NBER WORKING PAPER SERIES

FETAL POLLUTION EXPOSURE, COGNITIVE ABILITY, AND GENDER-SPECIFIC PARENTAL INVESTMENT

Xin Zhang
Yixuan Wang
Xingyi Hu
Xi Chen

Working Paper 35591
http://www.nber.org/papers/w35591

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2026

The authors acknowledge the Institute of Social Science Survey at Peking University for providing us with the China Family Panel Studies (CFPS) data. The authors acknowledge helpful comments by participants and discussants at various conferences, seminars, and workshops. Xin Zhang acknowledges financial support from the National Natural Science Foundation of China (72473011), the Major Program of Philosophy and Social Sciences Research of the Ministry of Education of China (2025JZDZ026), and the Fundamental Research Funds for the Central Universities. Yixuan Wang acknowledges financial support from the University Development Fund from The Chinese University of Hong Kong, Shenzhen (UDF01004394). Xi Chen acknowledges financial support from the Drazen scholarship and the Aden scholarship, which are dedicated to research on Chinese healthcare systems. The authors declare that they have no conflict of interest. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Xin Zhang, Yixuan Wang, Xingyi Hu, and Xi Chen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Fetal Pollution Exposure, Cognitive Ability, and Gender-specific Parental Investment
Xin Zhang, Yixuan Wang, Xingyi Hu, and Xi Chen
NBER Working Paper No. 35591
August 2026
JEL No. D13, I14, I24, J13, J16, Q53

ABSTRACT

We examine how prenatal exposure to air pollution interacts with gender-biased parental investments to shape long-run cognitive performance. Using nationally representative survey data from China, we show that fetal PM2.5 exposure significantly reduces cognitive ability for women, particularly among those with brothers. Our evidence suggests that the family investment channel operates primarily through education rather than health, with gender-biased educational resource allocation compounding negative outcomes for females. Specifically, conditional on the same level of fetal PM2.5 exposure, females receive less homework assistance from their families and attain lower levels of education.

Xin Zhang
Beijing Normal University
School of Statistics
xin.zhang614@bnu.edu.cn

Yixuan Wang
The Chinese University of Hong Kong, Shenzhen (CUHK-SZ)
School of Humanities and Social Science
yxwang@cuhk.edu.cn

Xingyi Hu
The Ohio State University
Department of Agricultural, Environmental, and Development Economics
hu.2285@osu.edu

Xi Chen
Yale University
Department of Health Policy and Management, and Department of Economics
and NBER
xi.chen@yale.edu

1 Introduction

Early-life shocks, whether positive or negative, can have profound and lasting influences on child development (e.g., Adhvaryu et al., 2019; Almond et al., 2009; Lundborg et al., 2022; Baker et al., 2019). Within this literature, a growing body of studies has estimated the effects of early-life exposure to air pollution on children's health, subsequent development, and adult outcomes.[1] Most of these studies estimate the overall effect of early-life exposure to air pollution on various outcomes. However, early-life shocks influence later human capital not only through biological channels but also through parental investments. An improved understanding of this latter channel may promote interventions and policies that improve social equity.

Numerous studies have examined how parental investments respond to early-life health shocks (e.g., Li et al., 2010; Yi et al., 2015; Adhvaryu and Nyshadham, 2016; Restrepo, 2016; Fan and Porter, 2020; Sanz-de Galdeano and Terskaya, 2022; Dizon-Ross, 2021).[2] However, estimating the average effect can obscure significant heterogeneities across population groups, such as by gender, particularly in countries where gender preference is prevalent. In such settings, girls and boys tend to receive a disproportionate share of resources, especially for credit-constrained families (Lei et al., 2017; Lin et al., 2021). In contrast, positive shocks may mitigate gender gaps in resources obtained, such as through released budget or time constraints. For example, Qin et al. (2025) show that economic development, reflected by both more household resources and a smaller gap in household resources between seasons, reduces the gender gap in birth quarter effects on neonatal investment.

This study explores gender-specific parental investment in response to children's fetal

[1] See Currie et al. (2014) for a synthesis of this literature. Most studies focus primarily on the impact of exposure in early life to air pollution on short-term outcomes, such as infant mortality (e.g., Chay and Greenstone, 2003; Currie and Neidell, 2005; Greenstone and Hanna, 2014; Tanaka, 2015; Knittel et al., 2016) and child health (e.g., Neidell, 2004; Lleras-Muney, 2010; Schlenker and Walker, 2016). However, relatively few studies make a direct connection between early-life exposure to air pollution and longer-term outcomes.

[2] See Almond and Mazumder (2013) for a comprehensive review of the literature on fetal origins and parental responses.

pollution exposure and its long-term impact on human capital formation, in a context where gender preference prevails. We contribute to policy formulation by highlighting the importance of gender equity starting from early life, and shed light on policies aimed at narrowing gender gaps. Specifically, we examine the causal impact of exposure to fetal PM2.5 on standardized verbal and math test scores across a wide age range of individuals in a nationally representative longitudinal survey in China. To address endogeneity in air pollution, we instrument PM2.5 concentrations using local wind direction and pollution of nearby upwind cities, a widely used instrumental variable (IV) strategy in recent studies (Zheng et al., 2019; Chen et al., 2021; Wang et al., 2022).

Our IV estimates suggest that a one standard deviation (SD) increase in fetal PM2.5 exposure leads to a decline in verbal and math test scores by 0.122 SD and 0.170 SD, respectively. Our findings reveal significant gender differences, with females being more vulnerable to the detrimental effects of PM2.5 exposure on both verbal and math abilities, especially when they have (male) siblings. To explain this gendered pattern, we test several potentially competing channels, including mortality selection, selection into motherhood, fertility decisions, and gender-specific investments by families. We find that gender-biased human capital investment by families limits girls' access to education and household resources, resulting in worse cognitive skills compared to boys when exposed to negative shocks *in utero*. Our estimates indicate that, when exposed to the same level of fetal PM2.5, females receive less assistance with homework from family members and achieve lower levels of education compared to their male counterparts.

This paper aims to contribute to the literature on several fronts. First, our study adds to the burgeoning literature on how parental investments respond to health endowments at birth. We document sex-selective postnatal educational investments as a mechanism linking prenatal PM2.5 exposure to gender disparities in cognitive abilities. The literature on gender-biased intra-household resource allocation primarily examines the short-term impact of external shocks (Behrman, 1988; Behrman and Deolalikar, 1990; Cameron and Worswick,

2001; Thomas et al., 2004; Coleman et al., 2025; Braido et al., 2012), while we focus on the long-term impacts. Moreover, larger female schooling responses to early-life health shocks have been documented in previous studies (Field et al., 2009; Maccini and Yang, 2009; Hoynes et al., 2016), but the reasons for these gender differences are relatively less well explored. Our study provides one of the first pieces of empirical evidence that gender preference can lead to gender disparities in intrahousehold educational investments in response to prenatal shocks, potentially affecting long-term outcomes.[3]

Second, studies examining the impact of air pollution on cognitive performance have mainly concentrated on its transient effects during the tests (e.g., Ebenstein et al., 2016; Zhang et al., 2018b; Cook et al., 2023; Krebs and Luechinger, 2024; Lai et al., 2022), early life exposure on educational attainment (Chen, 2025), or cumulative exposure over the years (Ham et al., 2014; Zhang et al., 2018b). One recent work by Ferro et al. (2024) studies the medium-term effect of prenatal exposure to air pollution, focusing on filled prescriptions and hospitalizations. However, evidence regarding the impact of prenatal exposure to air pollution on long-term cognitive performance is relatively limited. Our paper speaks directly to several economic studies on the relationship between *in utero* air pollution exposure and later-life cognitive abilities.[4] Closely related to this study, Molina (2021) estimates the long-term effect of *in utero* pollution exposure based on data from Mexico and finds that exposure

[3]Our paper is closely related to Wu et al. (2023) and Qin et al. (2025), who find that parents are more likely to compensate for girls under positive productivity shocks in early life. In contrast, our paper demonstrates that parents are more likely to invest in boys when faced with negative pollution shocks. Providing compensation to individuals with fewer advantages (females) during positive shocks acts as a safety net against economic volatility, while protecting individuals with advantages (males) during negative shocks helps mitigate risks and uncertainties associated with adverse events. Both of these actions can be viewed as risk-averse strategies, as they aim to minimize potential negative impacts on individuals, promote stability, and mitigate the risks associated with income inequality.

[4]For example, Sanders (2012) examines the impact of prenatal suspended particulate pollution on standardized test scores in high schools, using changes in relative manufacturing employment as an IV. However, their IV could affect test scores through channels other than air pollution, such as local educational funding and household income. Bharadwaj et al. (2017) analyze the impact of fetal exposure to CO and PM10 at the trimester level on math and language skills in the fourth grade in Chile. They employ a fixed-effect model and sibling comparisons to address concerns about residential sorting and time-invariant family characteristics. However, they neglect the effect of parents' differentiated investments in their offspring based on gender and birth order. Rosales-Rueda and Triyana (2019) estimate the effects of early-life exposure to air pollution using Indonesia's forest fires as a natural experiment. They find that children exposed to the fires *in utero* have worse health outcomes but do not suffer significant negative effects on cognitive function.

to pollution in the second trimester leads to significantly lower adult cognitive ability for both genders, but lower high school completion and income for women only. She further reveals that the gender differences in investment responses arise from gender-specific labor market expectations. Our paper provides an alternative explanation for gender differences in responses to early-life shocks.

Third, we estimate the impact of fetal exposure to air pollution on the low-stakes cognitive performance of adults and adolescents.[5] Existing studies mainly focus on the high-stakes exams taken by school-aged children (Sanders, 2012; Bharadwaj et al., 2017; Duque and Gilraine, 2022). However, the low-stakes cognitive tests in our study are close to our day-to-day cognitive activities. Our findings suggest that the quality of routine decision-making processes is compromised by fetal exposure to PM2.5.

Finally, few studies have investigated the impact of fetal exposure to air pollution on cognitive performance and parental responses in developing countries, primarily due to data limitations. However, examining these impacts in developing countries is crucial because a majority of the global population exposed to unhealthy air and most pollution-related premature deaths are in developing nations, where pollution levels are typically higher and overall health status is often poorer compared to more developed nations (World Health Organization, 2024a).[6] Moreover, parental responses to early shocks in children may be more important in developing countries with weaker health infrastructure, less developed credit markets, and inadequate social protection systems. Therefore, the role of the family must be taken into account when designing public policies to remedy inequality at birth or during early childhood.

The rest of the paper is organized as follows: Section 2 presents a conceptual framework to help guide the empirical analysis; Section 3 introduces the data we use to perform the

[5]As the data on air pollution is only available after 1980 and only interviewees above age 10 are required to take the cognition tests, the final datasets include respondents born between 1980 and 2000, who were aged between 10 and 30 at the time of the survey.

[6]For example, 98.6% of the population in China was exposed to PM2.5 at unsafe levels during 2013–2014 according to the World Health Organization (WHO) guideline (Long et al., 2018).

analysis; Section 4 presents our econometric strategy; Section 5 displays the main results, including the baseline results, robustness checks, and heterogeneous effects; Section 6 explores some potential mechanisms, and Section 7 draws conclusions.

2 Conceptual Framework

How might fetal exposure to ambient air pollution affect child and adult outcomes? This paper focuses on the impact of fetal exposure to PM2.5, i.e., particulate matter with a diameter smaller than 2.5 μm. PM2.5 is known to have more detrimental effects on health compared to larger particulates, which are typically trapped in the upper airways and can be cleared by mucociliary mechanisms. However, due to its minuscule size, PM2.5 can penetrate the lungs at the alveolar level and release toxic substances into the bloodstream, leading to respiratory and cardiovascular diseases. These detrimental effects are particularly pronounced *in utero*. The decreased oxygen supply or organ damage experienced by pregnant women results in reduced oxygen transfer to the fetus, which hinders fetal brain development. Moreover, the particulates can directly pass through the bloodstream to reach the fetus, posing a risk to the development of the fetal respiratory and cardiovascular systems.

All of these *in utero* physiological impacts can translate into damage to cognitive function as a child develops and transitions into adulthood. Consequently, fetal exposure to air pollution can affect long-term human capital development due to direct health impairments. Additionally, parental investments play a crucial role in shaping human capital formation, and these investments are influenced by the health endowments at birth. Hence, fetal exposure to air pollution may also indirectly affect long-term human capital formation through family investments.

To formalize these channels, we propose a simple framework based on the conceptual model presented in Currie et al. (2014). Our framework categorizes life into three distinct stages: fetus, childhood, and adulthood. The fetus stage begins *in utero* and ends at birth.

Childhood starts at birth and encompasses the entire duration of schooling. Adulthood can be perceived as the post-schooling period, during which individuals typically enter the workforce.

Assume that fetus (F) human capital H_F is dependent on fetal exposure P_F and family characteristics X, such as genetics, where $\partial f_F/\partial P_F < 0$:

$$H_F = f_F(P_F, X) \tag{1}$$

Childhood (C) human capital H_C is a function of childhood pollution exposure P_C and human capital accumulated *in utero* H_F, where $\partial f_C/\partial P_C < 0$, and $\partial f_C/\partial H_F > 0$. Importantly, families can make investments I_C in their children, with $\partial f_C/\partial I_C > 0$, indicating that an increase in investments made by the family results in higher levels of childhood human capital. It's worth noting that the investments made by families may rely on the initial human capital endowment at birth.

$$H_C = f_C(P_C, H_F, I_C(H_F)) \tag{2}$$

Finally, adult (A) outcomes H_A depend on adult pollution exposure P_A and on the childhood human capital H_C, where $\partial f_A/\partial P_A < 0$, and $\partial f_A/\partial H_C > 0$.

$$H_A = f_A(P_A, H_C) \tag{3}$$

Equations (1)-(3) suggest that the impact of fetal pollution exposure on childhood and adulthood human capital relies on both the direct effects of exposure and the transmission of those effects through parental investments. These two pathways can be demonstrated by taking the derivative of equations (2) and (3) with respect to P_F:

$$\frac{dH_C}{dP_F} = \frac{\partial f_C}{\partial H_F}\frac{\partial f_F}{\partial P_F} + \frac{\partial f_C}{\partial I_C}\frac{\partial I_C}{\partial H_F}\frac{\partial f_F}{\partial P_F} \tag{4}$$

$$\frac{dH_A}{dP_F} = \frac{\partial f_A}{\partial H_C}\frac{dH_C}{dP_F} = \frac{\partial f_A}{\partial H_C}\frac{\partial f_C}{\partial H_F}\frac{\partial f_F}{\partial P_F} + \frac{\partial f_A}{\partial H_C}\frac{\partial f_C}{\partial I_C}\frac{\partial I_C}{\partial H_F}\frac{\partial f_F}{\partial P_F} \tag{5}$$

In equations (4) and (5), the first term captures the direct biological effect of fetal exposure to air pollution, which is negative and reflects the physiological damage caused by pollution during fetal development. The second term captures the indirect effect operating through parental investments. To interpret this indirect channel, we consider two empirically relevant cases. First, if $\partial I_C/\partial H_F < 0$, parental investment is compensatory, implying that families allocate more resources to children with lower initial endowments. In this case, the indirect effect partially offsets the direct biological damage. Second, if $\partial I_C/\partial H_F > 0$, parental investment is reinforcing, implying that families invest less in children who are more adversely affected. In this case, the indirect effect amplifies the direct biological effect.[7]

Overall, the framework decomposes the long-run effect of fetal pollution exposure into a direct biological channel and an indirect channel operating through parental investment, thereby guiding our empirical analysis.[8] To complement this conceptual decomposition, we develop a two-child model with son preference in Appendix B. This model introduces additional structure that allows the sign of $\partial I_C/\partial H_F$ to be derived under explicit assumptions on parental beliefs and investment costs, and provides a structural interpretation of the mechanisms underlying the empirical patterns.

3 Data

We combine three datasets that collect information on cognitive test scores along with demographic variables, air pollution, and other weather measures.

[7] Our empirical results in Table 6 provide reduced-form evidence on which case is more consistent with the data. Importantly, the conceptual framework itself does not determine the sign of $\partial I_C/\partial H_F$ or impose compensatory or reinforcing parental investment. Rather, it serves as a decomposition device that clarifies how different parental responses may attenuate or amplify the biological consequences of prenatal pollution.

[8] Guided by this decomposition, we draw on three complementary sets of reduced-form evidence: birth outcomes, which to a large extent reflect gender differences in biological vulnerability (Table A10); parental responses, which shed light on the indirect investment channel (Table 6); and cognitive outcomes, which capture the overall long-run effect (Table 3). The results are consistent with greater biological vulnerability among boys but a stronger adverse investment response for girls.

3.1 Cognition data

Data on cognitive test scores are obtained from the China Family Panel Studies (CFPS), a nationally representative biennial longitudinal household survey of Chinese families and individuals. CFPS is funded by Peking University and carried out by the university's Institute of Social Science Survey.[9] We mainly rely on CFPS 2010 as it is the only wave that reveals the birthplace information for each respondent. The cognitive ability module in the questionnaire contains 24 mathematics questions and 34 word-recognition questions. All these questions are obtained from standard textbooks and are sorted in ascending order of difficulty. The starting question depends on the respondent's education level.[10] The test ends when the individual incorrectly answers three questions in succession. The final test score is defined as the rank of the hardest question a respondent can answer correctly. If the respondent fails to answer any question, the score is assigned as the rank of the starting question minus one.[11] Only individuals above the age of 10 are required to take the tests.[12] The cognitive test scores, along with demographic variables, are linked to air pollution and meteorological data based on the unique birth city identifier and the birth year and month information.

[9]The survey uses multistage probability proportional to size sampling with implicit stratification to represent Chinese society better. The 2010 CFPS baseline sample is drawn through three stages (i.e., county, village, and household) from 25 provinces. The 162 randomly chosen counties largely represent Chinese society (Xie and Hu, 2014).

[10]Specifically, those whose education level is primary school or below start with the 1st question; those who attended middle school begin with the 9th question in the verbal test and the 13th question in the math test; and those who finished high school or above start with the 21st question in the verbal test and the 19th question in the math test.

[11]For example, a respondent with a middle school education begins with the 9th question in the verbal test. If the hardest question one can correctly answer is the 14th question, then the verbal test score would be 14. However, if one fails the 9th, 10th, and 11th questions consecutively, the verbal test score would be 8. Since the respondents did not know the testing rules prior to the interviews, there should be no incentive to manipulate test performance on purpose.

[12]The verbal and math test scores were measured in the 2010 survey wave. Given that respondents in our sample were born between 1980 and 2000, their ages at the time of measurement range from 10 to 30 years.

3.2 Air pollution data

The concentration of PM2.5 is calculated from the satellite-based Aerosol Optical Depth (AOD) data following the formula provided by Buchard et al. (2016). The AOD data are drawn from the product M2T1NXAER version 5.12.4 from the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) released by the U.S. National Aeronautics and Space Administration (NASA). The data are available for each 0.5°× 0.625°(around 50 km × 60 km) grid at the hourly level since 1980. We aggregate all the grids within each city and calculate their mean values. We then average the data to a monthly level for each city. Figure A1 displays the spatial distribution of mean PM2.5 at the city level during 1980-2000. This dataset has been widely used in previous studies (Deschenes et al., 2020; Fu et al., 2021; Chen et al., 2022), and validated with air pollution data from ground-based monitoring stations in China (Chen et al., 2022).[13] We do not use ground-based pollution data because they only cover a few cities before 2012 and report the Air Pollution Index (API) instead of concentrations of specific air pollutants.

3.3 Weather data

The weather data are provided by the China National Meteorological Data Service Center (CMDC) under the National Meteorological Information Center of China. The dataset contains daily weather records of 824 monitoring stations along with their longitudes and latitudes in China. The weather measures include temperature, precipitation, wind speed, wind direction, sunshine duration, and relative humidity. To merge the birth cities with weather readings, we calculate the mean values of all the monitoring stations within each city. To account for possible non-linear effects of weather controls, we divide the spectrum of temperatures into 10 bins and calculate the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C, 16-20 °C, 20-24 °C, 24-28 °C, and

[13]Please see the online appendix of Chen et al. (2022) for a comprehensive comparison between AOD-based and station-based pollution data.

>28 °C) at the city-monthly level. For other weather measures, we use the monthly mean values with their quadratic terms for each city.

As the air pollution data are available after 1980 and only interviewees older than 10 years old are required to take the cognition tests, the final dataset includes 9,038 respondents born between 1980 and 2000. Table 1 provides summary statistics of key variables for the full sample, as well as the female and male subsamples, respectively. During our sample period, the 9-month mean exposure to PM2.5 exceeded 22 $\mu g/m^3$, which is more than four times the annual limit recommended by the WHO.[14] Figure A2 displays the time trend of monthly mean PM2.5 during 1980-2000.[15]

4 Empirical Strategy

4.1 OLS estimation

We use equation (6) to estimate the impact of prenatal exposure to air pollution on cognitive performance:

$$score_{ijym} = \beta_0 + \beta_1 PM2.5^G_{jym} + \beta_2 X_{ijym} + \beta_3 W_{jym} + \omega_j + \gamma_y + \tau_m + \varepsilon_{ijym}, \tag{6}$$

$$PM2.5^G_{jym} = \frac{1}{9} \sum_{t=ym-8}^{ym} PM2.5_{jt}, \tag{7}$$

where the outcome variable $score_{ijym}$ is the standardized verbal or math test scores for individual i born in city j in month m of year y. $PM2.5^G_{jym}$ denotes the average PM2.5

[14] According to the 2021 WHO Air Quality Guidelines, the recommended annual limit for PM2.5 is 5 $\mu g/m^3$. Source: https://www.who.int/news-room/feature-stories/detail/what-are-the-who-air-quality-guidelines.

[15] We also make a naive cross-country comparison of the PM2.5 concentration level in 1990 for China, India, Mexico, Thailand, and South Africa, using the Global Burden of Disease Study 2021 (GBD 2021) Air Pollution Exposure Estimates (Institute for Health Metrics and Evaluation, 2024). The comparison shows that India is the most polluted country with an average PM2.5 concentration of 62.00$\mu g/m^3$, followed by China (49.30$\mu g/m^3$), Thailand (44.30$\mu g/m^3$), Mexico (32.30$\mu g/m^3$), and South Africa (24.57$\mu g/m^3$). All of these numbers are significantly higher than the WHO air quality guideline. However, as noted by World Health Organization (2024b), this cross-country comparison should be conducted with caution due to measurement methods, data quality, and incomplete coverage.

concentration level during the full gestational age (the past 9 months before birth) as calculated by equation (7). X_{ijym} represents demographic controls, including gender, maternal age, and the mother's education years. W_{jym} contains a set of weather variables, including the total number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C, 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration and mean relative humidity with their square terms during the gestational period. ω_j indicates the birth city fixed effect; γ_y and τ_m denote the birth year and month fixed effects. The identification is from the monthly variation in PM2.5 pollution within each city after controlling for weather characteristics, annual shocks, and common seasonality effects. As a robustness check, we further add birth year-by-month FE, birth city-by-year FE, and birth city-by-month FE in equation (6). ε_{ijym} is the error term. Robust standard errors are clustered at the birth city level.[16] A robustness check with two-way clustering at the birth city and year-season levels is also performed.

4.2 IV estimation

OLS estimates of equation (6) are prone to bias for the following two reasons. First, time-varying omitted variables that correlate with PM2.5 pollution and cognitive test scores may bias the estimates despite the rich control variables and fixed effects included in the specifications. For example, local industrial development is correlated with the pollution level and also related to the outcome variables through two opposite channels. On the one hand, students may be lured to drop out of school to seek higher short-term income, which results in lower cognitive abilities. On the other hand, local industrial development may boost the local economy and provide better education quality to school-aged students, resulting in better cognitive abilities. Thus, the direction of omitted variable bias is ambiguous in our context. Second, there exist measurement errors in the mean PM2.5 readings at the city level matched to each respondent, because it is unknown where exactly the respondent lived

[16]The respondents were born in 767 counties across 214 cities, resulting in 214 clusters.

in their birth city. Therefore, we deal with the omitted variable issues and measurement errors by constructing IVs following a similar procedure employed by Wang et al. (2022).

PM2.5 could be easily transported from the nearby upwind cities to the focal city by wind (Zhang et al., 2018a; Chen et al., 2020). For each focal city j on month m, day d of year y, we construct the $P_{j,ymd}$ as a weighted average of the PM2.5 levels of all the nearby upwind cities within 50-200 km from city j using equation (8).[17] We omit cities less than 50 km from the focal city to alleviate concerns of endogeneity risk.

$$P_{j,ymd} = \sum_{n:\, within\ 50-200\ km\ of\ j,\ \cos\theta_{jn,ymd}>0} \omega_{jn,ymd} \times PM2.5_{n,ymd}, \tag{8}$$

$$\omega_{jn,ymd} = \frac{\frac{\cos(\theta_{jn,ymd})}{distance_{jn}} \times \mathbb{1}(\cos\theta_{jn,ymd} > 0)}{\sum_{k:\, within\ 50-200\ km\ of\ j} \frac{\cos(\theta_{jk,ymd})}{distance_{jk}} \times \mathbb{1}(\cos\theta_{jk,ymd} > 0)}, \tag{9}$$

$$IV_{j,ymd} = \sum_{t=ymd-6}^{ymd} P_{j,t}. \tag{10}$$

$PM2.5_{n,ymd}$ is the PM2.5 level of nearby city n on month m, day d of year y. We only keep the nearby cities in the upwind directions with $\cos\theta_{jn,ymd} > 0$, where $\theta_{jn,ymd}$ is the angle between the vector $\vec{nj}$ from the nearby city n to the focal city j and the prevailing wind direction on month m, day d of year y. Figure A3 provides a graphical illustration of the aforementioned parameters. According to equation (9), the weight $\omega_{jn,ymd}$ is larger if the distance between city n and j ($distance_{jn}$) is smaller and the angle $\theta_{jn,ymd}$ is closer to zero. Furthermore, considering that the transport of air pollutants takes time, we incorporate the impact of PM2.5 pollution from nearby cities in the preceding week. Thus, the IV for local PM2.5 pollution, denoted as $IV_{j,ymd}$, is defined in equation (10). After obtaining the $IV_{j,ymd}$ at the daily level, we aggregate it to the city-monthly level IV_{jym}. We then calculate IV^G_{jym} in the same way as PM2.5 in equation (7).

We use IV^G_{jym} to instrument for $PM2.5^G_{jym}$ in equation (6) and estimate the equation

[17] In our sample, the average area of a prefecture-level city is approximately 14.4 thousand km^2. The average distance between the focal city and its "upwind" counterpart is 136 km.

using two-stage least square (2SLS) method, where the first stage specification is:

$$PM2.5^{G}_{jym} = \beta_0 + \beta_1 IV^{G}_{jym} + \beta_2 X_{ijym} + \beta_3 W_{jym} + \omega_j + \gamma_y + \tau_m + \varepsilon_{ijym}. \tag{11}$$

We demonstrate the validity of the IVs with regard to the exclusion and relevance assumptions. The relevance assumption is reassured by the strong first-stage regression results shown in Table 2. For the exclusion restriction, since changes in the wind direction are a natural phenomenon, pollution transmitted from the nearby upwind cities to the focal city is uncorrelated with the unobservables in the focal city. Similar IV strategies have been widely adopted in recent studies (Deryugina et al., 2019; Bayer et al., 2009; Zheng et al., 2019; Chen et al., 2021; Barwick et al., 2021; Fu et al., 2022).[18] After controlling for meteorological variables, as well as spatial and temporal fixed effects, our identification assumption is that the transported air pollutants from nearby cities could only affect the cognitive performance of people born in the focal city by raising air pollutant concentration levels. We provide additional robustness checks on the IV construction method in Section 5.2.

5 Results

5.1 Baseline results

Table 2 displays the results for the impact of fetal exposure to PM2.5 on verbal and math test scores. All regressions control for demographic controls, weather controls, birth city fixed effects, as well as birth year and month fixed effects. Columns (1) and (2) of Table 2 first report the OLS estimates. The results show that the mean PM2.5 level during pregnancy is negatively correlated with both verbal and math test scores. Specifically, a one SD increase in PM2.5 (6.719 $\mu g/m^3$) during pregnancy leads to a reduction in verbal

[18] As a robustness check, we use the monthly average number of thermal inversions as an alternative instrument for air pollution exposure. The resulting estimates remain qualitatively similar to the baseline and heterogeneity results, indicating that our findings are robust to this alternative source of exogenous variation. These results are available upon request.

and math test scores by 0.813 (0.108 SD) and 0.867 (0.153 SD), respectively.[19] Column (3) of Table 2 presents the first-stage results of our IV estimation. The constructed IV is significantly and positively correlated with PM2.5. Overall, we find a strong first-stage relationship. The Kleibergen–Paap (KP) F-statistics are well above the Stock–Yogo critical value, which reassures the validity of the IVs. Columns (4) and (5) of Table 2 report the IV estimates. The results indicate that a one SD increase in fetal exposure to PM2.5 during the gestational period is associated with a reduction in verbal and math test scores by 0.921 (0.122 SD) and 0.968 (0.170 SD), respectively.

The magnitudes of the coefficients on PM2.5 in the IV estimates are greater than those in the OLS estimates, which is common in the literature (Barwick et al., 2021; Wang et al., 2022). From the previous analysis, two possible reasons exist for this downward bias. First, the benefit of omitted local industrial development through education quality outweighs its negative effect through better outside employment options. Since industrial development is also positively correlated with the PM2.5 level, the sign of the omitted variable bias is positive, biasing the OLS estimate towards zero. Second, measurement errors in PM2.5 could lead to attenuation bias. Given the potential endogeneity in the OLS regressions, we will focus on the interpretation of 2SLS results in the remaining sections.

Table A1 further examines the effect by trimester. We divide the 9-month gestational period into three pregnancy trimesters. Specifically, we assign months 1–3 to trimester 1, months 4–6 to trimester 2, and months 7–9 to trimester 3. The negative effects of exposure to PM2.5 are larger and more statistically significant in the first and third trimesters for both test scores, whereas PM2.5 in the second trimester is insignificant. Specifically, a one SD increase in PM2.5 (7.701 $\mu g/m^3$) in the third trimester is associated with a reduction in verbal and math test scores by 0.539 (0.072 SD) and 0.562 (0.099 SD), respectively. This finding is consistent with Bharadwaj et al. (2017), who also find that the test scores are sensitive to air pollution exposure in the first and third trimesters.[20]

[19]Note the SD of verbal and math test scores are 7.527 and 5.677, respectively.

[20]Some evidence in animal studies suggests that the first and third trimesters of pregnancy are particularly

Moreover, we estimate a non-linear relationship between prenatal air pollution and cognitive ability. According to the WHO global air quality guidelines,[21] we classify the PM2.5 concentrations into five categories: "$< 15\mu g/m^3$", "$15 - 25\mu g/m^3$", "$25 - 37.5\mu g/m^3$", "$37.5 - 50\mu g/m^3$", and "$> 50\mu g/m^3$". We assign each category a dummy variable and calculate the number of days falling in each category during the gestational period. We designate "$< 15\mu g/m^3$" as the reference group. As shown in Table A2, the marginal effects increase significantly and nonlinearly with increasing PM2.5 concentrations. Specifically, an additional day of exposure to polluted days with PM2.5 levels above 50 $\mu g/m^3$ during the gestational period, compared to days with good air quality ($< 15\mu g/m^3$), is associated with a decrease in verbal and math test scores by 0.355 (0.047 SD) and 0.244 (0.043 SD), respectively.

We compare the magnitude of our estimates with other similar studies that also examine the effect of early-life exposure to air pollution on cognitive performance. The relevant studies are summarized in Table A3. To facilitate comparisons across studies, we report the effect sizes in SD change of test scores per one SD higher in air pollutants. As for our estimates, a one SD increase in fetal exposure to PM2.5 is associated with a reduction in verbal and math test scores by 0.122 SD and 0.170 SD, respectively. Meanwhile, Sanders (2012) observes that a one SD increase in TSPs (7.27 $\mu g/m^3$) in the student's year of birth correlates with a 0.06 SD decline in high school scores. Bharadwaj et al. (2017) report that a one SD increase in PM10 (29.24 $\mu g/m^3$) during the third trimester is associated with a 0.029 SD reduction in math scores and a 0.037 SD reduction in language scores. While our estimate is marginally higher, it reflects a comparable magnitude to their findings.

important in brain development. For example, during the first trimester, exposure to air pollution can impair the migration of neuroblasts during neurogenesis and thus impede brain development (Woody and Brewster, 1990). Exposure to air pollution during the third trimester can block important receptors that regulate neuronal cell death, leading to neurodegeneration in the developing rat brain (Ikonomidou et al., 1999). Besides, an epidemiological study finds that third-trimester exposure to wood smoke leads to long-term deficits in neuropsychological performance (Dix-Cooper et al., 2012).

[21] Source: https://www.who.int/publications/i/item/9789240034228.

5.2 Robustness checks

In this section, we conduct a set of tests to check the robustness of our main results. We also repeat the full set of robustness checks separately for female and male respondents, with the corresponding results reported in Appendix C. To begin with, we employ alternative IV construction methods. First, we exclude nearby cities within 100 km instead of 50 km from the focal cities during the construction of IV to further address concerns of endogeneity. Second, following Rangel and Vogl (2019), we restrict the angle between the vector $\vec{nj}$ and the prevailing wind direction of the focal city j ($\theta_{jn,ymd}$) to 60°and 45°, respectively. Third, following Chen et al. (2021), as higher wind speeds in focal cities would transport PM2.5 from nearby cities more quickly, we add wind speed to the weight function as shown in equation (12). The estimation results are displayed in Table A4. Overall, the results remain largely unchanged, which establishes the robustness of our estimates across different IV construction methods.

$$\omega_{jn,ymd} = \frac{\frac{\cos(\theta_{jn,ymd})}{distance_{jn}} \times \mathbb{1}(\cos\theta_{jn,ymd} > 0) \times windspeed_{n,ymd}}{\sum_{k:\ within\ 50-200\ km\ of\ j} \frac{\cos(\theta_{jk,ymd})}{distance_{jk}} \times \mathbb{1}(\cos\theta_{jk,ymd} > 0) \times windspeed_{k,ymd}} \tag{12}$$

Another potential source of bias in our estimated results is that air pollution may affect migratory choices (Chen et al., 2022). It is possible that individuals who have experienced pollution shocks may decide to relocate to other cities.[22] To alleviate this concern, we exclude migrants from our sample. We employ two methods to identify migrants: (1) individuals who did not reside in the same city at birth and when they were three years old, and (2) individuals who did not reside in the same city at birth and during the survey period. The results, presented in Table A5, indicate that migration, and thus location sorting, is unlikely

[22]Due to data limitations, it is impossible to track migration behavior *in utero*, which may introduce bias into the environment-birthplace matching process. However, this migration cohort should constitute only a small proportion of the whole sample. In the sampling period (1980-2000), intercity migration was rare due to the strict *hukou* policy (e.g., Meng, 2012; Fan, 2019). Moreover, medical insurance in China was tightly bound with *hukou*, adding more difficulty for pregnant women to relocate to other cities (Müller, 2016; Ngai et al., 2019).

to bias our estimates significantly.

In addition, we conduct a placebo test to further support our identification assumptions. We forward the respondent's date of birth by one year and match the PM2.5 exposure, and construct IVs based on this "false birth date." The preconception exposure to PM2.5 should not affect cognitive performance unless the identified effect is driven by unobserved confounding factors or trends. In line with our expectations, the findings presented in Figure A4 demonstrate that exposure to PM2.5 before pregnancy has no significant impact on test scores, and the magnitudes of the coefficients are also small.

It is also plausible that postnatal exposure to air pollution, which may be associated with fetal exposure, could affect long-term cognitive performance. To account for this possibility, we simultaneously control for PM2.5 exposure during pregnancy, as well as at ages 0-1, 1-2, and 2-3 years, along with placebo exposure before conception, as shown in Figure A4.[23] As expected, the magnitudes and significance of prenatal PM2.5 exposure remain largely unchanged compared to our baseline IV estimates. In contrast, the effects of postnatal exposure are relatively small and tend to diminish with age.

Moreover, previous studies have shown that transitory exposure to PM2.5 can affect cognitive performance (Ebenstein et al., 2016; Zhang et al., 2018b). To address this, we further control for PM2.5 levels on the interview date and over the month preceding the interview. As shown in columns (1)–(4) of Table A6, our results remain robust. Finally, we control for the average monthly PM2.5 exposure from birth to the interview date. The results, presented in columns (5)–(6), remain robust, with the estimated effects of prenatal PM2.5 exposure remaining comparable in magnitude to the baseline estimates.

Our baseline results are also robust to a wide variety of specification checks. First, we add birth year-by-month and birth city-by-year fixed effects to the regressions in columns (1) and (2) of Table A7. Our main findings remain robust. Second, columns (3) and (4) show that the main estimation results still hold after controlling for the birth year-by-month

[23] We instrument placebo and postnatal PM2.5 exposure using wind directions, following the same approach as for fetal air pollution exposure.

and city-by-month fixed effects. Furthermore, in columns (5) and (6), standard errors are two-way clustered at the birth city and birth year-season levels (Cameron et al., 2011), which allows for serial correlation within each birth city and correlation across all birth cities in the same birth year-season. The standard errors are slightly larger compared to Table 2, but the significance level remains. These alternative fixed effects and two-way clustering further confirm the robustness of our results.

Finally, we examine the robustness of our results to alternative specifications of demographic controls. One potential concern is that some demographic variables may be endogenous, potentially biasing the estimated effects of the variables of interest. For example, maternal age at childbirth could be influenced by selective fertility decisions in response to air quality concerns. Columns (1) and (2) of Table A8 demonstrate that our results remain robust when demographic controls are excluded from the regression. Furthermore, we assess the robustness of our findings regarding family size and the gender composition of the children. In Columns (3) and (4), we include the number of siblings, while Columns (5) and (6) add dummy variables for the number of male and female siblings as control variables. The significance and magnitude of the estimated coefficients remain largely unchanged across these specifications.

5.3 Heterogeneous effects

We first examine the heterogeneous effects of PM2.5 exposure by gender. As shown in Panel A of Table 3, the coefficients on the PM2.5 level are statistically significant for females taking both tests and males taking the math test. For the math test, both females and males are affected, but the effect on females is larger and more significant. Therefore, our results suggest that females are more vulnerable to fetal PM2.5 exposure compared with their male counterparts. Moreover, the gender difference is statistically significant at the 5% level for the verbal test and at the 10% level for the math test.

Our identified stronger effect for females compared with males seems to contradict the

prevailing "fragile boys" hypothesis, which reveals that male fetuses are more vulnerable to detrimental influences *in utero* than female fetuses. The "fragile males" hypothesis has been widely documented in the existing literature (e.g., Kraemer, 2000; Eriksson et al., 2010; Currie and Schwandt, 2016; Almond and Mazumder, 2011; Ebenstein et al., 2016; Chen, 2025). We will investigate potential channels to solve this puzzle in the next section.

Moreover, we also test the heterogeneous effect by age cohort and socioeconomic status (SES). As revealed in Panel B of Table 3, the negative effect of prenatal exposure to air pollution on cognitive performance becomes more prominent with age. However, the differences across age groups are not statistically significant. To determine the SES of survey respondents, we categorize them into two groups based on their mothers' education level. Specifically, respondents whose mothers completed less than 9 years of education are assigned to the low SES group, while those whose mothers completed more than 9 years are assigned to the high SES group. The results in Panel C of Table 3 show that respondents with low SES, a measure of resource constraint, are more vulnerable to prenatal PM2.5 exposure. Furthermore, the difference in effects between SES groups is statistically significant.

Finally, we conduct heterogeneity analyses by rural versus urban residence and pollution levels at birth, with the results reported in Table A9. Panel A shows that the adverse effect of prenatal exposure to air pollution on cognitive performance is more pronounced in rural areas; however, the differences across subgroups are not statistically significant. Panel B indicates that respondents born in more polluted areas are more strongly affected, and the subgroup differences are statistically significant for math test scores.[24]

5.4 Effects of PM2.5 on birth outcomes and child health

Before exploring the underlying mechanisms, this subsection provides evidence on the impact of prenatal exposure to air pollution on birth outcomes and child health, thereby offering suggestive evidence on early-life health endowments that may shape subsequent

[24]Polluted and less polluted areas are defined based on the median of each county's average air pollution level over the period 1989–2000.

parental investment decisions.

The birth record data are drawn from the Chinese National Disease Surveillance Points (DSP) system, which covers 145 counties across 31 provinces over the period 1991–2000.[25] We use the logarithm of birth weight and an indicator for low birth weight (LBW) as measures of birth outcomes.[26] The results are reported in Table A10. We find that male infants are more adversely affected than females at birth, consistent with the "fragile boys" hypothesis. Specifically, for males, a 1 $\mu g/m^3$ increase in prenatal PM2.5 exposure is associated with a 0.183% (6.14 grams) reduction in birth weight and a 0.272 percentage point increase in the probability of LBW. Female birth weight is also negatively affected by prenatal air pollution exposure, although the magnitude is smaller. The gender difference in the effects is statistically significant for the LBW outcome.

Furthermore, we examine the impact of *in utero* exposure to PM2.5 on child health among individuals younger than 16 years old. Specifically, we analyze its effects on height, weight, and two standardized measures: height-for-age and BMI-for-age. Height-for-age is defined as height standardized by age and sex, and BMI-for-age is defined as body mass index standardized by age and sex, both measured relative to the WHO Child Growth Reference and expressed as z-scores.[27] As shown in Table A11, prenatal exposure to air pollution significantly reduces height and height-for-age among boys, while having no statistically significant effects on weight or BMI-for-age for either boys or girls. Specifically, a 1 $\mu g/m^3$ increase in PM2.5 exposure is associated with a 0.5% reduction in boys' height and a 0.054 decline in height-for-age. Moreover, the effect on height exhibits a statistically significant gender difference.

Taken together, these results suggest that prenatal pollution exposure generates observable differences in early-life health endowments, which may serve as signals to parents and

[25] The DSP system was established by the Chinese Academy of Preventive Medicine, with vital registration conducted by the "Disease Prevention Unit" of township hospitals at each DSP site. The system employs a multistage cluster probability sampling design to ensure that the data are nationally representative, covering approximately 1% of the Chinese population (Yang et al., 2005).

[26] LBW refers to infants with a birth weight below 2,500 grams.

[27] Data source: https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators.

potentially shape subsequent investment decisions.[28]

6 Mechanisms

In this section, we test several potential mechanisms that may explain the more pronounced effect of fetal PM2.5 exposure on females than on males. The first channel belongs to the biological pathway, specifically mortality selection, while the others are associated with the socioeconomic pathway, including selection into motherhood, strategic fertility decisions, and gender-specific parental investment.

6.1 Mortality selection and sex-selective abortion

Negative shocks may increase fetal mortality directly through a biological effect, also known as the culling effect. Meanwhile, the scarring effect may prevail when *in utero* adverse environment is not strong enough to cause fetal death but impacts surviving newborns. If the culling effect dominates the scarring effect for males, weak male fetuses are more likely to be selected out by PM2.5 exposure. This *in utero* mortality selection thus leads the surviving boys to be inherently healthier and perform better in later life.

We first test this channel by examining the effect of prenatal exposure to air pollution on being a male using the CFPS data. If male fetuses are at a greater risk of mortality compared to female fetuses, we would identify a significant negative effect on the individuals being male. Results in column (1) of Table 4 demonstrate that fetal exposure to air pollution has no distinguishable effects on the gender of respondents, which largely rules out the significance of mortality selection.

Additionally, pollution exposure may affect the likelihood of sex-selective abortion. We

[28]The birth outcome and cognitive estimates come from different data sources and sample frames: the birth-outcome estimates use DSP birth records over 1991-2000, whereas the cognitive estimates use CFPS respondents born between 1980 and 2000. We therefore treat cross-sample comparisons as *suggestive* evidence rather than as identification, and use them as motivation for the investment-response interpretation developed below.

calculate the sex ratio for individuals born between 1980 and 2010 within each birth-city-year-month cell using the 2000 and 2010 Population Census Data of China. As presented in column (2) of Table 4, the findings indicate that fetal exposure to PM2.5 has minimal effects on the sex ratio, thereby ruling out the possibility of mortality selection or sex-selection abortion.

6.2 Selection into motherhood and fertility decisions

Mothers who are aware of the risks associated with air pollution may choose to avoid exposure or plan their pregnancies during periods of better air quality. This awareness is likely correlated with the mother's education level and knowledge of air pollution. Additionally, exposure to air pollution can influence a couple's decision regarding subsequent pregnancies or the spacing between pregnancies, which may be associated with the gender of the previous child. If this happens, our findings could be susceptible to bias stemming from these selection processes and strategic pregnancy decisions. We test these potential channels by examining whether air pollution is associated with mothers' education level, age at birth, as well as birth spacing and the decision to have another child, by the gender of the previous child. Relevant results are shown in columns (3)-(8) of Table 4. We observe no significant effect of air pollution on any of these channels, suggesting that our results are not influenced by selection into motherhood or strategic fertility decisions.

6.3 Gender-specific parental investment

In this section, we test the pathway related to gender-specific parental investment. We first examine heterogeneous effects based on the number of siblings and gender composition. Table 5 reports the IV results. Panel A refers to the females, while Panel B corresponds to the males. For each outcome, we study the effect of PM2.5 for five subsamples, namely, only children, having siblings, having only brothers, having brothers & sisters, and having only sisters. The IV results suggest that females with siblings are more vulnerable to the

detrimental effects of PM2.5 exposure on both verbal and math abilities, particularly when they have only brothers. In contrast, for males, the adverse effects are not statistically significant for either verbal or math test scores, and the gender differences are statistically significant.[29]

For the female subsample, we further test the statistical significance of differences between only children and those with siblings, as well as between individuals with only male siblings and those with female siblings. As reported in Table 5, females with siblings are more adversely affected than only children, and females with only brothers are more adversely affected than those with sisters; both differences are statistically significant. Specifically, for a girl with siblings, an additional SD increase in fetal PM2.5 exposure (6.750 $\mu g/m^3$), on average, reduces her verbal and math test scores by 1.816 (0.239 SD) and 1.991 (0.346 SD), respectively.[30] The adverse effects are even larger for girls with only brothers, for whom the corresponding reductions are 3.186 (0.419 SD) in verbal test scores and 2.747 (0.478 SD) in math test scores.

A possible explanation could be the prevailing son preference in China (Wang, 2005; Lei et al., 2017). Deeply rooted in convention, sons are expected to inherit family names, provide economic support, and care for elderly parents. Daughters, on the other hand, are expected to marry out and join other families (e.g., Wei and Zhang, 2011). In credit-constrained families with multiple children, parents tend to allocate more resources to their sons to compensate for the damage caused by prenatal PM2.5 exposure.[31] Girls with fewer resources within the household and limited access to education, therefore, develop worse

[29] A potential concern is that China's "1.5-child policy," which permitted a second birth if the first child was a girl in some rural areas, may have affected sibling composition. However, the policy varied substantially across regions and over time (Gu et al., 2007), and imperfect compliance further weakened its influence on realized fertility outcomes. Excluding communities where the policy was in effect produces qualitatively similar baseline and heterogeneity results. These results are available upon request.

[30] Note the SD of verbal and math test scores for the female subsample are 7.597 and 5.753, respectively.

[31] The US Embassy began releasing the PM2.5 concentration levels in Beijing via Twitter in March 2008. Government and public awareness of PM2.5 significantly increased around 2013, following several severe haze and pollution events (Liu et al., 2022). Thus, it is unlikely that our respondents and their parents were aware of the health consequences of particulate matter during our sample period (1980-2010). The health consequences are more likely to have been inferred by parents through observing their children's health and intellectual development outcomes.

cognitive abilities than boys when exposed to negative shocks *in utero*.

To verify this assumption, we further test two parental human capital investment channels that could account for the observed phenomenon, namely, educational investment and health investment, by gender.[32] The first is that greater fetal exposure to PM2.5 leads females to receive less educational investment from their parents, such as spending fewer years in school or receiving less parental educational care, resulting in poorer cognitive performance. The second channel is that families may make health investments differently in their offspring based on gender preference.

6.3.1 Educational investment

In this section, we examine the impact of fetal PM2.5 exposure on educational investment, as reflected in educational attainment and parental time and financial inputs into education. We begin by estimating the effect of fetal PM2.5 exposure on years of schooling separately by gender, restricting the sample to individuals who have completed their education. As illustrated in columns (1)-(2) of Table 6, the coefficients on prenatal PM2.5 exposure are negative and statistically significant for females, but insignificant for males, with the gender difference being statistically significant. Specifically, a one SD increase in PM2.5 exposure (6.750 $\mu g/m^3$) during pregnancy results in a reduction of 1.060 years of education later in life for females.

For the full sample, we further introduce a dummy variable to denote age-appropriate grade completion.[33] As displayed in columns (3)-(4) of Table 6, the adverse effect is again more pronounced among females and statistically significant. Specifically, a one SD increase in PM2.5 exposure during pregnancy reduces the probability of age-appropriate grade completion by 6.08 percentage points for females, *ceteris paribus*.

[32] The CFPS collects data on these human capital investments only for children under the age of 16.

[33] For individuals aged between 10 and 20 years old, we define age-appropriate grade completion as the difference between age and years of education being less than 8 (implying that children aged 8 must have completed at least the first grade). For those aged 20 years or older, we define age-appropriate grade completion as having completed high school (12th grade).

We further test the heterogeneity by sibling gender composition for years of education and age-appropriate grade completion. The results are presented in Tables A12 and A13, respectively. The estimated patterns are broadly consistent with those observed for cognitive test scores in Table 5. Specifically, the adverse effects of air pollution exposure are more pronounced among females with siblings, with those having brothers experiencing the greatest impact.[34] In both Tables A12 and A13, females with only brothers are more adversely affected than those with sisters, and these differences are statistically significant. Quantitatively, for a girl with brothers, a one SD increase in prenatal PM2.5 exposure, on average, reduces her years of education by 2.369 years and decreases the probability of age-appropriate grade completion by 14.18 percentage points. These findings provide further evidence of the gendered consequences of fetal exposure to PM2.5 for educational outcomes.

Moreover, we examine the gender difference in how parental time and pecuniary investment in education respond to prenatal pollution shock. Specifically, we investigate the impact of fetal exposure to air pollution on family education spending and the time family members spend helping children with homework (measured in hours per week) across both male and female subsamples. The results are displayed in columns (5)-(8) of Table 6. Although we do not observe any gender difference in family education spending, we observe a noteworthy reduction in the time parents devote to helping their child with homework by 17.6% for females per one SD increase in PM2.5 exposure (6.750 $\mu g/m^3$).[35] We also compare the effects of prenatal exposure to air pollution on the time family members spend helping

[34] In Table A12, although the estimated coefficient for only children is larger than that for individuals with siblings, the difference is not statistically significant. Moreover, for the male subsample, the estimated coefficients are statistically insignificant and accompanied by relatively large standard errors. As a result, the point estimates are sensitive to sampling variation and should be interpreted with caution.

[35] Although we do not find statistically meaningful evidence of gender-biased pecuniary investment as measured by education spending, time investment is also an essential and equally important investment because time invested in tutoring could be at the expense of leisure and labor hours, which may translate into monetary terms (Yi et al., 2015). It is also important to note that a significant share of our sample comes from rural areas where the private tutoring market is nonexistent, particularly during the sample period 1980-2000 (most research that studies the Chinese private tutoring market focuses on the post-2010 period (e.g., Zhang, 2013; Liu and Bray, 2017; Huang et al., 2025; Meng et al., 2025)). This limited participation may have contributed to the absence of a significant difference in family education spending by gender, particularly given similar expenditures on compulsory schooling during that period.

children with homework for males and females, and find that the difference is statistically significant. Given the existing gender gap in parental investment (measured by time spent helping children with homework), the pollution shock of a 1 $\mu g/m^3$ increase in PM2.5 further exacerbates this gap by 17.03%.[36]

We further look at the heterogeneous effects on time spent helping children with homework by sibling gender composition. Table A14 displays a pattern similar to that in the educational attainment subsection discussed above. The IV results indicate that the adverse effects of prenatal exposure to PM2.5 on family assistance time with homework are only statistically significant for females, especially those with siblings.[37] In response to fetal exposure to air pollution, parents may allocate more resources to their sons as compensation. Girls, in turn, could receive less investment when they have brothers.[38]

Additionally, as a complement to time investments in education, we further explore parental monitoring behaviors that reflect attention to children's academic development. The CFPS data provide information on the self-reported frequency with which parents monitor their child's study. Specifically, we examine the impact of air pollution on two measures: how often parents ask their child to complete homework and how frequently they check the child's homework.[39] As reported in Table A15, higher levels of pollution exposure are associated

[36]This number is calculated following these steps: First, Table 6 shows that the gender gap (male minus female) in time spent helping children with homework is 0.364 hours. Second, Table 6 also reveals that the gender gap in the impact of prenatal exposure to air pollution represented in level is 0.062 hours (for female, the marginal effect in level is approximately -2.6% $\times$ 1.869 (the dependent variable mean) = -0.049 hours; for male, the marginal effect in level is approximately 0.6% $\times$ 2.233 = 0.013 hours). Third, by taking the ratio of the gender gap in marginal effect in step 2 and the gender gap in parental investment in step 1, we calculate that the pollution shock of a 1 $\mu g/m^3$ increase in PM2.5 exacerbates the gender gap in parental investment by 17.03% (0.062/0.364=0.1703).

[37]In Table A14, the estimated effect for only children is not statistically different from that for individuals with siblings, and although the coefficient in column (5) (sisters) is larger than that in column (4) (brothers), the difference is not statistically significant. This lack of statistical significance is likely driven by a limited sample size, as the analysis is restricted to children younger than 16 due to data availability constraints.

[38]Our results suggest that the observed gender gap in parental investment is more consistent with a general preference for sons rather than a reinforcing response to pollution-induced damage. As shown in Tables 5 and A12–A14, parents allocate more educational resources to boys in families with both sons and daughters, whereas we do not observe systematic differences across boys in families with only sons.

[39]The frequency is categorized into five levels: 5 = "very often" (6–7 times per week), 4 = "often" (2–3 times per week), 3 = "sometimes" (1–2 times per week), 2 = "rarely" (once per month), and 1 = "never".

with a decline in parental attention to girls' academic activities, which is consistent with the patterns observed for time investments discussed above. However, the estimated gender differences are not statistically significant.

6.3.2 Health investment

Since health and educational attainment are both key dimensions of individual capacity, disparities in early-life health investment within households may also contribute to gender differences in cognitive abilities later in life. Intriguingly, parental responses could differ across these two dimensions. For example, parents might prefer to compensate in health investment but reinforce in educational investment (Yi et al., 2015). In Table 7, we further examine the impact of fetal exposure to air pollution on health-related investments. Specifically, we analyze its effects on breastfeeding duration, medical expenses, and health insurance coverage for both male and female subgroups. The estimated gender differences are not statistically significant.[40] Overall, these results provide little evidence in support of this health investment channel.[41]

[40] The lack of gender differences in breastfeeding is not inconsistent with our educational investment results. The breastfeeding estimates are statistically indistinguishable from zero for both girls and boys, and their difference is also insignificant and economically small. Moreover, parental gender bias may vary across investment domains: parents may adopt different strategies when allocating educational versus health investments (Yi et al., 2015). Additionally, any true breastfeeding differences may also be obscured by recall error, and, in the 1990s Chinese context, limited awareness of health benefits of breastfeeding likely made it a less salient margin of intentional gendered investment.

[41] Our results do not contradict the literature documenting gender-biased parental health investments (e.g., Oster, 2009; Barcellos et al., 2014; Jayachandran and Kuziemko, 2011). Those studies primarily examine gender differences in the level of postnatal health investments, whereas our Table 7 estimates capture whether the effect of prenatal pollution exposure on such investments differs by child gender. These are distinct objects. In fact, the (unconditional) dependent variable means in Table 7 are consistent with the prior literature: boys in our sample receive longer breastfeeding duration than girls, and similar patterns hold for other health investments. Thus, the absence of a significant gender difference in the marginal response to prenatal pollution does not imply the absence of an underlying gender gap in health investments.

7 Conclusion

In this paper, we examine the long-term impact of prenatal exposure to PM2.5 on cognitive performance measured by standardized verbal and math test scores, using a nationally representative longitudinal survey in China. To address potential endogeneity in PM2.5, we construct IVs using the local wind direction and the pollution level of nearby upwind cities. Our IV estimates suggest that a one SD increase in fetal PM2.5 exposure leads to a decline in verbal and math test scores by 0.122 SD and 0.170 SD, respectively. We identify strong heterogeneity by gender. Females are more affected than males, which seemingly contradicts the "fragile males" hypothesis documented in the literature. It should be noted that, because these estimates reflect average pollution exposure at the city level, individual-level effects for certain demographic groups (e.g., girls) may be more pronounced. We examine the effect of the siblings' gender composition and find that females with siblings are more vulnerable to the detrimental effects of PM2.5 exposure on both verbal and math abilities, especially when they have brothers. Specifically, for a girl with only brothers in the family, a one SD increase in prenatal PM2.5 exposure is associated with a reduction in verbal and math test scores by 0.419 SD and 0.478 SD, respectively.

To explain this gendered pattern, we examine several competing hypotheses, including mortality selection, selection into motherhood, and fertility decisions, and find evidence consistent with gender-biased educational investments within families. Our findings suggest that gender-biased educational investments tend to amplify the adverse effects of early-life shocks on cognitive abilities for girls. Specifically, conditional on the same level of prenatal PM2.5 exposure, females receive less assistance with homework from family members and achieve lower levels of education compared to their male counterparts. Our back-of-the-envelope calculation shows that a 1 $\mu g/m^3$ increase in PM2.5 pollution exacerbates the gender gap in parental educational investment (measured by time spent helping children with homework) by 17.03%. This impact is particularly pronounced for females who have siblings, especially those with brothers.

In evaluating the long-term consequences of prenatal air pollution exposure, our study provides evidence on cognitive performance that may also be relevant to other developing countries. Our findings further suggest that parental investments may respond differently to adverse *in utero* shocks between boys and girls. Such gender-differentiated responses may amplify the long-term disadvantages associated with early-life environmental shocks and widen gender disparities in human capital outcomes. This finding highlights the importance of policies that both reduce prenatal exposure to air pollution and protect disadvantaged children from unequal postnatal investments, including maternal healthcare support, targeted educational assistance, and the effective enforcement of compulsory education. As urbanization continues, fertility policies evolve, and fertility rates decline, the gender-specific consequences of early-life environmental shocks and intra-household resource allocation warrant continued attention.

References

Adhvaryu, A., J. Fenske, and A. Nyshadham (2019): "Early life circumstance and adult mental health," *Journal of Political Economy*, 127, 1516–1549.

Adhvaryu, A. and A. Nyshadham (2016): "Endowments at birth and parents' investments in children," *The Economic Journal*, 126, 781–820.

Almond, D., L. Edlund, and M. Palme (2009): "Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden," *The Quarterly journal of economics*, 124, 1729–1772.

Almond, D. and B. Mazumder (2011): "Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy," *American Economic Journal: Applied Economics*, 3, 56–85.

——— (2013): "Fetal origins and parental responses," *Annu. Rev. Econ.*, 5, 37–56.

Baker, M., J. Gruber, and K. Milligan (2019): "The long-run impacts of a universal child care program," *American Economic Journal: Economic Policy*, 11, 1–26.

Barcellos, S. H., L. S. Carvalho, and A. Lleras-Muney (2014): "Child gender and parental investments in India: Are boys and girls treated differently?" *American Economic Journal: Applied Economics*, 6, 157–189.

Barwick, P. J., S. Li, D. Rao, and N. B. Zahur (2021): "The Healthcare Cost of Air Pollution: Evidence from the World's Largest Payment Network," *NBER Working Paper No. 24688*.

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

Behrman, J. R. (1988): "Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion?" *Oxford Economic Papers*, 40, 32–54.

Behrman, J. R. and A. B. Deolalikar (1990): "The intrahousehold demand for nutrients in rural south India: Individual estimates, fixed effects, and permanent income," *Journal of Human Resources*, 665–696.

Bharadwaj, P., M. Gibson, J. G. Zivin, and C. Neilson (2017): "Gray matters: Fetal pollution exposure and human capital formation," *Journal of the Association of Environmental and Resource Economists*, 4, 505–542.

Braido, L. H., P. Olinto, and H. Perrone (2012): "Gender bias in intrahousehold allocation: Evidence from an unintentional experiment," *Review of Economics and Statistics*, 94, 552–565.

Buchard, V., A. Da Silva, C. Randles, P. Colarco, R. Ferrare, J. Hair, C. Hostetler, J. Tackett, and D. Winker (2016): "Evaluation of the surface PM2.5 in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States," *Atmospheric Environment*, 125, 100–111.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011): "Robust inference with multiway clustering," *Journal of Business & Economic Statistics*, 29, 238–249.

Cameron, L. A. and C. Worswick (2001): "Education expenditure responses to crop loss in Indonesia: A gender bias," *Economic development and cultural change*, 49, 351–363.

Chay, K. Y. and M. Greenstone (2003): "The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession," *The Quarterly Journal of Economics*, 118, 1121–1167.

Chen, P. (2025): "Industrialization and pollution: The long-term Impact of early-life exposure on human capital formation," *Journal of Public Economics*, 241, 105270.

Chen, S., Y. Chen, Z. Lei, and J.-S. Tan-Soo (2021): "Chasing clean air: Pollution-

induced travels in China," *Journal of the Association of Environmental and Resource Economists*, 8, 59–89.

Chen, S., P. Oliva, and P. Zhang (2022): "The effect of air pollution on migration: evidence from China," *Journal of Development Economics*, 156, 102833.

Chen, Z., D. Chen, C. Zhao, M.-p. Kwan, J. Cai, Y. Zhuang, B. Zhao, X. Wang, B. Chen, J. Yang, et al. (2020): "Influence of meteorological conditions on PM2.5 concentrations across China: A review of methodology and mechanism," *Environment International*, 139, 105558.

Coleman, F. M., A. U. Ahmed, S. Roy, and J. Hoddinott (2025): "Understanding social safety nets and intra-household food allocation: Experimental evidence from Bangladesh," *Journal of Development Economics*, 103585.

Cook, N., A. Heyes, and N. Rivers (2023): "Clean Air and Cognitive Productivity: Effect and Adaptation," *Journal of the Association of Environmental and Resource Economists*, 10, 1265–1308.

Currie, J. and M. Neidell (2005): "Air pollution and infant health: what can we learn from California's recent experience?" *The Quarterly Journal of Economics*, 120, 1003–1030.

Currie, J. and H. Schwandt (2016): "The 9/11 dust cloud and pregnancy outcomes: A reconsideration," *Journal of Human Resources*, 51, 805–831.

Currie, J., J. G. Zivin, J. Mullins, and M. Neidell (2014): "What do we know about short-and long-term effects of early-life exposure to pollution?" *Annu. Rev. Resour. Econ.*, 6, 217–247.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor, and J. Reif (2019): "The

mortality and medical costs of air pollution: Evidence from changes in wind direction," *American Economic Review*, 109, 4178–4219.

Deschenes, O., H. Wang, S. Wang, and P. Zhang (2020): "The effect of air pollution on body weight and obesity: evidence from China," *Journal of Development Economics*, 145, 102461.

Dix-Cooper, L., B. Eskenazi, C. Romero, J. Balmes, and K. R. Smith (2012): "Neurodevelopmental performance among school age children in rural Guatemala is associated with prenatal and postnatal exposure to carbon monoxide, a marker for exposure to woodsmoke," *Neurotoxicology*, 33, 246–254.

Dizon-Ross, R. (2019): "Parents' beliefs about their children's academic ability: Implications for educational investments," *American Economic Review*, 109, 2728–2765.

——— (2021): "Using randomized information shocks to understand how parents' investments depend on their children's ability'," *Unpublished manuscript*.

Duque, V. and M. Gilraine (2022): "Coal use, air pollution, and student performance," *Journal of Public Economics*, 213, 104712.

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

Eriksson, J. G., E. Kajantie, C. Osmond, K. Thornburg, and D. J. Barker (2010): "Boys live dangerously in the womb," *American Journal of Human Biology*, 22, 330–335.

Fan, J. (2019): "Internal geography, labor mobility, and the distributional impacts of trade," *American Economic Journal: Macroeconomics*, 11, 252–288.

Fan, W. and C. Porter (2020): "Reinforcement or compensation? Parental responses to children's revealed human capital levels," *Journal of Population Economics*, 33, 233–270.

Ferro, S., A. Palma, C. Serra, and M. Stafoggia (2024): "Beyond birth: The medium-term health impact of prenatal exposure to air pollution," *Journal of Environmental Economics and Management*, 127, 103009.

Field, E., O. Robles, and M. Torero (2009): "Iodine deficiency and schooling attainment in Tanzania," *American Economic Journal: Applied Economics*, 1, 140–169.

Fu, S., V. B. Viard, and P. Zhang (2021): "Air pollution and manufacturing firm productivity: Nationwide estimates for China," *The Economic Journal*, 131, 3241–3273.

——— (2022): "Trans-boundary air pollution spillovers: Physical transport and economic costs by distance," *Journal of Development Economics*, 155, 102808.

Giannola, M. (2024): "Parental investments and intra-household inequality in child human capital: Evidence from a survey experiment," *The Economic Journal*, 134, 671–727.

Greenstone, M. and R. Hanna (2014): "Environmental regulations, air and water pollution, and infant mortality in India," *American Economic Review*, 104, 3038–3072.

Gu, B., F. Wang, Z. Guo, and E. Zhang (2007): "China's local and national fertility policies at the end of the twentieth century," *Population and development review*, 129–147.

Ham, J. C., J. S. Zweig, and E. Avol (2014): "Pollution, test scores and distribution of academic achievement: Evidence from California schools 2002–2008," *Manuscript, University of Maryland*.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016): "Long-run impacts of childhood access to the safety net," *American Economic Review*, 106, 903–934.

Huang, Z., Y. Liu, M. Ma, and L. Y. Yang (2025): "Biting the hand that teaches: Unraveling the economic impact of banning private tutoring in China," *Journal of Comparative Economics*.

Ikonomidou, C., F. Bosch, M. Miksa, P. Bittigau, J. Vockler, K. Dikranian, T. I. Tenkova, V. Stefovska, L. Turski, and J. W. Olney (1999): "Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain," *Science*, 283, 70–74.

Institute for Health Metrics and Evaluation (2024): "Global Burden of Disease Study 2021 (GBD 2021) Air Pollution Exposure Estimates and Risk Curves 1990–2021," Data set. Accessed 5 May 2026.

Jayachandran, S. and I. Kuziemko (2011): "Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India," *The Quarterly journal of economics*, 126, 1485–1538.

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

Kraemer, S. (2000): "The fragile male," *Bmj*, 321, 1609–1612.

Krebs, B. and S. Luechinger (2024): "Air pollution, cognitive performance, and the role of task proficiency," *Journal of the Association of Environmental and Resource Economists*, 11, 921–958.

Lai, W., S. Li, Y. Li, and X. Tian (2022): "Air pollution and cognitive functions: evidence from straw burning in China," *American Journal of Agricultural Economics*, 104, 190–208.

Lei, X., Y. Shen, J. P. Smith, and G. Zhou (2017): "Sibling gender composition's effect on education: Evidence from China," *Journal of Population Economics*, 30, 569–590.

Li, H., M. Rosenzweig, and J. Zhang (2010): "Altruism, favoritism, and guilt in the allocation of family resources: Sophie's choice in Mao's mass send-down movement," *Journal of Political Economy*, 118, 1–38.

Lin, C., Y. Sun, and C. Xing (2021): "Son Preference and Human Capital Investment among China's Rural-urban Migrant Households," *The Journal of Development Studies*, 57, 2077–2094.

Liu, H., J. Liu, M. Li, P. Gou, and Y. Cheng (2022): "Assessing the evolution of PM2.5 and related health impacts resulting from air quality policies in China," *Environmental Impact Assessment Review*, 93, 106727.

Liu, J. and M. Bray (2017): "Determinants of demand for private supplementary tutoring in China: Findings from a national survey," *Education Economics*, 25, 205–218.

Lleras-Muney, A. (2010): "The needs of the army using compulsory relocation in the military to estimate the effect of air pollutants on children's health," *Journal of Human Resources*, 45, 549–590.

Long, Y., J. Wang, K. Wu, and J. Zhang (2018): "Population exposure to ambient PM2.5 at the subdistrict level in China," *International Journal of Environmental Research and Public Health*, 15, 2683.

Lundborg, P., D.-O. Rooth, and J. Alex-Petersen (2022): "Long-term effects of childhood nutrition: evidence from a school lunch reform," *The Review of Economic Studies*, 89, 876–908.

Maccini, S. and D. Yang (2009): "Under the weather: Health, schooling, and economic consequences of early-life rainfall," *American Economic Review*, 99, 1006–1026.

Meng, J., H. Wang, Y. Yang, and M. Zhang (2025): "Education competition and fertility intention: Evidence from China's private tutoring ban," *The Economic Journal*, ueaf096.

Meng, X. (2012): "Impact of economic reform on labor market outcomes in China," *Journal of Economics Perspectives*, 26, 75–102.

Molina, T. (2021): "Pollution, ability, and gender-specific investment responses to shocks," *Journal of the European Economic Association*, 19, 580–619.

Müller, A. (2016): "Hukou and health insurance coverage for migrant workers," *Journal of Current Chinese Affairs*, 45, 53–82.

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

Ngai, L. R., C. A. Pissarides, and J. Wang (2019): "China's mobility barriers and employment allocations," *Journal of the European Economic Association*, 17, 1617–1653.

Oster, E. (2009): "Does increased access increase equality? Gender and child health investments in India," *Journal of development Economics*, 89, 62–76.

Qin, X., J. Yi, and H. Zhang (2025): "Quarter of Birth, Gender Inequality, and Economic Development," *Journal of Labor Economics*, forthcoming.

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

Restrepo, B. J. (2016): "Parental investment responses to a low birth weight outcome: who compensates and who reinforces?" *Journal of Population Economics*, 29, 969–989.

Rosales-Rueda, M. and M. Triyana (2019): "The persistent effects of early-life exposure to air pollution evidence from the indonesian forest fires," *Journal of Human Resources*, 54, 1037–1080.

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

Sanz-de Galdeano, A. and A. Terskaya (2022): "Sibling Differences in Genetic Propensity for Education: How do Parents React?" *The Review of Economics and Statistics*, 1–44.

Schlenker, W. and W. R. Walker (2016): "Airports, air pollution, and contemporaneous health," *The Review of Economic Studies*, 83, 768–809.

Tanaka, S. (2015): "Environmental regulations on air pollution in China and their impact on infant mortality," *Journal of Health Economics*, 42, 90–103.

Thomas, D., K. Beegle, E. Frankenberg, B. Sikoki, J. Strauss, and G. Teruel (2004): "Education in a Crisis," *Journal of Development Economics*, 74, 53–85.

Wang, C., Q. Lin, and Y. Qiu (2022): "Productivity loss amid invisible pollution," *Journal of Environmental Economics and Management*, 112, 102638.

Wang, W. (2005): "Son preference and educational opportunities of children in China—"I wish you were a boy!"," *Gender Issues*, 22, 3–30.

Wei, S.-J. and X. Zhang (2011): "The competitive saving motive: Evidence from rising sex ratios and savings rates in China," *Journal of political Economy*, 119, 511–564.

Woody, R. C. and M. A. Brewster (1990): "Telencephalic dysgenesis associated with presumptive maternal carbon monoxide intoxication in the first trimester of pregnancy," *Journal of Toxicology: Clinical Toxicology*, 28, 467–475.

World Health Organization (2024a): "Ambient (outdoor) air pollution and health," Fact sheet, accessed: 2025-08-10.

——— (2024b): "WHO Ambient Air Quality Database (Update Jan 2024)," Technical document (Excel version), Update 2024 (minor), 6th edition (V6.1). Geneva: World Health Organization.

Wu, J., J. Lin, and X. Han (2023): "Compensation for girls in early childhood and its long-run impact: family investment strategies under rainfall shocks," *Journal of Population Economics*, 36, 1225–1268.

Xie, Y. and J. Hu (2014): "An introduction to the China family panel studies (CFPS)," *Chinese Sociological Review*, 47, 3–29.

Yang, G., J. Hu, K. Q. Rao, J. Ma, C. Rao, and A. D. Lopez (2005): "Mortality registration and surveillance in China: history, current situation and challenges," *Population health metrics*, 3, 3.

Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2015): "Early health shocks, intra-household resource allocation and child outcomes," *The Economic Journal*, 125, F347–F371.

Zhang, B., L. Jiao, G. Xu, S. Zhao, X. Tang, Y. Zhou, and C. Gong (2018a): "Influences of wind and precipitation on different-sized particulate matter concentrations (PM2.5, PM10, PM2.5–10)," *Meteorology and Atmospheric Physics*, 130, 383–392.

Zhang, X., X. Chen, and X. Zhang (2018b): "The impact of exposure to air pollution on cognitive performance," *Proceedings of the National Academy of Sciences*, 115, 9193–9197.

Zhang, Y. (2013): "Does private tutoring improve students' National College Entrance Exam performance?—A case study from Jinan, China," *Economics of Education Review*, 32, 1–28.

Zheng, S., X. Zhang, W. Sun, and C. Lin (2019): "Air pollution and elite college

graduates' job location choice: Evidence from China," *The Annals of Regional Science*, 63, 295–316.

Table 1: Summary statistics

VARIABLES	**All** ($N = 9,038$)		**Female** ($N = 4,575$)		**Male** ($N = 4,463$)	
	mean	SD	mean	SD	mean	SD
Verbal test (point)	23.749	7.527	24.021	7.597	23.470	7.445
Math test (point)	13.692	5.677	13.518	5.753	13.870	5.592
PM2.5 level ($\mu g/m^3$)	22.444	6.719	22.461	6.750	22.427	6.687
Gender	0.494	0.500	0.000	0.000	1.000	0.000
Maternal age	25.590	4.438	25.504	4.395	25.677	4.480
Mother's education years	5.514	4.585	5.594	4.560	5.432	4.610
Number of days in temperature bins						
<-4 °C	17.876	30.727	18.188	30.995	17.557	30.450
-4-0 °C	14.241	15.600	14.697	15.823	13.774	15.355
0-4 °C	21.199	16.747	21.458	16.778	20.933	16.712
4-8 °C	25.797	16.362	25.788	16.113	25.806	16.615
8-12 °C	27.740	13.941	27.775	13.645	27.703	14.240
12-16 °C	32.574	14.340	32.576	14.530	32.573	14.143
16-20 °C	38.635	17.089	38.310	16.797	38.967	17.379
20-24 °C	42.817	21.209	42.547	21.699	43.093	20.694
24-28 °C	38.258	30.017	37.852	30.050	38.674	29.980
>28 °C	14.818	19.871	14.768	19.833	14.870	19.912
Precipitation (mm)	2.620	1.615	2.605	1.624	2.635	1.606
Wind speed (m/s)	2.300	0.762	2.296	0.764	2.304	0.760
Relative humidity (%)	70.543	8.715	70.400	8.835	70.690	8.588
Sunshine duration (hour)	5.715	1.293	5.725	1.300	5.704	1.286

Table 2: Baseline results

VARIABLES	OLS		first stage	2SLS	
	verbal test	math test	PM2.5 level	verbal test	math test
	(1)	(2)	(3)	(4)	(5)
PM2.5 level	-0.121***	-0.129***		-0.137**	-0.144***
	(0.046)	(0.039)		(0.059)	(0.051)
PM2.5 IV			0.131***		
			(0.005)		
Observations	9,038	9,038	9,038	9,038	9,038
KP first-stage F-statistic				720.6	720.6
Mean of Dep. Var.	23.75	13.69	22.44	23.75	13.69
Demographic controls	Y	Y	Y	Y	Y
Weather controls	Y	Y	Y	Y	Y
Birth city FE	Y	Y	Y	Y	Y
Birth year and month FE	Y	Y	Y	Y	Y

Notes: Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table 3: Heterogeneous effects by gender, age and SES (2SLS)

Panel A. *Gender*				
VARIABLES	verbal test		math test	
	female	male	female	male
	(1)	(2)	(3)	(4)
PM2.5 level	-0.212***	-0.037	-0.190***	-0.090*
	(0.078)	(0.082)	(0.070)	(0.054)
Observations	4,575	4,463	4,575	4,463
KP first-stage F-statistic	553.8	724.6	553.8	724.6
Mean of Dep. Var.	24.02	23.47	13.52	13.87
Subgroup difference	-0.175**		-0.100*	
	[0.040]		[0.100]	
Panel B. *Age*				
VARIABLES	verbal test		math test	
	age10-19	age20-30	age10-19	age20-30
	(1)	(2)	(3)	(4)
PM2.5 level	-0.091	-0.216*	-0.096*	-0.184**
	(0.067)	(0.124)	(0.056)	(0.092)
Observations	4,809	4,229	4,809	4,229
KP first-stage F-statistic	440.2	767.3	440.2	767.3
Mean of Dep. Var.	23.49	24.05	12.95	14.54
Subgroup difference	0.125		0.088	
	[0.130]		[0.110]	
Panel C. *SES*				
VARIABLES	verbal test		math test	
	low SES	high SES	low SES	high SES
	(1)	(2)	(3)	(4)
PM2.5 level	-0.202**	-0.042	-0.263***	-0.004
	(0.086)	(0.052)	(0.061)	(0.042)
Observations	5,501	3,537	5,501	3,537
KP first-stage F-statistic	596.1	826.3	596.1	826.3
Mean of Dep. Var.	22.29	26.02	12.55	15.47
Subgroup difference	-0.160*		-0.259***	
	[0.060]		[0.000]	

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the subgroup difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. SES=socioeconomic status. Dep. Var.=Dependent variable.

Table 4: Mechanism test: mortality selection, selection into motherhood and fertility decisions (2SLS)

VARIABLES	mortality selection		selection into motherhood		fertility decisions			
	CFPS male dummy	population census sex ratio (%)	mother's education years	mother's age at birth	birth spacing		having another child	
					female	male	female	male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5 level	-0.004	0.246	0.017	0.009	-0.005	-0.003	-0.003	-0.005
	(0.004)	(0.238)	(0.031)	(0.031)	(0.056)	(0.074)	(0.005)	(0.005)
Observations	9,038	107,538	9,038	9,038	2,165	1,434	4,575	4,463
KP first-stage F-statistic	720.9	712.1	719.8	720.4	805.5	422.5	553.8	724.6
Mean of Dep. Var.	0.494	129.2	5.514	25.59	4.278	4.477	0.473	0.321
Gender difference					-0.003		0.002	
					[0.500]		[0.400]	

Notes: Column (1) is regressed at the individual level using the CFPS data. Column (2) is regressed at the city-year-month level based on the 2000 and 2010 Population Census Data of China. All regressions include weather controls, birth city fixed effects, and birth year and month fixed effects. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. For demographic controls, gender is included in columns (3)-(4); maternal age is included in columns (1), (3), and (5)-(8); mother's education years is included in columns (1), (4), and (5)-(8). Robust standard errors, clustered at the city level, are presented in parentheses. P-values for the subgroup difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. CFPS=China Family Panel Studies. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table 5: Mechanism test: heterogeneous effect based on the number of siblings and their gender composition (2SLS)

VARIABLES	verbal test					math test				
	only children	having siblings	having only brothers	having brothers & sisters	having only sisters	only child	having siblings	having only brothers	having brothers & sisters	having only sisters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
				Panel A. *Female*						
PM2.5 level	-0.105	-0.269**	-0.472***	-0.115	-0.210	0.002	-0.295***	-0.407***	-0.156	-0.317**
	(0.093)	(0.105)	(0.138)	(0.243)	(0.201)	(0.078)	(0.080)	(0.097)	(0.172)	(0.149)
Observations	1,034	3,541	1,583	1,071	887	1,034	3,541	1,583	1,071	887
KP F-statistic	418.4	643.5	545.3	333.4	456.1	418.4	643.5	545.3	333.4	456.1
Mean of Dep. Var.	26.44	23.31	23.55	22.16	24.28	15.41	12.96	13.02	12.30	13.66
Subgroup difference	(2)-(1)=-0.164		(3)-(4)&(5)=-0.313*			(7)-(6)=-0.296**		(8)-(9)&(10)=-0.169*		
	[0.120]		[0.060]			[0.030]		[0.090]		
				Panel B. *Male*						
PM2.5 level	-0.093	0.010	0.139	0.276	-0.219	-0.099	-0.086	-0.096	-0.077	-0.091
	(0.122)	(0.106)	(0.186)	(0.396)	(0.144)	(0.085)	(0.070)	(0.108)	(0.219)	(0.091)
Observations	1,446	3,017	968	621	1,428	1,446	3,017	968	621	1,428
KP F-statistic	593.5	456.3	412.5	140.6	529.4	593.5	456.3	412.5	140.6	529.4
Mean of Dep. Var.	25.38	22.55	22.53	20.44	23.49	15.13	13.27	13.31	12.26	13.68
Gender difference	-0.012	-0.280**	-0.611***	-0.391*	0.009	0.101	-0.209**	-0.312***	-0.079	-0.226
	[0.340]	[0.020]	[0.000]	[0.100]	[0.450]	[0.330]	[0.020]	[0.010]	[0.420]	[0.140]

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the subgroup difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table 6: Mechanism test: effects of PM2.5 on educational investment (2SLS)

VARIABLES	educational attainment				time and pecuniary investment			
	years of education		age-appropriate grade completion		time spent helping children with homework (log form)		family education spending (log form)	
	female	male	female	male	female	male	female	male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5 level	-0.157**	0.054	-0.009**	-0.002	-0.026**	0.006	-0.002	0.059
	(0.073)	(0.070)	(0.005)	(0.004)	(0.012)	(0.016)	(0.023)	(0.039)
Observations	2,244	2,085	4,575	4,463	1,529	1,548	1,529	1,548
KP first-stage F-statistic	472.1	864.7	553.8	724.6	364.2	266	364.2	266
Mean of Dep. Var.	9.467	9.730	0.660	0.684	1.869	2.233	1080	1141
Gender difference	-0.211**		-0.008*		-0.032*		-0.061	
	[0.040]		[0.060]		[0.090]		[0.180]	

Notes: Columns (1)–(2) include only individuals who have completed their education. Columns (5)–(8) are restricted to children younger than 16, as information on time spent helping with homework is collected only for this age group. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the gender difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table 7: Mechanism test: effects of PM2.5 on health investment (2SLS)

VARIABLES	duration of breast-feeding (month)		medical expenses (log form)		health insurance coverage	
	female	male	female	male	female	male
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level	0.056	-0.106	-0.065*	-0.071	0.008	-0.003
	(0.111)	(0.181)	(0.036)	(0.056)	(0.007)	(0.009)
Observations	1,430	1,466	1,529	1,548	1,528	1,548
KP first-stage F-statistic	338.8	269.3	364.2	266	364.3	266
Mean of Dep. Var.	13.48	14.67	290.3	403.1	0.677	0.714
Gender difference	0.162		0.006		0.011	
	[0.210]		[0.460]		[0.310]	

Notes: Due to data availability, the sample only includes children younger than 16. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the gender difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

A Supplementary Figures and Tables

Figure A1: Spatial distribution of PM2.5 ($\mu g/m^3$) at city level during 1980-2000, China

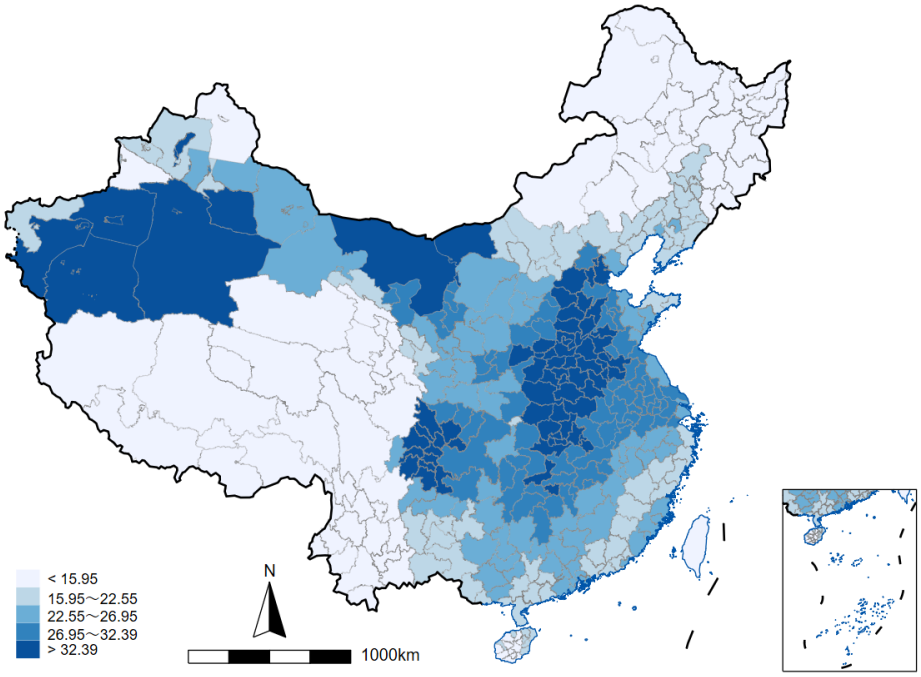


Figure A2: Monthly mean PM2.5 ($\mu g/m^3$) during 1980-2000, China

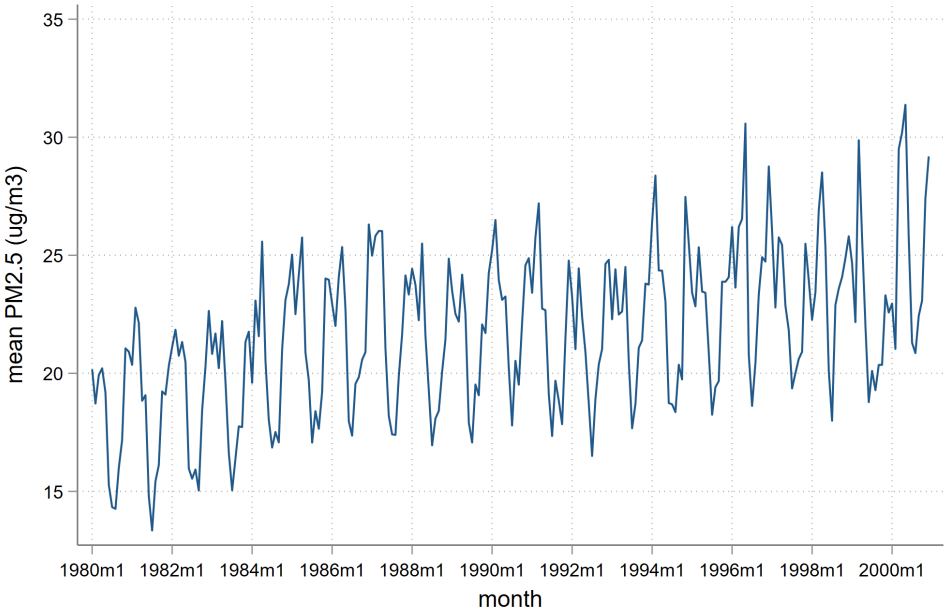
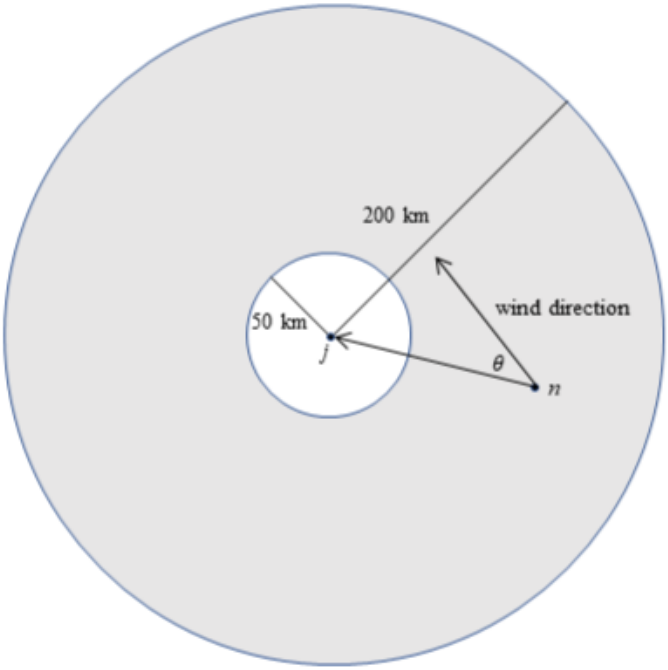
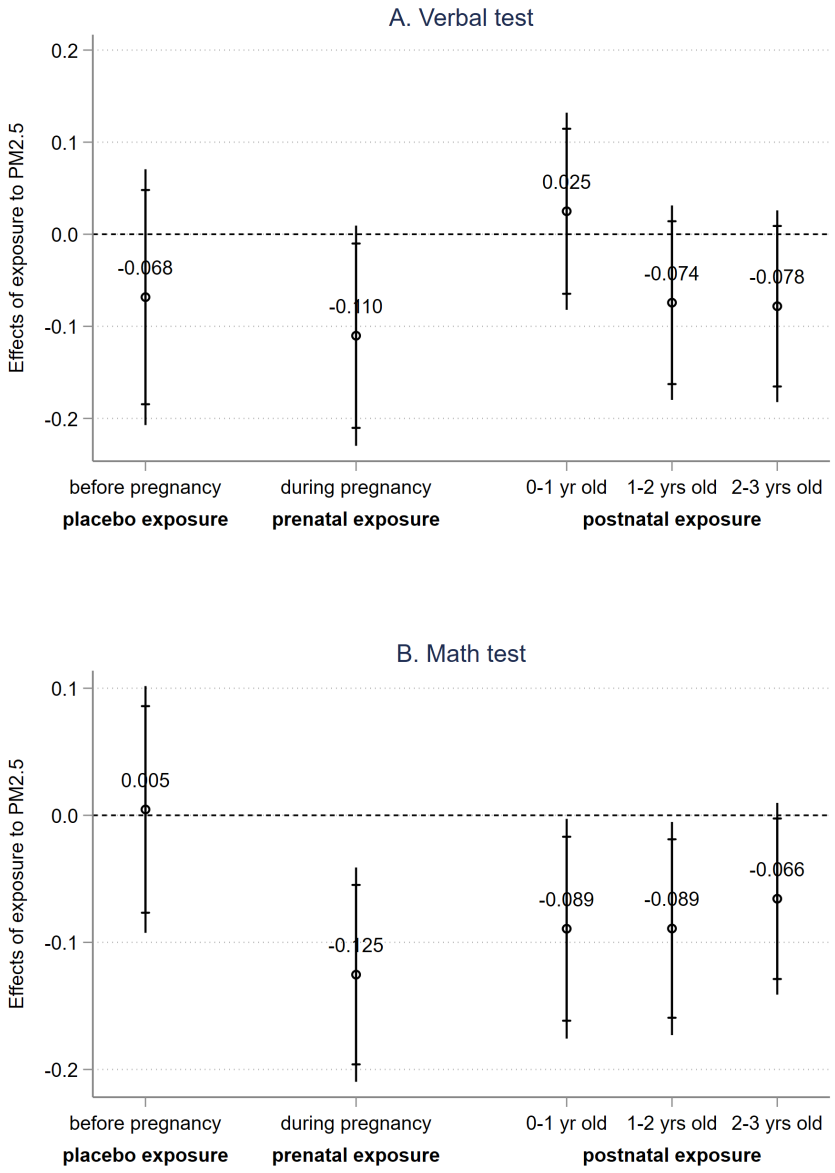


Figure A3: Illustration of the upwind instrumental variable strategy



Note: The figure illustrates the transported air pollution from a nearby upwind city n to the focal city j. θ denotes the angle between the vector $\vec{nj}$ and the wind direction.

Figure A4: Effects of placebo, prenatal, and postnatal exposure to PM2.5 on cognitive ability



Notes: The figure presents the estimated effects of placebo, prenatal, and postnatal exposure to PM2.5 on cognitive ability, with 90% and 95% confidence intervals. Panel A refers to verbal test scores, while panel B corresponds to math test scores. The sample is restricted to individuals who resided in the same city at birth and during the survey period. In each figure, we simultaneously control for PM2.5 exposure before pregnancy, during pregnancy, and at ages 0-1, 1-2, and 2-3 years. We instrument placebo and postnatal PM2.5 exposure using wind directions, following the same approach as for fetal air pollution exposure. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors are clustered at the birth city level.

Table A1: Effects of exposure to fetal pollution on cognitive ability by trimester

VARIABLES	OLS		2SLS	
	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)
PM2.5 level (trimester 1)	-0.066***	-0.072***	-0.081***	-0.094***
	(0.025)	(0.020)	(0.029)	(0.023)
PM2.5 level (trimester 2)	-0.007	-0.008	-0.007	-0.002
	(0.021)	(0.017)	(0.027)	(0.022)
PM2.5 level (trimester 3)	-0.065***	-0.068***	-0.070**	-0.073***
	(0.022)	(0.017)	(0.027)	(0.021)
Observations	9,038	9,038	9,038	9,038
KP first-stage F-statistic			286.5	286.5
Mean of Dep. Var.	23.75	13.69	23.75	13.69

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A2: Non-linear effects of exposure to fetal pollution on cognitive ability

VARIABLES	OLS	
	verbal test	math test
	(1)	(2)
PM2.5 level ($\mu g/m^3$)		
$(0, 15)$	–	–
	–	–
$[15, 25)$	-0.043	-0.038
	(0.051)	(0.039)
$[25, 37.5)$	-0.100*	-0.093**
	(0.056)	(0.047)
$[37.5, 50)$	-0.066	-0.175**
	(0.112)	(0.074)
$[50, +\infty)$	-0.355***	-0.244**
	(0.122)	(0.106)
Observations	9,038	9,038
Mean of Dep. Var.	23.75	13.69

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. Dep. Var.=Dependent variable.

Table A3: Relevant studies on the effect of early-life exposure to air pollution on cognitive performance

Study	Region	Years	Outcomes	Pollutants	Effects
Molina (2021)	Mexico	2002-2009 (survey year)	Raven's test score, height, years of schooling, annual income	CO, PM10	A one SD increase in average monthly thermal inversions per trimester (8.2) leads to a 0.106 SD decline in Raven's test scores.
Sanders (2012)	Texas, United States	1979-1985 (birth year)	High school math test scores	TSP	A one SD increase in TSPs (7.27 $\mu g/m^3$) in the student's year of birth is associated with a 0.06 SD decrease in high school scores.
Bharadwaj et al. (2017)	Santiago, Chile	1992-2001 (birth year)	fourth-grade SIMCE test scores	CO, PM10, AQI	A one SD increase in CO (0.93 ppm) in the third trimester is associated with a 0.034 SD reduction in fourth-grade math scores and a 0.040 SD reduction in language scores. A one SD increase in PM10 (29.24 $\mu g/m^3$) in the third trimester is associated with a 0.029 SD reduction in math scores and a 0.037 SD reduction in language scores.
Rosales-Rueda and Triyana (2019)	Indonesia	1997 (exposure year)	Height-for-age, height, lung capacity, cognitive test scores	aerosol index	They find no significant effect of *in utero* exposure to fires on cognitive function as measured by age-adjusted cumulative test scores.

Notes: To facilitate comparisons across studies, we report the effect sizes in SD change of test scores per one SD higher in air pollutants. Our IV estimates indicate that a one SD increase in fetal exposure to PM2.5 during the gestational period is associated with a reduction in verbal and math test scores by 0.122 SD and 0.170 SD, respectively. SD=standard deviation.

Table A4: Robustness checks for alternative IV constructions (2SLS)

VARIABLES	100 km range		60 degree		45 degree		wind speed	
	verbal test	math test	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5 level	-0.138**	-0.145***	-0.142**	-0.144***	-0.110*	-0.129***	-0.125*	-0.152***
	(0.059)	(0.051)	(0.060)	(0.050)	(0.058)	(0.044)	(0.064)	(0.053)
Observations	9,038	9,038	8,990	8,990	8,938	8,938	9,038	9,038
KP first-stage F-statistic	668.1	668.1	916.3	916.3	723.8	723.8	115	115
Mean of Dep. Var.	23.75	13.69	23.77	13.71	23.79	13.73	23.75	13.69

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A5: Robustness checks for migration (2SLS)

VARIABLES	excluding migrants at 3 years old		excluding migrants at the survey year	
	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)
PM2.5 level	-0.151**	-0.150***	-0.157**	-0.164***
	(0.061)	(0.050)	(0.066)	(0.054)
Observations	8,737	8,737	7,614	7,614
KP first-stage F-statistic	714.2	714.2	626.6	626.6
Mean of Dep. Var.	23.73	13.69	23.52	13.50

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A6: Robustness checks for postnatal exposure (2SLS)

VARIABLES	adding PM2.5 exposure on the interview date		adding PM2.5 exposure in the past month		adding PM2.5 exposure after birth	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level during pregnancy	-0.137**	-0.144***	-0.139**	-0.144***	-0.118**	-0.135***
	(0.060)	(0.051)	(0.060)	(0.051)	(0.059)	(0.051)
PM2.5 level on the interview date	0.029*	0.006				
	(0.016)	(0.012)				
PM2.5 level in the past month			0.038	0.015		
			(0.026)	(0.015)		
PM2.5 level after birth					-0.348***	-0.330***
					(0.131)	(0.116)
Observations	9,035	9,035	9,035	9,035	7,258	7,258
KP first-stage F-statistic	37.98	37.98	24.57	24.57	310.1	310.1
Mean of Dep. Var.	23.75	13.69	23.75	13.69	23.58	13.53

Notes: The sample in columns (5)-(6) is restricted to individuals who resided in the same city at birth and during the survey period. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A7: Robustness checks for alternative specifications (2SLS)

VARIABLES	adding birth year-by-month and city-by-year FE		adding birth year-by-month and city-by-month FE		two-way clustering	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level	-0.184*	-0.110*	-0.187**	-0.193***	-0.137**	-0.144***
	(0.096)	(0.063)	(0.078)	(0.073)	(0.060)	(0.051)
Observations	9,038	9,038	9,038	9,038	9,038	9,038
KP first-stage F-statistic	476.7	476.7	398.1	398.1	726.2	726.2
Mean of Dep. Var.	23.75	13.69	23.75	13.69	23.75	13.69
Birth city FE	N	N	N	N	Y	Y
Birth year and month FE	N	N	N	N	Y	Y
Birth year-by-month FE	Y	Y	Y	Y	N	N
Birth city-by-year FE	Y	Y	N	N	N	N
Birth city-by-month FE	N	N	Y	Y	N	N
Clustering	one-way	one-way	one-way	one-way	two-way	two-way

Notes: All regressions include demographic controls and weather controls. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level in columns (1)-(4) and two-way clustered at the birth city and birth year-season levels in columns (5)-(6), are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A8: Robustness checks for demographic controls (2SLS)

VARIABLES	excluding demographic controls		adding number of siblings		adding dummies for the number of male and female siblings	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level	-0.132**	-0.139***	-0.137**	-0.143***	-0.139**	-0.144***
	(0.060)	(0.052)	(0.058)	(0.050)	(0.058)	(0.049)
Observations	9,038	9,038	9,038	9,038	9,038	9,038
KP first-stage F-statistic	719.7	719.7	722.8	722.8	721.2	721.2
Mean of Dep. Var.	23.75	13.69	23.75	13.69	23.75	13.69
Demographic controls	N	N	Y	Y	Y	Y
Number of siblings	N	N	Y	Y	N	N
Dummies for the number of male and female siblings	N	N	N	N	Y	Y

Notes: All regressions include weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A9: Heterogeneous effects by region and pollution level (2SLS)

Panel A. *Region*				
VARIABLES	verbal test		math test	
	urban	rural	urban	rural
	(1)	(2)	(3)	(4)
PM2.5 level	-0.063	-0.105	-0.078*	-0.141**
	(0.063)	(0.086)	(0.043)	(0.066)
Observations	3,955	5,083	3,955	5,083
KP first-stage F-statistic	1229	403.4	1229	403.4
Mean of Dep. Var.	25.57	22.33	15.26	12.47
Subgroup difference	-0.042		-0.063	
	[0.329]		[0.230]	
Panel B. *Pollution level*				
VARIABLES	verbal test		math test	
	polluted	less polluted	polluted	less polluted
	(1)	(2)	(3)	(4)
PM2.5 level	-0.188*	-0.101	-0.225***	-0.080
	(0.101)	(0.083)	(0.072)	(0.068)
Observations	3,977	5,061	3,977	5,061
KP first-stage F-statistic	190.6	275.6	190.6	275.6
Mean of Dep. Var.	24.83	22.90	14.21	13.28
Subgroup difference	0.087		0.145**	
	[0.210]		[0.040]	

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the group difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A10: Effects of PM2.5 on health at birth (2SLS)

VARIABLES	log form of birth weight (g)		low birth weight (LBW)	
	female	male	female	male
	(1)	(2)	(3)	(4)
PM2.5 level	-0.118*	-0.183**	0.007	0.272**
	(0.066)	(0.080)	(0.090)	(0.110)
Observations	358,937	431,369	358,937	431,369
KP first-stage F-statistic	118.9	111.5	118.9	111.5
Mean of Dep. Var.	3277	3353	0.0295	0.0407
Gender difference	0.066		-0.266***	
	[0.180]		[0.000]	

Notes: The birth record data are drawn from the Chinese National Disease Surveillance Points (DSP) system. The coefficients on PM2.5 are scaled by 100 to make them more readable. All regressions include demographic controls, weather controls, day of birth year fixed effects, and county-by-birth-year fixed effects. Demographic controls include birth order, mother's age at birth and its square term, and dummies for mother's education level. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth county level, are presented in parentheses. P-values for the gender difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A11: Effects of PM2.5 on child health (2SLS)

VARIABLES	log form of height (cm)		log form of weight (kg)		height-for-age		BMI-for-age	
	female	male	female	male	female	male	female	male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5 level	0.102%	-0.516%**	-0.004%	0.300%	-0.025	-0.054*	-0.008	0.052
	(0.241)	(0.218)	(0.408)	(0.452)	(0.031)	(0.032)	(0.029)	(0.032)
Observations	3,460	3,905	3,375	3,826	3,460	3,905	3,375	3,826
KP first-stage F-statistic	1152	540.4	1110	541.2	1152	540.4	1110	541.2
Mean of Dep. Var.	117.4	116	25.63	25.77	-1.015	-1.042	0.0696	0.345
Gender difference	0.617**		-0.304		0.030		-0.060*	
	[0.050]		[0.350]		[0.140]		[0.060]	

Notes: The sample includes children younger than 16. Height-for-age is defined as height standardized by age and sex, and BMI-for-age is defined as body mass index (kg/m^2) standardized by age and sex, both measured relative to the WHO Child Growth Reference and expressed as z-scores. All regressions include demographic controls, weather controls, birth-year-by-month fixed effects, and birth-county-by-year fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth county level, are presented in parentheses. P-values for the gender difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable. BMI=body mass index.

Table A12: Mechanism test: heterogeneous effect on years of education based on the number of siblings and their gender composition (2SLS)

VARIABLE	years of education				
	full sample	**only children**	**having siblings**	**having only brothers**	**having sisters**
	(1)	(2)	(3)	(4)	(5)
			Panel A. *Female*		
PM2.5 level	-0.157**	-0.311	-0.142*	-0.351***	0.013
	(0.073)	(0.195)	(0.075)	(0.128)	(0.096)
Observations	2,244	354	1,890	765	1,125
KP first-stage F-statistic	472.1	127.1	543.6	764.8	308.3
Mean of Dep. Var.	9.467	12.33	8.931	9.050	8.850
Subgroup difference		(3)-(2)=0.168		(4)-(5)=-0.365**	
		[0.200]		[0.010]	
			Panel B. *Male*		
PM2.5 level	0.054	-0.026	0.047	0.015	0.023
	(0.070)	(0.130)	(0.086)	(0.179)	(0.121)
Observations	2,085	543	1,542	494	1,048
KP first-stage F-statistic	864.7	458.7	764.1	408.7	989.4
Mean of Dep. Var.	9.730	11.54	9.094	9.229	9.031
Gender difference	-0.211**	-0.285*	-0.189**	-0.367**	-0.010
	[0.040]	[0.060]	[0.020]	[0.040]	[0.500]

Notes: The sample includes only individuals who have completed their education. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the subgroup difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A13: Mechanism test: heterogeneous effect on age-appropriate grade completion based on the number of siblings and their gender composition (2SLS)

VARIABLE	age-appropriate grade completion				
	full sample	**only children**	**having siblings**	**having only brothers**	**having sisters**
	(1)	(2)	(3)	(4)	(5)
			Panel A. *Female*		
PM2.5 level	-0.009**	-0.001	-0.015***	-0.021**	-0.007
	(0.005)	(0.004)	(0.006)	(0.008)	(0.008)
Observations	4,575	1,034	3,541	1,583	1,958
KP first-stage F-statistic	553.8	418.4	643.5	545.3	532.7
Mean of Dep. Var.	0.660	0.890	0.593	0.629	0.564
Subgroup difference		(3)-(2)=-0.014*		(4)-(5)=-0.014*	
		[0.070]		[0.090]	
			Panel B. *Male*		
PM2.5 level	-0.002	0.001	-0.004	-0.008	-0.007
	(0.004)	(0.006)	(0.006)	(0.010)	(0.007)
Observations	4,463	1,446	3,017	968	2,049
KP first-stage F-statistic	724.6	593.5	456.3	412.5	346.1
Mean of Dep. Var.	0.684	0.833	0.612	0.615	0.611
Gender difference	-0.008*	-0.002	-0.011*	-0.013	0.000
	[0.060]	[0.300]	[0.090]	[0.150]	[0.490]

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the subgroup difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A14: Mechanism test: Heterogeneous effect on time spent helping child with homework based on the number of siblings and their gender composition (2SLS)

VARIABLE	log form of time spent helping child with homework (hours per week)				
	full sample	**only children**	**having siblings**	**having only brothers**	**having sisters**
	(1)	(2)	(3)	(4)	(5)
			Panel A. *Female*		
PM2.5 level	-0.026**	-0.032	-0.046**	-0.044*	-0.052
	(0.012)	(0.021)	(0.020)	(0.025)	(0.037)
Observations	1,529	431	1,098	554	544
KP first-stage F-statistic	364.2	298	525.7	342.6	260.7
Mean of Dep. Var.	1.869	2.596	1.584	1.446	1.724
Subgroup difference		(3)-(2)=-0.014		(4)-(5)=0.008	
		[0.410]		[0.420]	
			Panel B. *Male*		
PM2.5 level	0.006	-0.016	0.021	-0.038	0.049*
	(0.016)	(0.026)	(0.023)	(0.036)	(0.028)
Observations	1,548	582	966	319	647
KP first-stage F-statistic	266	197.2	164.1	85.10	126.6
Mean of Dep. Var.	2.233	2.899	1.832	1.912	1.793
Gender difference	-0.032*	-0.017	-0.067*	-0.006	-0.101
	[0.090]	[0.450]	[0.060]	[0.390]	[0.130]

Notes: The sample is restricted to children younger than 16, as information on time spent helping with homework is collected only for this age group. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the subgroup difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table A15: Effects of PM2.5 on the frequency of parents caring about their child's study (2SLS)

VARIABLES	frequency of asking the child to finish homework (1-5)		frequency of checking the child's homework (1-5)	
	female	male	female	male
	(1)	(2)	(3)	(4)
PM2.5 level	-0.043***	-0.013	-0.048**	-0.009
	(0.013)	(0.018)	(0.021)	(0.028)
Observations	1,516	1,538	1,516	1,536
KP first-stage F-statistic	367	262.8	367	262.5
Mean of Dep. Var.	3.889	3.951	2.790	2.915
Gender difference		-0.030		-0.039
		[0.250]		[0.230]

Notes: The sample is restricted to children younger than 16, as information on the frequency of parents caring about their child's study is collected only for this age group. The frequency is categorized into five levels: 5 = "very often" (6–7 times per week), 4 = "often" (2–3 times per week), 3 = "sometimes" (1–2 times per week), 2 = "rarely" (once per month), and 1 = "never". All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. P-values for the gender difference are displayed in brackets. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

B Model Illustration

B.1 Model setup

Human capital endowment A household has two children, indexed by $i \in \{M, F\}$, with a gender dummy g_i ($g_M = 1$ for a son, $g_F = 0$ for a daughter). Each child has a prenatal human capital endowment θ_i that is determined by an exogenous component μ_i and the prenatal pollution shock $e_i \geq 0$,

$$\theta_i \;=\; \mu_i + (\alpha + m\, g_i)\, e_i, \qquad \alpha < 0, \;\; m < 0. \tag{B1}$$

Here $\alpha < 0$ is the common biological damage of prenatal pollution and $m < 0$ is the additional damage borne by sons, so that

$$\frac{\partial \theta_M}{\partial e_M} \;<\; \frac{\partial \theta_F}{\partial e_F} \;<\; 0. \tag{B2}$$

Eq. (B2) reflects the *fragile males* channel: a given prenatal shock depresses a son's endowment by more than a daughter's (e.g., Kraemer, 2000; Eriksson et al., 2010).[1]

Perceived developmental potential Revealed human capital S_i is produced additively from the endowment and parental investment I_i:

$$S_i \;=\; \theta_i + b_I\, I_i, \qquad b_I > 0. \tag{B3}$$

Parents do not allocate investment solely on the basis of the realized endowment. Instead, they respond to the perceived marginal return to investment, which we allow to rise with the child's *perceived* developmental potential $\tilde{\theta}_i$. Thus, a negative shock that lowers perceived potential also lowers the perceived return to educational investment. Son preference enters this perception: parents partially insulate a son's perceived potential from a negative environmental shock by revising their expectations for a daughter downward after

[1] We assume that any surviving child has $\theta_i > 0$.

a shock, but holding their aspirations for a son closer to a reference level $\bar{\theta} > \theta_M(e_M)$,

$$\tilde{\theta}_i \;=\; \theta_i + s\, g_i\,(\bar{\theta} - \theta_i) \;=\; \begin{cases} \theta_F, & \text{daughter}, \\ (1-s)\,\theta_M + s\,\bar{\theta}, & \text{son}, \end{cases} \qquad s \in [0,1). \tag{B4}$$

The parameter s is the son preference insulation: $s = 0$ means parents perceive a son's potential at its realized value, while $s \to 1$ means a son's perceived potential is anchored to $\bar{\theta}$ and is unaffected by the shock. A daughter ($g_F = 0$) receives no insulation, $\tilde{\theta}_F = \theta_F$.

B.2 Parental objective

Parents choose (I_M, I_F) to maximize a perceived return to investment with son preference, net of a convex resource cost (e.g., Dizon-Ross, 2019):

$$\max_{I_M,\, I_F \geq 0} \sum_{i \in \{M,F\}} \left(b_I + \varphi\,\tilde{\theta}_i\right) I_i \;-\; \frac{c}{2}\left(I_M^2 + I_F^2\right), \qquad \varphi > 0,\; c > 0. \tag{B5}$$

The son preference parameter s enters the objective function through the child's perceived potential $\tilde{\theta}_i$.[2] The term $(b_I + \varphi\tilde{\theta}_i)$ is the marginal value parents attach to a unit of investment in child i, which rises with the perceived potential $\tilde{\theta}_i$ through $\varphi > 0$.[3] The quadratic term $\frac{c}{2} I_i^2$ is the convex cost of investment (including time and monetary costs) for child i, with c governing the curvature of investment costs.

[2] Combined with eq. (B4), one can interpret the son preference parameter s as the degree to which parents perceive the return to educational investment in sons as insulated from prenatal shocks. A larger s implies that parents *believe* an environmental shock almost does not affect their sons' marginal return to investment.

[3] The literature finds that parents perceive baseline ability and investment as complements, so they allocate more educational resources to higher-achieving children (e.g., Yi et al., 2015; Giannola, 2024). This corresponds to the $\varphi\tilde{\theta}$ term, which indicates that higher perceived potential raises the perceived return to parental investment.

B.3 Closed–form optimal investment

The first-order condition of eq. (B5) implies the closed-form interior optimum:

$$I_i^* \;=\; \frac{b_I + \varphi\,\tilde{\theta}_i}{c} \quad\Longrightarrow\quad S_i^* \;=\; \theta_i + b_I\, I_i^* \;=\; \theta_i + \frac{b_I\,(b_I + \varphi\tilde{\theta}_i)}{c}. \tag{B6}$$

Given the positive support of parameters $\{c, \varphi, b_I\}$, the solution is interior ($I_i^* > 0$) for both children. By anchoring a son's perceived potential closer to the reference level $\bar{\theta}$, a larger s makes I_M^* less sensitive to the prenatal environmental shock e_M.

Eq. (B6) shows that optimal investment increases with human capital endowment:

$$\frac{\partial I_i^*}{\partial \theta_i} = \frac{\varphi}{c}(1 - s g_i) \geq 0. \tag{B7}$$

Thus, investment is fully responsive to changes in a daughter's endowment, whereas the corresponding response for a son is attenuated by the son-biased insulation parameter s. A larger s therefore weakens the pass-through from changes in a son's underlying endowment to parental investment. At $s \to 1$, the son's investment is locally insensitive to his true endowment.

B.4 Comparative statics: investment by gender

Differentiating eq. (B4), $\partial\tilde{\theta}_i/\partial e_i = (1 - s\,g_i)(\alpha + m g_i)$, so from eq. (B6):

$$\frac{dI_i^*}{de_i} \;=\; \frac{\varphi}{c}\,(1 - s\,g_i)\,(\alpha + m g_i) < 0. \tag{B8}$$

Evaluated by gender,

$$\text{daughters:} \quad \frac{dI_F^*}{de_F} = \frac{\varphi\,\alpha}{c} \;<\; 0, \tag{B9}$$

$$\text{sons:} \quad \frac{dI_M^*}{de_M} = \frac{\varphi}{c}\,(1 - s)\,(\alpha + m) \;\xrightarrow[s\to 1]{}\; 0. \tag{B10}$$

A negative prenatal environmental shock lowers a daughter's perceived potential, hence the perceived return to her education, so parents withdraw investment ($dI_F^*/de_F < 0$). For a son, son preference s insulates perceived potential, so the decline in his perceived return and therefore, his investment response, is small; it vanishes as $s \to 1$.

Additionally, the male response is smaller in magnitude than the female response whenever son preference is sufficiently strong to offset the greater male vulnerability:

$$\left|\frac{dI_F^*}{de_F}\right| > \left|\frac{dI_M^*}{de_M}\right| \quad \Longleftrightarrow \quad s \;>\; 1 - \frac{|\alpha|}{|\alpha + m|}. \tag{B11}$$

B.5 Total effect of the shock on human capital

Using eqs. (B3) and (B8), the net effect of the prenatal shock on revealed human capital can be decomposed into a direct biological effect and an indirect investment effect:

$$\frac{dS_i}{de_i} = \underbrace{(\alpha + m g_i)}_{\text{direct biological } <0} + \underbrace{b_I \frac{dI_i^*}{de_i}}_{\text{indirect investment } <0} = (\alpha + m g_i)\left[1 + \frac{b_I\,\varphi\,(1 - s g_i)}{c}\right] < 0. \tag{B12}$$

By gender,

$$\text{daughters:} \quad \frac{dS_F}{de_F} = \alpha\left[1 + \frac{b_I \varphi}{c}\right] < 0, \tag{B13}$$

$$\text{sons:} \quad \frac{dS_M}{de_M} = (\alpha + m)\left[1 + \frac{b_I \varphi (1 - s)}{c}\right] < 0. \tag{B14}$$

Eqs. (B13) and (B14) show that a worse prenatal environmental shock reduces revealed human capital for both daughters and sons. For daughters, the direct biological effect is negative, and the indirect investment channel reinforces this loss: the shock lowers perceived developmental potential, reduces the perceived return to educational investment, and therefore leads parents to withdraw investment. This is reflected in the multiplier $1 + b_I\varphi/c > 1$, which makes the daughter's net loss larger in magnitude than the direct biological effect alone. For sons, the direct biological effect is more negative because $m < 0$, capturing the

fragile-males hypothesis. However, son-biased insulation dampens the investment response: the indirect channel is scaled by $(1 - s)$, so when s is large, sons' educational investment changes little and their net loss is driven primarily by the direct biological channel. Thus, the model can reconcile biological vulnerability among males with larger long-run losses for daughters: sons may suffer greater biological damage at birth, but daughters experience an additional loss through the withdrawal of parental educational investment.

The relative magnitude of the net effect between the daughter and the son hinges on the balance between sons' additional biological frangibility and daughters' investment-withdrawal channel induced by the son-biased insulation. When the son-biased insulation is sufficiently prominent:

$$s > \left[1 + \frac{c}{b_I \varphi}\right] \frac{|m|}{|\alpha + m|},$$

the net effect of prenatal environment shock on human capital for the daughter would exceed that for the son: $dS_F/de_F < dS_M/de_M < 0$.[4]

B.6 Summary

In this appendix, we do not attempt to structurally estimate parental preferences or the human-capital production function. Rather, the framework is intended to show that a simple parental investment problem, augmented with son-biased perceived returns, can rationalize the reduced-form patterns documented in the empirical analysis.

B.6.1 Correspondence with reduced form results

The main implications of the model correspond to the empirical findings along three dimensions. First, a negative prenatal environmental shock reduces daughters' perceived developmental potential, thereby lowering the perceived marginal return to educational investment, yielding $dI_F^*/de_F < 0$. This aligns with the empirical finding that fetal PM2.5

[4]To ensure that $s \in [0, 1)$, this condition is feasible only if $\frac{b_I \varphi}{c} > \frac{|m|}{|\alpha|}$.

exposure reduces family educational assistance and schooling outcomes for females (the odd columns in Table 6).

Second, son-biased belief insulation makes the perceived return to investing in sons less responsive to the same shock, so dI_M^*/de_M is small and can be arbitrarily close to zero as s approaches one, generally consistent with the weak or insignificant investment response estimated for males (the even columns in Table 6).

Third, because the direct biological effect of pollution remains negative for both children, the model implies $dS_F/de_F < 0$ and $dS_M/de_M < 0$, meaning that prenatal pollution lowers revealed human capital for both daughters and sons. Moreover, the model reconciles this pattern with the fragile-males hypothesis (Table A10). Sons may experience a larger direct biological loss, captured by $m < 0$, but son-biased insulation protects their educational investment, whereas daughters experience both the direct biological loss and an additional withdrawal of parental investment. When the daughters' investment-withdrawal channel is sufficiently strong relative to the sons' extra biological vulnerability, the model delivers $dS_F/de_F < dS_M/de_M < 0$, so the long-run human-capital loss is larger for daughters despite greater biological fragility among sons (Panel A of Table 3).

B.6.2 Correspondence with the conceptual framework

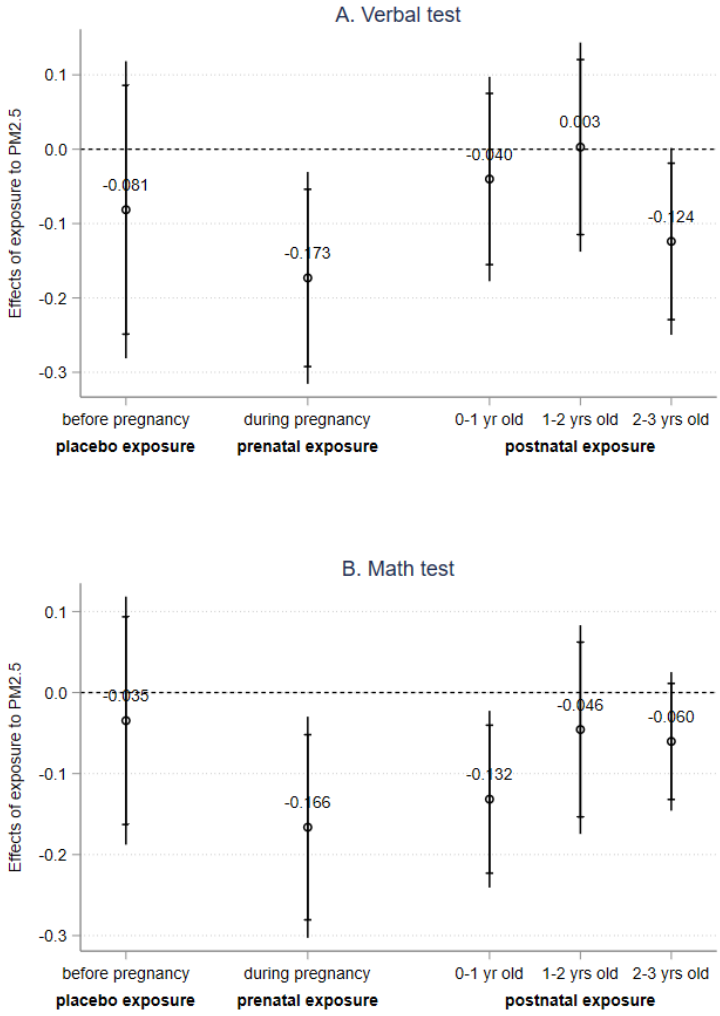
The two-child model in this Appendix provides a micro-foundation for the indirect parental-investment channel in the conceptual framework. Parental investment is complementary to perceived developmental potential, while son preference makes parents' perceptions of sons' developmental potential less responsive to adverse prenatal shocks. Consequently, parents optimally reduce investment in daughters following a negative environmental shock, while investment in sons changes only weakly, generating the general pattern observed empirically in Table 6. Using the notation of the conceptual framework, the indirect effect can be mapped into the model as follows:

$$\text{indirect effect} = \underbrace{\frac{\partial f_C}{\partial I_C}}_{b_I>0} \overbrace{\underbrace{\frac{\partial I_C}{\partial H_F}}_{\partial I_i^*/\partial\theta_i\geq 0} \underbrace{\frac{\partial f_F}{\partial P_F}}_{\alpha+mg_i<0}}^{\frac{\partial I_i^*}{\partial e_i}\leq 0 \text{ from Table 6}} \leq 0.$$

The first fraction corresponds to the marginal effect of investment on human capital, which is $b_I > 0$ in eq. (B3); the third fraction is the effect of prenatal pollution on endowment, which corresponds to $\alpha + mg_i < 0$ in eq. (B1). The second fraction corresponds to $\partial I_i^*/\partial\theta_i = \frac{\varphi}{c}(1 - sg_i) \geq 0$ in eq. (B7), indicating reinforcing parental investment or, at $s \to 1$, a neutral response for sons.

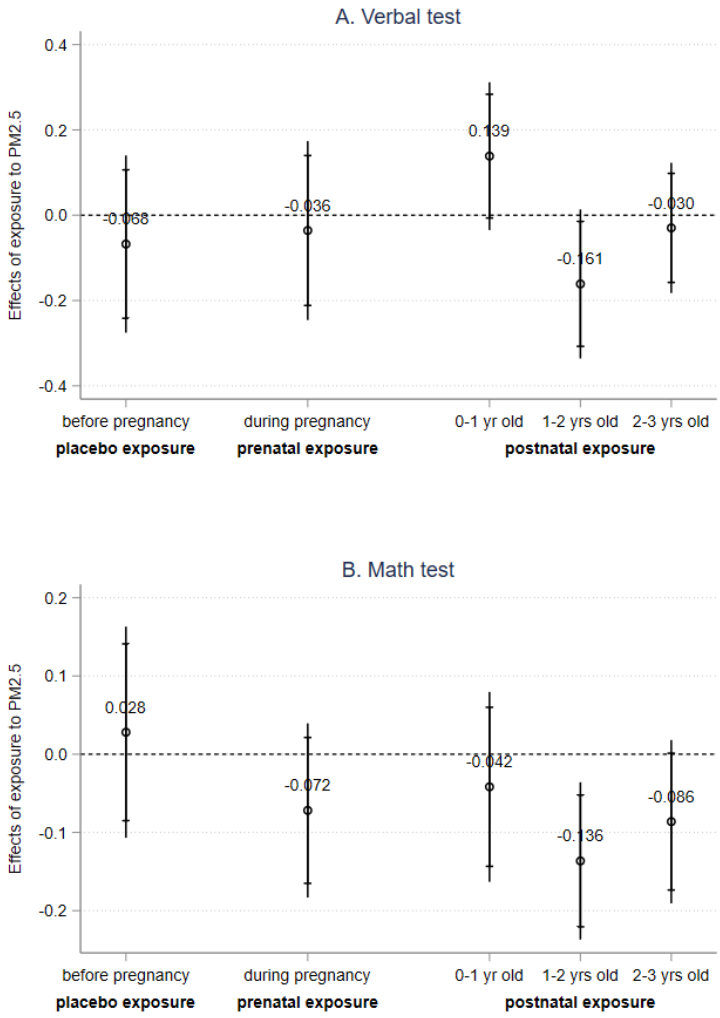
C Robustness Checks for Gender Heterogeneity

Figure C1: Effects of placebo, prenatal, and postnatal exposure to PM2.5 on cognitive ability (female subsample)



Notes: The figure presents the estimated effects of placebo, prenatal, and postnatal exposure to PM2.5 on cognitive ability, with 90% and 95% confidence intervals. Panel A refers to verbal test scores, while panel B corresponds to math test scores. The sample is restricted to individuals who resided in the same city at birth and during the survey period. In each figure, we simultaneously control for PM2.5 exposure before pregnancy, during pregnancy, and at ages 0-1, 1-2, and 2-3 years. We instrument placebo and postnatal PM2.5 exposure using wind directions, following the same approach as for fetal air pollution exposure. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors are clustered at the birth city level.

Figure C2: Effects of placebo, prenatal, and postnatal exposure to PM2.5 on cognitive ability (male subsample)



Notes: The figure presents the estimated effects of placebo, prenatal, and postnatal exposure to PM2.5 on cognitive ability, with 90% and 95% confidence intervals. Panel A refers to verbal test scores, while panel B corresponds to math test scores. The sample is restricted to individuals who resided in the same city at birth and during the survey period. In each figure, we simultaneously control for PM2.5 exposure before pregnancy, during pregnancy, and at ages 0-1, 1-2, and 2-3 years. We instrument placebo and postnatal PM2.5 exposure using wind directions, following the same approach as for fetal air pollution exposure. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors are clustered at the birth city level.

Table C1: Robustness checks for alternative IV constructions (female subsample)

VARIABLES	100 km range		60 degree		45 degree		wind speed	
	verbal test	math test	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5 level	-0.206***	-0.191***	-0.220***	-0.182***	-0.180**	-0.155***	-0.174**	-0.169**
	(0.078)	(0.071)	(0.078)	(0.068)	(0.071)	(0.059)	(0.082)	(0.073)
Observations	4,575	4,575	4,550	4,550	4,522	4,522	4,575	4,575
KP first-stage F-statistic	502.9	502.9	636.7	636.7	413.5	413.5	93.28	93.28
Mean of Dep. Var.	24.02	13.52	24.04	13.54	24.06	13.55	24.02	13.52

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C2: Robustness checks for alternative IV constructions (male subsample)

VARIABLES	100 km range		60 degree		45 degree		wind speed	
	verbal test	math test	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5 level	-0.049	-0.090	-0.031	-0.099*	-0.017	-0.099*	-0.066	-0.138**
	(0.083)	(0.055)	(0.085)	(0.053)	(0.087)	(0.053)	(0.094)	(0.068)
Observations	4,463	4,463	4,440	4,440	4,416	4,416	4,463	4,463
KP first-stage F-statistic	716.9	716.9	1035	1035	951.9	951.9	114.5	114.5
Mean of Dep. Var.	23.47	13.87	23.50	13.89	23.52	13.91	23.47	13.87

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C3: Robustness checks for migration (female subsample)

VARIABLES	excluding migrants at 3 years old		excluding migrants at the survey year	
	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)
PM2.5 level	-0.234***	-0.192***	-0.225**	-0.215**
	(0.082)	(0.071)	(0.088)	(0.083)
Observations	4,444	4,444	3,747	3,747
KP first-stage F-statistic	528.2	528.2	484.4	484.4
Mean of Dep. Var.	23.98	13.50	23.85	13.36

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C4: Robustness checks for migration (male subsample)

VARIABLES	excluding migrants at 3 years old		excluding migrants at the survey year	
	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)
PM2.5 level	-0.040	-0.103*	-0.053	-0.100*
	(0.087)	(0.057)	(0.093)	(0.059)
Observations	4,293	4,293	3,867	3,867
KP first-stage F-statistic	682.8	682.8	596.9	596.9
Mean of Dep. Var.	23.47	13.88	23.20	13.65

Notes: All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C5: Robustness checks for postnatal exposure (female subsample)

VARIABLES	adding PM2.5 exposure on the interview date		adding PM2.5 exposure in the past month		adding PM2.5 exposure after birth	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level during pregnancy	-0.210***	-0.190***	-0.212***	-0.190***	-0.190**	-0.206***
	(0.078)	(0.070)	(0.079)	(0.070)	(0.076)	(0.079)
PM2.5 level on the interview date	0.020	0.002				
	(0.019)	(0.014)				
PM2.5 level in the past month			0.003	0.002		
			(0.033)	(0.022)		
PM2.5 level after birth					-0.518***	-0.403***
					(0.179)	(0.129)
Observations	4,575	4,575	4,575	4,575	3,588	3,588
KP first-stage F-statistic	29.77	29.77	21.67	21.67	235.8	235.8
Mean of Dep. Var.	24.02	13.52	24.02	13.52	23.92	13.41

Notes: The sample in columns (5)-(6) is restricted to individuals who resided in the same city at birth and during the survey period. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C6: Robustness checks for postnatal exposure (male subsample)

VARIABLES	adding PM2.5 exposure on the interview date		adding PM2.5 exposure in the past month		adding PM2.5 exposure after birth	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level during pregnancy	-0.041	-0.092*	-0.042	-0.093*	-0.010	-0.033
	(0.083)	(0.055)	(0.082)	(0.054)	(0.099)	(0.064)
PM2.5 level on the interview date	0.038*	0.016				
	(0.022)	(0.015)				
PM2.5 level in the past month			0.073**	0.039*		
			(0.036)	(0.021)		
PM2.5 level after birth					-0.202	-0.309**
					(0.208)	(0.153)
Observations	4,460	4,460	4,460	4,460	3,670	3,670
KP first-stage F-statistic	38.43	38.43	24.11	24.11	342.7	342.7
Mean of Dep. Var.	23.47	13.87	23.47	13.87	23.25	13.64

Notes: The sample in columns (5)-(6) is restricted to individuals who resided in the same city at birth and during the survey period. All regressions include demographic controls, weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C7: Robustness checks for alternative specifications (female subsample)

VARIABLES	adding birth year-by-month and province-by-year FE		adding birth year-by-month and province-by-month FE		two-way clustering	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level	-0.226*	-0.180**	-0.254***	-0.226**	-0.212**	-0.190**
	(0.116)	(0.085)	(0.092)	(0.093)	(0.083)	(0.075)
Observations	4,575	4,575	4,575	4,575	4,575	4,575
KP first-stage F-statistic	172.4	172.4	308.5	308.5	570.5	570.5
Mean of Dep. Var.	24.02	13.52	24.02	13.52	24.02	13.52
Birth city FE	Y	Y	Y	Y	Y	Y
Birth year and month FE	N	N	N	N	Y	Y
Birth year-by-month FE	Y	Y	Y	Y	N	N
Birth province-by-year FE	Y	Y	N	N	N	N
Birth province-by-month FE	N	N	Y	Y	N	N
Clustering	one-way	one-way	one-way	one-way	two-way	two-way

Notes: All regressions include demographic controls and weather controls. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level in columns (1)-(4) and two-way clustered at the birth city and birth year-season levels in columns (5)-(6), are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C8: Robustness checks for alternative specifications (male subsample)

VARIABLES	adding birth year-by-month and province-by-year FE		adding birth year-by-month and province-by-month FE		two-way clustering	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level	-0.086	-0.031	-0.106	-0.115	-0.037	-0.090*
	(0.138)	(0.108)	(0.099)	(0.081)	(0.077)	(0.050)
Observations	4,460	4,460	4,460	4,460	4,463	4,463
KP first-stage F-statistic	194.8	194.8	403.1	403.1	673	673
Mean of Dep. Var.	23.47	13.87	23.47	13.87	23.47	13.87
Birth city FE	Y	Y	Y	Y	Y	Y
Birth year and month FE	N	N	N	N	Y	Y
Birth year-by-month FE	Y	Y	Y	Y	N	N
Birth province-by-year FE	Y	Y	N	N	N	N
Birth province-by-month FE	N	N	Y	Y	N	N
Clustering	one-way	one-way	one-way	one-way	two-way	two-way

Notes: All regressions include demographic controls and weather controls. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level in columns (1)-(4) and two-way clustered at the birth city and birth year-season levels in columns (5)-(6), are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C9: Robustness checks for demographic controls (female subsample)

VARIABLES	excluding demographic controls		adding number of siblings		adding dummies for the number of male and female siblings	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level	-0.208***	-0.186***	-0.210***	-0.188***	-0.211***	-0.186***
	(0.079)	(0.070)	(0.077)	(0.068)	(0.076)	(0.068)
Observations	4,575	4,575	4,575	4,575	4,575	4,575
KP first-stage F-statistic	554.2	554.2	558.5	558.5	557.4	557.4
Mean of Dep. Var.	24.02	13.52	24.02	13.52	24.02	13.52
Demographic controls	N	N	Y	Y	Y	Y
Number of siblings	N	N	Y	Y	N	N
Dummies for the number of male and female siblings	N	N	N	N	Y	Y

Notes: All regressions include weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.

Table C10: Robustness checks for demographic controls (male subsample)

VARIABLES	excluding demographic controls		adding number of siblings		adding dummies for the number of male and female siblings	
	verbal test	math test	verbal test	math test	verbal test	math test
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5 level	-0.034	-0.087	-0.041	-0.093*	-0.043	-0.097*
	(0.082)	(0.055)	(0.081)	(0.053)	(0.079)	(0.052)
Observations	4,463	4,463	4,463	4,463	4,463	4,463
KP first-stage F-statistic	721	721	722.8	722.8	719.9	719.9
Mean of Dep. Var.	23.47	13.87	23.47	13.87	23.47	13.87
Demographic controls	N	N	Y	Y	Y	Y
Number of siblings	N	N	Y	Y	N	N
Dummies for the number of male and female siblings	N	N	N	N	Y	Y

Notes: All regressions include weather controls, birth city fixed effects, and birth year and month fixed effects. Demographic controls include gender, maternal age, and mother's education years. Weather controls include the number of days falling in each temperature bin (<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C [reference group], 16-20 °C, 20-24 °C, 24-28 °C, >28 °C), mean precipitation, mean wind speed, mean sunshine duration, and mean humidity with their square terms during the gestation period. Robust standard errors, clustered at the birth city level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level. KP F-statistic=Kleibergen-Paap Wald rk F-statistic. Dep. Var.=Dependent variable.