From Missing Girls to Missing Boys: A Partial Feedback Loop between Gender Ratio Imbalance and Pollution*

Ran Song
National University of Singapore

Shang-jin Wei
Columbia University and NBER

Abstract

While discrimination against girls has been widely documented to generate a "missing girls" problem, this paper documents a new empirical pattern: air pollution can generate a "missing boys" problem among rural populations in China. Using individual-level birth data and an instrumental variable strategy leveraging variation in wind direction combined with the siting of distant coal-fired plants, we show that higher PM2.5 exposure significantly reduces the probability that a newborn is male among rural births. We then examine a reverse relationship and show that locations with more male-skewed early rural cohorts experience higher air pollution fifteen years later, when these cohorts reach marriageable age. This pattern is established using both cross-sectional variation in baseline sex ratios and a contiguous cross-province city-pair panel design exploiting differential changes in financial penalties under the One-Child Policy. Together, the results reveal an unintended demographic–environmental feedback loop operating through a harmful externality.

JEL Classification: J13, J16, Q53
Keywords: Sex Ratio Imbalance, Air Pollution, Rural *hukou*, China

* Ran Song, Department of Economics, National University of Singapore, Email: ran.song@nus.edu.sg.
Shang-jin Wei, Graduate School of Business and School of International and Public Affairs, Columbia University and NBER, Email: sw2446@gsb.columbia.edu.

1. Introduction

Skewed sex ratios at birth are among the most consequential demographic distortions of the past several decades. A large literature has examined how male-biased cohorts affect marriage markets, savings behavior, crime, and labor supply (Wei and Zhang, 2011; Tan, Wei and Zhang, 2021). Much less is known about whether environmental conditions themselves can shape demographic outcomes, and whether such effects interact with existing sex-ratio imbalances.

We begin by studying how air pollution affects the sex composition of births. A substantial biomedical and epidemiological literature suggests that male fetuses are more vulnerable to adverse environmental conditions. Establishing a causal relationship in an economic setting is challenging because pollution is endogenous to income, industrial activity, and health inputs. We therefore employ an instrumental-variable strategy that exploits plausibly exogenous variation in pollution exposure arising from the interaction of prevailing wind direction with the location and the historical placement of distant thermal power plants (as in Gao et al., (2023); Khanna et al., (2025)).

Using individual-level birth data from rural populations, we show that higher PM2.5 exposure significantly reduces the probability that a newborn is male. The effect is concentrated among first births and second births with an older male sibling, and is absent for urban *hukou* births. These patterns are consistent with biological evidence on differential vulnerability of male fetuses and suggest that pollution exposure can contribute to a "missing boys" problem among disadvantaged populations.

Our second empirical result is on the reverse relationship and asks whether demographic imbalance itself can shape environmental outcomes. When marriage markets become more competitive, households, firms, and local governments may respond by reallocating effort and resources toward faster growth and visible economic success. When growth is achieved through pollution-intensive activity, a male-biased sex ratio can translate into higher ambient pollution. Using prefecture-level data, we show that locations with more male-heavy rural cohorts experience significantly higher air pollution approximately fifteen years later, when these cohorts reach marriageable age.

Our empirical analysis is deliberately organized to speak to both directions of the relationship. First, we document systematic associations between earlier cohort-level sex ratios and later pollution outcomes, and we quantify the timing pattern that motivates an interpretation related to

marriage-market entry. Second, we estimate the causal effect of pollution exposure on the probability of a male birth using an instrumental-variable design that leverages meteorological transport of emissions from large point sources.

The paper's mechanism can be summarized as a marriage-market pressure channel with an environmental externality. Let local marriage-market tightness increase the marginal value of income or status for households with marriage-age children. This increase can operate through several margins that are observationally distinct but empirically related: (i) households may increase labor supply or accept higher-risk/higher-pollution employment; (ii) local entrepreneurs may expand activity in pollution-intensive sectors; (iii) local governments may place greater weight on short-run growth or employment relative to environmental enforcement; and (iv) households may increase consumption of energy-intensive goods. Each channel predicts that higher marriage-market pressure should be associated with greater pollution exposure, particularly in settings where the production structure is energy intensive and where enforcement is elastic.

The framework also yields timing and heterogeneity predictions that are central for credibility. If marriage-market pressure is the operative force, effects should be most pronounced when the relevant cohorts reach marriageable age, and should be stronger in environments where marriage markets are more local and where rural institutions (hukou restrictions, labor-market segmentation, and differential access to public services) make the pressure more binding. The draft's rural-versus-urban contrast and the emphasis on cohort timing are therefore not merely descriptive—they play the role of falsification and scope-condition tests.

It has been well documented that discrimination of girls/women has led to an unbalanced male-female ratio at birth, which Sen (1992) called "missing girls" problem. This problem is serious in Azerbaijan, China, India, and Vietnam with severe consequences. In this paper, we first document a previously unknown data regularity - air pollution tends to create a "missing boys" problem. Families in regions with severe local air pollution tend to have a lower boys-to-girls ratio at birth. By exploring exogenous variations in the local exposure to air pollution, we establish that the relationship is causal. We then document a second data pattern – greater local sex ratio imbalance at birth (too many boys) itself tend to create incentives for local economy to specialize in pollution intensive industries in subsequent years. These two data patterns collectively form a partial loop, suggesting pollution may play a (partial) role in natural correction of sex ratio imbalance.

Related findings in public health and epidemiology. Our results are consistent with findings from the public health and epidemiologic literature. First, lab and clinical studies show that Y-chromosome-bearing spermatozoa are more vulnerable to stressful environmental conditions (such as exposure to particulate matter, temperature stress, and prolonged incubation) compared to their X counterparts (You et al., 2017; Radwan et al., 2018; Oyeyipo et al., 2017). Indeed, X- and Y-bearing spermatozoa differ in various molecular, structural, and physiological aspects and these differences might contribute to the altered viability of Y-bearing spermatozoa under stressful conditions (Rahman et al., 2020).

Second, a large body of empirical evidence documents that the effects of prenatal exposure to negative environmental factors (such as extreme temperatures and air pollution) on birth outcomes are different between male and female fetuses. In particular, in-utero exposure to extreme temperatures increases stillbirth risk among male fetuses and results in fewer male births (Catalano, R., et al, 2008; Fukuda et al., 2014; Abdel Ghany et al., 2026; Basu, R., et al, 2016; Karlsson et al., 2021). Prenatal exposure to fine particles leads to birth weight and length deficiencies and preterm birth, with the effects being stronger for male newborns than for female newborns (Jedrychowski 2009; Lakshmanan 2021; Bekkar, B., et al.). Taken together, male fetuses are likely more sensitive to environmental stress exposure. Natural selection has conserved mechanisms by which women subjected to environmental stressors are more likely to abort frail male fetuses (Karlsson et al., 2021). Table C1 summarizes quantitative differences in vulnerability between X- and Y-bearing spermatozoa and between female and male fetuses, based on the studies reviewed above.

2. Data and Geographic Setup of Analysis

2.1 Pollution Data

We use satellite-based PM2.5 data to measure exposure to air pollution; according to the WHO, PM2.5 is considered the best indicator of health risk associated with air pollution.[1] The data cover periods before and after China began releasing official PM2.5 data in 2013. Prefecture-level annual PM2.5 concentrations are measured using the Global Annual PM2.5 Grids derived from satellite data by van Donkelaar et al. (2016). The raster grids in this ground-calibrated PM2.5 dataset have

[1] See WHO reports: http://www.who.int/mediacentre/news/releases/2014/air-quality/en/ .

a high grid-cell resolution of 0.01 degrees. This yields comprehensive and reliable measures of air quality for across all Chinese cities. Our satellite-based measure of PM2.5 is consistent with official PM2.5 measurements reported by the Ministry of Environmental Protection of China (MEP) (Appendix Figure A1). The correlation between satellite PM2.5 data and monitor-based PM2.5 data in China is as high as 0.8 (Gao, Song, and Timmins, 2023). [2]

2.2 Population Census of China

Individual Data on the Sex of Newborns The first part of our paper examines how pollution exposure is related to children's sex at birth. We use individual-level data drawn from China's 2015 National 1% Population Sample Survey to study how contemporaneous pollution exposure is associated with the sex of newborns. The survey records detailed information on children's sex and *hukou* status (rural or urban), as well as demographic characteristics of their mothers. We restrict the sample to children born in the 12 months preceding the survey month (October 2015) and classify these newborns into three groups: first-born children, second-born children with an older male sibling, and second-born children without an older male sibling.

Prefecture-level Data on Sex Ratios of Early Cohorts The second part of our analysis evaluates whether imbalanced sex ratios among early cohorts affect future pollution. We first leverage cross-sectional variation in early-cohort sex ratios across prefectures. Specifically, we use the 2000 Population Census of China to compute prefecture-level sex ratios for cohorts aged 3–10, separately for children with an urban *hukou* and those with a rural *hukou*, in 2000. We then link these early-cohort sex ratios to air pollution levels 15 years later, in 2015, when these individuals were aged 18–25. This allows us to assess how baseline sex ratios among early-age cohorts are associated with subsequent air pollution when these children have reached or are approaching marriageable age.

We next exploit intertemporal changes in early-cohort sex ratios within prefectures and employ changes in OCP financial penalties as an instrument for changes in early-cohort sex ratios.

[2] We focus on air pollution in our analysis, as water pollution has been controlled since 2000. Figure A2 shows that the level of water pollution emissions has declined since 2000, driven by China's water pollution controls under its 10th Five-Year Plan (2000–2005) and 11th Five-Year Plan (2006–2010). City-level water pollution is measured by chemical oxygen demand (COD) emissions and wastewater emissions, normalized by industrial output. In contrast, the Chinese government started reporting PM2.5 data and allocating more effort to controlling PM2.5 pollution only after 2012.

Specifically, we construct a two-period panel by combining early-age cohort sex ratios from 1990 and 2000. Sex ratio data for 1990 come from the 1990 Population Census. In this panel, we define children aged 3–10 in 1990 as the first cohort (period 1) and those aged 3–10 in 2000 as the second cohort (period 2). These cohorts reach or approach marriageable age in 2005 and 2015, respectively. We use this panel to examine how changes in early-age sex ratios between 1990 and 2000 lead to changes in air pollution between 2005 and 2015

2.3 Inputs Into Instrumental Variables and Controls

We construct an instrument for air pollution based on variation in wind direction and the historical placement of distant coal-fired power plants. Information on large-scale (capacity > 1 million KW) coal-fired power plants and their coal consumption comes from China's Electric Power Yearbooks and Energy Statistical Yearbooks. We manually collect information on each plant's establishment year and detailed location. Data on wind direction come from the China Meteorological Data Service Center.

We instrument for changes in early-cohort sex ratios using changes in OCP financial penalties for unauthorized births, following Wei and Zhang (2011). Data on OCP penalties are constructed by Ebenstein (2010). We obtain data on city attributes from China City Statistical Yearbooks. Appendix Table A1 reports summary statistics for the key variables used in the analysis.

2.4 Geographic Setup of Analysis

The geographic unit of our analysis is the Chinese prefecture-level city. Throughout the text, we use the terms "city" and "prefecture" interchangeably. China's administrative hierarchy comprises three levels: provinces, prefecture-level cities, and county-level units. A prefecture-level city encompasses an entire prefecture area, including both *urban* and *rural areas*. Importantly, a prefecture-level city is not a "city" in the conventional sense of a large, continuous urban settlement; rather, it is an administrative unit consisting of a central urban core (the core city) surrounded by rural areas (which together are divided into districts), as well as several surrounding county-level units (which also have their own urban areas surrounded by their own rural areas).

Within each prefecture, urban and rural areas are typically not clearly separated but are closely adjacent and spatially interspersed. Driven by China's rapid urbanization, a massive number of rural *hukou* holders work and reside in urban areas within prefectures.

Because different areas within the same prefecture lie within commuting distance and economic activities are highly integrated, a prefecture can be regarded as a labor market area and a commuting zone. Moreover, more than 80% of the spatial variation in air quality across China is driven by inter-prefecture differences (Gao et al., 2023).[3]

3. Descriptive Patterns of Pollution and Sex Ratios

This section describes the spatial and temporal patterns of pollution and sex ratios in the raw data. Figure 1a illustrates the geographic patterns of annual average PM2.5 concentration in 2015. Air pollution is unevenly distributed across space. The northeastern and eastern coastal regions experienced higher levels of air pollution. [4] Figure 1b illustrates the time trends of the mean and standard deviation of annual PM2.5 concentrations across Chinese cities since 1998. The average PM2.5 concentration in 2015 (39 $\mu g/m^3$) is approximately 40% greater than that in 1998 (27 $\mu g/m^3$); the mean level of PM2.5 pollution exceeds WHO air quality guidelines every year[5]. Moreover, increases in the mean coincided with increases in the cross-city variance of pollution.

Figure 2a plots the sex ratio among rural children aged 3–10 in 2000 against prefecture-level PM2.5 concentration in 2015. There is a clear positive association between early-cohort sex ratios and future pollution as the cohort reaches or approaches marriageable age. By contrast, Figure 2b shows that contemporaneous exposure to air pollution is negatively associated with the male-to-female ratio among rural newborns in 2015. If air pollution results in relatively fewer male births, these patterns suggest pollution may play a partial role in natural correction of sex ratio imbalance.

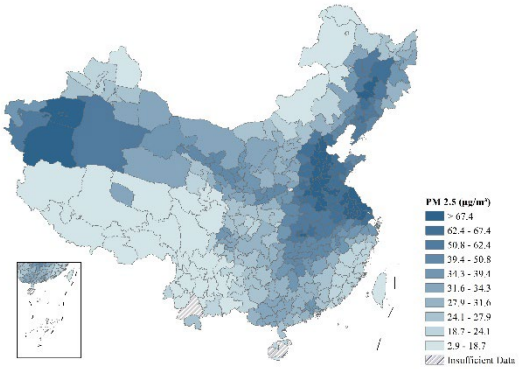
[3] We also include four centrally controlled municipalities (Beijing, Shanghai, Chongqing, and Tianjin), as their size is comparable to that of several large prefecture-level cities. Moreover, these municipalities also encompass both rural and urban areas within commuting distance. Our results are robust to excluding the centrally controlled municipalities along with several large prefecture-level cities.

[4] Manufacturing industries are concentrated in east China, and coal-burning for public heating contribute to severe pollution in the northeast.

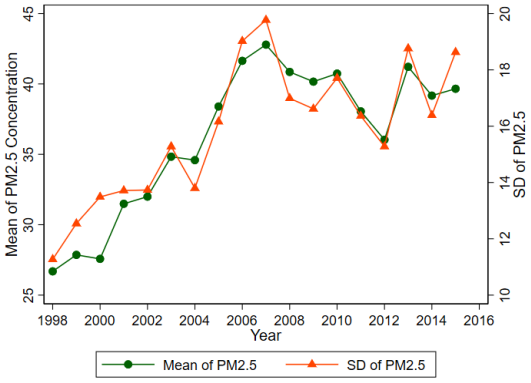
[5] See http://www.who.int/mediacentre/factsheets/fs313/en/ for information on WHO guidelines.

Figure 1: The Distribution in Pollution Across Cities and Over Time

(a) PM2. 5 Concentration in China



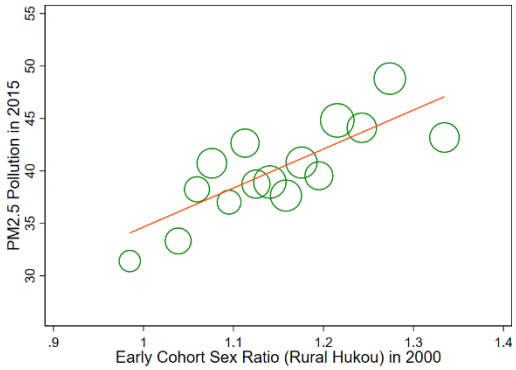
(b) Change in PM2.5 over time



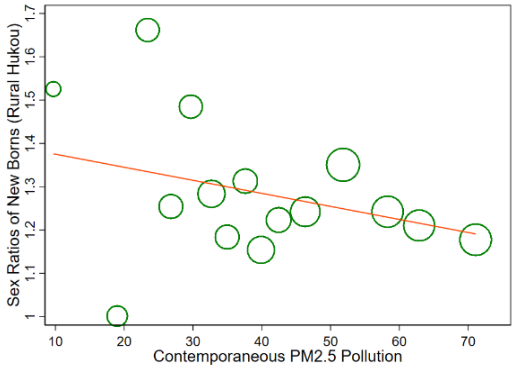
Notes: Spatial and temporal distribution of PM2.5 using the Global Annual PM2.5 Grids. The map shows the geographic spread in 2015. Figures 1b shows trends in the mean and the standard deviation of PM2.5 concentrations across Chinese cities from 1998 to 2016.

Figure 2: Pollution and Male-Female Ratios among Rural Children

(a) Early-Cohort Sex Ratios and Pollution 15 Years Later



(b) Contemporaneous Exposure and the Sex at Birth



N*otes:* Figure 2a plots the sex ratio among children (aged 3–10 with rural hukou) in 2000 against prefecture-level PM2.5 pollution in 2015. Cities are grouped into 15 groups according to the quantile of the sex ratio. The y-axis denotes the mean value of PM2.5 in each quantile in 2015, and the x-axis denotes the mean value of the sex ratio in each quantile in 2000. Figure 2b plots prefecture-level PM2.5 pollution in 2014 against the sex ratio among newborns (aged below 12 months by October 2015) with rural hukou. Cities are grouped into 15 groups according to the quantile of PM2.5 exposure. The y-axis denotes the mean value of the sex ratio of newborns in each quantile, and the x-axis denotes the mean value of PM2.5 pollution in 2014.

4. Does Pollution Creates "Missing Boys"?

4.1 Empirical Specification

We use individual-level data on newborn infants to assess whether pollution exposure affects the sex of children at birth. The specification is expressed as:

$$Male_{ij} = \beta_0 + \beta_1 PM2.5_j + \beta_2 Z_{ij} + \varepsilon_{ij} \quad (1)$$

The dependent variable $Male_{ij}$ is an indicator for whether the newborn infant i in city j is male or not. $PM2.5_j$ is city $j's$ mean PM2.5 concentration in the survey year of 2015 or the preceding year. Z_{ij} represents control variables, including the age and education level of infant *i*'s mother, as well as the male-to-female ratio among children aged 0–10 in the baseline year of 2010. We use the 2015 Population Survey to estimate equation 1 and limit the sample to infants born in the 12 months preceding the survey month (i.e., October 2015). We divide these newborns into three groups: first-born children, second-born children with an older male sibling, and second-born children without an older male sibling, and examine the heterogeneous effects of pollution across these groups.

Identification Strategy Pollution is likely to be correlated with local economic activity, so OLS estimates may be biased. To address this potential endogeneity concern, we create an instrument for air pollution based on how wind direction interacts with the siting of distant coal fired plants (as in Gao et al. (2023) and Khanna et al. (2025)).

The instrument quantifies the extent to which distant large-scale thermal power plants are located upwind of a city. Specifically, the IV's value is penalized if plants are not located directly upwind of the city and if it is farther away, using the following first-stage equation:

$$PM2.5_{jt} = \gamma_0 + \gamma_1 \sum_p^P \left(\frac{1}{\alpha_{pt}+1}\right)\left(\frac{1}{dist_{pj}}\right) C_{pt} + \mu_{jt} \quad (2)$$

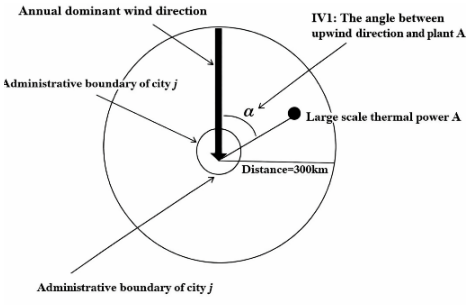
where α_{pt} denotes the angle between city $j's$ annual prevailing wind direction and plant p in year t, $dist_{pj}$ represents the distance from plant p to city j, and C_{pt} is the coal consumption of plant p in year t. We restrict the analysis to large-scale (capacity > 1 million KW) thermal power plants that are located outside city j but within 300km distance from the city. Figure 3a illustrates the

intuition behind the instrument. The underlying variation comes from how wind patterns transport air pollutants from distant coal-fired plants toward a city. Figure 3b shows a clear positive first-stage relationship between the wind direction IV and local air pollution, and Table A2 reports the first-stage results.

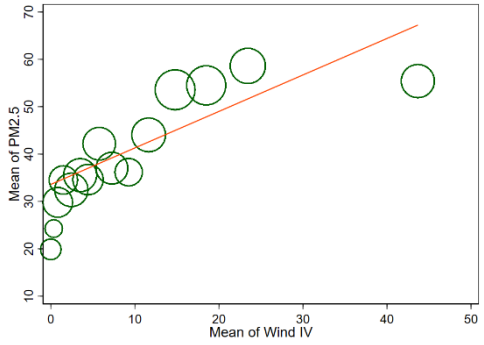
We expect our instrumental variable to be orthogonal to local economic activities. First, yearly wind patterns are plausibly exogenous to local economic activities. Second, these large-scale thermal power plants supply electricity to vast areas of China, including many remote regions; many do not supply electricity to their nearby cities but instead serve many remote provinces.[6] Further, the siting of large-scale power plants and the allocation of electricity supply from them are determined by the central government who owns the grid. It is difficult for local governments to exert influence on the setup of large-scale plants or on the allocation of their electricity supply.[7]

Figure 3: Wind direction, distance, and coal consumption in thermal power plants

(a) Depiction of instrument



(b) First-stage relationship



Notes: In the left panel, the thick arrow represents the annual dominant wind direction of city *j*. The dark dot represents a large-scale thermal power plant located outside city *j* and within 300km from the city. The angle α denotes the angle between the annual prevailing wind direction of city *j* and the large-scale power plant. Large-scale thermal power plants are defined as plants whose installed-capacities are larger than 1 million KW. In the right panel, cities are grouped into 15 groups according to the quantile of the wind direction IV measure. The y-axis denotes the mean value of PM2.5 in each quantile in 2015 and x-axis denotes the mean value of wind direction IV in each quantile.

We examine various potential concerns with this instrument in Section 4.3. We show the

[6] Inter-province electricity transmission is prevalent in China. For example, in 2016, Yunnan province transmitted about half of its power output to other provinces, and Sichuan province transmitted 40% of its power output to other provinces. Large-scale power plants are mainly responsible for the power generation for inter-province electricity transmission in China (Khanna et al, 2025).

[7] There are 333 large-scale (capacity>1 million KW) coal-fired plants in China in 2015.

robustness of our results to restricting power plants to those located far away from a given city and to those built before the establishment of the Ministry of Environmental Protection (MEP) in China. The siting of these plants is unlikely to be correlated with a city's local economic conditions or the government's potential environmental concerns. We also conduct falsification tests showing that baseline city characteristics (say, GDP per capita, population, and electricity consumption) do not predict the future placement of plants.

4.2 *Baseline Results*

Table 1 examines the effect of contemporaneous pollution exposure on the sex of infants born within one year preceding October 2015. The outcome variable is an indicator for male birth. We isolate a plausibly exogenous source of variation in pollution by leveraging the instrument based on wind direction and distant coal-fired power plants (Gao et al., 2023; Khanna et al., 2025). We limit the sample to newborns whose mothers hold a rural *hukou*.

These estimates suggest that pollution exposure reduces the share of male relative to female newborns among rural families. Panel A exploits pollution in 2014. Results imply that a 10 $\mu g/m^3$ increase in annual PM2.5 concentration in 2014 reduces the probability of a newborn being a boy by 3.8 percentage points (p.p.) overall, corresponding to 7% of the mean share of male newborns (53 p.p.) in our rural sample (Panel A, column 1). The effects of pollution are meaningfully stronger for first-born children and second-born children with a male sibling than second-born children without a male sibling (Panel A, columns 2–4). Panel B replicates our analysis by leveraging pollution exposure in 2015 and reveals a similar pattern. A 10 $\mu g/m^3$ rise in PM2.5 concentration in 2015 leads to a reduction in the share of male newborns by 2.2 percentage points overall (Panel B, column 1). Again, we observe a similar heterogeneous effect of pollution across three groups of newborns (Panel B, columns 2–4).

The intriguing heterogeneity may be related to pre-existing son preference among rural parents and fertility stopping rules under China's One Child Policy (OCP). Under the OCP, rural families in many provinces were allowed to have a second child if their first child was a girl. In 2013, the OCP was partially relaxed, and parents meeting certain criteria became eligible to have two children[8]. However, the OCP's provision based on the gender of the first child was still

[8] The relaxation of the OCP in 2013 permitted couples to have two children if at least one parent was an only child. In addition,

implemented in many parts of China. The gender ratio of first-born children in China matches biological expectations (Zeng et al., 1993; Almond et al., 2019). Nevertheless, the OCP created incentives for parents with strong son preference to engage in sex selection for their second-born child if their first-born child was a girl. This may explain why pollution has no effect on the sex of second-born children without a male sibling. Moreover, parents with conservative gender norms may invest more in their second-born boys if they only have an older daughter. Although we do not have data on the purchase of defensive measures against pollution, Appendix Table B1 provides suggestive evidence that second-born sons without an elder brother enjoy more months of breastfeeding than first-born children and second-born sons with an elder brother.

Table 1: Pollution and the Gender of Newborns: Children with Rural *Hukou*, 2015

	(1) Full sample	(2) First-order births	(3) Second-order births with a Male Sibling	(4) Second-order births without a male Sibling
	Dep. Var. Indicator for whether the new-born infant is a boy (=1)			
Panel A				
PM2.5, 2014	-0.00383***	-0.00466***	-0.00735***	0.00171
	(0.00126)	(0.00178)	(0.00267)	(0.00186)
Observations	6,739	3,384	1,482	1,873
F stat	50.69	49.70	54.61	45.24
Mean of Y	0.532	0.519	0.410	0.650
Mean of X	45.96	45.63	45.17	47.17
Panel B				
PM2.5, 2015	-0.00223***	-0.00289***	-0.00319**	0.000336
	(0.000724)	(0.00103)	(0.00147)	(0.00117)
Observations	6,739	3,384	1,482	1,873
F stat	69.43	68.39	81.90	60.87
Mean of Y	0.532	0.519	0.410	0.650
Mean of X	45.91	45.66	44.71	47.30

Notes: The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio for cohorts aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic.

remarried couples were also granted the right to have two children, regardless of whether either spouse was an only child.

We next turn our attention to urban *hukou* holders. Appendix Table A3 shows that pollution exposure has no meaningful effect on the sex of newborns whose mothers hold an urban *hukou*. Driven by rapid urbanization in China, a very large number of rural *hukou* holders co-reside with urban *hukou* holders in the urban areas of prefectural regions. However, *hukou* still determines one's socioeconomic status (SES); without an urban *hukou*, residents have only limited access to many government-provided services, such as public health insurance and public education. The environmental justice literature highlights that high-SES populations have considerably greater access to defensive resources against pollution and to information about exposure levels compared to low-SES groups (Hausman and Stolper, 2021; Luo et al., 2025). As a result, pollution may not affect the survival of male fetuses among urban populations in China.

A limitation of our study is that we cannot separately identify the effects of prenatal pollution exposure from those of pre-conception exposure on parental reproductive health. This identification challenge arises from the strong autocorrelation of local pollution levels, which renders these two channels empirically collinear in our setting. The public health and epidemiological literature has documented that stressful environmental conditions reduce the survival rate of Y-chromosome-bearing spermatozoa relative to their X counterparts during the pre-conception period, and increase stillbirth risk among male fetuses relative to female fetuses during the prenatal period. Our findings therefore capture the overall effects operating through these channels and provide nationally representative evidence on how pollution contributes to "missing" boys among disadvantaged populations and how the implications of pollution differ by SES.

4.3 Sensitivity Analysis

In this section, we conduct a wide range of robustness checks to examine threats to our identification strategy. We explore different specifications of the instrumental variable, alternative samples, and implement falsification tests.

4.3.1 Excluding Nearby or Newly Built Plants or Coal-Producing Regions

We first address concerns about the endogenous siting of power plants: policymakers may exploit the same function — the interaction of wind direction, distance to a city, and coal consumption — to determine where to locate new plants. For instance, such plants may be systematically built near poorer, less influential cities, raising the concern that the instrument is

correlated with unobserved city characteristics. Table A4 excludes any plants built within 50 km of a given city (column 1), and then within 100 km (column 3) and 150 km of the city (column 4) in the IV construction. Our estimates remain robust across these alternative specifications.

One might expect that newly built plants are subject to more regulatory scrutiny, as the Chinese government has only recently prioritized environmental regulation. In Table A5, we exclude newly built plants from the construction of the IV, and instead include cities with newly built plants in the "control" group. Although this empirical strategy is more conservative, we find a similar pattern. Importantly, when our IV is constructed from plants built at least a decade earlier, prior to the establishment of the MEP in China in 2008, the IV estimates remain qualitatively similar.

Another concern is that coal-fired power plants may be concentrated in coal-producing areas. Because Shanxi is the largest coal-producing province in China, we exclude it from our estimation sample. As reported in Table A6, our estimates remain largely unchanged.

4.3.2 An Alternative Version of the Wind IV

We create an alternative version of the IV that exploits only plausibly exogenous variation in wind patterns. Specifically, we use the coal consumption of power plants located upwind of a specific city as the IV, controlling for the total coal consumption of all other power plants located in the downwind area but within the same radius from the city. The first-stage specification is as follows:

$$PM2.5_j = \gamma_0 + \gamma_1 \sum_{p,Upwind}^{P} \left(\frac{1}{dist_{pj}}\right) C_{Pt} + \gamma_2 \sum_{p,down-upwind}^{P} \left(\frac{1}{dist_{pj}}\right) C_{pt+\mu_j} \qquad (3)$$

where $\sum_{p,upwind}^{P} \left(\frac{1}{dist_{pj}}\right) C_{pt}$ is the total coal consumption of plants located in the upwind area of city j, with more distant plants receiving lower weight. As illustrated in Figure A3, the upwind area is a section of a circular buffer drawn within a distance of 300km from city j, and the angle between left and right edges of the section and the prevailing wind direction of city j is 90 degrees. $\sum_{p,down-upwind}^{P} \left(\frac{1}{dist_{pj}}\right) C_{pt}$ is the total coal consumption of plants located within the same distance (300km) from city j but in the downwind area, with more distant plants similarly downweighted. Plants located within the city are excluded as before.

Appendix Table A7 presents the first-stage estimates of this alternative IV. Table A8 reports the second-stage IV results, yielding a similar empirical pattern. Again, pollution exposure reduces

the share of male newborns among the rural population. The variation in this alternative IV is driven only by how wind direction allocates coal consumption between upwind and downwind areas around a particular prefecture. Although we consider the distance to cities and the coal consumption of upwind plants in constructing the IV, we also account for the distance and coal consumption of those "controlled" plants located in the downwind area. Therefore, this IV specification allows us to isolate the effect of pollution from any remaining unobservables correlated with coal consumption and distance to cities.

4.3.3 Falsification Tests and Descriptive Evidence

We next examine the concern over the exogeneity of wind direction. One might expect the government to site coal-fired plants such that pollution would not travel to politically important or densely populated cities. If that were the case, thermal plants are unlikely to be located upwind of these "important" cities. Appendix Table A9 shows the number of large-scale thermal plants located upwind of China's five first-tier cities in 2014 along with their total coal consumption. Beijing and Tianjin are among the most densely populated and politically important cities in Northern China, but experience high pollution levels. There are 58 large scale coal-fired plants located upwind of Beijing, and 59 large scale thermal power plants located upwind of Tianjin (column 1). We also calculate the ratio of upwind large scale thermal plants to the total number of large scale plants. This ratio is 22.4% and 50.9% for Beijing and Tianjin, respectively (column 2). The mean ratio for all Chinese cities is 37.8%, falling between the values for Beijing and Tianjin.

Table A10 further implements falsification tests. We examine whether baseline city-level characteristics in 2004 can predict (a) the ratio of upwind plants built after 2005, and (b) the IV based on plants built after 2005. We find no meaningful associations between these variables and various indicators of a city's political or economic importance, including baseline population, GDP per capita, total electricity consumption, and industrial electricity consumption.

Taken together, these results suggest that the Chinese government does not intentionally site coal-fired plants away from populated or politically important cities, or avoid placing them upwind of such cities.

4.3.4 Additional Controls and Alternative Samples

Regional economic prosperity may confound the association between pollution and sex ratios. Appendix Table A11 shows that our estimates are robust to controlling for population and GDP per capita. Therefore, our IV appears to successfully isolate the effect of pollution from local economic activity. In Table A12, we further exclude densely populated and politically important cities[9], and our results remain similar. Table A13 examines the heterogeneity between big and small cities.

5. Does Male-biased Sex Ratio Imbalance Lead to More Pollution?

The second part of this paper investigates a reverse relationship: whether skewed sex ratios in China generate unintended environmental outcomes. Driven by a severe male-to-female imbalance, China's marriage markets have become more competitive for men, especially among rural and disadvantaged populations. As a result, households, firms, and local governments may reallocate effort and resources toward faster, more visible economic gains, potentially at higher environmental cost. For example, households may increase labor supply or sort into more pollution-intensive employment; local firms may expand activity in pollution-intensive sectors; and local governments may prioritize short-run growth over environmental enforcement. Thus, a male-biased sex ratio may translate into higher ambient pollution—an environmental externality of demographic imbalance. This part uses prefecture-level data to examine whether localities with skewed sex ratios among early cohorts experience systematically higher air pollution as those cohorts approach marriageable age.

5.1 Early-Cohort Sex Ratios and Subsequent Pollution Outcomes

OLS Regression Leveraging Baseline Sex Ratios. We first exploit cross-sectional variation in baseline sex ratios across cities. Specifically, we examine how the male-to-female ratio of the cohort aged 3–10 in the baseline year 2000 is associated with PM2.5 concentrations in 2015. Our empirical specification is:

$$Pollution_{j,2015} = \beta_0 + \beta_1 Gender\ Ratio_{j,2000} + \theta_{\ region} + \varepsilon_j \qquad (4)$$

[9] In particular, we exclude 4 four centrally controlled and several large prefecture-level cities.

The dependent variable is the amount of PM2.5 pollution in city j in 2015. $Gender\ Ratio_{j,2000}$ is the male-female ratio of the cohort aged 3-10 in city j in 2000. These children are at the age 18-25 in 2015, reaching or approaching marriageable age. θ_{region} represents macro-region fixed effects [10], accounting for regional differences in gender norms, marriage market norms and other socioeconomic factors. We use the 2000 Population Census to construct separate sex ratios for rural and urban children and examine potential heterogeneity between these groups.

IV Estimation using Contiguous City-pair Panel Data. We next exploit differences in male-to-female ratios across two groups of birth cohorts — those aged 3–10 in 1990 and those aged 3–10 in 2000 — to assess how changes in sex ratios across cohorts translate into subsequent changes in air pollution between 2005 and 2015. China's One-Child Policy (OCP) contributed substantially to sex ratio imbalance. To isolate plausibly exogenous variation in sex ratios from local confounders, we leverage differential intertemporal changes in provincial OCP policies across contiguous cities straddling a common provincial boundary. Cross-border neighboring cities are more similar in their underlying socioeconomic conditions than more distant locations (Dube et al., 2010).

Based on the stringency of the OCP, Chinese provinces are categorized into three zones: (1) the one-child zone, which permits only one child per family; (2) the 1.5-child zone, which allows a second child only if the first child is a girl; and (3) the two-child zone, which permits two or three children. Once a city is classified into one of these OCP regimes, it remains in the same zone throughout the study period. Provinces in different OCP zones are likely to experience differential intertemporal changes in financial penalties for unauthorized births. [11]

We define two cities as a pair if they share a border and are located in different provinces that belong to different OCP zones. [12] Figure 4a illustrates the distribution of Chinese cities across OCP zones, while Figure 4b shows the contiguous city-pair sample for our analysis. Within these city

[10] There are seven macro-regions in China: East China, North China, Central China, South China, Southwest China, Northwest China and Northeast China.

[11] The percent change in local OCP fine rates between the two periods (1990–1997 and 2000–2007) ranges from −23% to 462% in one-child provinces, from −76% to 296% in 1.5-child provinces, and from 57% to 199% in two-child provinces.

[12] By limiting our sample to neighboring cities located across provincial borders and in different OCP zones, we ensure substantial variation in treatment intensity within each city pair.

pairs, differences in sex ratios among early cohorts across paired cities are positively associated with differences in subsequent pollution (Figure 4c).

We assign a pair id p to each contiguous city pair and our empirical specification is:

$$Pollution_{jp,t+15} = \beta_0 + \beta_1 Gender\ Ratio_{jp,t} + \mu_{p,t} + \theta_{region,t} + \varepsilon_{jp,t} \quad (5)$$

The dependent variable is the amount of PM2.5 pollution in city j of pair p in 2005 and 2015, respectively. $Gender\ Ratio_{jp,t}$ is the male-female ratio of the city among cohorts aged 3-10 in 1990 and 2000, respectively. The first cohort is at or approaching marriageable age by 2005, while the second cohort reaches a similar stage around 2015. $\mu_{p,t}$ represents city-pair by year fixed effects, and $\theta_{region,t}$ represents macro-region by year fixed effects. [13]

We use province-specific fine rates (share of fines in average annual income) for violating the local OCP as an instrument for the male-to-female ratios among rural children and allow the first-stage relationship to differ by OCP zones. Specifically, we exploit variation in the change in 8-year average fine rates for OCP violations between two periods (1990–1997 and 2000–2007) across cities within contiguous city-pairs that straddle a common provincial border. By including pair-specific time effects ($\mu_{pair,t}$), we exploit only within-pair variation in OCP penalty changes between the two periods, ensuring that cross-pair differences in levels or trends do not confound the estimates.

This exercise reduces the endogeneity concern. First, we control for time-varying unobserved factors using city-pair-time fixed effects. Second, contiguous regions are relatively similar with respect to underlying economic conditions and hence form better controls (Dube, 2010). Last, cities on province borders tend to be further away from the province government which sets the OCP penalties, thus OCP fine rates implemented in those border cities are unlikely to be affected by their local economic conditions. [14]

Our baseline specification uses robust standard errors clustered at the province level to account for serial correlation of residuals within provinces. Our results are robust to using two-way cluster-robust standard errors at the province and province-border-segment levels to account for potential

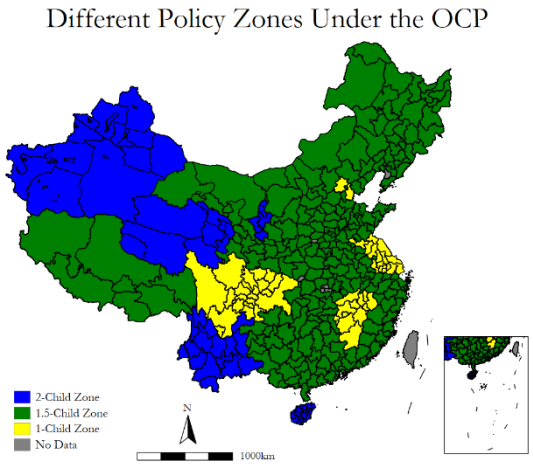
[13] The city-pair panel used to estimate Equation 5 consists of 468 observations from 234 city pairs, involving 110 unique cities.

[14] However, those border cities are required to comply with the OCP guidelines of their respective provinces.

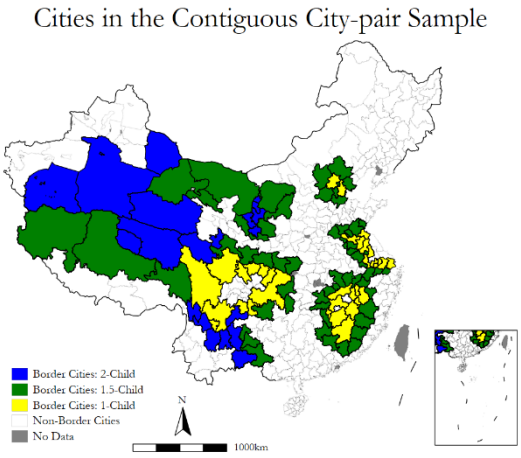
correlation of residuals within provinces or within border segments.[15]

Figure 4: OCP Zones and Neighboring City Samples

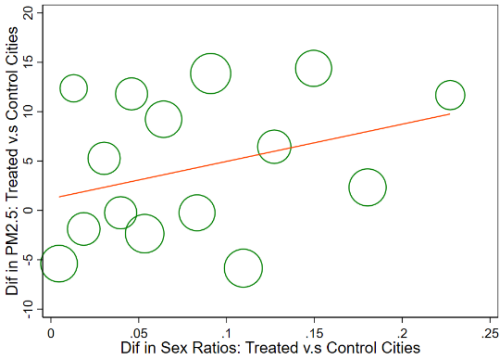
(a)Policy Zones Under the OCP



(b) Cites in the Neighboring City-pair Sample



(c) Differences in Baseline Sex Ratios in 2000 and Pollution in 2015: Within Contiguous City-Pair



Notes: Figure 4a illustrates the spatial distribution of OCP policy zones (one-child, 1.5-child, and two-child policies) together with the administrative boundaries of all prefecture-level cities. Figure 4b shows the cities in the contiguous-city-pair sample used in our analysis. Figure 4c plots differences in PM2.5 pollution in 2015 against differences in the baseline sex ratio in 2000 between treated and control cities within contiguous city pairs, conditional on macro-region fixed effects. Within a particular city pair, treated cities are those with relatively higher sex ratios, and control cities are those with relatively lower sex ratios. City pairs are grouped into 15 groups according to the quantiles of sex ratio differences. The y-axis denotes the mean value of the differences in PM2.5 in each quantile, and the x-axis denotes the mean value of baseline sex ratio differences.

[15] The latter addresses the possibility that residuals may not be independent across city pairs that share the same border segment between two provinces.

5.2 OLS Results Leveraging Sex Ratios at Baseline

5.21 Baseline Results

Table 2 presents baseline estimates of the effects of male-female ratios among the early cohort on subsequent pollution levels 15 years later. Macro-region fixed effects are included across all specifications. Column 1 leverages sex ratios among all children, regardless of *hukou* status. The full-sample sex ratios among those aged 3–10 are positively and significantly associated with future PM2.5 exposure in each city in 2015. Columns 2–4 break the sample by children's *hukou* type (rural vs. urban) and document that the effect of early sex ratios on future pollution is driven only by children with rural *hukou*, while sex ratios among their urban counterparts appear unrelated to subsequent pollution.

We control for city-tier fixed effects in columns 5–7 to further account for differential economic prosperity and industrial structures between large and small cities. Again, the relationship between early sex ratios and future ambient pollution is statistically significant only for rural children. Based on the point estimate in column 7, the effect of a one-standard-deviation (0.09) increase in sex ratios at baseline among rural children corresponds to 11 percent of a standard-deviation (18.3) increase in PM2.5 pollution 15 years later ($0.09 \times 22.4 / 18.3 = 0.11$).

Skewed, male-biased sex ratios among early cohorts would increase the competitiveness of marriage markets when these cohorts attain marriageable ages. Moreover, rural men are in a more disadvantaged position in marriage markets compared to their urban counterparts. These two factors may induce rural households to increase labor supply toward faster-growing, higher-earning economic activities, potentially at a higher environmental cost. Driven by increased supply of rural workers, firms and the government may also expand production in pollution-intensive sectors (e.g., the manufacturing sector). In contrast, urban hukou holders are much less likely to work in polluting sectors which are more labor-intensive and therefore have a higher demand for rural workers, and their *hukou* status also grants more advantages in marriage markets. Indeed, Table B2 divides sectors into two groups based on pollution intensity and highlights that rural workers are disproportionately concentrated in these polluting sectors.

Table 2: Early-Cohort Sex Ratios in 2000 and Future Pollution in 2015

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var.	PM2.5 Pollution 15 years later						
Male-female Ratio	26.07***						
	(8.534)						
Rural Male-female Ratio		24.58***		24.17***	23.76***		22.39***
		(7.934)		(7.866)	(7.844)		(7.701)
Urban Male-female Ratio			4.628	1.046		6.846	3.305
			(5.876)	(5.540)		(6.092)	(5.725)
Observations	306	306	306	306	306	306	306
Adjusted R-sq	0.536	0.552	0.541	0.552	0.565	0.556	0.566
Region FEs	X	X	X	X	X	X	X
City-Tier FEs					X	X	X

Notes: The dependent variable is prefecture-level PM2.5 concentration in 2015. The key independent variables are prefecture-level male–female ratios for cohorts aged 3–10 in 2000. We include sex ratios separately for cohorts with rural hukou and for those with urban hukou. All specifications control for the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai). Robust standard errors clustered at the city level are reported in parentheses. $p < 0.10$, *$p < 0.05$, **$p < 0.01$.

5.2.2 Sensitivity Analysis

Alternative Measures and Other Pollutants Appendix Table A14 repeats the exercise and uses the log of PM2.5 pollution as the outcome variable. Again, the association between sex ratio imbalances and future pollution is driven by rural *hukou* holders. A one-standard-deviation (0.09) increase in sex ratios among rural children would lead to an approximately 7 percent (0.09×0.8) increase in future PM2.5 pollution. Table A15 shows that baseline rural sex ratios are also significantly positively associated with city-level SO2 concentrations in the future. Nevertheless, we do not observe any meaningful relationship between rural sex ratios and subsequent water pollution (Table A16). [16]

Placebo Test Appendix Table A17 conducts the analysis using a "placebo" age group—those aged between 23 and 30 in 2000, and shows that their sex ratios are unrelated to subsequent pollution in 2015. The "placebo" group was aged 38–45 in 2015; since the majority of them would have

[16] City-level water pollution is measured as chemical oxygen demand (COD) emissions and wastewater emissions, normalized by city population or industrial output. COD and wastewater emissions have both declined since 2000 (Appendix Figure A2).

been married for many years by that time, their sex ratios are unlikely to affect the local marriage market. Therefore, the relationship between early-cohort sex imbalance and future pollution that we documented is indeed driven by increased marriage-market pressures on men.

5.3 IV Results using Contiguous City-pair Panel Data

Sex ratios may be correlated with certain local characteristics. This section uses contiguous city-pair panel data and exploits policy discontinuities at province borders. Specifically, we leverage variations in sex ratios driven by differential intertemporal changes in OCP financial penalties between cities within these neighboring city-pairs.

Appendix Table A18 presents the first-stage results. The monetary penalties for unauthorized births are positively associated with rural sex ratios among early age cohorts, with the effects being stronger for one-child zones than for non-one-child zones[17]. The joint F-statistics for IVs are around 18.[18]

Table 3 reports the second-stage estimates. All specifications control for cohort size, since the OCP may affect future labor supply by altering the size of each birth cohort. We sometimes control for the standard deviation of OCP fines as well as the share of minority children in each cohort.[19]Although the IV-estimation based on contiguous city-pairs sweeps out variation across different city-pairs, the coefficient estimates are qualitatively similar to those from our baseline OLS specification. The point estimates of rural sex ratios are around 22 and robust across various controls. Accordingly, among cities in our neighboring city-pair sample, a one-standard-deviation (0.08) increase in rural sex ratios in the early cohort aged 3–10 is associated with 9 percent of a standard-deviation (19.6) increase in city-level PM2.5 concentration as they enter adulthood ($0.08 \times 22 / 19.6 = 0.09$).

Some cities appear in more than one contiguous city-pair. Our baseline analysis treats all observations in the sample equally, and Appendix Table A20 assigns less weight to duplicate observations, weighting observations according to the inverse of the number of duplicates. The

[17] None-one-child zones includes cities in the 1.5-child zone and the 2-child zone.

[18] In the contiguous city-pair analysis, we focus on rural sex ratios because the first-stage relationship between OCP penalties and urban sex ratios is weak, with F-statistics below conventional thresholds (Table A19).

[19] Adding the standard deviation of OCP fines accounts for the uncertainties of the OCP, which may affect prospective parents' expectations of future fertility policies. Fertility restrictions under the OCP are less binding for ethnic minorities, thus we examine the robustness of our results to controlling for minority shares.

results from weighted IV regressions are quantitatively and qualitatively similar. Our estimates are also robust to using an alternative error term structure assumed for inference—two-way robust standard errors clustered at the province and province-border-segment levels (Appendix Table A21), which further account for potential correlation of residuals within common border segments between two provinces (as in Dube et al., 2010).[20]

Table 3: Sex Ratio and Pollution: IV Results using City-Pair Panel Sample

	(1)	(2)	(3)	(4)
Dep. Var.	PM2.5 Pollution 15 years later			
Rural Male-female Ratio	21.93***	22.25**	21.06***	21.38***
	(6.490)	(7.960)	(5.835)	(7.239)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share			X	X
Observations	468	468	468	468
F-Stat for IVs	18.72	17.95	18.05	18.88

Notes: The analysis is based on city-pair panel data linking male-to-female ratios (among rural children aged 3–10) in 1990 and 2000 to PM2.5 pollution levels in 2005 and 2015. We treat those aged 3-10 in 1990 as the first cohort and those aged 3-10 in 2000 as the second cohort. All specifications control for region by year fixed effects and city-pair by year fixed effects, interactions between the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai) and cohort fixed effects, and the interaction between the distance to nearest big cities and cohort fixed effects. Robust standard errors clustered at the province level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic.

5.4 Bridging the Two Parts of the Paper

The two parts of our paper highlight a surprising demographic–environmental feedback loop driven by marriage-market competition and adverse environmental externalities. Due to strong gender discrimination, China exhibits severe male-biased sex ratio imbalances. We document that skewed male-female ratios among early cohorts translate into higher levels of future pollution as these cohorts enter marriage markets. As pollution rises, contemporaneous exposure reduces the share of surviving male newborns. These relationships act in opposite directions and are meaningful only for rural *hukou* holders.

[20] Table A22 reports the corresponding OLS results using the contiguous city-pair panel. Again, baseline rural sex ratios are significantly positively associated with subsequent pollution. The coefficient estimates of rural sex ratios are smaller in magnitude. One possible reason is that certain wind patterns and meteorological conditions may transport pollutants from the "treated" to the "control" city within a neighboring city pair. Our IV based on changes in OCP penalties accounts for the effects of these potential confounders.

How much of the cross-city gaps in sex ratios and ambient pollution do these relationships explain? We illustrate using Yangzhou and Linyi (two comparable medium-sized cities) and Haikou and Xining (two comparable provincial capitals). Gaps in early rural cohort sex ratios in 2000 contribute to 8.5% and 9.1%, respectively, of exposure gaps in 2015 between Yangzhou and Linyi, and between Haikou and Xining. However, pollution exposure in 2015 narrows differences in the fraction of male newborns among rural households by 7.5% and 9.9%, respectively, between Yangzhou and Linyi, and between Haikou and Xining. Taken together, pollution plays a partial but non-trivial role in the natural correction of sex ratio imbalance.

6.Conclusion

Broader implications. The results highlight that demographic distortions can propagate into domains that are not typically studied in the sex-ratio literature, including environmental externalities. At the same time, the paper's feedback result underscores that biological or health mechanisms can feed back into demographic structure in ways that are economically meaningful but socially costly.

A disciplined takeaway is therefore twofold. First, policies or norms that generate persistent demographic imbalance may carry hidden externality costs, beyond the well-known marriage-market and welfare consequences. Second, even if an imbalance partially "self-adjusts" through adverse health channels, such adjustment is not a substitute for policy: it operates through a negative externality that society would typically want to mitigate.

These conclusions naturally motivate future work: identifying the most important behavioral and policy channels that map marriage-market pressure into production and enforcement decisions; quantifying welfare tradeoffs more precisely; and studying whether similar mechanisms operate in other settings where sex ratios are distorted and pollution-intensive growth is prevalent.

This paper studies a new role of pollution in the adjustment process of sex ratio imbalance based on Chinese data. It is well known that persistent discrimination against women has resulted in highly skewed male-to-female ratios among children with rural *hukou*. We document two new patterns. The first is a robust pattern in which higher male-to-female ratios among early-age rural cohorts lead to higher levels of air pollution when these cohorts reach marriage age. This relationship is evident both in cross-sectional variation across prefectures exploiting baseline sex-ratio differences and in panel variation that leverages differential changes in the penalties under

the One-Child Policy between neighboring prefectures across different provinces. We further show that contemporaneous pollution exposure reduces the number of rural boys at birth, contributing to a partial rebalancing of sex ratios among rural children. These pollution-induced gender selection effects are stronger for first-born children and for second-born children with an elder male sibling. In contrast, neither urban cohort sex ratios predict future pollution, nor does contemporaneous pollution affect the sex of urban births.

References

Abdel Ghany, Jasmin, Joshua Wilde, Anna Dimitrova, Ridhi Kashyap, and Raya Muttarak. "Temperature and Sex Ratios at Birth." *Proceedings of the National Academy of Sciences* 123, no. 8 (2026): e2422625123.

Almond, Douglas, Hongbin Li, and Shuang Zhang. "Land Reform and Sex Selection in China." *Journal of Political Economy* 127, no. 2 (2019): 560-585.

Basu, Rupa, Varada Sarovar, and Brian J. Malig. "Association between High Ambient Temperature and Risk of Stillbirth in California." *American Journal of Epidemiology* 183, no. 10 (2016): 894–901.

Bekkar, Bruce, Susan Pacheco, Rupa Basu, and Nathaniel DeNicola. "Association of Air Pollution and Heat Exposure with Preterm Birth, Low Birth Weight, and Stillbirth in the US." *JAMA Network Open* 3, no. 6 (2020): e208243.

Catalano, Ralph, Tim Bruckner, and Kirk R. Smith. "Ambient Temperature Predicts Sex Ratios and Male Longevity." *Proceedings of the National Academy of Sciences* 105, no. 6 (2008): 2244–2247.

Dube, Arindrajit, T. William Lester, and Michael Reich. "Minimum Wage Effects across State borders: Estimates Using Contiguous Counties." *The Review of Economics and Statistics* 92, no. 4 (2010): 945-964

Ebenstein, Avraham. "The 'Missing Girls' of China and the Unintended Consequences of the One Child Policy." *Journal of Human Resources* 45, no. 1 (2010): 87–115.

Fukuda, Misao, Kiyomi Fukuda, Takashi Shimizu, Miho Nobunaga, Linn Salto Mamsen, and Claus Yding Andersen. "Climate Change Is Associated with Male:Female Ratios of Fetal Deaths and Newborn Infants in Japan." *Fertility and Sterility* 102, no. 5 (2014): 1364–1370.

Gao, Xuwen, Ran Song, and Christopher Timmins. "Information, Migration, and the Value of Clean Air." *Journal of Development Economics* 163 (2023): 103079.

Hausman, Catherine, and Samuel Stolper. "Inequality, Information Failures, and Air Pollution." *Journal of Environmental Economics and Management* 110 (2021): 102552.

Jedrychowski, Wieslaw, Frederica Perera, Dorota Mrozek-Budzyn, Elzbieta Mroz, Elzbieta Flak,

Jack D. Spengler, Susan Edwards, Ryszard Jacek, Irena Kaim, and Zbigniew Skolicki. “Gender Differences in Fetal Growth of Newborns Exposed Prenatally to Airborne Fine Particulate Matter.” *Environmental Research* 109, no. 4 (2009): 447–456.

Karlsson, Lena, Johan Junkka, Erling Häggström Lundevaller, and Barbara Schumann. “Ambient Temperature and Stillbirth Risks in Northern Sweden, 1880–1950.” *Environmental Epidemiology* 5, no. 6 (2021): e176.

Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song. “The Productivity Consequences of Pollution-Induced Migration in China.” *American Economic Journal: Applied Economics* 17, no. 2 (2025): 184–224.

Lakshmanan, Ashwini, Yueh-Hsiu Mathilda Chiu, Brent A. Coull, Allan C. Just, Sarah L. Maxwell, Joel Schwartz, Alexandros Gryparis, Itai Kloog, Rosalind J. Wright, and Robert O. Wright. “Associations between Prenatal Traffic-Related Air Pollution Exposure and Birth Weight: Modification by Sex and Maternal Pre-pregnancy Body Mass Index.” *Environmental Research* 137 (2015): 268–277.

Luo, Zhonghui, Lala Ma, Rui Xie, and Ran Song. “A Spatially Differentiated Water Pollution Policy Leads to Economic and Health Inequity.” *Proceedings of the National Academy of Sciences* 122, no. 39 (2025): e2421404122.

Oyeyipo, Ibukun P., Michelle van der Linde, and Stefan S. du Plessis. “Environmental Exposure of Sperm Sex-Chromosomes: A Gender Selection Technique.” *Toxicological Research* 33, no. 4 (2017): 315–323.

Radwan, Michał, Emila Dziewirska, Paweł Radwan, Lucjusz Jakubowski, Wojciech Hanke, and Joanna Jurewicz. “Air Pollution and Human Sperm Sex Ratio.” *American Journal of Men’s Health* 12, no. 4 (2018): 907–912.

Rahman, Md Saidur, and Myung-Geol Pang. “New Biological Insights on X and Y Chromosome-Bearing Spermatozoa.” *Frontiers in Cell and Developmental Biology* 7 (2020): 388.

Sen, Amartya. “Missing Women.” *BMJ* 304, no. 6827 (1992): 587–588.

Tan, Zhibo, Shang-Jin Wei, and Xiaobo Zhang. “Deadly Discrimination: Implications of ‘Missing Girls’ for Workplace Safety.” *Journal of Development Economics* 152 (2021): 102678.

Wei, Shang-Jin, and Xiaobo Zhang. “The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China.” *Journal of Political Economy* 119, no. 3 (2011): 511–564.

You, Young-Ah, Woo-Sung Kwon, Md Saidur Rahman, Yoo-Jin Park, Young-Ju Kim, and Myung-Geol Pang. “Sex Chromosome-Dependent Differential Viability of Human Spermatozoa during Prolonged Incubation.” *Human Reproduction* 32, no. 6 (2017): 1183–1191.

Zeng, Yi, Tu Ping, Gu Baochang, Xu Yi, Li Bohua, and Li Yongpiing. "Causes and Implications of the Recent Increase in the Reported Sex ratio at Birth in China." *Population and development review* (1993): 283-302.

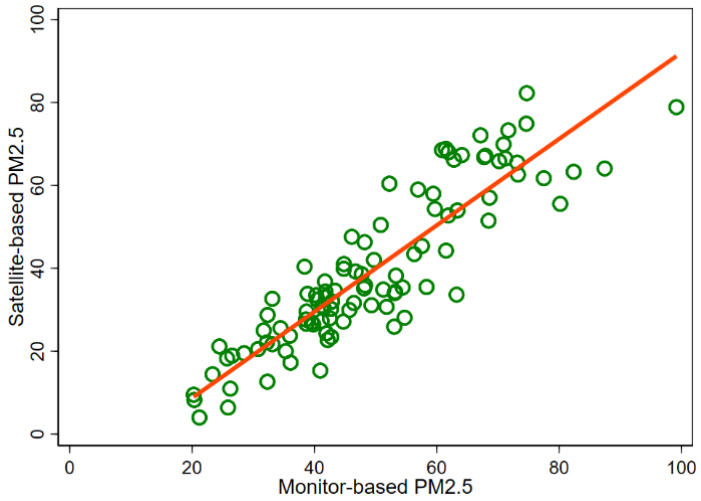
Appendix

Table of Contents

A. Additional Results, Tests and Figures A2
A.1 Additional Figures and Tables A2
A.2 Additional Results: Contemporaneous Pollution and Sex at Birth A4
A.2.1 First-Stage Results and the Effect on Urban Newborns A4
A.2.2 Tests on the Wind-Based IV A5
A.2.3 Additional Controls and Alternative Samples A9
A.3 Additional Results: Sex-Ratio and Future Pollution A12
A.3.1 Additional OLS Results Leveraging Baseline Variations A12
A.3.2 Additional IV Results Using Contiguous City-pair Panel A15
B. Suggestive Evidence A18
C. Related Studies in Public Health and Epidemiology A20

A. Additional Results, Tests and Figures
A.1 Additional Figures and Tables

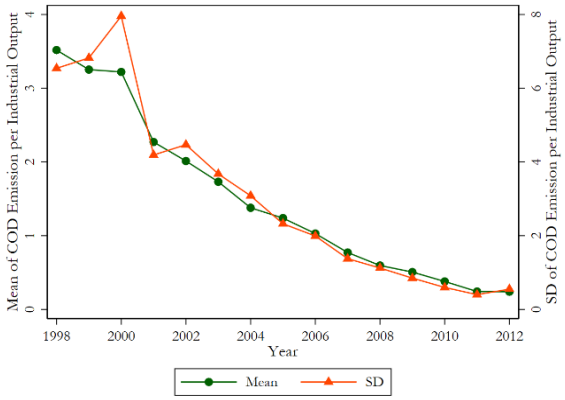
Figure A1: Monitor-based PM2.5 and Satellite-based PM2.5



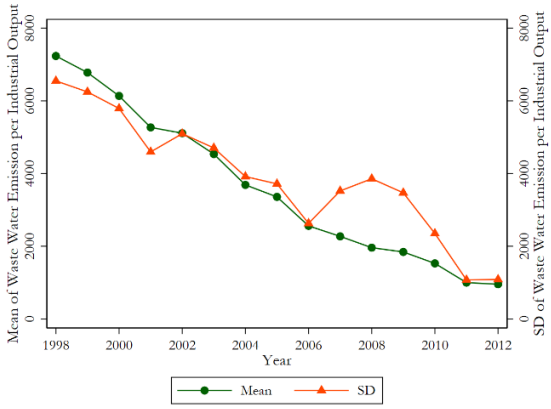
Notes: Cities are grouped into one hundred groups according to the percentile of PM2.5 concentration measured by ground monitors. The y-axis denotes the mean value of satellite-based PM2.5 in each percentile, and x-axis denotes the mean value of monitor-based PM2.5 in each percentile. Monitor-based data come from the official website of the MEP of China, and satellite-based PM2.5 data are drawn from the Global Annual PM2.5 Grids.

Figure A2: Trend of Water Pollution Emissions

(a) COD Emissions

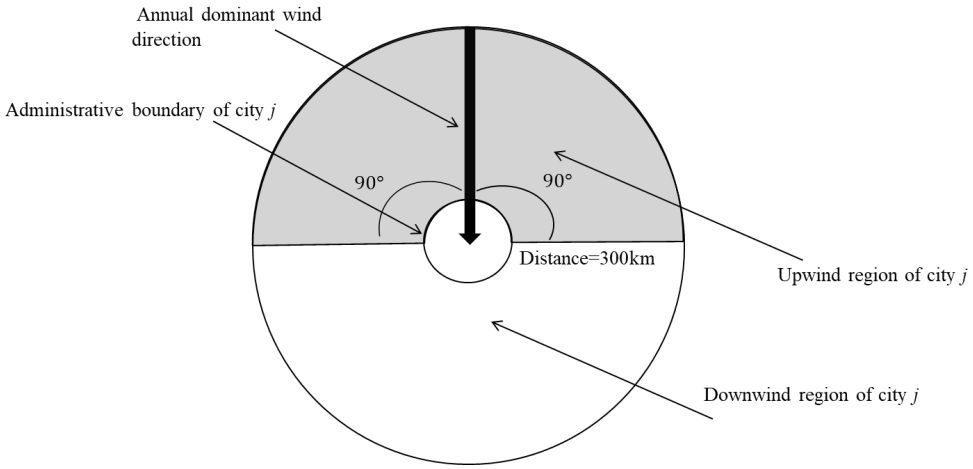


(b) Waster Water Emissions



Notes: Figures show trends in average COD emissions and average wastewater emissions across Chinese cities, as well as the standard deviations of these emissions from 1998 to 2012. Emissions are normalized by industrial output. Data come from the Environmental Survey and Reporting (ESR).

Figure A3 Alternative Wind Direction IV



Notes: The thick arrow denotes the annual dominant wind direction of city j. The dark grey area is the upwind area of city j, which is defined as a section of a circular buffer drawn at a distance of 0 km-300km from the city, and the angle between the left/right side of the section and the annual dominant wind direction of city j is 90 degree. We exclude the upwind area from the 0km-300km 'loop' around city j, and the white area is defined as the down-upwind area around the city.

Table A1: Summary Statistics of Key Variables

Variable name	Mean	Std. dev
Panel A: Contemporaneous Exposure and Newborn Sex		
PM2.5 Pollution in 2014	45.96	16.19
PM2.5 Pollution in 2015	45.91	19.17
Indicator for a newborn male (=1): Rural Sample	0.53	0.50
Indicator for a newborn male (=1): Urban Sample	0.53	0.50
Panel B: Early Cohort Sex Ratios and Future Pollution		
PM2.5 Pollution in 2005	39.45	15.72
PM2.5 Pollution in 2015	40.86	18.26
Early Cohort Sex Ratios: Rural Sample in 1990	1.09	0.06
Early Cohort Sex Ratios: Rural Sample in 2000	1.15	0.09
Early Cohort Sex Ratios: Urban Sample in 1990	1.08	0.14
Early Cohort Sex Ratios: Urban Sample in 2000	1.15	0.15

Notes: This table reports summary statistics for the outcome variables and key independent variables using the baseline regression samples. The summary statistics for PM2.5 pollution in 2025 differ slightly between Panels A and B because Panel A matches PM2.5 data with individual-level data on newborns, whereas Panel B uses city-level data.

A.2 Additional Results: Contemporaneous Pollution and Sex at Birth
A.2.1 First-Stage Results and the Effect on Urban Newborns

Table A2: First-Stage Results: Wind-Based IV

	(1)	(2)	(3)	(4)
Dep. Var	PM2.5, 2014	PM2.5, 2015	PM2.5, 2014	PM2.5, 2015
	Rural Children Sample		Urban Children Sample	
Wind IV, 2014	0.552***		0.432***	
	(0.0775)		(0.0664)	
Wind IV, 2015		0.923***		0.617***
		(0.111)		(0.103)
F stat of IVs	50.69	69.43	42.35	35.99
Observations	6,739	6,739	5,198	5,198

Notes: The key independent variable is the instrumental variable based on the interaction between wind direction, distance to coal plant, and coal consumption at power plant. We merge prefecture-level data on PM2.5 pollution and the wind-based IV with individual-level data on newborns. We control for baseline male-female ratios of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic.

Table A3: Pollution and the Gender of Newborns: Children with Urban *Hukou*, 2015

	(1)	(2)	(3)	(4)
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	0.00101		0.000579	
	(0.00135)		(0.00136)	
PM2.5, 2015		0.000147		-0.000124
		(0.000838)		(0.000826)
Observations	5,198	5,198	5,198	5,198
F stat of IVs	42.35	35.99	46.39	38.40
Mean of Y	0.532	0.532	0.532	0.532
Mean of X	43.05	42.85	43.05	42.85
City Tier FE			X	X

Notes: The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic

A.2.2 Tests on the Wind-Based IV

Table A4: Excluding Nearby Power Plants: Full Sample of Rural children

	(1)	(2)	(3)	(4)
Distance	0-300km	50-300 km	100-300km	150-300km
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	-0.00383***	-0.00348***	-0.00318***	-0.00311**
	(0.00126)	(0.00115)	(0.00117)	(0.00131)
Observations	6,739	6,739	6,739	6,739
F stat of IVs	50.69	51.76	43.97	27.20
Mean of Y	0.532	0.532	0.532	0.532
Mean of X	45.96	45.96	45.96	45.96
PM2.5, 2015	-0.00223***	-0.00205***	-0.00174**	-0.00160**
	(0.000724)	(0.000726)	(0.000740)	(0.000811)
Observations	6,739	6,739	6,739	6,739
F stat of IVs	69.43	70.16	64.98	42.75
Mean of Y	0.532	0.532	0.532	0.532
Mean of X	45.91	45.91	45.91	45.91

Notes: The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic.

Table A5: Excluding Newly Built Power Plants: Full Sample of Rural children

	(1) Plants> 5 yrs	(2) Plants> 10 yrs	(3) Plants> 15 yrs	(4) Plants> 20 yrs
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	-0.00376***	-0.00434***	-0.00262***	-0.00265**
	(0.00119)	(0.00134)	(0.000934)	(0.00105)
Observations	6,739	6,739	6,739	6,739
F stat of IVs	51.57	34.78	47.98	36.92
Mean of Y	0.532	0.519	0.410	0.650
Mean of X	45.96	45.63	45.17	47.17
PM2.5, 2015	-0.00215***	-0.00223***	-0.00141**	-0.00158**
	(0.000704)	(0.000714)	(0.000582)	(0.000625)
Observations	6,739	6,739	6,739	6,739
F stat of IVs	71.74	55.52	59.26	45.80
Mean of Y	0.532	0.519	0.410	0.650
Mean of X	45.91	45.66	44.71	47.30

Notes: The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. We report Kleibergen-Paap rk Wald F statistic.

Table A6: Excluding Coal Producing Regions: Full Sample of Rural children

	(1)	(2)	(3)	(4)
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	-0.00344***		-0.00348***	
	(0.00116)		(0.00115)	
PM2.5, 2015		-0.00255***		-0.00256***
		(0.000843)		(0.000839)
Observations	6,553	6,553	6,553	6,553
F stat of IVs	52.80	61.81	52.41	61.81
Mean of Y	0.532	0.532	0.532	0.532
Mean of X	46.32	46.26	46.32	46.26
City Tier FE			X	X

Notes: The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. We exclude cites located in Shanxi Province. Standard errors that are clustered at the city level are reported in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. We report Kleibergen-Paap rk Wald F statistic.

Table A7: Alternative Wind IV: First-Stage Results

	(1)	(2)	(3)	(4)
Dep. Var	PM2.5, 2014	PM2.5, 2015	PM2.5, 2014	PM2.5, 2015
Upwind Coal Consumption, 2014	0.260***		0.258***	
	(0.0424)		(0.0430)	
Upwind Coal Consumption, 2015		0.351***		0.354***
		(0.0529)		(0.0542)
Observations	6,739	6,739	6,739	6,739
Downwind Coal Consumption	X	X	X	X
City Tier FE			X	X

Notes: The independent variable is the coal consumption of power plants located upwind of the city. We merge prefecture-level data on PM2.5 pollution and the wind-based IV with individual-level data on newborns. We control for the coal consumption of power plants located downwind of the city, baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Alternative Wind IV: Second-Stage Results

	(1)	(2)	(3)	(4)
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	-0.00287**		-0.00296**	
	(0.00136)		(0.00139)	
PM2.5, 2015		-0.00218**		-0.00226**
		(0.000954)		(0.000958)
Observations	6,739	6,739	6,739	6,739
F stat of IVs	37.79	43.96	36.09	42.70
Downwind Coal Consumption	X	X	X	X
City Tier FE			X	X

Notes: The instrumental variables specification uses the coal consumption of power plants located upwind of the city. Individual data on newborns born within 12 months of the census year and month. We control for the coal consumption of power plants located downwind of the city, baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic

Table A9: The Coal-fired Plants Located Upwind of Large metropolitans

City	Number of Upwind Plants	Ratio of Upwind Plants	Coal Consumption of Upwind Plants	Smallest Angle of Plants
Beijing	58	22.41%	2372	25
Tianjin	59	50.85%	2123	26
Shanghai	62	1.61%	2067	16
Guangzhou	30	16.67%	899	22
Shenzhen	25	44.00%	785	26
National mean	43.16	37.82%	511	25

Notes: The statistics are calculated using the large-scale thermal power plants located outside a given city and within 500km. Following Freeman, Liang, Song, and Timmins (2019), we define the upwind region as a section of a circular buffer drawn at a distance of 500km from a given city, and the angle between the section and the annual dominant wind direction of the city being at least 45 degrees

Table A10: Falsification Test: Baseline Economy and the Wind Direction IV

Dependent variable:	The ratio of upwind plants		Wind direction and coal plants IV	
	(1)	(2)	(3)	(4)
Baseline Population	-0.0106	-0.0132	0.412	0.226
	(0.0309)	(0.0279)	(0.489)	(0.440)
Baseline GDP per Capita	-0.0151	-0.0231	-0.132	-0.414
	(0.0302)	(0.0289)	(0.526)	(0.522)
Industrial Elec. Cons.	0.0225		0.210	
	(0.0224)		(0.347)	
Baseline Elec. Cons.		0.0255		0.414
		(0.0178)		(0.269)
Observations	281	281	281	281
City Controls	Yes	Yes	Yes	Yes

Notes: City level regressions for 281 cities. We drop cities with missing values in baseline characteristics. Dependent variables are based on plants built post 2005, independent variables are measured in the year 2004. Standard errors clustered at the hukou city level are reported in parentheses. City controls include distance to Shanghai, Tianjin and Shenzhen seaports.

A.2.3 Additional Controls and Alternative Samples

Table A11: Controlling for Population and GDP per Capita

	(1) Full sample	(2) First-order Births	(3) Second-order Births with Male Siblings	(4) Second-order Births without Male Siblings
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	-0.00529***	-0.00679***	-0.0104***	0.00304
	(0.00185)	(0.00261)	(0.00381)	(0.00249)
Observations	6,313	3,191	1,355	1,767
F stat of IVs	31.04	31.77	30.38	26.63
Mean of Y	0.532	0.519	0.410	0.650
Mean of X	45.96	45.63	45.17	47.17
PM2.5, 2015	-0.00283***	-0.00395***	-0.00400**	0.000868
	(0.000951)	(0.00137)	(0.00186)	(0.00144)
Observations	6,313	3,191	1,355	1,767
F stat of IVs	53.82	54.58	56.72	46.45
Mean of Y	0.532	0.519	0.410	0.650
Mean of X	45.91	45.66	44.71	47.30

Notes: The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio for cohorts aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), baseline city population, baseline GDP per capita, mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic

Table A12: Excluding Politically Important Cities

	(1) Full sample	(2) First-order Births	(3) Second-order Births with Male Siblings	(4) Second-order Births without Male Siblings
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	-0.00379***	-0.00417***	-0.00785***	0.00136
	(0.00114)	(0.00157)	(0.00259)	(0.00176)
Observations	6,156	3,049	1,376	1,731
F stat	-0.013	-0.007	-0.049	0.003
Mean of Y	54.36	53.50	58.85	48.13
Mean of X	0.532	0.523	0.408	0.648
PM2.5, 2015	-0.00224***	-0.00269***	-0.00362**	0.000262
	(0.000670)	(0.000936)	(0.00145)	(0.00113)
Observations	6,156	3,049	1,376	1,731
F stat of IVs	-0.006	-0.001	-0.010	0.004
Mean of Y	67.91	65.88	85.31	60.05
Mean of X	0.532	0.523	0.408	0.648

Notes: Provincial-level cities and sub-provincial level cities are excluded from the sample. The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A13: Pollution and the Gender of Rural Newborns: Heterogeneity by City GDP

	(1)	(2)	(3)	(4)
	City GDP≥ Mean		City GDP< Mean	
Dep. Var:	**Indicator for whether the new-born infant is a boy (=1)**			
PM2.5, 2014	-0.00645**		-0.00347*	
	(0.00301)		(0.00211)	
PM2.5, 2015		-0.00372***		-0.00151
		(0.00133)		(0.00138)
Observations	3,996	3,996	2,743	2,743
F stat of IVs	12.07	25.60	32.73	39.38
Mean of Y	0.528	0.528	0.538	0.538
Mean of X	50.91	51.60	38.74	37.62

Notes: Cites are divided into two groups based on average city-level GDP. The instrumental variables specification uses the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Individual data on newborns born within 12 months of the census year and month. We control for baseline male-female ratio of the cohort aged 0-10 in 2010, the minimum distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai), mother's age, an indicator for whether the mother has a college degree, and an indicator for whether the mother has a high school degree. Standard errors that are clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.3 Additional Results: Sex-Ratio and Future Pollution
A.3.1 Additional OLS Results Leveraging Baseline Variations

Table A14: Log PM2.5 in 2015 as the Dependent Variable

	(1)	(2)	(3)	(2)	(3)	(4)
Dep. Var.	Log PM2.5 Pollution 15 years later					
Rural Male-female Ratio	0.800***		0.797***	0.779***		0.751***
	(0.269)		(0.251)	(0.267)		(0.249)
Urban Male-female Ratio		0.126	0.00789		0.186	0.0675
		(0.210)	(0.191)		(0.220)	(0.201)
Observations	306	306	306	306	306	306
Adjusted R-sq	0.522	0.507	0.522	0.533	0.520	0.533
Region FEs	X	X	X	X	X	X
City-Tier FEs				X	X	X

Notes: The dependent variable is the log of prefecture-level PM2.5 concentration in 2015. The key independent variables are prefecture-level male–female ratios for cohorts aged 3–10 in 2000. We include sex ratios separately for cohorts with rural hukou and urban hukou. All specifications control for the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai). Robust standard errors clustered at the city level are reported in parentheses. $p < 0.10$, *$p < 0.05$, **$p < 0.01$.

Table A15: SO2 Pollution in 2015

	(1)	(2)	(3)	(2)	(3)	(4)
Dep. Var.	SO2 Concentration 15 Years Later					
Rural Male-female Ratio	16.49***		16.33***	15.78**		14.85**
	(6.285)		(6.171)	(6.216)		(6.135)
Urban Male-female Ratio		2.824	0.429		4.642	2.314
		(3.836)	(3.599)		(3.871)	(3.643)
Observations	306	306	306	306	306	306
Adjusted R-sq	0.522	0.507	0.522	0.533	0.520	0.533
Region FEs	X	X	X	X	X	X
City-Tier FEs				X	X	X

Notes: The dependent variable is prefecture-level SO2 concentration in 2015. The key independent variables are prefecture-level male–female ratios for cohorts aged 3–10 in 2000. We include sex ratios separately for cohorts with rural hukou and urban hukou. All specifications control for the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai). Robust standard errors clustered at the city level are reported in parentheses. $p < 0.10$, *$p < 0.05$, **$p < 0.01$.

Table A16: Early Cohorts Sex Ratio is Unrelated to Future Water Pollution

	(1)	(2)	(3)	(4)
Panel A Dep. Var:	COD Emissions /Population		Waste Water Emissions/Population	
Rural Male-Female Ratio	-0.159	-0.129	1.298	1.159
	(0.242)	(0.239)	(2.083)	(1.947)
Observations	277	277	277	277
R-squared	0.155	0.153	0.0891	0.108
Panel B Dep. Var:	COD Emissions /Ind Output		Waste Water Emissions/ Ind Output	
Rural Male-Female Ratio	-0.0679	-0.0620	0.0200	0.0408
	(0.0470)	(0.0461)	(0.0969)	(0.0957)
Observations	277	277	277	277
Adjusted R-sq	0.0950	0.102	0.113	0.129
Region FE	X	X	X	X
City Tier			X	X

Notes: The dependent variable is prefecture-level water pollution (COD emissions/industrial output, waste water emissions/industrial output, COD emissions/population, waster water emissions/population) in 2012. The key independent variable is prefecture-level male–female ratios for rural cohorts aged 6–13 in 2000. All specifications control for the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai). Robust standard errors clustered at the city level are reported in parentheses. $p < 0.10$, *$p < 0.05$, **$p < 0.01$.

Table A17: Placebo Tests: Sex Ratios among Cohorts Aged 23-40

	(1)	(2)	(3)	(4)
Dep. Var.	PM2.5, 15 years later		Log PM2.5, 15 years later	
Rural Male-Female Ratio	-2.916	0.0336	0.122	0.201
	(8.916)	(8.715)	(0.298)	(0.298)
Observations	305	305	305	305
Adjusted R-sq	0.526	0.535	0.484	0.491
Region FE	X	X	X	X
City Tier			X	X

Notes: The dependent variable is PM2.5 concentration and the log of PM2.5 concentration at the prefecture level in 2015. The key independent variables are prefecture-level male–female ratios for cohorts aged 3–10 in 2000. We include sex ratios separately for cohorts with rural hukou and urban hukou. All specifications control for the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai). Robust standard errors clustered at the city level are reported in parentheses. $p < 0.10$, *$p < 0.05$, **$p < 0.01$.

A.3.2 Additional IV Results Using Contiguous City-pair Panel

Table A18: First-Stage Results using Contiguous City-Pair Panel Sample

	(1)	(2)	(3)	(4)
Dep. Var.	Rural Male-female Ratio			
Panel A: SEs Clustered at the Province Level				
OCP Fine × 1-child Regions(=1)	0.0397***	0.0395***	0.0429***	0.0425***
	(0.00650)	(0.00655)	(0.00713)	(0.00689)
OCP Fine in Other Regions(=1)	0.0303*	0.0308*	0.0319*	0.0329*
	(0.0156)	(0.0167)	(0.0168)	(0.0183)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share			X	X
Cohort Size	X	X	X	X
F-stat of IVs	18.72	17.95	18.05	18.88
Observations	468	468	468	468
Panel B: Two-Way Clustered SEs				
OCP Fine × 1-child Regions(=1)	0.0397***	0.0395***	0.0429***	0.0425***
	(0.00762)	(0.00752)	(0.00805)	(0.00762)
OCP Fine in Other Regions(=1)	0.0303*	0.0308*	0.0319*	0.0329*
	(0.0162)	(0.0174)	(0.0169)	(0.0187)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share			X	X
Cohort Size	X	X	X	X
F-stat of IVs	13.57	13.73	14.29	15.70
Observations	468	468	468	468

Notes: We use contiguous city-pair panel data on male-to-female ratios (among rural-*hukou* cohorts aged 3–10) in 1990 and 2000. We treat those aged 3-10 in 1990 as the first cohort and those aged 3-10 in 2000 as the second cohort. All specifications control for region by year fixed effects and city-pair by year fixed effects, interactions between the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai) and cohort fixed effects, and the interaction between the distance to nearest big cities and cohort fixed effects. In panel A, robust standard errors clustered at the province level are reported in parentheses. In panel B, two-way robust standard errors clustered at the province and border-segment levels are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic.

Table A19: First-Stage Results: Sex Ratios in the Full Sample and Rural Sample

	(1)	(2)	(3)	(4)
Dep. Var.	Full Sample Male-female Ratio		Urban Male-female Ratio	
OCP Fine × 1-child Regions (=1)	0.0309***	0.0319***	0.0362**	0.0423***
	(0.00834)	(0.00754)	(0.0160)	(0.0130)
OCP Fine in Other Regions (=1)	0.0277**	0.0241*	-0.0293	-0.0523
	(0.0108)	(0.0137)	(0.0420)	(0.0489)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share	X	X	X	X
Cohort Size	X	X	X	X
F-stat of IVs	6.760	8.816	3.869	5.903
Observations	468	468	404	404

Notes: We use contiguous city-pair panel data on male-to-female ratios (among cohorts aged 3–10) in 1990 and 2000. We treat those aged 3-10 in 1990 as the first cohort and those aged 3-10 in 2000 as the second cohort. All specifications control for region by year fixed effects and city-pair by year fixed effects, interactions between the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai) and cohort fixed effects, and the interaction between the distance to nearest big cities and cohort fixed effects. In panel A, robust standard errors clustered at the province level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic.

Table A20: Weighted IV Estimates: Inverse Number of Cities as Weights

	(1)	(2)	(3)	(4)
Dep. Var.	PM2.5 Pollution 15 years later			
Rural Male-female Ratio	24.08***	24.01**	23.20***	23.27**
	(8.353)	(10.56)	(7.858)	(9.914)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share			X	X
Observations	468	468	468	468
F-stat of IVs	14.11	13.91	12.61	13.26

Notes: The analysis is based on city-pair panel data linking male-to-female ratios (among rural-*hukou* cohorts aged 3–10) in 1990 and 2000 to PM2.5 pollution levels in 2005 and 2015. We treat those aged 3-10 in 1990 as the first cohort and those aged 3-10 in 2000 as the second cohort. All specifications control for region by year fixed effects and city-pair by year fixed effects, interactions between the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai) and cohort fixed effects, and the interaction between the distance to nearest big cities and cohort fixed effects. We conduct weighted IV regressions, where the weights are based on the inverse of the number of cities in the city-pair sample. Robust standard errors clustered at the province level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We report Kleibergen-Paap rk Wald F statistic.

Table A21: Two-way SEs Clustered at the Province and Border-Segment Levels

	(1)	(2)	(3)	(4)
Dep. Var.	PM2.5 Pollution 15 years later			
Rural Male-female Ratio	21.93***	22.25**	21.06***	21.38**
	(7.708)	(9.011)	(7.015)	(8.249)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share			X	X
Observations	468	468	468	468
F-stat for IVs	13.57	13.73	14.29	15.70

Notes: The analysis is based on city-pair panel data linking male-to-female ratios (among rural-*hukou* cohorts aged 3–10) in 1990 and 2000 to PM2.5 pollution levels in 2005 and 2015. We treat those aged 3-10 in 1990 as the first cohort and those aged 3-10 in 2000 as the second cohort. All specifications control for region by year fixed effects and city-pair by year fixed effects, interactions between the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai) and cohort fixed effects, and the interaction between the distance to nearest big cities and cohort fixed effects. Two-way robust standard errors clustered at the province and border-segment levels are reported in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A22: OLS Regressions: Contiguous City-Pair Panel Sample

	(1)	(2)	(3)	(4)
Dep. Var.	Rural Male-female Ratio			
Full Sample Male-female Ratio	3.384	3.844	3.338	3.791
	(3.009)	(2.969)	(2.970)	(2.914)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share			X	X
Cohort Size	X	X	X	X
Observations	468	468	468	468
Rural Male-female Ratio	7.809***	7.805***	7.777***	7.735***
	(2.194)	(2.163)	(2.004)	(2.101)
Cohort Size	X	X	X	X
SD of OCP Fines		X		X
Minority Share			X	X
Cohort Size	X	X	X	X
Observations	468	468	468	468

Notes: The analysis is based on city-pair panel data linking male-to-female ratios (among cohorts aged 3–10) in 1990 and 2000 to PM2.5 pollution levels in 2005 and 2015. We treat those aged 3-10 in 1990 as the first cohort and those aged 3-10 in 2000 as the second cohort. All specifications control for region by year fixed effects and city-pair by year fixed effects, interactions between the distance to the three major seaports in China (Shenzhen, Tianjin, and Shanghai) and cohort fixed effects, and the interaction between the distance to nearest big cities and cohort fixed effects. Two-way robust standard errors clustered at the province level are reported in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

B. Suggestive Evidence

Rural parents have stronger son preferences if their first-born children are girls.

Table B1: Duration of Breastfeeding by Gender and Birth Order

	(1)	(2)	(3)	(4)
Dep. Var.	The Number of Months of Breastfeeding for Children			
	Urban Sample		Rural Sample	
First-born Boys	0.460	0.383	0.323	0.424
	(0.367)	(0.365)	(0.339)	(0.326)
Higher-order Girls	-1.281	-1.165	0.324	0.440
with elder brothers	(0.885)	(1.105)	(0.533)	(0.563)
Higher-order Boys	-0.219	-0.118	0.0547	0.265
with elder brothers	(0.569)	(0.632)	(0.555)	(0.599)
Higher-order Girls	-0.403	-0.682	0.665	0.596
without elder brothers	(1.102)	(1.083)	(0.520)	(0.549)
Higher-order Boys	-0.471	-0.213	1.382*	1.425*
without elder brothers	(0.558)	(0.554)	(0.790)	(0.804)
Observations	1,072	1,072	1,536	1,535
R-squared	0.150	0.209	0.167	0.250
Province FE	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Mean of Y	10.80	10.80	11.31	11.31

Notes: Data come from the China Health and Nutrition Survey (CHNS). We use first-born girls as the base group. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Number of Workers in High- and low-Pollution Sectors (Unit:1 Million)

	(1) Low-Pollution Sectors	(2) High-Pollution Sectors
	Panel A: Indicator: Waste Gas Emissions/Output	
Workers with *Rural* Hukou, 2000	96.241	460.853
Workers with *Urban* Hukou, 2000	72.404	70.258
Workers with *Rural* Hukou, 2015	115.368	243.339
Workers with *Urban* Hukou, 2015	70.217	75.513
	Panel B: Indicator: SO2 Emissions/Output	
Workers with *Rural* Hukou, 2000	93.313	463.781
Workers with *Urban* Hukou, 2000	65.672	76.990
Workers with *Rural* Hukou, 2015	95.751	262.956
Workers with *Urban* Hukou, 2015	61.030	84.700

Notes: The table shows the number of workers in high- and low-pollution sectors, respectively. Panel A uses waste gas emissions per output to divide sectors into two groups, and panel B uses SO2 emissions per output to categorize sectors.

C. Related Studies in Public Health and Epidemiology

Table C1: Findings from Public Health and Epidemiological Studies

	Panel A : Differential vulnerability between X- and Y-bearing spermatozoa
Oyeyipo et al. (2017)	Environmental stress (acidic media, increased temperatures, and elevated H_2O_2) selectively reduces Y-sperm survival, resulting in relative enrichment of X-sperm. Example: at pH 5.5, the proportion of X-bearing sperm increases from 55% (neat proportion) to 62% (+12.7 p.p.), while the proportion of Y-bearing sperm declines from 45% (neat proportion) to 38% (−15.5 p.p.). At 41 °C, X-bearing sperm accounts for approximately 59% of surviving sperm, compared with 41% Y-bearing sperm.
Radwan et al. (2018)	Air pollution shifts the sperm sex ratio against Y-bearing sperm. Conditional on characteristics of the study population, a 1 µg/m³ increase in PM2.5 exposure reduces the proportion of Y- to X-chromosome–bearing sperm by 0.56% ($p < 0.01$). Similarly, a 1 µg/m³ increase in PM10 exposure reduces the Y/X sperm ratio by 0.55% ($p < 0.03$).
You et al. (2017)	Y-bearing sperm exhibit lower survival than X-bearing sperm under extreme temperatures and non-neutral pH conditions. At 22 °C, under pH 6.5 (slightly acidic) or pH 8.5 (alkaline) conditions, the Y:X chromosome ratio, which is initially approximately 1, declines to about 0.86 after 5 days of incubation. Differences in survival rates under non-neutral pH conditions widen further at both lower (4 °C) and higher (37 °C) temperatures.
	Panel B: Differential vulnerability between female and male fetuses
Basu et al. (2016)	Heat-related stillbirth risk is higher for male fetuses. A 10 °F (5.6 °C) increase in apparent temperature raises stillbirth risk by 10.4% overall, with larger effects for male fetuses (13.3%) than for female fetuses (9.8%).
Fukuda et al. (2014)	Temperature shocks are associated with higher male fetal loss. A statistically significant positive association was found between yearly mean temperature differences (the annual average of monthly mean temperature deviations from the corresponding 30-year historical monthly means) and sex ratios of fetal deaths (Pearson correlation coefficients(r)=0.670, P<.0001). And a statistically significant negative association was found between.
Abdel Ghany et al. (2026)	In sub-Saharan Africa, high temperatures during the first trimester are associated with fewer male births. A one-SD increase in the number of first-trimester days with a daily maximum temperature of 20–25 °C reduces the probability of a male birth by 0.33 percentage points; corresponding reductions are 0.46 percentage points for a one-SD increase in 25–30 °C days and 0.59 percentage points for a one-SD increase in days exceeding 30 °C.
Jedrychowski et al. (2009)	Male fetuses are more growth-restricted by air pollution than female fetuses. A 30 µg/m³ increase in prenatal

	PM2.5 exposure (a shift from the 25th to the 75th percentile of exposure in the sample) leads to a birth-weight deficit of 189 g in male newborns, compared with 17 g in female newborns; the corresponding reduction in birth length among male infants amounts to 1.1 cm, compared with 0.4 cm among female infants.
Karlsson et al. (2021)	Extreme temperatures disproportionately increase stillbirth risk among male fetuses. When mean temperatures during the seven days prior to birth are around −10 °C, the incidence rate ratio (IRR) of stillbirth is 2.01 for girls and 2.69 for boys, relative to the minimum-mortality temperature of +14 °C. At high temperatures (+25 °C), an adverse effect is observed only for boys, with an IRR of 2.41 relative to +14 °C, while no meaningful effect is detected for girls.
Lakshmanan et al. (2015)	Associations of prenatal air pollution exposure and reduced birth weight were most evident in males born to obese mothers. On average, a 1 $\mu g/m^3$ increase in PM2.5 is associated with a −0.26 SD decline in birth weight for gestational age (BWGA) z-scores among male infants of obese mothers; no comparable effect for females.