education could potentially bias the estimates of air pollution effects. However, in a recent review, only 7 out of 25 studies on air pollution and birth outcomes in China control for maternal education (Jacobs et al., 2017). On the other hand, people choose where to live based on their social economic status (SES) and preferences. People with lower SES may live in areas with higher pollution. Hajat, Hsia and O'Neill (2015)'s global review of socioeconomic disparities in air pollution exposure shows that areas where low SES communities dwell experience higher concentrations of criteria air pollutants. Since SES, usually imperfectly measured, also has effects on health, failing to take into account this self-selection may over- or underestimate the effect of the environment. For example, Grafova et al. (2014) discuss the selection bias in the effects of neighborhood environment on health, and find conventional estimates underestimate the effect of the economic environment on health.

Beyond these limitations, there is not yet agreement in existing studies regarding the important possibility of heterogeneous vulnerability to pollution effects across key demographic groups characterized child sex and maternal age. DiPietro and Voegtline (2017) review converging evidence that suggests that infant and early childhood developmental outcomes of male fetuses exposed to prenatal adversities are more highly impaired than those of female fetuses. Moreover, there is some suggestive evidence that the effect of air pollution on low birthweight differs by sex of the child (Ghosh et al., 2007). While risks for adverse birth outcomes increase with advanced maternal age (for a summary, see Sauer (2015)), the question of whether there is heightened vulnerability to air pollution exposures for women at advanced maternal age is not yet well studied.[5]

[5] Young maternal age is also a risk factor for poor pregnancy outcomes, though for different reasons. For evidence

Studies from China have shown varied results regarding sex differences in vulnerability to prenatal air pollution exposure. For example, Lu et al. (2019) find heterogeneous associations between preterm risk and perinatal exposure to indoor mold/damp stains by sex using retrospective recall data from a single city, but do not find any significant difference in associations with NO_2 exposure by sex. Wang et al. (2021) find that O_3 exposures are associated with low birthweight of female – but not male – babies using birth-certificate data in Guangzhou between January 2015 and July 2017. Chen and Ho (2016) find that incense burning during pregnancy is associated with significantly lower birthweight and smaller head circumference in boys, but not in girls, using a term-birth sample born in 2005 from the Taiwan Birth Cohort Study. On the other hand, Liang et al. (2021) do not find significant differences between males and females in stillbirth risk in response to $PM_{2.5}$ exposure using birth-cohort data in seven cities in Southern China between 2014 and 2017.

A few studies have assessed the presence of heterogeneous effects of air pollution exposure on birth outcomes by maternal age and show mixed patterns. An associational study in Wuxi, China provides evidence of a stronger association of PM_{10} with preterm birth among those with advanced maternal age, compared to others (Han et al., 2018). Wang et al. (2021) find increased risk of term low birthweight with maternal exposure to ozone among women older than 35 or younger than 25, compared to others, using data from Guangzhou, China between January 2015 and July 2017. However, Li, Guo and Williams (2016) do not find that an acute impact of hourly ambient air pollution on preterm birth is modified by maternal age, using data from Brisbane, Australia.

In this paper, we estimate the effects of prenatal air pollution on birth outcomes by 1) using unique birth-certificate data from two cities in South China that are similar in climate and social and economic development, but different in pollution levels; 2) addressing the issues of pregnancy timing and self-selection by exploiting the exogenous improvement of air quality during the Guangzhou 2010 Asian Games and controlling for quarterly trends in birth outcomes; and, 3) examining the heterogeneity of effects across child sex and maternal age groups. Given high-frequency daily air pollution level data, and properly controlled seasonal trends, our strategy exploits variation in the timing of conception within the same season, and compares the birth outcomes in Guangzhou, where air quality was significantly improved during the Asian Games, with those in Shenzhen. Although the city of Shenzhen was also

from China, see Xie et al. (2021).

in the area of regional air quality control during the Asian Games, we will show below that pollutant concentrations were significantly reduced in Guangzhou relative to Shenzhen during the Asian Games.

Our methodology is preferable to that used in much of the prior literature in that it not only addresses self-selection endogeneity bias – the problem that individuals choose where to live based on their health preferences and other observed characteristics – but also provides a way to disentangle the impacts of air pollution from weather, as air quality was exogenously improved during the Asian Games by stringent air-quality-control policies. Exploiting exogenous air-quality improvement during a large sports event to study its health impact does have precedent in the literature. For example, Rich et al. (2015) exploited the natural experiment of air-pollution decline during the Beijing Olympics to evaluate whether having specific months of pregnancy overlap with the 2008 Beijing Olympics was associated with larger birthweights, using individual data on term births, compared with pregnancies during the same dates in 2007 or 2009. However, there is no control city in the sample. In addition, exploiting the same 2008 Beijing Olympic Games, He, Fan and Zhou (2016) analyzed monthly mortality data from nationally representative surveillance points to estimate the effect of air pollution on mortality in China.

3 Air-Quality Regulations Related to the 2010 Guangzhou Asian Games

The 2010 Guangzhou Asian Games were the second Asian Games held in China, after the 1990 Beijing Asian Games. To prepare for the Asian Games, the Guangdong government adopted systematic, multi-stage improvement measures and implemented a series of regulations on air-quality control starting with the announcement in 2004 of Guangzhou as the host city for the 16th Asian Games. Since Beijing had successfully hosted the Olympics in 2008, the Guangdong government learned a great deal from the Beijing government's experiences with air-quality control, especially with regard to regional cooperation and multi-city control actions.

The key area targeted for air-quality improvement included the Pearl River Delta (PRD) Metropolitan Region, where the cities of Guangzhou and Shenzhen are situated. The straight-line distance between the two cities is about 100km. Although this regional air pollution management strategy may have improved Shenzhen's air quality during the same period, we later will show that relative to Shenzhen, air quality in Guangzhou was improved significantly during the Asian Games. On February 24, 2010, the Guangdong Province Environment Protection Department issued the "Notice on Air Quality Regulations during the Pre-Asian Games Period" (Department of Environmental Protection of Guangdong Government, 2010). It lists a series of measures for quality control during the last year before the Asian Games, for example, installation of desulfurization and denitrification by Guangzhou Papermaking Plant and power plants and converting to clean fuels from coal furnaces by the Guangzhou Steel Company. Low-steam-capacity industrial boilers were required to be phased out, all gas stations, oil depots and oil tankers were required to complete oil and gas recovery management, volatile-organic-compounds (VOCs) related industries were required to reduce emissions, and the building-materials industry was required to phase out low-capacity facilities and to install dust removal and denitrification facilities and meet emission requirements. As part of regional air-quality-control measures, PRD cities were required to set up complete vehicle-emissions inspection systems and monitoring and data-sharing networks, and upgrade the quality standard of gasoline for motor vehicles. Ambient air-quality monitoring and surveillance mechanism also were set-up in PRD cities.

While the previous measures were taken steadily and progressively, the most strict control measures were implemented during the Asian Games period (i.e., between October 20, 2010 and December 20, 2010) – the period on which we are focusing.[6] These measures involved temporary production shutdowns for plants that did not meet emission standards and traffic controls.[7] Efforts were made to make sure concentrations of SO_2, NO_2 and PM_{10} were lower than National Standard Grade 2 levels.

All of these regulations were strictly binding. Non-compliers faced the risk of temporary shutdown as punishment during the two-month Asian Games. As a result, air pollutant levels dropped significantly during the two months of the Guangzhou Asian Games (Liu et al., 2013; Xu et al., 2013). According to Liu et al. (2013), the Asian Games abatement strategy reduced emissions by 41.1% for SO_2, 41.9% for NO_x, 26.5% for PM_{10}, 25.8% for $PM_{2.5}$, and 39.7% for

[6] The Asian Games were held between November 1, 2010 and December 20, 2010. Because the strict measures were taken starting on October 20, 2010, we define the Asian Games Air Pollution Control period as October 20 - December 20, 2010.

[7] Similar trends were observed in lockdown periods during the Covid-19 Pandemic, in which most affected countries adopted partial or complete lockdown policies. As a result, air quality largely improved (National Aeronautics and Space Administration, U.S.A., 2020).

VOC. The concentrations of SO_2, NO_2, PM_{10} and $PM_{2.5}$ were reduced by 66.8%, 51.3%, 21.5% and 17.1%, respectively.

4 Data

4.1 Data on Birth Outcomes

We use birth-certificate data collected by one of the authors from one district in the city of Guangzhou and all of the city of Shenzhen for the period between January 2009 and February 2012. The birth-certificate data cover all births during the period in these districts and cities as required by law. All locations report basic birth outcomes including estimated gestational age based on reported last menstrual period, birthweight, birth length, sex, and parity, as well as maternal age and education.[8] Neither of these locations reported maternal height and marital status[9] in their birth-certificate systems. We focus on three outcomes: prematurity (defined as gestational age at birth less than 37 weeks), birthweight, and low birthweight (defined as birthweight less than 2500g). Table A1 lists the basic statistics of these outcomes by city. All three of these birth outcomes are associated with subsequent outcomes in childhood and adulthood. But these associations may reflect a range of factors, such as maternal health and family background, not just the impacts of the birth outcomes. For birthweight, in addition, there is evidence of casual effects through using monozygotic (MZ, identical) twins to control for all family background factors including genetics that the MZ twins share in common. Previous studies using MZ twins find that birthweights have impacts on outcomes ranging from schooling attainment to adolescent behaviors to adult earnings (Behrman and Rosenzweig, 2004; Conley, Strully and Bennett, 2003; Møllegaard, 2020; Torche and Conley, 2016). This means that if pollution affects birthweights, there is evidence of effects through birthweights over the life cycle.

We only include singleton live births in our sample (819,619). We exclude observations with less than 28 weeks or above 46 weeks of gestational age or with missing information on gestational age (1553 observations dropped), with birthweight less than 500g (51 observations dropped) or birth length less than 28cm or longer than 60cm (13 observations dropped).

[8] Although maternal job information is also available in both cities, 40% observations in Shenzhen sample has missing values. As there was no uniform birth certificate in the region at that time, maternal job categories were not collected in a consistent way in these two cities. We have tried to aggregate jobs into broad, comparable categories and include them in robustness checks in Appendix 2.

[9] Children born to single mothers are not culturally acceptable in China and are rare.

We further drop those with maternal age under 15 (927 observations dropped) or above 60 years old (5 observations dropped). To avoid fixed-cohort bias,[10] we delete births within this period with conception dates earlier than 18 June 2008 and those with conception dates later than 13 April 2011 (as gestational age varies between 196 and 322 days in the sample), which leaves 545,703 observations. We further drop 23,000 observations that do not have any of the following information: birthweight, gestational age at birth, gender of the baby, maternal age, maternal education or parity. In the end, 129,131 , 183,959 , 201,889, and 7,724 birth observations from 2009, 2010, 2011, 2012, respectively, constitute the main sample for analysis.

The distribution of the number of births throughout the year in this data (as shown in Figure A3) is consistent with the observation of Yang (2021) using national census data that the peak of number of births is in October, when those who get married during the Chinese New Year are most likely to give birth. Labor migrants may be most likely to show this pattern. Birthweight also shows strong seasonality in this sample, as shown in Figure A1. Shenzhen has higher average birthweight than Guangzhou throughout the year. The two cities show very similar seasonality trends: children born in the summer have higher birthweight than those born in the winter. These similar patterns help validate Shenzhen as a control city.

4.2 Data on Air Pollution and Weather

The data on air pollution come from the Guangzhou and Shenzhen environmental bureaus that report the daily average levels of three monitored air pollutants (NO_2, PM_{10}, and SO_2) at all monitoring stations in each city during 2008 to 2012. These pollutants are measured according to the National Standard GB3095—1996. For the very few missing observations in our data, we replace them with moving averages for the most recent 5 days. Data such as these are very rare in China for the period under study, as China only published an air pollution index (API) before 2014 and individual pollutant levels were generally not publicly available. As the API is an index score based on the most dominant pollutant, using the API alone does not reveal information on all "criteria" pollutants.

Figure A2 shows the daily average levels of the three air pollutants over the duration of the study period. Three patterns merit note. First, Guangzhou and Shenzhen have dif-

[10] Fixed-cohort bias emerges when a sample consists of births during a fixed period—this approach will include only the longer pregnancies at the start of the study and only the shorter pregnancies at the end of the study. This has the potential to bias studies of environmental exposures (Strand, Barnett and Tong, 2011).

ferent compositions of air pollutants and air quality in Shenzhen is on average better than Guangzhou. Second, in general, air pollution is worse in the winter than in the summer in both cities. The average levels of NO_2, PM_{10}, and SO_2 in Guangzhou are 52.22, 92, and 35$\mu g/m^3$ in November - January, compared to 27.91, 52.56, and 32.21 $\mu g/m^3$ in June - August. Third, SO_2 shows a decreasing trend throughout the entire period (conditional on season), which is consistent with the ongoing desulfurization effort by the Guangdong government during the last decade.

Guangzhou and Shenzhen have typical subtropical climates, with very mild winters and hot, rainy, and humid summers due to the Asian monsoon. The temperature throughout the year varies between 11°C and 33°C. Rainfall is abundant, at around 1700 millimeters a year, concentrated from May to September. We obtain hourly data on temperatures and dew points at two meters above sea-level as well as precipitation from reanalysis data at single levels by the European Centre for Medium-Range Weather Forecasts (ECMWF) (European Centre for Medium-Range Weather Forecasts, 2020*a*) for both cities for the period between January 1, 2008 and February 29, 2012 and calculate the daily mean temperature and daily mean relative humidity and 24-hr precipitation.[11]

We also obtain universal thermal climate indices (UTCI) from the same source. The UTCI, a thermal comfort indicator based on human heat-balance models, is designed to be applicable in all seasons and climates and for all spatial and temporal scales (European Centre for Medium-range Weather Forecasts, 2020*b*). The UTCI is a one-dimensional index that reflects "the human physiological reaction to the multidimensionally defined actual outdoor thermal environment" (Bröde et al., 2012, p.2). Scores can be classified into ten thermal stress categories, ranging from extreme cold stress to extreme heat stress (European Centre for Medium-range Weather Forecasts, 2020*b*). In a subtropical, humid environment such as Guangzhou and Shenzhen, this index provides a wider range of variation than does ambient temperature.[12] Table A2 lists the distribution statistics of UTCI in both cities between 2000 and 2008, which serves as the reference period based on which we define extreme weather in section 4.4.

[11] We calculate relative humidity (rh) by $rh = exp(5423 * ((1/273) - (1/d2m)))/exp(5423 * ((1/273) - (1/t2m)))$, where $d2m$ is 2-meter above sea level dew point, and $t2m$ is 2-meter above sea level temperature.

[12] We provide details on how we acquire and process UTCI data in Liu et al. (2021).

4.3 Prenatal Air-Pollution-Exposure Measurement

Due to data constraints, we only know individuals' cities of residence but not their home addresses. We therefore calculate the daily mean of air pollutants' levels for each city and assign each individual their accumulative prenatal exposure based on their gestational periods. However, this measure suffers from a potential problem – it generates a spurious inverse correlation between adverse birth outcomes and accumulated air pollution levels because a shorter gestational age implies a shorter third trimester and therefore smaller accumulated air-pollution exposure. We avoid this spurious relationship by calculating the daily average exposure level (i.e. dividing the accumulated exposure by the number of days of gestational age).

4.4 Extreme Weather Exposure Measurement

Besides air pollutants, weather conditions such as temperature and humidity have been shown to be related with birth outcomes (Beltran, Wu and Laurent, 2014; He et al., 2016; Murray et al., 2000; Siniarska and Kozieł, 2010). In this paper, we control for weather with two different specifications. In one specification, we control for daily mean average temperature and relative humidity during pregnancy. In the other specification, we use the universal thermal climate indices (UTCI) to define extreme cold and extreme hot weather in order to control for the possible effects of extreme temperatures on both ends. More specifically, we define a threshold for extreme cold days as a daily mean UTCI below $0°C$, and for extreme hot days as a daily mean UTCI above $34°C$.[13] It can be seen that UTCI has a larger variance than ambient temperature and is more likely to fall into extreme temperature ranges. We calculate the percentage of time during pregnancy exposed to extreme cold (number of days with the daily lowest apparent temperature under $0°C$ divided by the number of pregnancy days) and extreme heat (number of days with the daily highest apparent temperature over $34°C$ divided by the number of pregnancy days) for each pregnancy and use these two variables to control for extreme weather conditions.[14]

[13]The reason that we use 0°C and 34°C as cutoffs to define extreme cold and hot is because these two cutoffs are close to the bottom 1% and top 1% of historical UTCI data of both cities between 2000 and 2008 as shown in Table A2.

[14]We do not use the absolute number of extremely cold or hot days in regressions due to the inverse causality concern already discussed.

5 Methods

The method that is applied in this paper exploits the change in air pollution during the Asian Games in Guangzhou, relative to Shenzhen, to identify the effect on birth outcomes. As seen in the data section, Guangzhou and Shenzhen are similar in climate as well as social and economic development, with Shenzhen having slightly better air quality prior to the Asian Games. Since what counts for the analysis is the changes that occurred during the Asian Games, this slight difference in air quality before the Asian Games does not jeopardize our approach.

In order to investigate whether Guangzhou air quality was significantly improved relative to Shenzhen during the Asian Games, we first run the following difference-in-difference regression with daily monitored air pollutants levels data from monitoring stations as well as meteorological data:

$$\begin{aligned} P_{yds} = a * AsianGames_{yd} + b * Guangzhou + c * Guangzhou * AsianGames_{yd} + y + d + dow \\ + spline(temp, df = 3) + spline(rainfall, df = 3) + u_{yds} \end{aligned} \tag{1}$$

P_{yds} is the level of air pollutant P on date d in year y at monitoring site s. We model it as a function of year fixed effects y, date-in-a-year fixed effects d, day-of-week fixed effect dow, city fixed effects $Guangzhou$, and spline functions of observed daily mean temperature and rainfall each with 3 degrees of freedom, as well as an error term u_{yds}. We define $AsianGames_{yd}$ as a binary variable equal to 1 if year y is 2010 and date d is between October 20th and December 20th. This term estimates the average change in the air pollutant in both cities during the Asian Games. Parameter c is the coefficient of interest, which estimates the effect of the Asian Games on air quality in Guangzhou city during the Asian Games, in addition to the average effect in both cities.

Next, we estimate the effects of air pollution on birth outcomes. Observed associations between air pollution and birth outcomes may suffer from two types of bias that work in opposite directions: 1) attenuation bias caused by measurement error of exposure level as we do not have accurate information on where in the cities people spent their time; 2) location selection bias as people with more resources (and perhaps, better health) may choose in what city to live in part according to air quality. Ignoring the selection bias may over- or under-estimate the effect of environment on health or other outcomes. As these two types of biases may work in different directions, it is hard to predict the direction of bias in the aggregate. We address this issue by exploiting the exogenous air quality improvement during the Guangzhou Asian Games and treat the event as a "quasi-experiment" that indirectly affected birth outcomes through air-quality improvements. Thus, the methodology we undertake to identify the causal impacts of air pollution on adverse birth outcomes is to use Asian Games exposure as an instrument for prenatal air-pollution exposure. The estimated effect is the Local Average Treatment Effect on the Treated (LATE). More specifically, we estimate the correlation between birth outcomes and individual prenatal air pollution exposure, with the number of days overlapping between a woman's pregnancy and the air-quality control period for the Asian Games as an instrumental variable that brings about the exogenous change. The empirical framework is described in Equations (2) and (3).

$$P_i = \sigma X_i + \rho M_i + \gamma * 1(c_i = "Guangzhou") * G_i + y_i + q_i * c_i + c_i + \epsilon_i \tag{2}$$

$$Y_{icyq} = \alpha X_i + \theta P_i + \beta M_i + y_i + q_i * c_i + c_i + u_i \tag{3}$$

Eq.(2) is the first-stage estimation of pregnant woman i's pollution exposure P_i, as a function of individual characteristics X_i, meteorological conditions M_i during pregnancy, as well as a policy-exposure term $\gamma * 1(c_i = "Guangzhou") * G_i$, where, G_i is pregnant woman i's time exposed to the Asian Games during pregnancy. Here, G_i satisfies the exclusion restriction in that, as long as pregnant women are not timing pregnancy relative to the Asian Games, it only affects one's exposure to air pollution and not directly affects birth outcomes. It serves as the instrumental variable. We also control for conception-year fixed effects y_i, city-specific conception-quarter[15] fixed effects $q_i * c_i$ and city fixed effects c_i.

Eq. (3) is the outcome estimation, in which Y_{icyq} is the birth-outcome variable of individual i from city c conceived in year y and quarter q. The birth outcome can be a continuous variable – birthweight, or a binary variable – low birthweight or prematurity. The birth outcome is a function of individual-specific characteristics (X), as well as exposure variables including prenatal exposure to ambient air pollution (P) and meteorological conditions (M). As in the first stage, we also control for conception-year fixed effects y_i, city-specific conception-quarter fixed effects $q_i * c_i$ and city fixed effects c_i.

The measurement of birth outcomes, air pollution and weather, and exposure variables have been described in Section 4. In addition, confounding factors include a quadratic

[15]We define quarters based on local climate: quarter 1 - spring: March-May, quarter 2 - summer: June-August, quarter 3 - autumn: September - November, quarter 4 - winter: December - February.

term in maternal age, a set of dummy variables for maternal-schooling categories, a set of dummy variables for parity, a binary variable indicating child's sex as in the vector of X, and conception-year fixed effects, city fixed effects as well as city-specific conception-quarter fixed effects.

We consider two candidates for G_i: 1) a continuous variable measuring how many days one's pregnancy duration overlapped with the Asian Games. 2) a continuous variable measuring how many days one's *expected* pregnancy duration overlapped with the Asian Games, by assuming everyone has 39 weeks of gestation. We argue that the first of these is not a valid instrument because one is likely to have more overlap with Asian Games if one has longer gestation given the same gestational date. Therefore, this variable may not satisfy the exclusion restriction because it is correlated with outcome variables not only through air pollution exposure. In contrast, the second instrumental-variable candidate simply exploits the variation in timing of conception and its exogenous co-variance with the timing of Asian Games to instrument for variation in exposure. This instrumental variable satisfies the exclusion restriction if we believe that individual's timing of conception was not affected by the Asian Games, conditional on conception quarter-in-a-year fixed effects. Our following analysis adopts the second instrumental variable.

Besides pollution and weather, we also control for conception-year fixed effects y_i and a city-specific conception quarter-in-a-year effect $q_i * c_i$ to control for city-specific seasonal patterns and the possible weather preference that may introduce selection bias. This city-specific quarter-in-a-year term controls for all unobserved time-varying city fixed effects, including population composition change due to pregnancy timing and the common shocks to birth outcomes experienced by individuals who conceived in the same season of a year within the same city[16]. The reason for controlling for conception quarter-in-a-year trends instead of month trends of birth outcomes in our study is that the instrumental variable we use is defined based on city and timing of conception (basically covering November and December of 2010), i.e., the exogenous variation in exposure variable comes from interaction between a city fixed effect and conception-month fixed effect. Controlling for month trends in birth outcomes likely would cause collinearity with this variation and therefore over-control

[16] In more recent papers examining environment on birth outcomes, some studies control for quarterly trends (Wang et al., 2021), or monthly trends in a year (He et al., 2016), or even day-in-a-year trends (Chen et al., 2020). Since air pollution also varies across time in a year, there could be a tradeoff between under-control and over-control of these trends.

the model. Note that with fixed effects rather than only random effects, the other types of shocks that may also affect birth outcomes (such as typhoons) are allowed to be correlated with the right-side variables, including air pollution and weather. u_i is the error term that represents the effects of random unobserved factors and measurement error. As we have controlled for year fixed effects and city-specific conception quarter-in-a-year fixed effects, we assume u_i is independent and identically distributed (i.i.d.). In regression results, we report heteroscedasticity-consistent or robust standard errors.

6 Results

6.1 Effects of Asian Games on Air-Pollution Reduction

Our first set of results in regressions of daily monitored level of each pollutant between 2008 and 2012 show that during the Asian Games, the three air pollutants' levels were significantly reduced in Guangzhou. The coefficients on $1.Guangzhou * 1.policy$ in Table 1 correspond to coefficient c in Equation 1, i.e., the coefficient of interest. As shown in the first panel, Column (1)-(3), after controlling for meteorological conditions including temperature and rainfall, city fixed effects, year fixed effects and day-in-a-year fixed effects, we find PM_{10}, NO_2 and SO_2 in Guangzhou were significantly reduced during the Asian Games compared to the same period in 2009 and 2011. More specifically, compared to Shenzhen, Guangzhou witnessed a further reduction of PM_{10} by 33.91 µg/m^3 (s.d.=6.44), SO_2 by 8.24 µg/m^3 (s.d.=2.45) and NO_2 by 24.64 µg/m^3 (s.d.=4.62).

We also implemented placebo tests that defined three imaginary, alternative dates for the Asian Games: 1) one year before (Scenario A. $Policy = [20Oct2009 - 20Dec2009]$), 2) one year after (Scenario B. $Policy = [20Oct2011 - 20Dec2011]$), and, 3) five months before (Scenario C. $Policy = [20May2010 - 20Jul2010]$). The significant positive coefficients on $Guangzhou * policy$ term in Placebo Test Scenario A implies air pollutant levels in Guangzhou in 2009 were significantly higher than in the same period in 2010 and 2011, compared to the control city Shenzhen. In Placebo Test Scenario B, we find PM_{10} and NO_2 are reduced in both Guangzhou and Shenzhen in 2011 compared to 2009 and 2010, while SO_2 is more reduced in Guangzhou compared to Shenzhen between 20 October and 20 December in 2011, thanks to persistent desulphurisation efforts by the Guangzhou government. In Placebo Test Scenario C, we do not see significant differences in air pollutants' levels between May and July in

2010 compared to other years between two cities. These placebo tests lay the foundation for our estimation strategy below of using the overlap period between pregnancy and the Asian Games to instrument for one's prenatal air pollution exposure.

6.2 Effects of Air Pollution on Birth Outcomes

We estimate Equations (2) & (3) with a linear continuous model for birthweight and probit models for low birthweight and preterm birth using the two-stage least squares method (2SLS and IV-Probit separately). Due to high correlations among pollutants, we estimate the equation with each pollutant separately.

As mentioned in Section 5, we use number of days overlapped between *expected* pregnancy (by assuming each has 39 weeks of pregnancy) and the Asian Games to instrument for the exposure measurement. Consistent with difference-in-difference estimation results on air quality improvement during Asian Games, the first-stage results (Table A3) show the instrument is highly significantly correlated with all three air pollutant exposures. One more day of overlap of pregnancy with the Asian Games period reduces daily average exposure to PM_{10}, NO_2 and SO_2 by 0.16, 0.26, 0.13 $\mu g/m^3$ respectively, in the specification with linear control of temperature and relative humidity. The results are similar in the other specification with extreme temperatures as weather controls.

Marginal effects of PM_{10} exposure on birth outcomes are presented in Table 2. For comparison, we present results of OLS or probits in the left 6 columns ((1)-(6)), while the right 6 columns (7) to (12) are for 2SLS or IV-probits. Specifications A and B vary by the weather controls. Specification A controls for daily mean temperature and relative humidity over one's pregnancy, while specification B controls for percentage of time during pregnancy that a woman was exposed to extreme temperatures, defined by UTCI.

Although the effects of PM_{10} on mean birthweights are similar between OLS and IV estimates, the marginal effects of PM_{10} on the risk of low birthweight and preterm are much smaller with IV – nearly half of the size of those without IV, given the same specification. With IV estimates, a 10 $\mu g/m^3$ increase in PM_{10} reduces birthweight by around 14 grams, and increases the risk of preterm by 0.45% to 0.65%, while the effect on the risk of low birthweight is not statistically significant in a specification controlling for extreme temperatures, and only weakly significantly increases low birthweight by 0.3% in the specification controlling for linear temperature and relative humidity.

Table 1: Difference-in-Difference Estimation of Effects on Air Pollution of Asian Games

	I. Real Policy Effect (Policy = [20Oct2010 - 20Dec2010])			II. Placebo Test Scenario A (Policy = [20Oct2009-20Dec2009])			III. Placebo Test Scenario B (Policy = [20Oct2011-20Dec2011])			IV. Placebo Test Scenario C (Policy = [20May2010-20JUL2020])		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	PM_{10}	SO_2	NO_2	PM_{10}	SO_2	NO_2	PM_{10}	SO_2	NO_2	PM_{10}	SO_2	NO_2
1.policy	15.77***	0.927	5.500	0.599	2.843	0.299	-12.34**	2.289	-7.299**	4.892	1.366	3.456
	(5.138)	(1.954)	(3.688)	(5.170)	(1.959)	(3.699)	(5.179)	(1.965)	(3.718)	(4.378)	(1.663)	(3.148)
1.Guangzhou*1.policy	-33.91***	-8.244***	-24.64***	13.64**	5.502**	20.64***	7.106	-6.733***	-4.040	4.177	-2.815	-4.190
	(6.441)	(2.450)	(4.624)	(6.463)	(2.448)	(4.624)	(6.479)	(2.458)	(4.652)	(5.365)	(2.037)	(3.858)
Observations	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576
R-squared	0.457	0.595	0.476	0.452	0.597	0.483	0.453	0.594	0.472	0.452	0.593	0.472

Notes: Estimation of Daily air pollutant level from all monitoring stations in Guangzhou and Shenzhen city between 2008 and 2012. All specifications include Year FE, City FE, Day in a Year FE. Justified models also justify meteorological variables including temperature and rainfall. Standard errors in parentheses. *** significant at 1%, ** significant at 5%, *significant at 10%.

Table 2: Marginal Effects of PM_{10} (by 10 μg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm.

	Probit Model/OLS Model						IV-Probit Model/2SLS Model					
VARIABLES	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PM_{10}*10μg/m^3	0.007***	0.006***	-16.37***	-13.70***	0.012***	0.009***	0.0033*	0.0022	-14.18***	-13.39***	0.0065***	0.0045**
	(0.001)	(0.001)	(2.60)	(2.31)	(0.001)	(0.001)	(0.002)	(0.002)	(4.97)	(5.20)	(0.002)	(0.002)
Temp. (°C)	-0.0038***		10.59***		-.0044 ***		-0.0041***		10.76***		-0.0048***	
	(0.000)		(0.948)		(0.000)		(0.000)		(1.01)		(0.000)	
Rel. hum. (%)	0.0015***		-3.83***		0.0023***		0.0012***		-3.63***		0.0017***	
	(0.000)		(0.44)		(0.000)		(0.000)		(0.59)		(0.000)	
Ext. cold (%)		0.0023***		-6.41***		0.0032***		0.002***		-6.39***		0.0028***
		(0.001)		(1.25)		(0.001)		(0.001)		(1.30)		(0.001)
Ext. hot (%)		0.0003		-0.095		0.0007		0.000		-0.10		0.00086
		(0.000)		(1.15)		(0.001)		(0.000)		(1.16)		(0.001)
Obs.	522,703						522,703					
Spec.	A	B	A	B	A	B	A	B	A	B	A	B

Notes: Each column represents a separate regression. (1)-(6) are from Probit/OLS regressions without applying IV, while (7)-(12) are from IV-Probit/2SLS regressions. (1),(2),(5),(6),(7),(8),(11),(12) are marginal effects of 10 μg/m^3 PM_{10} on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4),(9),(10) are marginal effects of 10μg/m^3 PM_{10} on Mean Birthweight. Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Results for the other two pollutants, NO_2 and SO_2, are presented in Tables A4 and A5 in the Appendix. The effects of SO_2 are strongest: a 10 µg/m^3 increase in SO_2 reduces average birthweight by 18 to 20 grams, and increases the preterm rate by around 0.76 to 0.83%. Among the three pollutants, NO_2 shows the smallest effects: 10 µg/m^3 increases in NO_2 reduce average birthweight by 8.7 to 9.3 grams, and increase the preterm rate by 0.33 to 0.38%.

6.3 Estimation Results by Trimester

To test the sensitivity of effects within various subperiods of pregnancy, we also estimate the impacts of each pollutant by trimester, with the coefficients displayed in Table 3. Each column represents results from a regression, i.e., different pollutants enter different regressions, while the exposure to the three trimesters are from the same regression. It is clear that third trimester exposure to PM_{10} plays a key role in reducing average birthweight. A 10 µg/m^3 increase in PM_{10} in the third trimester decreases average birthweight by 13-15 grams, while exposures in the first and second trimester don't have significant effects on birthweight. The most sensitive trimester to PM_{10} in terms of preterm risk depends on the weather control: in specification A with linear control of mean temperature and humidity, the third trimester is most sensitive to PM_{10}, while in specification B with extreme temperature defined by UTCI, the second trimester is the only sensitive period. The statistical significance is much lower for the preterm risk outcome. The estimated effects of SO_2 and NO_2 by trimester do not show sensitive periods, although three trimesters combined have significant effects on birthweight and preterm. To save space, the estimated coefficients are not displayed, but are available on request.

7 Heterogeneity of Effects by Sex and Maternal Age

Previous research shows evidence of gender/sex differences in response to environmental exposures (Keitt, Fagan and Marts, 2004; Clougherty, 2010; Kim et al., 2017). To test the hypothesis that different sexes may have different responses to air pollution in the fetal period, we estimate heterogeneous effects by infants' sex. We do this by adding interaction terms between infant's sex and air-pollutant exposures. The estimated coefficients of these interaction terms for the three birth outcome variables are shown in the first line in Table 4.

Table 3: Marginal Effects of Prenatal Exposure to PM_{10} on Low Birthweight, Average Birthweight, and Preterm Risk by Trimester

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
PM_{10} in the first trimester *10 μg/m^3	-0.0003	0.00061	1.63	1.22	0.0002	0.00087
	(0.002)	(0.001)	(3.97)	(3.76)	(0.002)	(0.002)
PM_{10} in the second trimester*10 μg/m^3	0.001	-.00033	-0.56	-1.17	0.002	0.0027*
	(0.001)	(0.001)	(2.67)	(3.30)	(0.001)	(0.001)
PM_{10} in the third trimester*10 μg/m^3	0.003	0.002	-15.03***	-13.12***	0.0035*	0.001
	(0.002)	(0.001)	(4.60)	(3.86)	(0.002)	(0.002)
Weather control	A	B	A	B	A	B
Observations	522,703					

Notes: Each column represents a separate regression. Instrumental variables are a set of variables that measure the percentage of time overlapped with Asian Games period in each trimester. Besides air pollution exposure, other control variables include a binary variable indicating baby is female, parity, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Males are more vulnerable to PM_{10} in terms of average birthweight at a significance level of 10%: on average, males have a 1.51 gram reduction in birthweight in correspondence to a 1 μg/m^3 increase in PM_{10} (i.e., they experience a 15.1 gram extra reduction in birthweight in correspondence to a 10 μg/m^3 increase in PM_{10} compared to females). Nevertheless, males are not more likely than females to experience low birthweight or preterm birth because of air pollution in particular.

The impacts of air pollution may also vary by maternal age. Mothers aged at least 35 deliver babies with lower birthweights – in response to a 10 μg/m^3 increase in PM_{10} exposure, lower by 17.3 grams, on average (for both Specifications A and B). However, they are not more likely than other mothers to have low birthweight babies nor to have preterm births in response to air pollution.

Table A6 and A7 give the coefficient estimates for interaction terms for NO_2 and SO_2. SO_2 shows similar interaction term coefficients (males show an 8 more grams reduction in birthweight in response to a 10 μg/m^3 increase in SO_2, and mothers older than 35 years old deliver babies with 22 grams lower birthweight, on average, in response to a 10 μg/m^3 increase in SO_2). However, NO_2 shows a different pattern. Males do not experience disadvantage, and indeed show an advantage, in birthweight in response to NO_2 exposure (i.e., female fetuses are more vulnerable to NO_2 than are males). But babies of mothers aged over 35 are even more vulnerable than others to NO_2: they experience 77.6 grams lower birthweights, on average. Neither of these pollutants have sex heterogeneous effects on the risk of low birthweight nor preterm.

8 Discussion and Conclusion

This study seeks to provide plausibly causal estimates of the impact of prenatal air pollution exposure on birth outcomes by exploiting a large-scale sports event – the Guangzhou 2010 Asian Games – during which air quality control was strictly enforced and air quality was exogenously improved. Using daily air-quality data from monitoring stations in Guangzhou and the control city, Shenzhen, we first show that air quality indeed improved during the Asian Games period in Guangzhou relative to Shenzhen. To demonstrate the improvement, we presented difference-in-difference estimation results and implemented placebo tests by assigning hypothetical alternative counterfactual dates for the Asian Games period. With these findings established, we use the overlap between pregnancy and the actual Asian Games period (given 39 weeks of gestational age) to instrument for ambient air-pollution exposure. We find that air pollutants PM_{10}, NO_2 and SO_2 significantly decrease birthweights and increase preterm delivery risks. 10 µg/m^3 increases in PM_{10}, NO_2, SO_2, significantly decrease mean birthweight by around 14, 9, and 18-20 grams, respectively. The same changes increase preterm risk by 0.45-0.65%, 0.33-0.38% and 0.76-0.83%, respectively. For average birthweight, we also find that the third trimester is an especially vulnerable period for exposure to PM_{10}.

We compare our findings first with associations summarized in a review of epidemiological studies using Chinese data: Jacobs et al. (2017) summarizes estimates that show a 10 $\mu g/m^3$ increase in PM_{10} reduces birthweights by 0-9.1 grams. Our estimates on birthweight are of the same order of magnitude but slightly bigger than the findings in these association studies. If we run logit regressions of preterm birth without instrumenting on PM_{10}, the odds ratio of preterm, 1.027 (s.d.= 0.00283) for specification A, is in the range between 1 and 1.05, as in most epidemiological studies. Nevertheless, the marginal effects on preterm birth are roughly halved in the instrumental-variable model, compared to a standard probit model without instrumenting, and the statistical significance is much weaker for low birthweight. This pattern suggests possible selection bias in the simple association studies of the form that couples who have resources and knowledge to choose lower pollution also have better birth

Table 4: Heterogeneous Effects of PM_{10} on Birth Outcomes by Sex and Maternal Age Group

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
PM_{10} * male	-.000326	-0.000315	-1.508*	-1.506*	0.0008897	0.0008034
	(0.003)	(0.003)	(0.772)	(0.772)	(0.003)	(0.003)
PM_{10} *(mom age>=35)	0.000377	0.000376	-1.729***	-1.727***	0.000444	0.000444
	(0.000280)	(0.000280)	(0.533)	(0.533)	(0.000315)	(0.000315)
Specification	A	B	A	B	A	B

Notes: IV estimates of the interaction term between PM_{10} and male or age35+ on birth outcomes, with the same set of control variables as in Table 2. *** significant at 1%, ** significant at 5%, *significant at 10%.

outcomes. The timing of the sensitivity window at the third trimester to PM_{10} is also consistent with Rich et al. (2015) who find month 8 in pregnancy is the critical window for air pollution.

We next compare our pollution-effect estimates to estimates from developed countries, where air pollution is generally lower, in order to illuminate the possible non-linearity of dose-response curves (Fleischer et al., 2014; Arceo, Hanna and Oliva, 2016). There have been a large number of associational studies. For example, Liu et al. (2003) found that the risk of preterm delivery was associated with third-trimester exposure to SO_2 (OR=1.09 for 5.0 ppb (or 13.1 $\mu g/m^3$) increase, with daily average at 12.8 $\mu g/m^3$) using live-birth datasets from Vancouver, Canada.

However, causal studies of the impact of air pollution exposure on birthweight and prematurity are limited. Exceptions include Coneus and Spiess (2012), who find no significant effect from NO_2, or SO_2 in estimates of a mother fixed-effects model using a small sample in Germany[17], where air quality was much better compared to Guangzhou. Currie, Neidell and Schmieder (2009) also find no effects for PM_{10} in New Jersey in the United States[18], where average concentration of PM_{10} is only about 30 $\mu g/m^3$–less than half of the average level in our study. Together with the null findings for PM_{10} and NO_2 in Coneus and Spiess (2012) and Currie, Neidell and Schmieder (2009), the significant findings reported here suggest the possibility that air pollutants at higher concentrations, as in this study, are more likely to cause lower birthweights and prematurity. This contrast across studies with quite different pollution levels suggests the possible existence of threshold or other nonlinear effects of air pollution on birth outcomes.

This study also contributes to a literature on heterogeneous vulnerability to air pollution by maternal age and child sex. In terms of average birthweight, males are more vulnerable to PM_{10} and SO_2 than are females, but females are more vulnerable to NO_2. With regard to average birthweight, children of mothers over 35 years old are also more vulnerable than others when exposed to higher levels of all three types of air pollutants. NO_2 is associated with the greatest harm to babies of mothers over 35 years in comparison with SO_2 and PM_{10}. We do not find sex and maternal age differences in the effect of air pollution on preterm risk.

[17] Coneus and Spiess (2012) find that CO significantly reduces birthweight, but find no significant effects from O_3, NO_2 and SO_2.

[18] Currie, Neidell and Schmieder (2009) find negative effects of CO on birth outcomes, but no effects for O_3 or PM_{10}.

There are several limitations in our approach. First, and most important, is exposure measurement. We do not have residential addresses of mothers to permit personalized pollution-exposure measures by matching mothers' addresses to nearby monitoring stations or a high spatial resolution pollution map that uses more sophisticated approaches, such as inverse distance weighting (such as used in Lu et al. (2020)), kriging methods, and land-use regression methods (compared in Mercer et al. (2011)). Of course, if women move around the city for work, socializing or shopping, a broader geographical representation of pollution than their local residential neighborhood may be appropriate. Second, also due to data limitations, we cannot control for additional confounders such as smoking and passive smoking, cooking smoke exposure, and avoidance behavior. However, there is no reason to think that such confounders are correlated with the changes in pollution due to the Asian Games. For this reason, the fact that we cannot control for these confounders is not likely to affect our point estimates but could affect the precision of our estimates. Despite these limitations, this paper offers an original contribution by exploiting a unique setting of exogenous change in air quality to estimate impacts of pollution on birth outcomes.

In summary, our paper identifies significant impacts of prenatal-air-pollution exposure on birth outcomes by exploiting a natural experiment – the Asian Games – during which air quality was exogenously improved. This approach is advantageous in addressing self-selection related to exposure and disentangling confounding meteorological factors from air pollution, as the Asian Games event provided an excellent opportunity to observe exogenous change in air pollution but not in weather conditions. Finally, maternal age and child sex differences in vulnerability to air pollution in our findings suggest that these heterogeneities deserve future study.

References

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature*, 56(4): 1360–1446.

Almond, Douglas, Lena Edlund, and Mårten Palme. 2009. "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden." *The Quarterly Journal of Economics*, 124(4): 1729–1772.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *The Economic Journal*, 126(591): 257–280.

Behrman, Jere R., and Mark R. Rosenzweig. 2004. "Returns to Birthweight." *The Review of Economics and Statistics*, 86(2): 586–601.

Beltran, Alyssa J., Jun Wu, and Olivier Laurent. 2014. "Associations of Meteorology with Adverse Pregnancy Outcomes: A Systematic Review of Preeclampsia, Preterm Birth and Birth Weight." *International Journal of Environmental Research and Public Health*, 11(1): 91–172.

Bharadwaj, Prashant, Matthew Gibson, Joshua Graff Zivin, and Christopher Neilson. 2017. "Gray Matters: Fetal Pollution Exposure and Human Capital Formation." *Journal of the Association of Environmental and Resource Economists*, 4(2): 505–542. Publisher: The University of Chicago Press.

Bosetti, Cristina, Mark J Nieuwenhuijsen, Silvano Gallus, Sonia Cipriani, Carlo La Vecchia, and Fabio Parazzini. 2010. "Ambient particulate matter and preterm birth or birth weight: a review of the literature." *Archives of toxicology*, 84(6): 447–460.

Bröde, Peter, Dusan Fiala, Krzysztof Błażejczyk, Ingvar Holmér, Gerd Jendritzky, Bernhard Kampmann, Birger Tinz, and George Havenith. 2012. "Deriving the operational procedure for the Universal Thermal Climate Index (UTCI)." *International Journal of Biometeorology*, 56(3): 481–494.

Chen, Le-Yu, and Christine Ho. 2016. "Incense Burning during Pregnancy and Birth Weight and Head Circumference among Term Births: The Taiwan Birth Cohort Study." *Environmental Health Perspectives*, 124(9): 1487–1492.

Chen, Xi, Chih Ming Tan, Xiaobo Zhang, and Xin Zhang. 2020. "The effects of prenatal exposure to temperature extremes on birth outcomes: the case of China." *Journal of Population Economics*.

Clougherty, Jane E. 2010. "A growing role for gender analysis in air pollution epidemiology." *Environmental health perspectives*, 118(2): 167.

Coneus, Katja, and C. Katharina Spiess. 2012. "Pollution exposure and child health: Evidence for infants and toddlers in Germany." *Journal of Health Economics*, 31(1): 180–196.

Conley, Dalton, Kate W Strully, and Neil G Bennett. 2003. *The Starting Gate: Birth Weight and Life Chances.* University of California Press.

Currie, Janet, and Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics*, 3(1): 65–90.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What Do We Know About Short-and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics*, 6(1): 217–247.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder. 2009. "Air pollution and infant health: Lessons from New Jersey." *Journal of Health Economics*, 28(3): 688–703.

Department of Environmental Protection of Guangdong Government. 2010. "Notice on Air Quality Regulations during the Pre-Asian Games Period (Yue Environment Notice[2010] No.22)." http://gdee.gd.gov.cn/shbtwj/content/post_2303946.html.

DiPietro, J. A., and K. M. Voegtline. 2017. "The gestational foundation of sex differences in development and vulnerability." *Neuroscience*, 342: 4–20.

Dong, Guang-Hui, Zhengmin (Min) Qian, Pamela K. Xaverius, Edwin Trevathan, Salwa Maalouf, Jamaal Parker, Laiji Yang, Miao-Miao Liu, Da Wang, Wan-Hui Ren, Wenjun Ma, Jing Wang, Alan Zelicoff, Qiang Fu, and Maayan Simckes. 2013. "Association Between Long-Term Air Pollution and Increased Blood Pressure and Hypertension in China." *Hypertension*, 61(3): 578–584.

European Centre for Medium-Range Weather Forecasts. 2020*a*. "Climate Data Store." https://cds.climate.copernicus.eu/cdsapp#!/home(Accessed at 2020-05-24).

European Centre for Medium-range Weather Forecasts. 2020*b*. "Thermal comfort indices derived from ERA5 reanalysis." https://cds.climate.copernicus.eu/doi/10.24381/cds.553b7518 (Accessed at 2020-04-20).

Fleischer, Nancy L., Mario Merialdi, Aaron van Donkelaar, Felipe Vadillo-Ortega, Randall V. Martin, Ana Pilar Betran, João Paulo Souza, and Marie S. O´Neill. 2014. "Outdoor Air Pollution, Preterm Birth, and Low Birth Weight: Analysis of the World Health Organization Global Survey on Maternal and Perinatal Health." *Environmental Health Perspectives*, 122(4): 425–430.

Fu, Li, Yi Chen, Xinyi Yang, Zuyao Yang, Sha Liu, Lei Pei, Baixiang Feng, Ganxiang Cao, Xin Liu, Hualiang Lin, Xing Li, Yufeng Ye, Bo Zhang, Jiufeng Sun, Xiaojun Xu, Tao Liu, and Wenjun Ma. 2019. "The associations of air pollution exposure during pregnancy with fetal growth and anthropometric measurements at birth: a systematic review and meta-analysis." *Environmental Science and Pollution Research*, 26(20): 20137–20147.

Ghosh, Rakesh, Judith Rankin, Tanja Pless-Mulloli, and Svetlana Glinianaia. 2007. "Does the effect of air pollution on pregnancy outcomes differ by gender? A systematic review." *Environmental Research*, 105(3): 400–408.

Grafova, Irina B., Vicki A. Freedman, Nicole Lurie, Rizie Kumar, and Jeannette Rogowski. 2014. "The difference-in-difference method: Assessing the selection bias in the effects of neighborhood environment on health." *Economics & Human Biology*, 13: 20–33.

Guan, Wei-Jie, Xue-Yan Zheng, Kian Fan Chung, and Nan-Shan Zhong. 2016. "Impact of air pollution on the burden of chronic respiratory diseases in China: time for urgent action." *The Lancet*, 388(10054): 1939–1951.

Hajat, Anjum, Charlene Hsia, and Marie S. O'Neill. 2015. "Socioeconomic Disparities and Air Pollution Exposure: a Global Review." *Current Environmental Health Reports*, 2(4): 440–450.

Han, Yingying, Panhua Jiang, Tianyu Dong, Xinliang Ding, Ting Chen, Gro Dehli Villanger, Heidi Aase, Lu Huang, and Yankai Xia. 2018. "Maternal air pollution exposure and preterm birth in Wuxi, China: Effect modification by maternal age." *Ecotoxicology and Environmental Safety*, 157: 457–462.

Health Effects Institute. 2019. "State of Global Air 2019." Health Effects Institute Special Report ISSN 2578-6873, Boston, MA.

He, Guojun, Maoyong Fan, and Maigeng Zhou. 2016. "The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games." *Journal of Environmental Economics and Management*, 79: 18–39.

He, Jian-Rong, Yu Liu, Xiao-Yan Xia, Wen-Jun Ma, Hua-Liang Lin, Hai-Dong Kan, Jin-Hua Lu, Qiong Feng, Wei-Jian Mo, Ping Wang, Hui-Min Xia, Xiu Qiu, and Louis J.

Muglia. 2016. "Ambient Temperature and the Risk of Preterm Birth in Guangzhou, China (2001–2011)." *Environmental Health Perspectives*, 124(7): 1100–1106.

Huang, Ivory, Diane Mak, Prudence Cheung, Miriam Abraham, Tom Clemens, and Steve Turner. 2019. "A systematic review of associations between maternal exposures during pregnancy other than smoking and antenatal fetal measurements." *Environmental Research*, 173: 528–538.

Ito, Kazuhiko, George D Thurston, and Robert A Silverman. 2007. "Characterization of PM2.5, gaseous pollutants, and meteorological interactions in the context of time-series health effects models." *Journal of Exposure Science & Environmental Epidemiology*, 17(S2): S45–S60.

Jacobs, Milena, Guicheng Zhang, Shu Chen, Ben Mullins, Michelle Bell, Lan Jin, Yuming Guo, Rachel Huxley, and Gavin Pereira. 2017. "The association between ambient air pollution and selected adverse pregnancy outcomes in China: A systematic review." *Science of The Total Environment*, 579: 1179–1192.

Jayachandran, Seema. 2009. "Air quality and early-life mortality evidence from Indonesia's wildfires." *Journal of Human Resources*, 44(4): 916–954.

Keitt, Sarah K, Thomas F Fagan, and Sherry A Marts. 2004. "Understanding sex differences in environmental health: a thought leaders' roundtable." *Environmental Health Perspectives*, 112(5): 604–609.

Kim, Younoh, Scott Knowles, James Manley, and Vlad Radoias. 2017. "Long-run health consequences of air pollution: Evidence from Indonesia's forest fires of 1997." *Economics & Human Biology*, 26: 186–198.

Liang, Zhijiang, Yin Yang, Jing Yi, Zhengmin Qian, Zilong Zhang, Stephen Edward McMillin, Echu Liu, Hualiang Lin, and Guocheng Liu. 2021. "Maternal PM2.5 exposure associated with stillbirth: A large birth cohort study in seven Chinese cities." *International Journal of Hygiene and Environmental Health*, 236: 113795.

Li, Shanshan, Yuming Guo, and Gail Williams. 2016. "Acute Impact of Hourly Ambient Air Pollution on Preterm Birth." *Environmental Health Perspectives*, 124(10): 1623–1629.

Liu, Huan, Xuemei Wang, Jinpu Zhang, Kebin He, Ye Wu, and Jiayu Xu. 2013. "Emission controls and changes in air quality in Guangzhou during the Asian Games." *Atmospheric Environment*, 76: 81–93.

Liu, Shiliang, Daniel Krewski, Yuanli Shi, Yue Chen, and Richard T Burnett. 2003. "Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada." *Environmental Health Perspectives*, 111(14): 1773–1778.

Liu, Xiaoying, Jere Behrman, Emily Hannum, Fan Wang, and Qingguo Zhao. 2021. "Same Environment, Stratified Impacts? Air Pollution, Extreme Temperatures, and Birth Weight in Southeast China." PIER Working Paper No. 21-001. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3761257.

Liu, Ying, Jihong Xu, Dian Chen, Pei Sun, and Xu Ma. 2019. "The association between air pollution and preterm birth and low birth weight in Guangdong, China." *BMC Public Health*, 19(1): 1–10.

Lu, Chan, Lanqin Cao, Dan Norbäck, Yuguo Li, Jing Chen, and Qihong Deng. 2019. "Combined effects of traffic air pollution and home environmental factors on preterm birth in China." *Ecotoxicology and Environmental Safety*, 184: 109639.

Lu, Chan, Weishe Zhang, Xiangrong Zheng, Jingchi Sun, Lv Chen, and Qihong Deng. 2020. "Combined effects of ambient air pollution and home environmental factors on low birth weight." *Chemosphere*, 240: 124836.

Luechinger, Simon. 2014. "Air pollution and infant mortality: A natural experiment from power plant desulfurization." *Journal of Health Economics*, 37: 219–231.

Mercer, Laina D., Adam A. Szpiro, Lianne Sheppard, Johan Lindström, Sara D. Adar, Ryan W. Allen, Edward L. Avol, Assaf P. Oron, Timothy Larson, L. J. Sally Liu, and Joel D. Kaufman. 2011. "Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air)." *Atmospheric Environment*, 45(26): 4412–4420.

Murray, L. J, D. P. J O'Reilly, N Betts, C. C Patterson, G Davey Smith, and A. E Evans. 2000. "Season and outdoor ambient temperature: effects on birth weight." *Obstetrics & Gynecology*, 96(5, Part 1): 689–695.

Møllegaard, Stine. 2020. "The effect of birth weight on behavioral problems in early adolescence: New evidence from monozygotic twins." *Economics Human Biology*, 36: 100828.

National Aeronautics and Space Administration, U.S.A. 2020. "Airborne Nitrogen Dioxide Plummets Over China." https://earthobservatory.nasa.gov/images/146362/airborne-nitrogen-dioxide-plummets-over-china.

Qian, Zhengmin, Shengwen Liang, Shaoping Yang, Edwin Trevathan, Zhen Huang, Rong Yang, Jing Wang, Ke Hu, Yiming Zhang, Michael Vaughn, Longjiao Shen, Wenjin Liu, Pu Li, Patrick Ward, Li Yang, Wei Zhang, Wei Chen, Guanghui Dong, Tongzhang Zheng, Shunqing Xu, and Bin Zhang. 2016. "Ambient air pollution and preterm birth: A prospective birth cohort study in Wuhan, China." *International Journal of Hygiene and Environmental Health*, 219(2): 195–203.

Rangel, Marcos A, and Tom S Vogl. 2019. "Agricultural fires and health at birth." *The Review of Economics and Statistics*, 101(4): 616–630.

Rich, David Q., Kaibo Liu, Jinliang Zhang, Sally W. Thurston, Timothy P. Stevens, Ying Pan, Cathleen Kane, Barry Weinberger, Pamela Ohman-Strickland, Tracey J. Woodruff, Xiaoli Duan, Vanessa Assibey-Mensah, and Junfeng Zhang. 2015. "Differences in Birth Weight Associated with the 2008 Beijing Olympics Air Pollution Reduction: Results from a Natural Experiment." *Environmental Health Perspectives*, 123(9): 880–887.

Sanders, Nicholas J. 2012. "What doesn't kill you makes you weaker: prenatal pollution exposure and educational outcomes." *Journal of Human Resources*, 47(3): 826–850.

Sauer, Mark V. 2015. "Reproduction at an advanced maternal age and maternal health." *Fertility and Sterility*, 103(5): 1136–1143.

Shah, Prakesh S., Taiba Balkhair, and Knowledge Synthesis Group on Determinants of Preterm/LBW births. 2011. "Air pollution and birth outcomes: a systematic review." *Environment International*, 37(2): 498–516. 00142.

Siniarska, A., and S. Kozieł. 2010. "Association of birth weight and length with air temperature, sunlight, humidity and rainfall in the city of Warsaw, Poland." *HOMO*, 61(5): 373–380.

Šrám, Radim J, Blanka Binková, Jan Dejmek, and Martin Bobak. 2005. "Ambient air pollution and pregnancy outcomes: a review of the literature." *Environmental Health Perspectives*, 375–382.

Strand, Linn Beate, Adrian G. Barnett, and Shilu Tong. 2011. "Methodological challenges when estimating the effects of season and seasonal exposures on birth outcomes." *BMC Medical Research Methodology*, 11(1): 1–8.

Sun, Zhe, Liyang Yang, Xiaoxia Bai, Wei Du, Guofeng Shen, Jie Fei, Yonghui Wang, An Chen, Yuanchen Chen, and Meirong Zhao. 2019. "Maternal ambient air pollution exposure with spatial-temporal variations and preterm birth risk assessment during 2013–2017 in Zhejiang Province, China." *Environment International*, 133: 105242.

Tang, Deliang, Ting Yu Li, Judith C Chow, Sanasi U Kulkarni, John G Watson, Steven Sai Hang Ho, Zhang Y Quan, LR Qu, and Frederica Perera. 2014. "Air pollution effects on fetal and child development: A cohort comparison in China." *Environmental Pollution*, 185: 90–96.

Torche, Florencia, and Dalton Conley. 2016. "A Pound of Flesh: The Use of Birthweight as a Measure of Human Capital Endowment in Economics Research." *The Oxford Handbook of Economics and Human Biology*.

Tsoli, Stergiani, George B. Ploubidis, and Olga-Ioanna Kalantzi. 2019. "Particulate air pollution and birth weight: A systematic literature review." *Atmospheric Pollution Research*, 10(4): 1084–1122.

Wang, Qiong, Huazhang Miao, Joshua L. Warren, Meng Ren, Tarik Benmarhnia, Luke D. Knibbs, Huanhuan Zhang, Qingguo Zhao, and Cunrui Huang. 2021. "Association of maternal ozone exposure with term low birth weight and susceptible window identification." *Environment International*, 146: 106208.

World Health Organization. 2019. "Ambient air pollution: Health impacts." http://www.who.int/airpollution/ambient/health-impacts/en/.

Xie, Yanxia, Xiaodong Wang, Yi Mu, Zheng Liu, Yanping Wang, Xiaohong Li, Li Dai, Qi Li, Mingrong Li, Peiran Chen, Jun Zhu, and Juan Liang. 2021. "Characteristics and adverse outcomes of Chinese adolescent pregnancies between 2012 and 2019." *Scientific Reports*, 11(1): 12508.

Xu, H.M., J. Tao, Steven Sai Hang Ho, K.F. Ho, J.J. Cao, N. Li, Judith C. Chow, G.H. Wang, Y.M. Han, R.J. Zhang, John G. Watson, and J.Q. Zhang. 2013. "Characteristics of fine particulate non-polar organic compounds in Guangzhou during the 16th Asian Games: Effectiveness of air pollution controls." *Atmospheric Environment*, 76: 94–101.

Yang, Yuxiang. 2021. "Analysing the seasonality of births in mainland China." *Journal of Biosocial Science*, 53(2): 233–246. Publisher: Cambridge University Press.

Zhu, P. F., Y. Zhang, J. Ban, T. T. Li, and X. M. Shi. 2017. "Air pollution and adverse birth outcome in China: a comprehensive review." *Zhonghua Liu Xing Bing Xue Za Zhi (Chinese)*, 38(3): 393–399.

Appendix 1: Tables and Figures

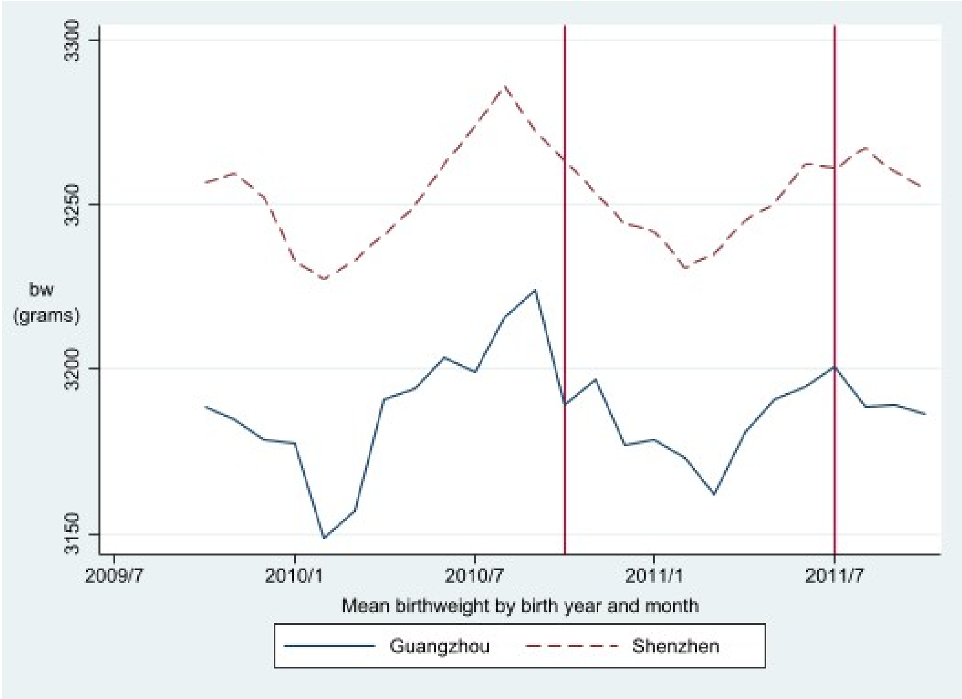

Figure A1: Time Trend of Mean Birthweight by City and Month

Note: Monthly average birth weight is calculated based on birth certificate for each city at each birth month. As Asian Games were held during Oct 01, 2010 and Dec 20, 2010, any births with prenatal periods overlapped with this period (or born between October 2010 and September 2011) are included in the area between the two vertical lines.

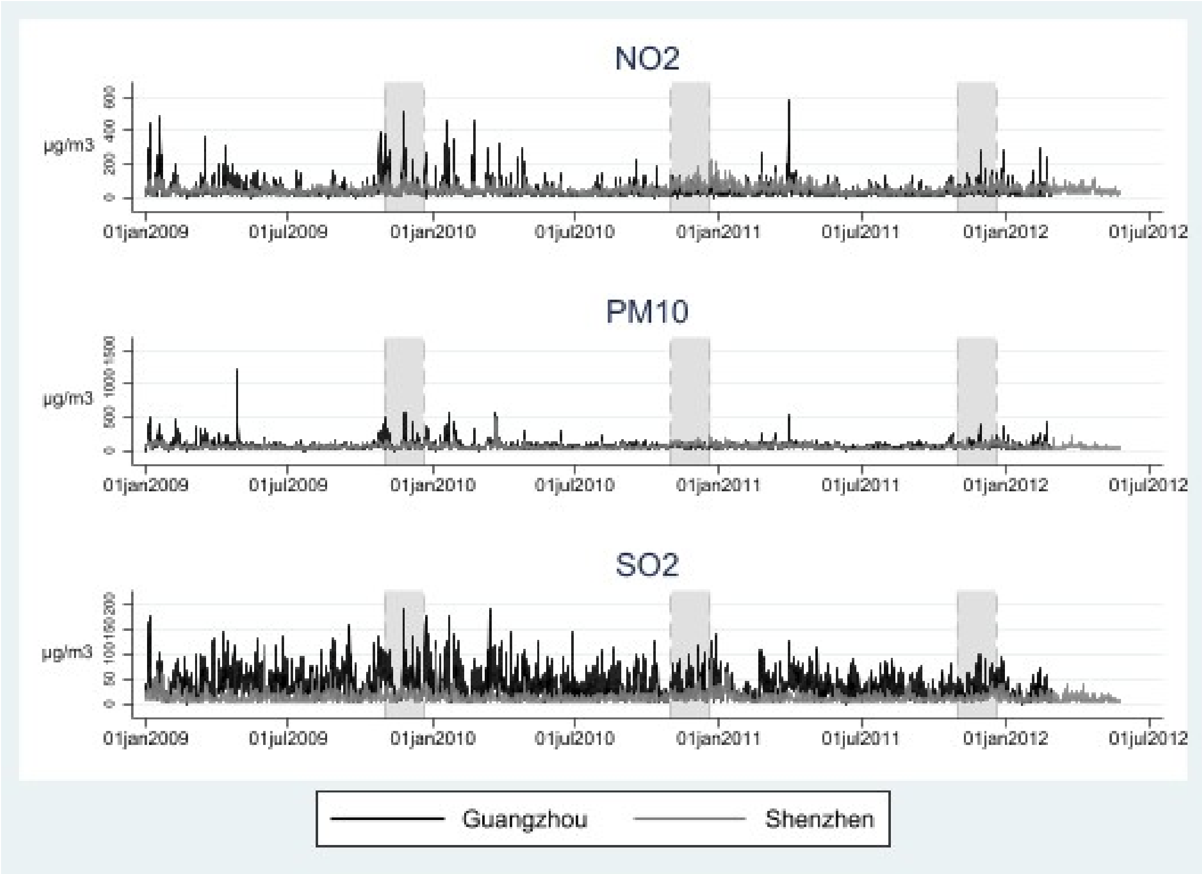

Figure A2: The Daily Concentrations of Three Air Pollutants in Guangzhou and Shenzhen Cities During 2009-2011

Note: The daily concentrations of each pollutant (PM_{10}, NO_2 and SO_2 are the average concentration across the whole city. The three periods shaded in grey between two neighboring vertical dashed lines are the same season as the Asian Games in each year, i.e., October 20-December 20 in 2009, 2010, and 2011, separately.

Figure A3: The Number of Live Single Births by Month

Table A1: Study Population Characteristics and Exposure Measurements by City [n (%) or mean(sd)]

	(1) Sample		(2) Guangzhou		(3) Shenzhen		(4) p value for
	mean/n	sd/%	mean/n	sd/%	mean/n	sd/%	t or Chi-square test
Sex (female = 1)	0.46	(0.50)	0.47	(0.50)	0.46	(0.50)	0.57
Preterm	0.05	(0.22)	0.07	(0.26)	0.05	(0.21)	0.00
Birthweight	3240.44	(456.73)	3181.36	(469.68)	3248.15	(454.45)	0.00
Low Birthweight	0.04	(0.20)	0.06	(0.23)	0.04	(0.19)	0.00
Maternal age	27.31	(4.74)	29.06	(4.19)	27.08	(4.76)	0.00
Parity	1.46	(0.72)	1.25	(0.52)	1.49	(0.74)	0.00
Mother's education, n (%)							0.00
< high school	194,990	(37.3%)	3,641	(6.03%)	191,349	(41.38%)	
high school graduate or equivalent	198,140	(37.91%)	36,895	(61.15%)	161,245	(34.87%)	
3 year college	63,226	(12.1%)	8,252	(13.68%)	54,974	(11.89%)	
4 year college graduate	66,347	(12.69%)	11,546	(19.14%)	54,801	(11.85%)	
Mother's job category, n (%)							0.00
Staff	56,596	(10.83 %)	56,596	(93.80 %)	73,438	(15.88 %)	
Agricultural related	5,983	(1.14%)	276	(0.46%)	5,707	(1.23%)	
Housewife	129,749	(24.82%)			129,749	(28.06%)	
Professional and technical personnel	2,050	(0.39%)	497	(0.82%)	1,553	(0.34%)	
Business	26,109	(4.99%)	202	(0.33%)	25,907	(5.60%)	
Worker	35,587	(6.81%)	76	(0.13%)	35,511	(7.68%)	
Other	2,687	(0.51%)	2,687	(4.45%)			
Missing	190,504	(36.45%)			190,504	(41.20%)	
Daily mean PM_{10} exposure(μg/m^3)	60.41	(7.37)	72.33	(7.08)	58.85	(5.82)	0.00
Daily mean NO_2 exposure(μg/m^3)	41.93	(5.09)	40.34	(5.98)	42.14	(4.92)	0.00
Daily mean SO_2 exposure(μg/m^3)	14.25	(7.86)	35.29	(5.17)	11.50	(1.05)	0.00
Daily mean temperature(°C)	23.18	(1.57)	22.76	(1.80)	23.23	(1.52)	0.00
Daily mean relative humidity(%)	75.37	(3.34)	72.17	(2.91)	75.79	(3.16)	0.00
% Time exposed to UTCI $<= 0°C$	1.39	(0.87)	1.96	(1.19)	1.31	(0.80)	0.00
% Time exposed to UTCI $>= 34°C$	0.43	(0.95)	2.19	(1.87)	0.20	(0.32)	0.00
N	522703		60334		462369		

Notes: Birth weight and gestational age are obtained from birth records from each city. Gestational age is calculated based on self-reported last menstrual period. The remaining variables of individual characteristics are self-reported.

Table A2: Distribution of Daily Mean Universal Thermal Climate Indices (UTCI) in Guangzhou and Shenzhen Between 2000 and 2008.

	Guangzhou	Shenzhen
Top 1% (°C)	34.23	33.35
Top 2.5% (°C)	33.56	32.74
Bottom 2.5% (°C)	2.52	3.82
Bottom 1% (°C)	-0.87	-0.24
Mean (°C)	22.61	22.46
s.d.	8.74	8.05
Number of days above 34 °C	45	12
Number of days above 33 °C	172	54

Source: Copernicus and the European Centre for Medium-Range Weather Forecasts to combine measurements of temperature and humidity (European Centre for Medium-range Weather Forecasts, 2020*b*).

Table A3: First-stage coefficients of different definitions of instrumental variables on prenatal exposure levels to different pollutants

	(1) PM_{10}	(2) NO_2	(3) SO_2
Specification A: temperature + relative humidity	-0.161***	-0.262***	-0.126***
	(0.000552)	(0.000643)	(0.000388)
Specification B: extreme cold + extreme hot defined by UTCI	-0.164***	-0.236***	-0.108***
	(0.000567)	(0.000732)	(0.000483)

Notes: Each estimate represents a regression coefficient of the first-stage regression of prenatal exposure to each pollutant on the instrumental variable, i.e., a continuous variable summarizing the number of days during one's expected pregnancy (given 39 weeks of gestation) overlapping with the Asian Games period with two specifications: Specification A controls for linear term of daily mean temperature and daily mean relative humidity; Specification B controls for extreme cold and hot weather during pregnancy defined by UTCI. Besides meteorological variables, other control variables include sex, quadratic term of maternal age, parity, education categories, day-of-a-week dummies, conception year fixed effects, city-specific conception season fixed effects. *** significant at 1%. Robust standard errors are displayed in parentheses.

Table A4: Marginal Effects of NO_2 (by 10 μg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm.

	Probit Model/OLS Model						IV-Probit Model/2SLS Model					
VARIABLES	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NO_2*μg/m^3	0.002***	0.002***	-8.56***	-7.074***	0.004***	0.003***	0.0020*	0.0016	-8.70***	-9.28**	0.0038***	0.0033**
	(0.001)	(0.001)	(2.311)	(1.866)	(0.001)	(0.001)	(0.001)	(0.001)	(3.16)	(3.72)	(0.001)	(0.002)
Temp. (°C)	-.044 ***		115.42***		-0.054 ***		-0.004***		11.536***		-.0053691 ***	
	(0.004)		(9.297)		(0.004)		(0.000)		(0.93)		(0.00041)	
Rel. hum. (%)	0.011***		-32.79***		0.016***		0.001***		-3.296***		.0015667***	
	(0.002)		(4.523)		(0.002)		(0.000)		(0.51)		(0.00023)	
Ext. cold (%)		0.021***		-55.506***		.029***		0.002***		-5.581***		0.0029***
		(0.005)		(1.247)		(0.006)		(0.001)		(1.25)		(0.001)
Ext. hot (%)		9.86E-05		10.052		0.003		0.000		1.49		0.00026
		(0.005)		(12.226)		(0.005)		(0.001)		(1.40)		(0.001)
Obs.	522,703						522,703					
Spec.	A	B	A	B	A	B	A	B	A	B	A	B

Notes: Each column represents a separate regression. (1)-(6) are from Probit/OLS regressions without applying IV, while (7)-(12) are from IVProbit/2SLS regressions. (1),(2),(5),(6),(7),(8),(11),(12) are marginal effects of 10 μg/m^3 NO_2 on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4),(9),(10) are marginal effects of 10μg/m^3 No_2 on Mean Birthweight. Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Table A5: Marginal Effects of SO_2 (by 10 μg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm.

	Probit Model/OLS Model						IV-Probit Model/2SLS Model					
VARIABLES	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SO_2* 10μg/m^3	0.01***	0.009***	-15.97***	-14.36***	0.014***	0.0133***	0.0045*	0.0031	-18.12***	-20.34**	.0083 ***	0.0076**
	(0.002)	(0.002)	(4.213)	(4.741)	(0.002)	(0.002)	(0.002)	(0.003)	(6.570)	(8.160)	(0.003)	(0.003)
Temp. (°C)	-0.0048***		12.298***		-0.00589***		-0.005***		12.354***		-0.0057 ***	
	(0.000)		(0.931)		(0.000)		(0.000)		(0.94)		(0.00415)	
Rel. hum. (%)	0.0011***		-2.531***		0.00137***		0.001***		-2.564***		0.0013***	
	(0.000)		(0.370)		(0.000)		(0.000)		(0.38)		(0.00178)	
Ext. cold (%)		0.002***		-5.639***		0.003***		0.002***		-5.716***		0.003***
		(0.001)		(1.248)		(0.001)		(0.001)		(1.251)		(0.001)
Ext. hot (%)		-0.0007		1.267		-0.00072		0.000		2.023		0.000
		(0.001)		(1.299)		(0.001)		(0.001)		(1.524)		(0.001)
Obs.	522,703						522,703					
Spec.	A	B	A	B	A	B	A	B	A	B	A	B

Notes: Each column represents a separate regression. (1)-(6) are from Probit/OLS regressions without applying IV, while (7)-(12) are from IVProbit/2SLS regressions. (1),(2),(5),(6),(7),(8),(11),(12) are marginal effects of 10 μg/m^3 SO_2 on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4),(9),(10) are marginal effects of 10μg/m^3 SO_2 on Mean Birthweight. Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Table A6: Heterogeneous Effects of NO_2 on Birth Outcomes by Sex and Maternal Age Group

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
NO_2 * male	0.000414	0.000419	2.086**	2.086**	-.0012	-.0011
	(0.005)	(0.005)	(1.066)	(1.066)	(0.004)	(0.004)
NO_2 * (momage>=35)	0.00169	0.00169	-7.757***	-7.747***	0.00199	0.00199
	(0.00126)	(0.00126)	(2.417)	(2.417)	(0.00142)	(0.00142)
Specification	A	B	A	B	A	B

Notes: IV estimates of the interaction term between NO_2 and male or age35+ on birth outcomes, with the same set of control variables as in Table 2. *** significant at 1%, ** significant at 5%, *significant at 10%.

Table A7: Heterogeneous Effects of SO_2 on Birth Outcomes by Sex and Maternal Age Group

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
SO_2 * male	-.000151	-0.000151	-0.806**	-0.808*	0.000450	0.000452
	(0.002)	(0.002)	(0.412)	(0.412)	(0.002)	(0.002)
SO_2 * (momage>=35)	0.000483	0.000483	-2.218***	-2.214***	0.000569	0.000569
	(0.000359)	(0.000359)	(0.684)	(0.684)	(0.000405)	(0.000405)
Specification	A	B	A	B	A	B

Notes: IV estimates of the interaction term between SO_2 and male or age35+ on birth outcomes, with the same set of control variables as in Table 2. *** significant at 1%, ** significant at 5%, *significant at 10%.

Appendix 2: Robustness check: missing observations of labor migrants

There still may be concern over labor migrants in the sample, as these two cities are the most popular destinations for labor migrants in China. Labor migrants may return to their hometowns for child birth. Although this should not be a concern for this study as long as labor migrants do not leave the city for child birth specifically during the Asian Games period, we do find that the number of live deliveries is lower during thr Asian Games period in Guangzhou (20 October 2010 - 20 December 2010) compared to the same period in other years, relative to Shenzhen (see Figure A3.). As we do not observe migration status of women in the Guangzhou sample, we try to control for mothers' job types as a sensitivity check as labor migrants are more likely to work as "workers". The main results remain basically the same, as shown in Table A8. Because mothers' job type information was not collected in a consistent framework between the two cities, and 40.3% of the population have missing information on job, [19] we do not present these results as the main results. The similar results on air pollution exposure after controlling for mothers' job categories should help alleviate concern over missing labor migrants in the sample.

[19] In the regression that controls for maternal job category, we treat missing job information as a job category because it is likely that missing is not random.

Table A8: Marginal Effects of PM_{10}, NO_2 and SO_2 (by 10 μg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm (by IV-probit/2SLS Estimates, with Control of Mother's Job Category).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	PM_{10}						NO_2						SO_2					
VARIABLES	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
Pollutant(10μg/m^3)	0.003	0.0016	-11.69**	-11.33**	0.0061***	0.0040*	0.000	0.0012	-7.16**	-7.84**	0.0035***	0.0028*	0.000	0.0029	-14.94**	-17.19**	0.0077***	0.0068**
	(0.002)	(0.002)	(0.50)	(0.52)	(0.000)	(0.002)	(0.000)	(0.001)	(3.06)	(3.62)	(0.001)	(0.002)	(0.000)	(0.000)	(6.38)	(7.93)	(0.003)	(0.003)
Temp.(°C)	-.0041***		11.09***		-.0047***		-.0044***		11.73***		-.0053		-0.0046***		12.41***		-0.0056***	
	(0.000)		(1.01)		(0.000)		(0.000)		(0.94)		(0.000)		(0.000)		(9.45)		(0.000)	
Rel. hum.(%)	0.0012***		-3.54***		0.0017***		0.0011***		-3.26***		0.0016***		0.001***		-2.66***		0.0013***	
	(0.000)		(0.60)		(0.000)		(0.000)		(0.51)		(0.000)		(0.000)		(0.38)		(0.000)	
Ext. cold (%)		0.0019***		-6.56***		0.0027***		0.002***		-5.88***		0.0029***		0.0021***		-5.99***		0.0029***
		(0.001)		(1.305)		(0.001)		(0.001)		(1.25)		(0.001)		(0.001)		(1.26)		(0.001)
Ext. hot (%)		0.000		0.16		0.00090*		0.000		1.51		0.00035		0.000		1.96		0.000
		(0.000)		(1.169)		(0.001)		(0.001)		(1.40)		(0.001)		(0.001)		(1.53)		(0.001)
Obs.	516,720						516,720						516,720					
Specification	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Mother's job	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: The sample removes individuals who conduct agricultural related job. Each column represents a separate regression from IV-Probit/2SLS regressions. (1),(2),(5),(6) are marginal effects of 10 μg/m^3 PM_{10} on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4) are marginal effects of 10μg/m^3 PM_{10} on Mean Birthweight (similar for other two pollutants). Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, maternal job category, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Highlights:

1. Effects of prenatal exposure to PM_{10}, NO_2 and SO_2 on low birthweights, birthweights and prematurity risks in South China are assessed.
2. Identification comes from exogenous improvement in air quality during the 2010 Guangzhou Asian Games.
3. Air pollutants significantly reduce average birthweights and increase preterm risks, especially in the third trimester for PM_{10}.
4. Male babies are more vulnerable to PM_{10} and SO_2, but birthweights of female babies are more sensitive to NO_2.
5. For birthweights, babies of mothers who are at least 35 years old show more vulnerability to all three air pollutants.

The Asian Games, Air Pollution and Birth Outcomes in South China: an Instrumental Variable Approach[*,†]

Xiaoying Liu[‡] Huazhang Miao[§] Jere R. Behrman[¶] Emily Hannum[||]
Zhijiang Liang[**] Qingguo Zhao[††]

September 7, 2021

[*]Liu, Behrman and Hannum acknowledge funding for this project from the National Science Foundation (No:1756738) and the China Research and Engagement Fund and School of Arts and Sciences Dean's Integrative Global Inquiries Fund, both at the University of Pennsylvania. We thank Janet Currie, Guy Grossman, Heather Schofield, Herbert Smith, Harsha Thirumurthy, Fan Wang, and two anonymous referees and the editor for helpful suggestions. The authors all declare they have no actual or potential competing financial interests.

[†]This research has been reviewed by the Research Ethics Committee at Guangdong Women and Children Hospital [201601032] and the Institutional Review Board at the University of Pennsylvania [#825326].

[‡]*Corresponding author.* Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104 (email: xiaoyliu@sas.upenn.edu)

[§]Guangdong Women and Children Hospital, Guangzhou, Guangdong Province, China

[¶]Department of Economics and Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA

[||]Department of Sociology and Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, USA

[**]Guangdong Women and Children Hospital, Guangzhou, Guangdong Province, China

[††]*Corresponding author.* Epidemiological Research Office of Key Laboratory of Male Reproduction and Genetics (National Health and Family Planning Commission), Family Planning Research Institute of Guangdong Province, Guangzhou, Guangdong Province, China (email: zqgfrost@126.com).

Abstract

We estimate the effects of air-pollution exposure on low birthweight, birthweight, and prematurity risk in South China, for all expectant mothers and by maternal age group and child sex. We do so by exploiting exogenous improvement in air quality during the 2010 Guangzhou Asian Games, when strict regulations were mandated to assure better air quality. We use daily air-pollution levels collected from monitoring stations in Guangzhou, the Asian Games host city, and Shenzhen, a nearby control city, between 2009 and 2011. We first show that air quality during the Asian Games significantly improved in Guangzhou, relative to Shenzhen. Next, using birth-certificate data for both cities for 2009 to 2011 and using expected pregnancy overlap with the Asian Games as an instrumental variable, we study the effects of three pollutants (PM_{10}, SO_2, and NO_2) on birth outcomes. Four main conclusions emerge: 1) air pollutants significantly reduce average birthweight and increase preterm risk; 2) for birthweight, late pregnancy is most sensitive to PM_{10} exposure, but there is not consistent evidence of a sensitive period for other pollutants and outcomes; 3) for birthweight, babies of mothers who are at least 35 years old show more vulnerability to all three air pollutants; and 4) male babies show more vulnerability than female babies to PM_{10} and SO_2, but birthweights of female babies are more sensitive than those of male babies to NO_2.

Key words: ambient air pollution, birthweights, preterm births, instrumental variable, China

JEL codes: I1, I12, I14, I18, Q51, Q52, Q53, Q54

1 Introduction

In 2016, 91% of the world's population lived in areas with air pollution that exceeded the World Health Organization air-quality guidelines for ambient fine particulate matter $PM_{2.5}$ – airborne particles less than or equal to 2.5 micrometers in aerodynamic diameter.[1] However, nearly all (86%) of the most extreme concentrations – above 75 $\mu g/m^3$ – were experienced by populations in China, India, Pakistan, and Bangladesh (Health Effects Institute, 2019). The World Health Organization estimates that 4.2 million premature deaths every year are linked to ambient air pollution (World Health Organization, 2019). Pregnant women and children may be particularly vulnerable to the ill effects of air pollution. Despite natural barriers protecting infants in utero, prenatal air pollution exposures have been linked to adverse birth outcomes (Fu et al., 2019; Huang et al., 2019; Šrám et al., 2005; Tsoli, Ploubidis and Kalantzi, 2019; Currie et al., 2014), and effects may persist into later childhood (Almond, Edlund and Palme, 2009; Almond, Currie and Duque, 2018; Currie et al., 2014; Tang et al., 2014; Bharadwaj et al., 2017; Sanders, 2012).[2] However, reviews suggest that evidence from epidemiological studies is inconsistent and causal links are hard to identify (Šrám et al., 2005; Bosetti et al., 2010).

In a separate line of research, economists have capitalized on novel natural experiments to come closer to estimating effects of air pollution during the prenatal period on infant health. For example, Jayachandran (2009), Rangel and Vogl (2019), and Kim et al. (2017) exploit incidents of natural fire or biomass fire and wind direction to identify the effect of air-pollution exposure on birth outcomes or longer-term health in Indonesia and Brazil. Currie and Walker (2011) exploit the introduction of electronic toll collection in the US and show reduced traffic pollution brought by electronic toll collection reduced prematurity and low birthweight among mothers within two kilometers of a toll plaza relative to mothers two to ten kilometers away. Luechinger (2014) exploits mandated desulfurization at power plants and prevailing wind directions to study the effects on infant mortality in Germany.[3]

Our paper follows this strand of literature by exploiting an international sports event: the

[1] The standard is an annual mean concentration of 10 $\mu g/m^3$.

[2] Mechanisms of air pollution effects on fetal development have been studied in animal models. For example, prenatal exposure to SO_2 leads to developmental and functional toxicities; NO_2 suppresses antioxidant defense systems; CO interferes with oxygen delivery to fetuses by displacement of oxygen from hemoglobin; and early fetal exposure to PM can alter trophoblast formation and vascularization of the placenta (Shah, Balkhair and Knowledge Synthesis Group on Determinants of Preterm/LBW births, 2011).

[3] See Currie et al. (2014) for a review of earlier studies.

2010 Guangzhou Asian Games, during which air quality was exogenously improved. We estimate, for the first time to our knowledge, air-quality impact on birth outcomes in South China. Specifically, assuming 39 weeks of gestation, we use the time during pregnancy that overlapped with the Asian Games as an instrumental variable to identify exogenous variation in air-pollution exposure and estimate the causal impacts on birth outcomes in response to such an exogenous change. Our approach is also novel in assessing heterogeneity in the impact of air pollution on birth outcomes by child sex and by maternal age group.

This paper is structured as follows. In Section 2, we summarize several issues commonly seen in the literature on air pollution and birth outcomes. In Section 3, we describe the air-quality regulations implemented before and during the 2010 Guangzhou Asian Games. We introduce the data we use and measurements of air pollution and weather exposures in Section 4. The empirical methods and main results are presented in Sections 5 and 6, respectively. Section 7 examines heterogeneity of effects by child sex and maternal age. Section 8 provides discussion and conclusions.

2 Background and Research Objectives

Deterioration in China's air quality since the 1990s and increased public availability of air quality data since 2014 have prompted a proliferation of research on the health effects of air pollution (Dong et al., 2013; Guan et al., 2016), including studies on air pollution and fetal growth and pregnancy outcomes (Qian et al., 2016; Zhu et al., 2017; Fu et al., 2019; Lu et al., 2019). Much of this literature studied associations between air pollution and birth outcomes, including infant mortality, low birthweight, and prematurity risk. However, most existing studies using individual birth information from China have suffered from one or both of the following limitations.

First, most of these studies use individual-level birth data from a single city within a relatively short time period–two to three years.[4] But multi-city and multi-year studies have statistical power advantages (Ito, Thurston and Silverman, 2007). In the single-city design, moreover, it is difficult to disentangle air pollution from meteorological factors as they usually co-vary and both can have impacts on birth outcomes. Multi-city samples with bigger variation in both air pollution and weather therefore have advantages because different cities

[4] Two recent papers from China are exceptions: Sun et al. (2019) covers all of Zhejiang Province, though this work does not control for weather, and Liu et al. (2019) covers the whole of Guangdong Province.

usually have different compositions of air pollution and patterns of interactions between air pollution and weather.

Second, most epidemiological studies fail to consider pregnancy timing and self-selection of residence. Pregnancy timing, either conscious or unconscious, is a significant phenomenon in many parts of the world. For example, using China's Fifth National Population Census, Yang (2021) finds there is a salient peak of birth numbers in October, which is mainly driven by labor migrants who often get married and conceive in the Chinese New Year period (usually between late January and mid February). Since labor migrants, on average, have lower educational attainment than non-migrants, ceteris paribus, failing to consider changes over time in a year in the composition of the pregnant population by controlling for maternal education could potentially bias the estimates of air pollution effects. However, in a recent review, only 7 out of 25 studies on air pollution and birth outcomes in China control for maternal education (Jacobs et al., 2017). On the other hand, people choose where to live based on their social economic status (SES) and preferences. People with lower SES may live in areas with higher pollution. Hajat, Hsia and O'Neill (2015)'s global review of socioeconomic disparities in air pollution exposure shows that areas where low SES communities dwell experience higher concentrations of criteria air pollutants. Since SES, usually imperfectly measured, also has effects on health, failing to take into account this self-selection may over- or underestimate the effect of the environment. For example, Grafova et al. (2014) discuss the selection bias in the effects of neighborhood environment on health, and find conventional estimates underestimate the effect of the economic environment on health.

Beyond these limitations, there is not yet agreement in existing studies regarding the important possibility of heterogeneous vulnerability to pollution effects across key demographic groups characterized child sex and maternal age. DiPietro and Voegtline (2017) review converging evidence that suggests that infant and early childhood developmental outcomes of male fetuses exposed to prenatal adversities are more highly impaired than those of female fetuses. Moreover, there is some suggestive evidence that the effect of air pollution on low birthweight differs by sex of the child (Ghosh et al., 2007). While risks for adverse birth outcomes increase with advanced maternal age (for a summary, see Sauer (2015)), the question of whether there is heightened vulnerability to air pollution exposures for women at advanced maternal age is not yet well studied.[5]

[5] Young maternal age is also a risk factor for poor pregnancy outcomes, though for different reasons. For evidence

Studies from China have shown varied results regarding sex differences in vulnerability to prenatal air pollution exposure. For example, Lu et al. (2019) find heterogeneous associations between preterm risk and perinatal exposure to indoor mold/damp stains by sex using retrospective recall data from a single city, but do not find any significant difference in associations with NO_2 exposure by sex. Wang et al. (2021) find that O_3 exposures are associated with low birthweight of female – but not male – babies using birth-certificate data in Guangzhou between January 2015 and July 2017. Chen and Ho (2016) find that incense burning during pregnancy is associated with significantly lower birthweight and smaller head circumference in boys, but not in girls, using a term-birth sample born in 2005 from the Taiwan Birth Cohort Study. On the other hand, Liang et al. (2021) do not find significant differences between males and females in stillbirth risk in response to $PM_{2.5}$ exposure using birth-cohort data in seven cities in Southern China between 2014 and 2017.

A few studies have assessed the presence of heterogeneous effects of air pollution exposure on birth outcomes by maternal age and show mixed patterns. An associational study in Wuxi, China provides evidence of a stronger association of PM_{10} with preterm birth among those with advanced maternal age, compared to others (Han et al., 2018). Wang et al. (2021) find increased risk of term low birthweight with maternal exposure to ozone among women older than 35 or younger than 25, compared to others, using data from Guangzhou, China between January 2015 and July 2017. However, Li, Guo and Williams (2016) do not find that an acute impact of hourly ambient air pollution on preterm birth is modified by maternal age, using data from Brisbane, Australia.

In this paper, we estimate the effects of prenatal air pollution on birth outcomes by 1) using unique birth-certificate data from two cities in South China that are similar in climate and social and economic development, but different in pollution levels; 2) addressing the issues of pregnancy timing and self-selection by exploiting the exogenous improvement of air quality during the Guangzhou 2010 Asian Games and controlling for quarterly trends in birth outcomes; and, 3) examining the heterogeneity of effects across child sex and maternal age groups. Given high-frequency daily air pollution level data, and properly controlled seasonal trends, our strategy exploits variation in the timing of conception within the same season, and compares the birth outcomes in Guangzhou, where air quality was significantly improved during the Asian Games, with those in Shenzhen. Although the city of Shenzhen was also

from China, see Xie et al. (2021).

in the area of regional air quality control during the Asian Games, we will show below that pollutant concentrations were significantly reduced in Guangzhou relative to Shenzhen during the Asian Games.

Our methodology is preferable to that used in much of the prior literature in that it not only addresses self-selection endogeneity bias – the problem that individuals choose where to live based on their health preferences and other observed characteristics – but also provides a way to disentangle the impacts of air pollution from weather, as air quality was exogenously improved during the Asian Games by stringent air-quality-control policies. Exploiting exogenous air-quality improvement during a large sports event to study its health impact does have precedent in the literature. For example, Rich et al. (2015) exploited the natural experiment of air-pollution decline during the Beijing Olympics to evaluate whether having specific months of pregnancy overlap with the 2008 Beijing Olympics was associated with larger birthweights, using individual data on term births, compared with pregnancies during the same dates in 2007 or 2009. However, there is no control city in the sample. In addition, exploiting the same 2008 Beijing Olympic Games, He, Fan and Zhou (2016) analyzed monthly mortality data from nationally representative surveillance points to estimate the effect of air pollution on mortality in China.

3 Air-Quality Regulations Related to the 2010 Guangzhou Asian Games

The 2010 Guangzhou Asian Games were the second Asian Games held in China, after the 1990 Beijing Asian Games. To prepare for the Asian Games, the Guangdong government adopted systematic, multi-stage improvement measures and implemented a series of regulations on air-quality control starting with the announcement in 2004 of Guangzhou as the host city for the 16th Asian Games. Since Beijing had successfully hosted the Olympics in 2008, the Guangdong government learned a great deal from the Beijing government's experiences with air-quality control, especially with regard to regional cooperation and multi-city control actions.

The key area targeted for air-quality improvement included the Pearl River Delta (PRD) Metropolitan Region, where the cities of Guangzhou and Shenzhen are situated. The straight-line distance between the two cities is about 100km. Although this regional air pollution management strategy may have improved Shenzhen's air quality during the same period, we later will show that relative to Shenzhen, air quality in Guangzhou was improved significantly during the Asian Games. On February 24, 2010, the Guangdong Province Environment Protection Department issued the "Notice on Air Quality Regulations during the Pre-Asian Games Period" (Department of Environmental Protection of Guangdong Government, 2010). It lists a series of measures for quality control during the last year before the Asian Games, for example, installation of desulfurization and denitrification by Guangzhou Papermaking Plant and power plants and converting to clean fuels from coal furnaces by the Guangzhou Steel Company. Low-steam-capacity industrial boilers were required to be phased out, all gas stations, oil depots and oil tankers were required to complete oil and gas recovery management, volatile-organic-compounds (VOCs) related industries were required to reduce emissions, and the building-materials industry was required to phase out low-capacity facilities and to install dust removal and denitrification facilities and meet emission requirements. As part of regional air-quality-control measures, PRD cities were required to set up complete vehicle-emissions inspection systems and monitoring and data-sharing networks, and upgrade the quality standard of gasoline for motor vehicles. Ambient air-quality monitoring and surveillance mechanism also were set-up in PRD cities.

While the previous measures were taken steadily and progressively, the most strict control measures were implemented during the Asian Games period (i.e., between October 20, 2010 and December 20, 2010) – the period on which we are focusing.[6] These measures involved temporary production shutdowns for plants that did not meet emission standards and traffic controls.[7] Efforts were made to make sure concentrations of SO_2, NO_2 and PM_{10} were lower than National Standard Grade 2 levels.

All of these regulations were strictly binding. Non-compliers faced the risk of temporary shutdown as punishment during the two-month Asian Games. As a result, air pollutant levels dropped significantly during the two months of the Guangzhou Asian Games (Liu et al., 2013; Xu et al., 2013). According to Liu et al. (2013), the Asian Games abatement strategy reduced emissions by 41.1% for SO_2, 41.9% for NO_x, 26.5% for PM_{10}, 25.8% for $PM_{2.5}$, and 39.7% for

[6] The Asian Games were held between November 1, 2010 and December 20, 2010. Because the strict measures were taken starting on October 20, 2010, we define the Asian Games Air Pollution Control period as October 20 - December 20, 2010.

[7] Similar trends were observed in lockdown periods during the Covid-19 Pandemic, in which most affected countries adopted partial or complete lockdown policies. As a result, air quality largely improved (National Aeronautics and Space Administration, U.S.A., 2020).

VOC. The concentrations of SO_2, NO_2, PM_{10} and $PM_{2.5}$ were reduced by 66.8%, 51.3%, 21.5% and 17.1%, respectively.

4 Data

4.1 Data on Birth Outcomes

We use birth-certificate data collected by one of the authors from one district in the city of Guangzhou and all of the city of Shenzhen for the period between January 2009 and February 2012. The birth-certificate data cover all births during the period in these districts and cities as required by law. All locations report basic birth outcomes including estimated gestational age based on reported last menstrual period, birthweight, birth length, sex, and parity, as well as maternal age and education.[8] Neither of these locations reported maternal height and marital status[9] in their birth-certificate systems. We focus on three outcomes: prematurity (defined as gestational age at birth less than 37 weeks), birthweight, and low birthweight (defined as birthweight less than 2500g). Table A1 lists the basic statistics of these outcomes by city. All three of these birth outcomes are associated with subsequent outcomes in childhood and adulthood. But these associations may reflect a range of factors, such as maternal health and family background, not just the impacts of the birth outcomes. For birthweight, in addition, there is evidence of casual effects through using monozygotic (MZ, identical) twins to control for all family background factors including genetics that the MZ twins share in common. Previous studies using MZ twins find that birthweights have impacts on outcomes ranging from schooling attainment to adolescent behaviors to adult earnings (Behrman and Rosenzweig, 2004; Conley, Strully and Bennett, 2003; Møllegaard, 2020; Torche and Conley, 2016). This means that if pollution affects birthweights, there is evidence of effects through birthweights over the life cycle.

We only include singleton live births in our sample (819,619). We exclude observations with less than 28 weeks or above 46 weeks of gestational age or with missing information on gestational age (1553 observations dropped), with birthweight less than 500g (51 observations dropped) or birth length less than 28cm or longer than 60cm (13 observations dropped).

[8] Although maternal job information is also available in both cities, 40% observations in Shenzhen sample has missing values. As there was no uniform birth certificate in the region at that time, maternal job categories were not collected in a consistent way in these two cities. We have tried to aggregate jobs into broad, comparable categories and include them in robustness checks in Appendix 2.

[9] Children born to single mothers are not culturally acceptable in China and are rare.

We further drop those with maternal age under 15 (927 observations dropped) or above 60 years old (5 observations dropped). To avoid fixed-cohort bias,[10] we delete births within this period with conception dates earlier than 18 June 2008 and those with conception dates later than 13 April 2011 (as gestational age varies between 196 and 322 days in the sample), which leaves 545,703 observations. We further drop 23,000 observations that do not have any of the following information: birthweight, gestational age at birth, gender of the baby, maternal age, maternal education or parity. In the end, 129,131 , 183,959 , 201,889, and 7,724 birth observations from 2009, 2010, 2011, 2012, respectively, constitute the main sample for analysis.

The distribution of the number of births throughout the year in this data (as shown in Figure A3) is consistent with the observation of Yang (2021) using national census data that the peak of number of births is in October, when those who get married during the Chinese New Year are most likely to give birth. Labor migrants may be most likely to show this pattern. Birthweight also shows strong seasonality in this sample, as shown in Figure A1. Shenzhen has higher average birthweight than Guangzhou throughout the year. The two cities show very similar seasonality trends: children born in the summer have higher birthweight than those born in the winter. These similar patterns help validate Shenzhen as a control city.

4.2 Data on Air Pollution and Weather

The data on air pollution come from the Guangzhou and Shenzhen environmental bureaus that report the daily average levels of three monitored air pollutants (NO_2, PM_{10}, and SO_2) at all monitoring stations in each city during 2008 to 2012. These pollutants are measured according to the National Standard GB3095—1996. For the very few missing observations in our data, we replace them with moving averages for the most recent 5 days. Data such as these are very rare in China for the period under study, as China only published an air pollution index (API) before 2014 and individual pollutant levels were generally not publicly available. As the API is an index score based on the most dominant pollutant, using the API alone does not reveal information on all "criteria" pollutants.

Figure A2 shows the daily average levels of the three air pollutants over the duration of the study period. Three patterns merit note. First, Guangzhou and Shenzhen have dif-

[10] Fixed-cohort bias emerges when a sample consists of births during a fixed period—this approach will include only the longer pregnancies at the start of the study and only the shorter pregnancies at the end of the study. This has the potential to bias studies of environmental exposures (Strand, Barnett and Tong, 2011).

ferent compositions of air pollutants and air quality in Shenzhen is on average better than Guangzhou. Second, in general, air pollution is worse in the winter than in the summer in both cities. The average levels of NO_2, PM_{10}, and SO_2 in Guangzhou are 52.22, 92, and $35\mu g/m^3$ in November - January, compared to 27.91, 52.56, and 32.21 $\mu g/m^3$ in June - August. Third, SO_2 shows a decreasing trend throughout the entire period (conditional on season), which is consistent with the ongoing desulfurization effort by the Guangdong government during the last decade.

Guangzhou and Shenzhen have typical subtropical climates, with very mild winters and hot, rainy, and humid summers due to the Asian monsoon. The temperature throughout the year varies between 11°C and 33°C. Rainfall is abundant, at around 1700 millimeters a year, concentrated from May to September. We obtain hourly data on temperatures and dew points at two meters above sea-level as well as precipitation from reanalysis data at single levels by the European Centre for Medium-Range Weather Forecasts (ECMWF) (European Centre for Medium-Range Weather Forecasts, 2020*a*) for both cities for the period between January 1, 2008 and February 29, 2012 and calculate the daily mean temperature and daily mean relative humidity and 24-hr precipitation.[11]

We also obtain universal thermal climate indices (UTCI) from the same source. The UTCI, a thermal comfort indicator based on human heat-balance models, is designed to be applicable in all seasons and climates and for all spatial and temporal scales (European Centre for Medium-range Weather Forecasts, 2020*b*). The UTCI is a one-dimensional index that reflects "the human physiological reaction to the multidimensionally defined actual outdoor thermal environment" (Bröde et al., 2012, p.2). Scores can be classified into ten thermal stress categories, ranging from extreme cold stress to extreme heat stress (European Centre for Medium-range Weather Forecasts, 2020*b*). In a subtropical, humid environment such as Guangzhou and Shenzhen, this index provides a wider range of variation than does ambient temperature.[12] Table A2 lists the distribution statistics of UTCI in both cities between 2000 and 2008, which serves as the reference period based on which we define extreme weather in section 4.4.

[11]We calculate relative humidity (rh) by $rh = exp(5423 * ((1/273) - (1/d2m)))/exp(5423 * ((1/273) - (1/t2m)))$, where $d2m$ is 2-meter above sea level dew point, and $t2m$ is 2-meter above sea level temperature.

[12]We provide details on how we acquire and process UTCI data in Liu et al. (2021).

4.3 Prenatal Air-Pollution-Exposure Measurement

Due to data constraints, we only know individuals' cities of residence but not their home addresses. We therefore calculate the daily mean of air pollutants' levels for each city and assign each individual their accumulative prenatal exposure based on their gestational periods. However, this measure suffers from a potential problem – it generates a spurious inverse correlation between adverse birth outcomes and accumulated air pollution levels because a shorter gestational age implies a shorter third trimester and therefore smaller accumulated air-pollution exposure. We avoid this spurious relationship by calculating the daily average exposure level (i.e. dividing the accumulated exposure by the number of days of gestational age).

4.4 Extreme Weather Exposure Measurement

Besides air pollutants, weather conditions such as temperature and humidity have been shown to be related with birth outcomes (Beltran, Wu and Laurent, 2014; He et al., 2016; Murray et al., 2000; Siniarska and Kozieł, 2010). In this paper, we control for weather with two different specifications. In one specification, we control for daily mean average temperature and relative humidity during pregnancy. In the other specification, we use the universal thermal climate indices (UTCI) to define extreme cold and extreme hot weather in order to control for the possible effects of extreme temperatures on both ends. More specifically, we define a threshold for extreme cold days as a daily mean UTCI below $0°C$, and for extreme hot days as a daily mean UTCI above $34°C$.[13] It can be seen that UTCI has a larger variance than ambient temperature and is more likely to fall into extreme temperature ranges. We calculate the percentage of time during pregnancy exposed to extreme cold (number of days with the daily lowest apparent temperature under $0°C$ divided by the number of pregnancy days) and extreme heat (number of days with the daily highest apparent temperature over $34°C$ divided by the number of pregnancy days) for each pregnancy and use these two variables to control for extreme weather conditions.[14]

[13] The reason that we use 0°C and 34°C as cutoffs to define extreme cold and hot is because these two cutoffs are close to the bottom 1% and top 1% of historical UTCI data of both cities between 2000 and 2008 as shown in Table A2.

[14] We do not use the absolute number of extremely cold or hot days in regressions due to the inverse causality concern already discussed.

5 Methods

The method that is applied in this paper exploits the change in air pollution during the Asian Games in Guangzhou, relative to Shenzhen, to identify the effect on birth outcomes. As seen in the data section, Guangzhou and Shenzhen are similar in climate as well as social and economic development, with Shenzhen having slightly better air quality prior to the Asian Games. Since what counts for the analysis is the changes that occurred during the Asian Games, this slight difference in air quality before the Asian Games does not jeopardize our approach.

In order to investigate whether Guangzhou air quality was significantly improved relative to Shenzhen during the Asian Games, we first run the following difference-in-difference regression with daily monitored air pollutants levels data from monitoring stations as well as meteorological data:

$$\begin{aligned} P_{yds} = a * AsianGames_{yd} + b * Guangzhou + c * Guangzhou * AsianGames_{yd} + y + d + dow \\ + spline(temp, df = 3) + spline(rainfall, df = 3) + u_{yds} \end{aligned} \tag{1}$$

P_{yds} is the level of air pollutant P on date d in year y at monitoring site s. We model it as a function of year fixed effects y, date-in-a-year fixed effects d, day-of-week fixed effect dow, city fixed effects $Guangzhou$, and spline functions of observed daily mean temperature and rainfall each with 3 degrees of freedom, as well as an error term u_{yds}. We define $AsianGames_{yd}$ as a binary variable equal to 1 if year y is 2010 and date d is between October 20th and December 20th. This term estimates the average change in the air pollutant in both cities during the Asian Games. Parameter c is the coefficient of interest, which estimates the effect of the Asian Games on air quality in Guangzhou city during the Asian Games, in addition to the average effect in both cities.

Next, we estimate the effects of air pollution on birth outcomes. Observed associations between air pollution and birth outcomes may suffer from two types of bias that work in opposite directions: 1) attenuation bias caused by measurement error of exposure level as we do not have accurate information on where in the cities people spent their time; 2) location selection bias as people with more resources (and perhaps, better health) may choose in what city to live in part according to air quality. Ignoring the selection bias may over- or under-estimate the effect of environment on health or other outcomes. As these two types of biases may work in different directions, it is hard to predict the direction of bias in the aggregate. We address this issue by exploiting the exogenous air quality improvement during the Guangzhou Asian Games and treat the event as a "quasi-experiment" that indirectly affected birth outcomes through air-quality improvements. Thus, the methodology we undertake to identify the causal impacts of air pollution on adverse birth outcomes is to use Asian Games exposure as an instrument for prenatal air-pollution exposure. The estimated effect is the Local Average Treatment Effect on the Treated (LATE). More specifically, we estimate the correlation between birth outcomes and individual prenatal air pollution exposure, with the number of days overlapping between a woman's pregnancy and the air-quality control period for the Asian Games as an instrumental variable that brings about the exogenous change. The empirical framework is described in Equations (2) and (3).

$$P_i = \sigma X_i + \rho M_i + \gamma * 1(c_i = "Guangzhou") * G_i + y_i + q_i * c_i + c_i + \epsilon_i \tag{2}$$

$$Y_{icyq} = \alpha X_i + \theta P_i + \beta M_i + y_i + q_i * c_i + c_i + u_i \tag{3}$$

Eq.(2) is the first-stage estimation of pregnant woman i's pollution exposure P_i, as a function of individual characteristics X_i, meteorological conditions M_i during pregnancy, as well as a policy-exposure term $\gamma * 1(c_i = "Guangzhou") * G_i$, where, G_i is pregnant woman i's time exposed to the Asian Games during pregnancy. Here, G_i satisfies the exclusion restriction in that, as long as pregnant women are not timing pregnancy relative to the Asian Games, it only affects one's exposure to air pollution and not directly affects birth outcomes. It serves as the instrumental variable. We also control for conception-year fixed effects y_i, city-specific conception-quarter[15] fixed effects $q_i * c_i$ and city fixed effects c_i.

Eq. (3) is the outcome estimation, in which Y_{icyq} is the birth-outcome variable of individual i from city c conceived in year y and quarter q. The birth outcome can be a continuous variable – birthweight, or a binary variable – low birthweight or prematurity. The birth outcome is a function of individual-specific characteristics (X), as well as exposure variables including prenatal exposure to ambient air pollution (P) and meteorological conditions (M). As in the first stage, we also control for conception-year fixed effects y_i, city-specific conception-quarter fixed effects $q_i * c_i$ and city fixed effects c_i.

The measurement of birth outcomes, air pollution and weather, and exposure variables have been described in Section 4. In addition, confounding factors include a quadratic

[15]We define quarters based on local climate: quarter 1 - spring: March-May, quarter 2 - summer: June-August, quarter 3 - autumn: September - November, quarter 4 - winter: December - February.

term in maternal age, a set of dummy variables for maternal-schooling categories, a set of dummy variables for parity, a binary variable indicating child's sex as in the vector of X, and conception-year fixed effects, city fixed effects as well as city-specific conception-quarter fixed effects.

We consider two candidates for G_i: 1) a continuous variable measuring how many days one's pregnancy duration overlapped with the Asian Games. 2) a continuous variable measuring how many days one's *expected* pregnancy duration overlapped with the Asian Games, by assuming everyone has 39 weeks of gestation. We argue that the first of these is not a valid instrument because one is likely to have more overlap with Asian Games if one has longer gestation given the same gestational date. Therefore, this variable may not satisfy the exclusion restriction because it is correlated with outcome variables not only through air pollution exposure. In contrast, the second instrumental-variable candidate simply exploits the variation in timing of conception and its exogenous co-variance with the timing of Asian Games to instrument for variation in exposure. This instrumental variable satisfies the exclusion restriction if we believe that individual's timing of conception was not affected by the Asian Games, conditional on conception quarter-in-a-year fixed effects. Our following analysis adopts the second instrumental variable.

Besides pollution and weather, we also control for conception-year fixed effects y_i and a city-specific conception quarter-in-a-year effect $q_i * c_i$ to control for city-specific seasonal patterns and the possible weather preference that may introduce selection bias. This city-specific quarter-in-a-year term controls for all unobserved time-varying city fixed effects, including population composition change due to pregnancy timing and the common shocks to birth outcomes experienced by individuals who conceived in the same season of a year within the same city[16]. The reason for controlling for conception quarter-in-a-year trends instead of month trends of birth outcomes in our study is that the instrumental variable we use is defined based on city and timing of conception (basically covering November and December of 2010), i.e., the exogenous variation in exposure variable comes from interaction between a city fixed effect and conception-month fixed effect. Controlling for month trends in birth outcomes likely would cause collinearity with this variation and therefore over-control

[16]In more recent papers examining environment on birth outcomes, some studies control for quarterly trends (Wang et al., 2021), or monthly trends in a year (He et al., 2016), or even day-in-a-year trends (Chen et al., 2020). Since air pollution also varies across time in a year, there could be a tradeoff between under-control and over-control of these trends.

the model. Note that with fixed effects rather than only random effects, the other types of shocks that may also affect birth outcomes (such as typhoons) are allowed to be correlated with the right-side variables, including air pollution and weather. u_i is the error term that represents the effects of random unobserved factors and measurement error. As we have controlled for year fixed effects and city-specific conception quarter-in-a-year fixed effects, we assume u_i is independent and identically distributed (i.i.d.). In regression results, we report heteroscedasticity-consistent or robust standard errors.

6 Results

6.1 Effects of Asian Games on Air-Pollution Reduction

Our first set of results in regressions of daily monitored level of each pollutant between 2008 and 2012 show that during the Asian Games, the three air pollutants' levels were significantly reduced in Guangzhou. The coefficients on $1.Guangzhou * 1.policy$ in Table 1 correspond to coefficient c in Equation 1, i.e., the coefficient of interest. As shown in the first panel, Column (1)-(3), after controlling for meteorological conditions including temperature and rainfall, city fixed effects, year fixed effects and day-in-a-year fixed effects, we find PM_{10}, NO_2 and SO_2 in Guangzhou were significantly reduced during the Asian Games compared to the same period in 2009 and 2011. More specifically, compared to Shenzhen, Guangzhou witnessed a further reduction of PM_{10} by 33.91 μg/m^3 (s.d.=6.44), SO_2 by 8.24 μg/m^3 (s.d.=2.45) and NO_2 by 24.64 μg/m^3 (s.d.=4.62).

We also implemented placebo tests that defined three imaginary, alternative dates for the Asian Games: 1) one year before (Scenario A. $Policy = [20Oct2009 - 20Dec2009]$), 2) one year after (Scenario B. $Policy = [20Oct2011 - 20Dec2011]$), and, 3) five months before (Scenario C. $Policy = [20May2010 - 20Jul2010]$). The significant positive coefficients on $Guangzhou * policy$ term in Placebo Test Scenario A implies air pollutant levels in Guangzhou in 2009 were significantly higher than in the same period in 2010 and 2011, compared to the control city Shenzhen. In Placebo Test Scenario B, we find PM_{10} and NO_2 are reduced in both Guangzhou and Shenzhen in 2011 compared to 2009 and 2010, while SO_2 is more reduced in Guangzhou compared to Shenzhen between 20 October and 20 December in 2011, thanks to persistent desulphurisation efforts by the Guangzhou government. In Placebo Test Scenario C, we do not see significant differences in air pollutants' levels between May and July in

2010 compared to other years between two cities. These placebo tests lay the foundation for our estimation strategy below of using the overlap period between pregnancy and the Asian Games to instrument for one's prenatal air pollution exposure.

6.2 Effects of Air Pollution on Birth Outcomes

We estimate Equations (2) & (3) with a linear continuous model for birthweight and probit models for low birthweight and preterm birth using the two-stage least squares method (2SLS and IV-Probit separately). Due to high correlations among pollutants, we estimate the equation with each pollutant separately.

As mentioned in Section 5, we use number of days overlapped between *expected* pregnancy (by assuming each has 39 weeks of pregnancy) and the Asian Games to instrument for the exposure measurement. Consistent with difference-in-difference estimation results on air quality improvement during Asian Games, the first-stage results (Table A3) show the instrument is highly significantly correlated with all three air pollutant exposures. One more day of overlap of pregnancy with the Asian Games period reduces daily average exposure to PM_{10}, NO_2 and SO_2 by 0.16, 0.26, 0.13 $\mu g/m^3$ respectively, in the specification with linear control of temperature and relative humidity. The results are similar in the other specification with extreme temperatures as weather controls.

Marginal effects of PM_{10} exposure on birth outcomes are presented in Table 2. For comparison, we present results of OLS or probits in the left 6 columns ((1)-(6)), while the right 6 columns (7) to (12) are for 2SLS or IV-probits. Specifications A and B vary by the weather controls. Specification A controls for daily mean temperature and relative humidity over one's pregnancy, while specification B controls for percentage of time during pregnancy that a woman was exposed to extreme temperatures, defined by UTCI.

Although the effects of PM_{10} on mean birthweights are similar between OLS and IV estimates, the marginal effects of PM_{10} on the risk of low birthweight and preterm are much smaller with IV – nearly half of the size of those without IV, given the same specification. With IV estimates, a 10 $\mu g/m^3$ increase in PM_{10} reduces birthweight by around 14 grams, and increases the risk of preterm by 0.45% to 0.65%, while the effect on the risk of low birthweight is not statistically significant in a specification controlling for extreme temperatures, and only weakly significantly increases low birthweight by 0.3% in the specification controlling for linear temperature and relative humidity.

Table 1: Difference-in-Difference Estimation of Effects on Air Pollution of Asian Games

	I. Real Policy Effect (Policy = [20Oct2010 - 20Dec2010])			II. Placebo Test Scenario A (Policy = [20Oct2009-20Dec2009])			III. Placebo Test Scenario B (Policy = [20Oct2011-20Dec2011])			IV. Placebo Test Scenario C (Policy = [20May2010-20JUL2020])		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	PM_{10}	SO_2	NO_2	PM_{10}	SO_2	NO_2	PM_{10}	SO_2	NO_2	PM_{10}	SO_2	NO_2
1.policy	15.77***	0.927	5.500	0.599	2.843	0.299	-12.34**	2.289	-7.299**	4.892	1.366	3.456
	(5.138)	(1.954)	(3.688)	(5.170)	(1.959)	(3.699)	(5.179)	(1.965)	(3.718)	(4.378)	(1.663)	(3.148)
1.Guangzhou*1.policy	-33.91***	-8.244***	-24.64***	13.64**	5.502**	20.64***	7.106	-6.733***	-4.040	4.177	-2.815	-4.190
	(6.441)	(2.450)	(4.624)	(6.463)	(2.448)	(4.624)	(6.479)	(2.458)	(4.652)	(5.365)	(2.037)	(3.858)
Observations	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576	4,576
R-squared	0.457	0.595	0.476	0.452	0.597	0.483	0.453	0.594	0.472	0.452	0.593	0.472

Notes: Estimation of Daily air pollutant level from all monitoring stations in Guangzhou and Shenzhen city between 2008 and 2012. All specifications include Year FE, City FE, Day in a Year FE. Justified models also justify meteorological variables including temperature and rainfall. Standard errors in parentheses. *** significant at 1%, ** significant at 5%, *significant at 10%.

Table 2: Marginal Effects of PM_{10} (by 10 µg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm.

	Probit Model/OLS Model						IV-Probit Model/2SLS Model					
VARIABLES	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PM_{10}*10µg/m^3	0.007***	0.006***	-16.37***	-13.70***	0.012***	0.009***	0.0033*	0.0022	-14.18***	-13.39***	0.0065***	0.0045**
	(0.001)	(0.001)	(2.60)	(2.31)	(0.001)	(0.001)	(0.002)	(0.002)	(4.97)	(5.20)	(0.002)	(0.002)
Temp. (°C)	-0.0038***		10.59***		-.0044 ***		-0.0041***		10.76***		-0.0048***	
	(0.000)		(0.948)		(0.000)		(0.000)		(1.01)		(0.000)	
Rel. hum. (%)	0.0015***		-3.83***		0.0023***		0.0012***		-3.63***		0.0017***	
	(0.000)		(0.44)		(0.000)		(0.000)		(0.59)		(0.000)	
Ext. cold (%)		0.0023***		-6.41***		0.0032***		0.002***		-6.39***		0.0028***
		(0.001)		(1.25)		(0.001)		(0.001)		(1.30)		(0.001)
Ext. hot (%)		0.0003		-0.095		0.0007		0.000		-0.10		0.00086
		(0.000)		(1.15)		(0.001)		(0.000)		(1.16)		(0.001)
Obs.	522,703						522,703					
Spec.	A	B	A	B	A	B	A	B	A	B	A	B

Notes: Each column represents a separate regression. (1)-(6) are from Probit/OLS regressions without applying IV, while (7)-(12) are from IV-Probit/2SLS regressions. (1),(2),(5),(6),(7),(8),(11),(12) are marginal effects of 10 µg/m^3 PM_{10} on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4),(9),(10) are marginal effects of 10µg/m^3 PM_{10} on Mean Birthweight. Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Results for the other two pollutants, NO_2 and SO_2, are presented in Tables A4 and A5 in the Appendix. The effects of SO_2 are strongest: a 10 µg/m^3 increase in SO_2 reduces average birthweight by 18 to 20 grams, and increases the preterm rate by around 0.76 to 0.83%. Among the three pollutants, NO_2 shows the smallest effects: 10 µg/m^3 increases in NO_2 reduce average birthweight by 8.7 to 9.3 grams, and increase the preterm rate by 0.33 to 0.38%.

6.3 Estimation Results by Trimester

To test the sensitivity of effects within various subperiods of pregnancy, we also estimate the impacts of each pollutant by trimester, with the coefficients displayed in Table 3. Each column represents results from a regression, i.e., different pollutants enter different regressions, while the exposure to the three trimesters are from the same regression. It is clear that third trimester exposure to PM_{10} plays a key role in reducing average birthweight. A 10 µg/m^3 increase in PM_{10} in the third trimester decreases average birthweight by 13-15 grams, while exposures in the first and second trimester don't have significant effects on birthweight. The most sensitive trimester to PM_{10} in terms of preterm risk depends on the weather control: in specification A with linear control of mean temperature and humidity, the third trimester is most sensitive to PM_{10}, while in specification B with extreme temperature defined by UTCI, the second trimester is the only sensitive period. The statistical significance is much lower for the preterm risk outcome. The estimated effects of SO_2 and NO_2 by trimester do not show sensitive periods, although three trimesters combined have significant effects on birthweight and preterm. To save space, the estimated coefficients are not displayed, but are available on request.

7 Heterogeneity of Effects by Sex and Maternal Age

Previous research shows evidence of gender/sex differences in response to environmental exposures (Keitt, Fagan and Marts, 2004; Clougherty, 2010; Kim et al., 2017). To test the hypothesis that different sexes may have different responses to air pollution in the fetal period, we estimate heterogeneous effects by infants' sex. We do this by adding interaction terms between infant's sex and air-pollutant exposures. The estimated coefficients of these interaction terms for the three birth outcome variables are shown in the first line in Table 4.

Table 3: Marginal Effects of Prenatal Exposure to PM_{10} on Low Birthweight, Average Birthweight, and Preterm Risk by Trimester

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
PM_{10} in the first trimester *10 µg/m^3	-0.0003	0.00061	1.63	1.22	0.0002	0.00087
	(0.002)	(0.001)	(3.97)	(3.76)	(0.002)	(0.002)
PM_{10} in the second trimester*10 µg/m^3	0.001	-.00033	-0.56	-1.17	0.002	0.0027*
	(0.001)	(0.001)	(2.67)	(3.30)	(0.001)	(0.001)
PM_{10} in the third trimester*10 µg/m^3	0.003	0.002	-15.03***	-13.12***	0.0035*	0.001
	(0.002)	(0.001)	(4.60)	(3.86)	(0.002)	(0.002)
Weather control	A	B	A	B	A	B
Observations	522,703					

Notes: Each column represents a separate regression. Instrumental variables are a set of variables that measure the percentage of time overlapped with Asian Games period in each trimester. Besides air pollution exposure, other control variables include a binary variable indicating baby is female, parity, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Males are more vulnerable to PM_{10} in terms of average birthweight at a significance level of 10%: on average, males have a 1.51 gram reduction in birthweight in correspondence to a 1 µg/m^3 increase in PM_{10} (i.e., they experience a 15.1 gram extra reduction in birthweight in correspondence to a 10 µg/m^3 increase in PM_{10} compared to females). Nevertheless, males are not more likely than females to experience low birthweight or preterm birth because of air pollution in particular.

The impacts of air pollution may also vary by maternal age. Mothers aged at least 35 deliver babies with lower birthweights – in response to a 10 µg/m^3 increase in PM_{10} exposure, lower by 17.3 grams, on average (for both Specifications A and B). However, they are not more likely than other mothers to have low birthweight babies nor to have preterm births in response to air pollution.

Table A6 and A7 give the coefficient estimates for interaction terms for NO_2 and SO_2. SO_2 shows similar interaction term coefficients (males show an 8 more grams reduction in birthweight in response to a 10 µg/m^3 increase in SO_2, and mothers older than 35 years old deliver babies with 22 grams lower birthweight, on average, in response to a 10 µg/m^3 increase in SO_2). However, NO_2 shows a different pattern. Males do not experience disadvantage, and indeed show an advantage, in birthweight in response to NO_2 exposure (i.e., female fetuses are more vulnerable to NO_2 than are males). But babies of mothers aged over 35 are even more vulnerable than others to NO_2: they experience 77.6 grams lower birthweights, on average. Neither of these pollutants have sex heterogeneous effects on the risk of low birthweight nor preterm.

8 Discussion and Conclusion

This study seeks to provide plausibly causal estimates of the impact of prenatal air pollution exposure on birth outcomes by exploiting a large-scale sports event – the Guangzhou 2010 Asian Games – during which air quality control was strictly enforced and air quality was exogenously improved. Using daily air-quality data from monitoring stations in Guangzhou and the control city, Shenzhen, we first show that air quality indeed improved during the Asian Games period in Guangzhou relative to Shenzhen. To demonstrate the improvement, we presented difference-in-difference estimation results and implemented placebo tests by assigning hypothetical alternative counterfactual dates for the Asian Games period. With these findings established, we use the overlap between pregnancy and the actual Asian Games period (given 39 weeks of gestational age) to instrument for ambient air-pollution exposure. We find that air pollutants PM_{10}, NO_2 and SO_2 significantly decrease birthweights and increase preterm delivery risks. 10 µg/m^3 increases in PM_{10}, NO_2, SO_2, significantly decrease mean birthweight by around 14, 9, and 18-20 grams, respectively. The same changes increase preterm risk by 0.45-0.65%, 0.33-0.38% and 0.76-0.83%, respectively. For average birthweight, we also find that the third trimester is an especially vulnerable period for exposure to PM_{10}.

We compare our findings first with associations summarized in a review of epidemiological studies using Chinese data: Jacobs et al. (2017) summarizes estimates that show a 10 $\mu g/m^3$ increase in PM_{10} reduces birthweights by 0-9.1 grams. Our estimates on birthweight are of the same order of magnitude but slightly bigger than the findings in these association studies. If we run logit regressions of preterm birth without instrumenting on PM_{10}, the odds ratio of preterm, 1.027 (s.d.= 0.00283) for specification A, is in the range between 1 and 1.05, as in most epidemiological studies. Nevertheless, the marginal effects on preterm birth are roughly halved in the instrumental-variable model, compared to a standard probit model without instrumenting, and the statistical significance is much weaker for low birthweight. This pattern suggests possible selection bias in the simple association studies of the form that couples who have resources and knowledge to choose lower pollution also have better birth

Table 4: Heterogeneous Effects of PM_{10} on Birth Outcomes by Sex and Maternal Age Group

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
PM_{10} * male	-.000326	-0.000315	-1.508*	-1.506*	0.0008897	0.0008034
	(0.003)	(0.003)	(0.772)	(0.772)	(0.003)	(0.003)
PM_{10} *(mom age>=35)	0.000377	0.000376	-1.729***	-1.727***	0.000444	0.000444
	(0.000280)	(0.000280)	(0.533)	(0.533)	(0.000315)	(0.000315)
Specification	A	B	A	B	A	B

Notes: IV estimates of the interaction term between PM_{10} and male or age35+ on birth outcomes, with the same set of control variables as in Table 2. *** significant at 1%, ** significant at 5%, *significant at 10%.

outcomes. The timing of the sensitivity window at the third trimester to PM_{10} is also consistent with Rich et al. (2015) who find month 8 in pregnancy is the critical window for air pollution.

We next compare our pollution-effect estimates to estimates from developed countries, where air pollution is generally lower, in order to illuminate the possible non-linearity of dose-response curves (Fleischer et al., 2014; Arceo, Hanna and Oliva, 2016). There have been a large number of associational studies. For example, Liu et al. (2003) found that the risk of preterm delivery was associated with third-trimester exposure to SO_2 (OR=1.09 for 5.0 ppb (or 13.1 $\mu g/m^3$) increase, with daily average at 12.8 $\mu g/m^3$) using live-birth datasets from Vancouver, Canada.

However, causal studies of the impact of air pollution exposure on birthweight and prematurity are limited. Exceptions include Coneus and Spiess (2012), who find no significant effect from NO_2, or SO_2 in estimates of a mother fixed-effects model using a small sample in Germany[17], where air quality was much better compared to Guangzhou. Currie, Neidell and Schmieder (2009) also find no effects for PM_{10} in New Jersey in the United States[18], where average concentration of PM_{10} is only about 30 $\mu g/m^3$–less than half of the average level in our study. Together with the null findings for PM_{10} and NO_2 in Coneus and Spiess (2012) and Currie, Neidell and Schmieder (2009), the significant findings reported here suggest the possibility that air pollutants at higher concentrations, as in this study, are more likely to cause lower birthweights and prematurity. This contrast across studies with quite different pollution levels suggests the possible existence of threshold or other nonlinear effects of air pollution on birth outcomes.

This study also contributes to a literature on heterogeneous vulnerability to air pollution by maternal age and child sex. In terms of average birthweight, males are more vulnerable to PM_{10} and SO_2 than are females, but females are more vulnerable to NO_2. With regard to average birthweight, children of mothers over 35 years old are also more vulnerable than others when exposed to higher levels of all three types of air pollutants. NO_2 is associated with the greatest harm to babies of mothers over 35 years in comparison with SO_2 and PM_{10}. We do not find sex and maternal age differences in the effect of air pollution on preterm risk.

[17] Coneus and Spiess (2012) find that CO significantly reduces birthweight, but find no significant effects from O_3, NO_2 and SO_2.

[18] Currie, Neidell and Schmieder (2009) find negative effects of CO on birth outcomes, but no effects for O_3 or PM_{10}.

There are several limitations in our approach. First, and most important, is exposure measurement. We do not have residential addresses of mothers to permit personalized pollution-exposure measures by matching mothers' addresses to nearby monitoring stations or a high spatial resolution pollution map that uses more sophisticated approaches, such as inverse distance weighting (such as used in Lu et al. (2020)), kriging methods, and land-use regression methods (compared in Mercer et al. (2011)). Of course, if women move around the city for work, socializing or shopping, a broader geographical representation of pollution than their local residential neighborhood may be appropriate. Second, also due to data limitations, we cannot control for additional confounders such as smoking and passive smoking, cooking smoke exposure, and avoidance behavior. However, there is no reason to think that such confounders are correlated with the changes in pollution due to the Asian Games. For this reason, the fact that we cannot control for these confounders is not likely to affect our point estimates but could affect the precision of our estimates. Despite these limitations, this paper offers an original contribution by exploiting a unique setting of exogenous change in air quality to estimate impacts of pollution on birth outcomes.

In summary, our paper identifies significant impacts of prenatal-air-pollution exposure on birth outcomes by exploiting a natural experiment – the Asian Games – during which air quality was exogenously improved. This approach is advantageous in addressing self-selection related to exposure and disentangling confounding meteorological factors from air pollution, as the Asian Games event provided an excellent opportunity to observe exogenous change in air pollution but not in weather conditions. Finally, maternal age and child sex differences in vulnerability to air pollution in our findings suggest that these heterogeneities deserve future study.

References

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature*, 56(4): 1360–1446.

Almond, Douglas, Lena Edlund, and Mårten Palme. 2009. "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden." *The Quarterly Journal of Economics*, 124(4): 1729–1772.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *The Economic Journal*, 126(591): 257–280.

Behrman, Jere R., and Mark R. Rosenzweig. 2004. "Returns to Birthweight." *The Review of Economics and Statistics*, 86(2): 586–601.

Beltran, Alyssa J., Jun Wu, and Olivier Laurent. 2014. "Associations of Meteorology with Adverse Pregnancy Outcomes: A Systematic Review of Preeclampsia, Preterm Birth and Birth Weight." *International Journal of Environmental Research and Public Health*, 11(1): 91–172.

Bharadwaj, Prashant, Matthew Gibson, Joshua Graff Zivin, and Christopher Neilson. 2017. "Gray Matters: Fetal Pollution Exposure and Human Capital Formation." *Journal of the Association of Environmental and Resource Economists*, 4(2): 505–542. Publisher: The University of Chicago Press.

Bosetti, Cristina, Mark J Nieuwenhuijsen, Silvano Gallus, Sonia Cipriani, Carlo La Vecchia, and Fabio Parazzini. 2010. "Ambient particulate matter and preterm birth or birth weight: a review of the literature." *Archives of toxicology*, 84(6): 447–460.

Bröde, Peter, Dusan Fiala, Krzysztof Błażejczyk, Ingvar Holmér, Gerd Jendritzky, Bernhard Kampmann, Birger Tinz, and George Havenith. 2012. "Deriving the operational procedure for the Universal Thermal Climate Index (UTCI)." *International Journal of Biometeorology*, 56(3): 481–494.

Chen, Le-Yu, and Christine Ho. 2016. "Incense Burning during Pregnancy and Birth Weight and Head Circumference among Term Births: The Taiwan Birth Cohort Study." *Environmental Health Perspectives*, 124(9): 1487–1492.

Chen, Xi, Chih Ming Tan, Xiaobo Zhang, and Xin Zhang. 2020. "The effects of prenatal exposure to temperature extremes on birth outcomes: the case of China." *Journal of Population Economics*.

Clougherty, Jane E. 2010. "A growing role for gender analysis in air pollution epidemiology." *Environmental health perspectives*, 118(2): 167.

Coneus, Katja, and C. Katharina Spiess. 2012. "Pollution exposure and child health: Evidence for infants and toddlers in Germany." *Journal of Health Economics*, 31(1): 180–196.

Conley, Dalton, Kate W Strully, and Neil G Bennett. 2003. *The Starting Gate: Birth Weight and Life Chances.* University of California Press.

Currie, Janet, and Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics*, 3(1): 65–90.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What Do We Know About Short-and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics*, 6(1): 217–247.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder. 2009. "Air pollution and infant health: Lessons from New Jersey." *Journal of Health Economics*, 28(3): 688–703.

Department of Environmental Protection of Guangdong Government. 2010. "Notice on Air Quality Regulations during the Pre-Asian Games Period (Yue Environment Notice[2010] No.22)." http://gdee.gd.gov.cn/shbtwj/content/post_2303946.html.

DiPietro, J. A., and K. M. Voegtline. 2017. "The gestational foundation of sex differences in development and vulnerability." *Neuroscience*, 342: 4–20.

Dong, Guang-Hui, Zhengmin (Min) Qian, Pamela K. Xaverius, Edwin Trevathan, Salwa Maalouf, Jamaal Parker, Laiji Yang, Miao-Miao Liu, Da Wang, Wan-Hui Ren, Wenjun Ma, Jing Wang, Alan Zelicoff, Qiang Fu, and Maayan Simckes. 2013. "Association Between Long-Term Air Pollution and Increased Blood Pressure and Hypertension in China." *Hypertension*, 61(3): 578–584.

European Centre for Medium-Range Weather Forecasts. 2020*a*. "Climate Data Store." https://cds.climate.copernicus.eu/cdsapp#!/home(Accessed at 2020-05-24).

European Centre for Medium-range Weather Forecasts. 2020*b*. "Thermal comfort indices derived from ERA5 reanalysis." https://cds.climate.copernicus.eu/doi/10.24381/cds.553b7518 (Accessed at 2020-04-20).

Fleischer, Nancy L., Mario Merialdi, Aaron van Donkelaar, Felipe Vadillo-Ortega, Randall V. Martin, Ana Pilar Betran, João Paulo Souza, and Marie S. O´Neill. 2014. "Outdoor Air Pollution, Preterm Birth, and Low Birth Weight: Analysis of the World Health Organization Global Survey on Maternal and Perinatal Health." *Environmental Health Perspectives*, 122(4): 425–430.

Fu, Li, Yi Chen, Xinyi Yang, Zuyao Yang, Sha Liu, Lei Pei, Baixiang Feng, Ganxiang Cao, Xin Liu, Hualiang Lin, Xing Li, Yufeng Ye, Bo Zhang, Jiufeng Sun, Xiaojun Xu, Tao Liu, and Wenjun Ma. 2019. "The associations of air pollution exposure during pregnancy with fetal growth and anthropometric measurements at birth: a systematic review and meta-analysis." *Environmental Science and Pollution Research*, 26(20): 20137–20147.

Ghosh, Rakesh, Judith Rankin, Tanja Pless-Mulloli, and Svetlana Glinianaia. 2007. "Does the effect of air pollution on pregnancy outcomes differ by gender? A systematic review." *Environmental Research*, 105(3): 400–408.

Grafova, Irina B., Vicki A. Freedman, Nicole Lurie, Rizie Kumar, and Jeannette Rogowski. 2014. "The difference-in-difference method: Assessing the selection bias in the effects of neighborhood environment on health." *Economics & Human Biology*, 13: 20–33.

Guan, Wei-Jie, Xue-Yan Zheng, Kian Fan Chung, and Nan-Shan Zhong. 2016. "Impact of air pollution on the burden of chronic respiratory diseases in China: time for urgent action." *The Lancet*, 388(10054): 1939–1951.

Hajat, Anjum, Charlene Hsia, and Marie S. O'Neill. 2015. "Socioeconomic Disparities and Air Pollution Exposure: a Global Review." *Current Environmental Health Reports*, 2(4): 440–450.

Han, Yingying, Panhua Jiang, Tianyu Dong, Xinliang Ding, Ting Chen, Gro Dehli Villanger, Heidi Aase, Lu Huang, and Yankai Xia. 2018. "Maternal air pollution exposure and preterm birth in Wuxi, China: Effect modification by maternal age." *Ecotoxicology and Environmental Safety*, 157: 457–462.

Health Effects Institute. 2019. "State of Global Air 2019." Health Effects Institute Special Report ISSN 2578-6873, Boston, MA.

He, Guojun, Maoyong Fan, and Maigeng Zhou. 2016. "The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games." *Journal of Environmental Economics and Management*, 79: 18–39.

**He, Jian-Rong, Yu Liu, Xiao-Yan Xia, Wen-Jun Ma, Hua-Liang Lin, Hai-Dong Kan, Jin-Hua Lu, Qiong Feng, Wei-Jian Mo, Ping Wang, Hui-Min Xia, Xiu Qiu, and Louis J.

Muglia. 2016. "Ambient Temperature and the Risk of Preterm Birth in Guangzhou, China (2001–2011)." *Environmental Health Perspectives*, 124(7): 1100–1106.

Huang, Ivory, Diane Mak, Prudence Cheung, Miriam Abraham, Tom Clemens, and Steve Turner. 2019. "A systematic review of associations between maternal exposures during pregnancy other than smoking and antenatal fetal measurements." *Environmental Research*, 173: 528–538.

Ito, Kazuhiko, George D Thurston, and Robert A Silverman. 2007. "Characterization of PM2.5, gaseous pollutants, and meteorological interactions in the context of time-series health effects models." *Journal of Exposure Science & Environmental Epidemiology*, 17(S2): S45–S60.

Jacobs, Milena, Guicheng Zhang, Shu Chen, Ben Mullins, Michelle Bell, Lan Jin, Yuming Guo, Rachel Huxley, and Gavin Pereira. 2017. "The association between ambient air pollution and selected adverse pregnancy outcomes in China: A systematic review." *Science of The Total Environment*, 579: 1179–1192.

Jayachandran, Seema. 2009. "Air quality and early-life mortality evidence from Indonesia's wildfires." *Journal of Human Resources*, 44(4): 916–954.

Keitt, Sarah K, Thomas F Fagan, and Sherry A Marts. 2004. "Understanding sex differences in environmental health: a thought leaders' roundtable." *Environmental Health Perspectives*, 112(5): 604–609.

Kim, Younoh, Scott Knowles, James Manley, and Vlad Radoias. 2017. "Long-run health consequences of air pollution: Evidence from Indonesia's forest fires of 1997." *Economics & Human Biology*, 26: 186–198.

Liang, Zhijiang, Yin Yang, Jing Yi, Zhengmin Qian, Zilong Zhang, Stephen Edward McMillin, Echu Liu, Hualiang Lin, and Guocheng Liu. 2021. "Maternal PM2.5 exposure associated with stillbirth: A large birth cohort study in seven Chinese cities." *International Journal of Hygiene and Environmental Health*, 236: 113795.

Li, Shanshan, Yuming Guo, and Gail Williams. 2016. "Acute Impact of Hourly Ambient Air Pollution on Preterm Birth." *Environmental Health Perspectives*, 124(10): 1623–1629.

Liu, Huan, Xuemei Wang, Jinpu Zhang, Kebin He, Ye Wu, and Jiayu Xu. 2013. "Emission controls and changes in air quality in Guangzhou during the Asian Games." *Atmospheric Environment*, 76: 81–93.

Liu, Shiliang, Daniel Krewski, Yuanli Shi, Yue Chen, and Richard T Burnett. 2003. "Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada." *Environmental Health Perspectives*, 111(14): 1773–1778.

Liu, Xiaoying, Jere Behrman, Emily Hannum, Fan Wang, and Qingguo Zhao. 2021. "Same Environment, Stratified Impacts? Air Pollution, Extreme Temperatures, and Birth Weight in Southeast China." PIER Working Paper No. 21-001. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3761257.

Liu, Ying, Jihong Xu, Dian Chen, Pei Sun, and Xu Ma. 2019. "The association between air pollution and preterm birth and low birth weight in Guangdong, China." *BMC Public Health*, 19(1): 1–10.

Lu, Chan, Lanqin Cao, Dan Norbäck, Yuguo Li, Jing Chen, and Qihong Deng. 2019. "Combined effects of traffic air pollution and home environmental factors on preterm birth in China." *Ecotoxicology and Environmental Safety*, 184: 109639.

Lu, Chan, Weishe Zhang, Xiangrong Zheng, Jingchi Sun, Lv Chen, and Qihong Deng. 2020. "Combined effects of ambient air pollution and home environmental factors on low birth weight." *Chemosphere*, 240: 124836.

Luechinger, Simon. 2014. "Air pollution and infant mortality: A natural experiment from power plant desulfurization." *Journal of Health Economics*, 37: 219–231.

Mercer, Laina D., Adam A. Szpiro, Lianne Sheppard, Johan Lindström, Sara D. Adar, Ryan W. Allen, Edward L. Avol, Assaf P. Oron, Timothy Larson, L. J. Sally Liu, and Joel D. Kaufman. 2011. "Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air)." *Atmospheric Environment*, 45(26): 4412–4420.

Murray, L. J, D. P. J O'Reilly, N Betts, C. C Patterson, G Davey Smith, and A. E Evans. 2000. "Season and outdoor ambient temperature: effects on birth weight." *Obstetrics & Gynecology*, 96(5, Part 1): 689–695.

Møllegaard, Stine. 2020. "The effect of birth weight on behavioral problems in early adolescence: New evidence from monozygotic twins." *Economics Human Biology*, 36: 100828.

National Aeronautics and Space Administration, U.S.A. 2020. "Airborne Nitrogen Dioxide Plummets Over China." https://earthobservatory.nasa.gov/images/146362/airborne-nitrogen-dioxide-plummets-over-china.

Qian, Zhengmin, Shengwen Liang, Shaoping Yang, Edwin Trevathan, Zhen Huang, Rong Yang, Jing Wang, Ke Hu, Yiming Zhang, Michael Vaughn, Longjiao Shen, Wenjin Liu, Pu Li, Patrick Ward, Li Yang, Wei Zhang, Wei Chen, Guanghui Dong, Tongzhang Zheng, Shunqing Xu, and Bin Zhang. 2016. "Ambient air pollution and preterm birth: A prospective birth cohort study in Wuhan, China." *International Journal of Hygiene and Environmental Health*, 219(2): 195–203.

Rangel, Marcos A, and Tom S Vogl. 2019. "Agricultural fires and health at birth." *The Review of Economics and Statistics*, 101(4): 616–630.

Rich, David Q., Kaibo Liu, Jinliang Zhang, Sally W. Thurston, Timothy P. Stevens, Ying Pan, Cathleen Kane, Barry Weinberger, Pamela Ohman-Strickland, Tracey J. Woodruff, Xiaoli Duan, Vanessa Assibey-Mensah, and Junfeng Zhang. 2015. "Differences in Birth Weight Associated with the 2008 Beijing Olympics Air Pollution Reduction: Results from a Natural Experiment." *Environmental Health Perspectives*, 123(9): 880–887.

Sanders, Nicholas J. 2012. "What doesn't kill you makes you weaker: prenatal pollution exposure and educational outcomes." *Journal of Human Resources*, 47(3): 826–850.

Sauer, Mark V. 2015. "Reproduction at an advanced maternal age and maternal health." *Fertility and Sterility*, 103(5): 1136–1143.

Shah, Prakesh S., Taiba Balkhair, and Knowledge Synthesis Group on Determinants of Preterm/LBW births. 2011. "Air pollution and birth outcomes: a systematic review." *Environment International*, 37(2): 498–516. 00142.

Siniarska, A., and S. Kozieł. 2010. "Association of birth weight and length with air temperature, sunlight, humidity and rainfall in the city of Warsaw, Poland." *HOMO*, 61(5): 373–380.

Šrám, Radim J, Blanka Binková, Jan Dejmek, and Martin Bobak. 2005. "Ambient air pollution and pregnancy outcomes: a review of the literature." *Environmental Health Perspectives*, 375–382.

Strand, Linn Beate, Adrian G. Barnett, and Shilu Tong. 2011. "Methodological challenges when estimating the effects of season and seasonal exposures on birth outcomes." *BMC Medical Research Methodology*, 11(1): 1–8.

Sun, Zhe, Liyang Yang, Xiaoxia Bai, Wei Du, Guofeng Shen, Jie Fei, Yonghui Wang, An Chen, Yuanchen Chen, and Meirong Zhao. 2019. "Maternal ambient air pollution exposure with spatial-temporal variations and preterm birth risk assessment during 2013–2017 in Zhejiang Province, China." *Environment International*, 133: 105242.

Tang, Deliang, Ting Yu Li, Judith C Chow, Sanasi U Kulkarni, John G Watson, Steven Sai Hang Ho, Zhang Y Quan, LR Qu, and Frederica Perera. 2014. "Air pollution effects on fetal and child development: A cohort comparison in China." *Environmental Pollution*, 185: 90–96.

Torche, Florencia, and Dalton Conley. 2016. "A Pound of Flesh: The Use of Birthweight as a Measure of Human Capital Endowment in Economics Research." *The Oxford Handbook of Economics and Human Biology*.

Tsoli, Stergiani, George B. Ploubidis, and Olga-Ioanna Kalantzi. 2019. "Particulate air pollution and birth weight: A systematic literature review." *Atmospheric Pollution Research*, 10(4): 1084–1122.

Wang, Qiong, Huazhang Miao, Joshua L. Warren, Meng Ren, Tarik Benmarhnia, Luke D. Knibbs, Huanhuan Zhang, Qingguo Zhao, and Cunrui Huang. 2021. "Association of maternal ozone exposure with term low birth weight and susceptible window identification." *Environment International*, 146: 106208.

World Health Organization. 2019. "Ambient air pollution: Health impacts." http://www.who.int/airpollution/ambient/health-impacts/en/.

Xie, Yanxia, Xiaodong Wang, Yi Mu, Zheng Liu, Yanping Wang, Xiaohong Li, Li Dai, Qi Li, Mingrong Li, Peiran Chen, Jun Zhu, and Juan Liang. 2021. "Characteristics and adverse outcomes of Chinese adolescent pregnancies between 2012 and 2019." *Scientific Reports*, 11(1): 12508.

Xu, H.M., J. Tao, Steven Sai Hang Ho, K.F. Ho, J.J. Cao, N. Li, Judith C. Chow, G.H. Wang, Y.M. Han, R.J. Zhang, John G. Watson, and J.Q. Zhang. 2013. "Characteristics of fine particulate non-polar organic compounds in Guangzhou during the 16th Asian Games: Effectiveness of air pollution controls." *Atmospheric Environment*, 76: 94–101.

Yang, Yuxiang. 2021. "Analysing the seasonality of births in mainland China." *Journal of Biosocial Science*, 53(2): 233–246. Publisher: Cambridge University Press.

Zhu, P. F., Y. Zhang, J. Ban, T. T. Li, and X. M. Shi. 2017. "Air pollution and adverse birth outcome in China: a comprehensive review." *Zhonghua Liu Xing Bing Xue Za Zhi (Chinese)*, 38(3): 393–399.

Appendix 1: Tables and Figures

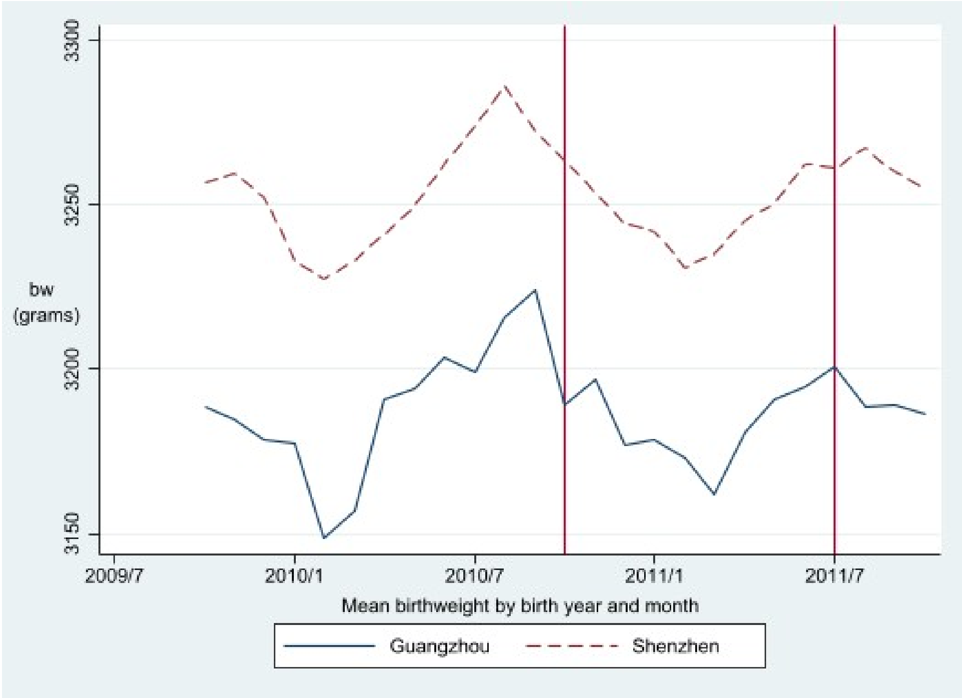

Figure A1: Time Trend of Mean Birthweight by City and Month

Note: Monthly average birth weight is calculated based on birth certificate for each city at each birth month. As Asian Games were held during Oct 01, 2010 and Dec 20, 2010, any births with prenatal periods overlapped with this period (or born between October 2010 and September 2011) are included in the area between the two vertical lines.

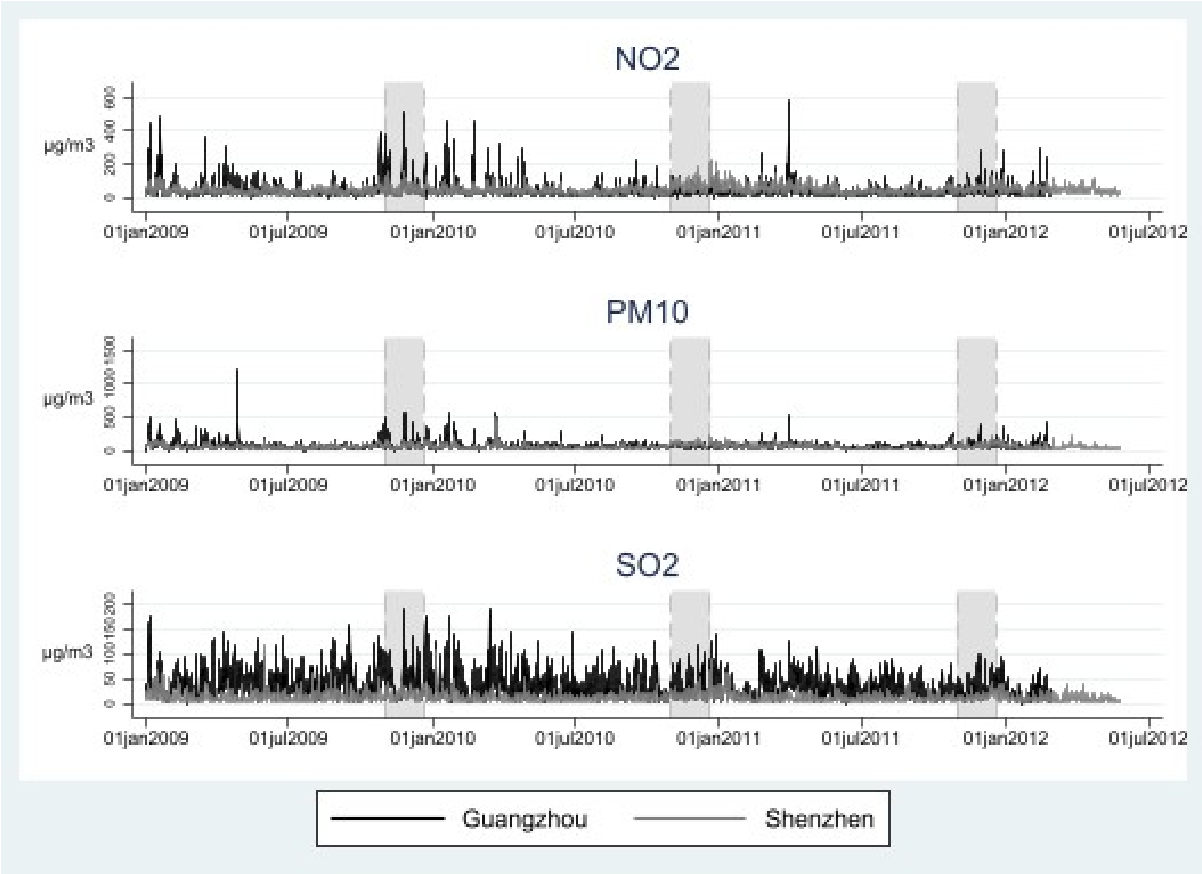

Figure A2: The Daily Concentrations of Three Air Pollutants in Guangzhou and Shenzhen Cities During 2009-2011

Note: The daily concentrations of each pollutant (PM_{10}, NO_2 and SO_2 are the average concentration across the whole city. The three periods shaded in grey between two neighboring vertical dashed lines are the same season as the Asian Games in each year, i.e., October 20-December 20 in 2009, 2010, and 2011, separately.

Figure A3: The Number of Live Single Births by Month

Table A1: Study Population Characteristics and Exposure Measurements by City [n (%) or mean(sd)]

	(1) Sample		(2) Guangzhou		(3) Shenzhen		(4) p value for
	mean/n	sd/%	mean/n	sd/%	mean/n	sd/%	t or Chi-square test
Sex (female = 1)	0.46	(0.50)	0.47	(0.50)	0.46	(0.50)	0.57
Preterm	0.05	(0.22)	0.07	(0.26)	0.05	(0.21)	0.00
Birthweight	3240.44	(456.73)	3181.36	(469.68)	3248.15	(454.45)	0.00
Low Birthweight	0.04	(0.20)	0.06	(0.23)	0.04	(0.19)	0.00
Maternal age	27.31	(4.74)	29.06	(4.19)	27.08	(4.76)	0.00
Parity	1.46	(0.72)	1.25	(0.52)	1.49	(0.74)	0.00
Mother's education, n (%)							0.00
< high school	194,990	(37.3%)	3,641	(6.03%)	191,349	(41.38%)	
high school graduate or equivalent	198,140	(37.91%)	36,895	(61.15%)	161,245	(34.87%)	
3 year college	63,226	(12.1%)	8,252	(13.68%)	54,974	(11.89%)	
4 year college graduate	66,347	(12.69%)	11,546	(19.14%)	54,801	(11.85%)	
Mother's job category, n (%)							0.00
Staff	56,596	(10.83 %)	56,596	(93.80 %)	73,438	(15.88 %)	
Agricultural related	5,983	(1.14%)	276	(0.46%)	5,707	(1.23%)	
Housewife	129,749	(24.82%)			129,749	(28.06%)	
Professional and technical personnel	2,050	(0.39%)	497	(0.82%)	1,553	(0.34%)	
Business	26,109	(4.99%)	202	(0.33%)	25,907	(5.60%)	
Worker	35,587	(6.81%)	76	(0.13%)	35,511	(7.68%)	
Other	2,687	(0.51%)	2,687	(4.45%)			
Missing	190,504	(36.45%)			190,504	(41.20%)	
Daily mean PM_{10} exposure(µg/m^3)	60.41	(7.37)	72.33	(7.08)	58.85	(5.82)	0.00
Daily mean NO_2 exposure(µg/m^3)	41.93	(5.09)	40.34	(5.98)	42.14	(4.92)	0.00
Daily mean SO_2 exposure(µg/m^3)	14.25	(7.86)	35.29	(5.17)	11.50	(1.05)	0.00
Daily mean temperature(°C)	23.18	(1.57)	22.76	(1.80)	23.23	(1.52)	0.00
Daily mean relative humidity(%)	75.37	(3.34)	72.17	(2.91)	75.79	(3.16)	0.00
% Time exposed to UTCI <= 0°C	1.39	(0.87)	1.96	(1.19)	1.31	(0.80)	0.00
% Time exposed to UTCI >= 34°C	0.43	(0.95)	2.19	(1.87)	0.20	(0.32)	0.00
N	522703		60334		462369		

Notes: Birth weight and gestational age are obtained from birth records from each city. Gestational age is calculated based on self-reported last menstrual period. The remaining variables of individual characteristics are self-reported.

Table A2: Distribution of Daily Mean Universal Thermal Climate Indices (UTCI) in Guangzhou and Shenzhen Between 2000 and 2008.

	Guangzhou	Shenzhen
Top 1% ($^\circ$C)	34.23	33.35
Top 2.5% ($^\circ$C)	33.56	32.74
Bottom 2.5% ($^\circ$C)	2.52	3.82
Bottom 1% ($^\circ$C)	-0.87	-0.24
Mean ($^\circ$C)	22.61	22.46
s.d.	8.74	8.05
Number of days above 34 $^\circ$C	45	12
Number of days above 33 $^\circ$C	172	54

Source: Copernicus and the European Centre for Medium-Range Weather Forecasts to combine measurements of temperature and humidity (European Centre for Medium-range Weather Forecasts, 2020*b*).

Table A3: First-stage coefficients of different definitions of instrumental variables on prenatal exposure levels to different pollutants

	(1) PM_{10}	(2) NO_2	(3) SO_2
Specification A: temperature + relative humidity	-0.161***	-0.262***	-0.126***
	(0.000552)	(0.000643)	(0.000388)
Specification B: extreme cold + extreme hot defined by UTCI	-0.164***	-0.236***	-0.108***
	(0.000567)	(0.000732)	(0.000483)

Notes: Each estimate represents a regression coefficient of the first-stage regression of prenatal exposure to each pollutant on the instrumental variable, i.e., a continuous variable summarizing the number of days during one's expected pregnancy (given 39 weeks of gestation) overlapping with the Asian Games period with two specifications: Specification A controls for linear term of daily mean temperature and daily mean relative humidity; Specification B controls for extreme cold and hot weather during pregnancy defined by UTCI. Besides meteorological variables, other control variables include sex, quadratic term of maternal age, parity, education categories, day-of-a-week dummies, conception year fixed effects, city-specific conception season fixed effects. *** significant at 1%. Robust standard errors are displayed in parentheses.

Table A4: Marginal Effects of NO_2 (by 10 μg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm.

	Probit Model/OLS Model						IV-Probit Model/2SLS Model					
VARIABLES	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NO_2*μg/m^3	0.002***	0.002***	-8.56***	-7.074***	0.004***	0.003***	0.0020*	0.0016	-8.70***	-9.28**	0.0038***	0.0033**
	(0.001)	(0.001)	(2.311)	(1.866)	(0.001)	(0.001)	(0.001)	(0.001)	(3.16)	(3.72)	(0.001)	(0.002)
Temp. (°C)	-.044 ***		115.42***		-0.054 ***		-0.004***		11.536***		-.0053691 ***	
	(0.004)		(9.297)		(0.004)		(0.000)		(0.93)		(0.00041)	
Rel. hum. (%)	0.011***		-32.79***		0.016***		0.001***		-3.296***		.0015667***	
	(0.002)		(4.523)		(0.002)		(0.000)		(0.51)		(0.00023)	
Ext. cold (%)		0.021***		-55.506***		.029***		0.002***		-5.581***		0.0029***
		(0.005)		(1.247)		(0.006)		(0.001)		(1.25)		(0.001)
Ext. hot (%)		9.86E-05		10.052		0.003		0.000		1.49		0.00026
		(0.005)		(12.226)		(0.005)		(0.001)		(1.40)		(0.001)
Obs.	522,703						522,703					
Spec.	A	B	A	B	A	B	A	B	A	B	A	B

Notes: Each column represents a separate regression. (1)-(6) are from Probit/OLS regressions without applying IV, while (7)-(12) are from IVProbit/2SLS regressions. (1),(2),(5),(6),(7),(8),(11),(12) are marginal effects of 10 μg/m^3 NO_2 on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4),(9),(10) are marginal effects of 10μg/m^3 No_2 on Mean Birthweight. Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Table A5: Marginal Effects of SO_2 (by 10 μg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm.

VARIABLES	Probit Model/OLS Model						IV-Probit Model/2SLS Model					
	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SO_2* 10μg/m^3	0.01***	0.009***	-15.97***	-14.36***	0.014***	0.0133***	0.0045*	0.0031	-18.12***	-20.34**	.0083 ***	0.0076**
	(0.002)	(0.002)	(4.213)	(4.741)	(0.002)	(0.002)	(0.002)	(0.003)	(6.570)	(8.160)	(0.003)	(0.003)
Temp. (°C)	-0.0048***		12.298***		-0.00589***		-0.005***		12.354***		-0.0057 ***	
	(0.000)		(0.931)		(0.000)		(0.000)		(0.94)		(0.00415)	
Rel. hum. (%)	0.0011***		-2.531***		0.00137***		0.001***		-2.564***		0.0013***	
	(0.000)		(0.370)		(0.000)		(0.000)		(0.38)		(0.00178)	
Ext. cold (%)		0.002***		-5.639***		0.003***		0.002***		-5.716***		0.003***
		(0.001)		(1.248)		(0.001)		(0.001)		(1.251)		(0.001)
Ext. hot (%)		-0.0007		1.267		-0.00072		0.000		2.023		0.000
		(0.001)		(1.299)		(0.001)		(0.001)		(1.524)		(0.001)
Obs.	522,703						522,703					
Spec.	A	B	A	B	A	B	A	B	A	B	A	B

Notes: Each column represents a separate regression. (1)-(6) are from Probit/OLS regressions without applying IV, while (7)-(12) are from IVProbit/2SLS regressions. (1),(2),(5),(6),(7),(8),(11),(12) are marginal effects of 10 μg/m^3 SO_2 on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4),(9),(10) are marginal effects of 10μg/m^3 SO_2 on Mean Birthweight. Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Table A6: Heterogeneous Effects of NO_2 on Birth Outcomes by Sex and Maternal Age Group

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
NO_2 * male	0.000414	0.000419	2.086**	2.086**	-.0012	-.0011
	(0.005)	(0.005)	(1.066)	(1.066)	(0.004)	(0.004)
NO_2 * (momage>=35)	0.00169	0.00169	-7.757***	-7.747***	0.00199	0.00199
	(0.00126)	(0.00126)	(2.417)	(2.417)	(0.00142)	(0.00142)
Specification	A	B	A	B	A	B

Notes: IV estimates of the interaction term between NO_2 and male or age35+ on birth outcomes, with the same set of control variables as in Table 2. *** significant at 1%, ** significant at 5%, *significant at 10%.

Table A7: Heterogeneous Effects of SO_2 on Birth Outcomes by Sex and Maternal Age Group

VARIABLES	Low Birthweight		Birthweight		Preterm	
	(1)	(2)	(3)	(4)	(5)	(6)
SO_2 * male	-.000151	-0.000151	-0.806**	-0.808*	0.000450	0.000452
	(0.002)	(0.002)	(0.412)	(0.412)	(0.002)	(0.002)
SO_2 * (momage>=35)	0.000483	0.000483	-2.218***	-2.214***	0.000569	0.000569
	(0.000359)	(0.000359)	(0.684)	(0.684)	(0.000405)	(0.000405)
Specification	A	B	A	B	A	B

Notes: IV estimates of the interaction term between SO_2 and male or age35+ on birth outcomes, with the same set of control variables as in Table 2. *** significant at 1%, ** significant at 5%, *significant at 10%.

Appendix 2: Robustness check: missing observations of labor migrants

There still may be concern over labor migrants in the sample, as these two cities are the most popular destinations for labor migrants in China. Labor migrants may return to their hometowns for child birth. Although this should not be a concern for this study as long as labor migrants do not leave the city for child birth specifically during the Asian Games period, we do find that the number of live deliveries is lower during thr Asian Games period in Guangzhou (20 October 2010 - 20 December 2010) compared to the same period in other years, relative to Shenzhen (see Figure A3.). As we do not observe migration status of women in the Guangzhou sample, we try to control for mothers' job types as a sensitivity check as labor migrants are more likely to work as "workers". The main results remain basically the same, as shown in Table A8. Because mothers' job type information was not collected in a consistent framework between the two cities, and 40.3% of the population have missing information on job, [19] we do not present these results as the main results. The similar results on air pollution exposure after controlling for mothers' job categories should help alleviate concern over missing labor migrants in the sample.

[19] In the regression that controls for maternal job category, we treat missing job information as a job category because it is likely that missing is not random.

Table A8: Marginal Effects of PM_{10}, NO_2 and SO_2 (by 10 µg/m^3) on the Probability of Low Birthweight, Mean Birthweight and Risk of Preterm (by IV-probit/2SLS Estimates, with Control of Mother's Job Category).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
			PM_{10}						NO_2						SO_2			
VARIABLES	Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm		Low Birthweight		Birthweight		Preterm	
Pollutant(10µg/m^3)	0.003	0.0016	-11.69**	-11.33**	0.0061***	0.0040*	0.000	0.0012	-7.16**	-7.84**	0.0035***	0.0028*	0.000	0.0029	-14.94**	-17.19**	0.0077***	0.0068**
	(0.002)	(0.002)	(0.50)	(0.52)	(0.000)	(0.002)	(0.000)	(0.001)	(3.06)	(3.62)	(0.001)	(0.002)	(0.000)	(0.000)	(6.38)	(7.93)	(0.003)	(0.003)
Temp.(°C)	-.0041***		11.09***		-.0047***		-.0044***		11.73***		-.0053		-0.0046***		12.41***		-0.0056***	
	(0.000)		(1.01)		(0.000)		(0.000)		(0.94)		(0.000)		(0.000)		(9.45)		(0.000)	
Rel. hum.(%)	0.0012***		-3.54***		0.0017***		0.0011***		-3.26***		0.0016***		0.001***		-2.66***		0.0013***	
	(0.000)		(0.60)		(0.000)		(0.000)		(0.51)		(0.000)		(0.000)		(0.38)		(0.000)	
Ext. cold (%)		0.0019***		-6.56***		0.0027***		0.002***		-5.88***		0.0029***		0.0021***		-5.99***		0.0029***
		(0.001)		(1.305)		(0.001)		(0.001)		(1.25)		(0.001)		(0.001)		(1.26)		(0.001)
Ext. hot (%)		0.000		0.16		0.00090*		0.000		1.51		0.00035		0.000		1.96		0.000
		(0.000)		(1.169)		(0.001)		(0.001)		(1.40)		(0.001)		(0.001)		(1.53)		(0.001)
Obs.			516,720						516,720						516,720			
Specification	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Mother's job	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: The sample removes individuals who conduct agricultural related job. Each column represents a separate regression from IV-Probit/2SLS regressions. (1),(2),(5),(6) are marginal effects of 10 µg/m^3 PM_{10} on the probability of Low Birthweight and Preterm at the mean values of all control variables. (3),(4) are marginal effects of 10µg/m^3 PM_{10} on Mean Birthweight (similar for other two pollutants). Control variables include parity, sex of the baby, quadratic function of maternal age, maternal education categories, maternal job category, conception year fixed effects, city-specific conception season fixed effects and weather variables. Specification A and B vary with the weather variables. Specification A controls for linear terms of mean temperature and relative humidity during one's pregnancy, while specification B controls for the percentage of pregnancy exposed to extreme cold and hot defined by UTCI. *** significant at 1%, ** significant at 5%, *significant at 10%. Robust standard errors are in parentheses.

Author contributions:

Xiaoying Liu: Conceptualization, Data curation, Formal analysis, Software, Funding acquisition, Methodology, Writing –original draft, Writing –review & editing.

Huazhang Miao: Data curation, Writing –review & editing.

Jere Behrman: Funding acquisition, Methodology, Writing –review & editing.

Emily Hannum: Funding acquisition, Methodology, Writing –review & editing.

Zhijiang Liang: Data curation, Writing –review & editing.

Qingguo Zhao: Conceptualization, Data curation, Resources, Writing –review & editing.