When the Wind Blows: Agricultural Fire Exposure, Parental Investment, and Long-term Outcomes

Hai Hong; Kevin Chen[§]

Abstract:

This paper examines the medium and long-term human capital consequences of in-utero exposure to agricultural fires in rural China. We exploit exogenous variations in birth month, fire intensity, and wind direction to identify the causal effect of fetal exposure to fire. Our findings suggest that in-utero exposure to agricultural fires significantly reduces individuals' health, cognitive, and non-cognitive performance in adolescence, with the effect primarily driven by the male sample. Tracking these cohorts into adulthood, we show that fire exposure decreases the number of education years and earnings. Besides the transmission of adverse conditions in early life, a key mechanism driving the persistent effect of fetal exposure is that liquidity-constrained households reinforce the negative impacts by reallocating resources *away* from exposed children. Using the rollout of China's New Cooperative Medical Scheme (NCMS) as a quasi-experiment, we present evidence that health insurance coverage can largely offset the effects of agricultural fire exposure by easing financial constraints and promoting parental investments. Our findings underscore the disproportionate cost of pollution on vulnerable rural families and have significant policy implications for mitigating the adverse effects of pollution exposure.

Key Words: Agricultural Fire; In-utero Exposure; Parental Investment; Health Insurance.
JEL Codes: Q51; Q53; D13; I13; I14.

[§] Hong: China Academy for Rural Development, School of Public Affairs, Zhejiang University (email address: haihong@zju.edu.cn). Chen: China Academy for Rural Development, School of Public Affairs, Zhejiang University; International Food Policy Research Institute (email address: kzchen@zju.edu.cn). Kevin Chen is the corresponding author.

1. Introduction

Agricultural fire is a prevalent practice in developing countries for disposing of crop residues after harvest. While it benefits agricultural production by reducing the cost of residue disposal and eliminating potential pests and diseases, it also contributes significantly to seasonal air pollution, leading to severe health and cognitive consequences (Rangel and Vogl 2019; Graff Zivin et al. 2020; He et al. 2020; Lai et al. 2022; Pullabhotla and Souza 2022; Ayesh 2023; Du et al. 2024; Garg et al. 2024). This issue is particularly salient for infants, who are more vulnerable and susceptible to environmental pollutants (Currie and Neidell 2005; Currie et al. 2009; Almond and Currie 2011; Almond, Edlund, and Palme 2009; Currie et al. 2014), and the potential costs of pollution exposure (e.g., cognitive deficits, chronic diseases) may not become apparent until years after birth, leading policymakers to underestimate their long-term effects. Moreover, due to lower income and weaker health infrastructure in the rural counterparts of developing countries, fetal exposure to agricultural fires may further affect the decision of intra-household resource reallocation (Yi et al. 2015), which could even exacerbate the potential costs of pollution exposure.

In this paper, we examine the long-term effects of in-utero exposure to agricultural fires on adolescent and adult outcomes in rural China and explore potential mechanisms. To carry out the empirical analysis, we face two empirical challenges. First, since valid measures of agricultural fires (e.g., satellite observations) are only available after the 2000s, to estimate the long-term effects of fetal agricultural fire exposure for individuals born before the 2000s, we need effective proxy measures for agricultural fires. Second, to shed light on the long-term effects of pollution exposure, as well as potential mechanisms (e.g., intra-household resource allocation), we need detailed data that records individuals' health and non-health outcomes (e.g., cognitive and labor market performance) and parental investment.

To address the first challenge, we resort to exogenous agricultural potential yield and borrow insights from previous studies to construct measures for upwind/downwind potential yield (Rangel and Vogl 2019; He et al. 2020; Lai et al. 2022). We provide a set of verification tests in Section 4 to show that the upwind/downwind potential yield measures are valid proxies for agricultural fire intensity and air pollution (e.g., $PM_{2.5}$) across different counties. To address the second challenge, we exploit a nationally representative household dataset, the China Family Panel Studies (CFPS), which records detailed health and (non-)cognitive measures for adolescents and the corresponding parental investment (e.g., health and education expenditure). The CFPS also tracks adolescents into adulthood, allowing for an examination of the labor market impacts of in-utero agricultural fire exposure. In addition, CFPS records retrospective questions on early-life conditions (e.g., health at birth and age 1), which enables us to investigate how in-utero agricultural fire exposure affects early-life outcomes.

We begin by estimating the effects of in-utero agricultural fire exposure on the health, cognitive, and non-cognitive outcomes of adolescents. Our identification hinges on three sources of plausibly exogenous variations. The first is variations in individuals' birth month, which we exploit to determine during which trimester the individual is exposed to agricultural fires. The second is variations in fire intensity across different counties, proxied by agricultural potential yields. We augment our identification by including a third variation in wind direction, which allows us to implement an upwind-downwind specification that eliminates potential income effects that may confound our identification. Our results suggest that in-utero agricultural fire exposure can have significant deleterious effects on adolescent development. Specifically, we find that agricultural fire exposure leads to worsened health outcomes (measured by a composite index combining information on self-reported health, hospitalization, and chronic respiratory disease), cognitive outcomes (measured by standardized word test scores), and non-cognitive outcomes. Additionally, we show that the estimated effects are more pronounced if individuals are exposed to agricultural fires during the first and third trimesters, which is in line with previous literature in economics and epidemiology (Glinianaia et al. 2004; Šrám et al. 2005; Currie and Neidell 2005; Kannan et al. 2006; Currie et al. 2014; Rangel and Vogl 2019). The effects are stronger for boys than girls, possibly because male fetuses are more susceptible to in-utero pollution exposure. We find no effects on urban adolescents.

Due to the panel structure of the CFPS, we are able to track the same adolescents into their adulthood, which allows us to explore the effects of in-utero agricultural fire exposure on educational and labor market outcomes. Our findings reveal that fire exposure can significantly lower the number of years of education. Conditional on entering the labor market, we show that exposed cohorts earn lower annual wages and are more likely to work in the agricultural sector.

We then proceed to explore the potential mechanisms through which in-utero exposure to agricultural fires leads to persistent effects. As suggested by Currie et al. (2014), there are at least two channels through which early-life exposure to air pollution translates into long-term consequences. The first is the direct channel through the transmission of adverse early life outcomes, and the second is the indirect channel through intra-household resource reallocation, i.e., parental investment. While the sign of the first channel is theoretically unambiguous, parental investment in responding to the adverse health shocks could be either reinforcing or compensating.[1] To shed light on these potential mechanisms, we first examine how agricultural fire exposure worsens early-life health conditions. We find consistent evidence that in-utero agricultural fire exposure increases the number of illnesses at age 1, and leads to shortened gestation months and lower birth weight.

Next, we investigate how parents respond to negative health shocks induced by agricultural fire exposure. We show that parents reduce both health and education investment in children who are exposed to agricultural fires, consistent with reinforcement behavior. Exploring the potential heterogeneity, we find that the reduction in parental investments is primarily driven by mothers with lower education levels and households with lower income. This suggests that liquidity constraints may be an important driver that explains why rural households reduce their investment in exposed children. Moreover, our results also imply that the consequences of agricultural fire exposure are unevenly distributed across rural households, with more disadvantaged households being more severely affected.

Given that agricultural fires can have significant adverse long-term effects on rural adolescents and that parental responses could even reinforce such negative impacts, a critical policy question is what measures can be taken to mitigate the adverse pollution effects. In the last part of our empirics, we investigate the effects of the provision of public health insurance on mitigating the effects of agricultural fire exposure. To this end, we leverage the sequential rollout of the New Cooperative Medical Scheme (NCMS) in rural China (see Section 2.3 for a more detailed description of the program), which is the largest insurance program in history (Gruber et al. 2023). The NCMS program is financed by low individual contributions and high government payments and offers generous subsidies for inpatient expenses. We follow Huang and Liu (2023) and denote individuals who were less than 5 years old when the NCMS program was implemented as those exposed to the policy. We find that, for individuals who are exposed to the NCMS program, in-utero exposure to agricultural fires has no significant effect on adolescent outcomes, which suggests a mitigating role of public health insurance coverage. Moreover, we provide evidence that the mitigating role of the NCMS program is mainly through increasing parental investments, especially for those more disadvantaged households, and find no evidence that NCMS exposure can mitigate adverse health outcomes at birth.

This paper speaks to three strands of literature in environmental and health economics. First, we add to

[1] The compensatory channel suggests that family would invest more on children who are more exposed to pollution, while the reinforcing channel suggests that family would invest more on children who are less exposed to pollution, as the return to human capital is higher.

the burgeoning literature that examines the consequences of air pollution from agricultural fires (Rangel and Vogl 2019; Graff Zivin et al. 2020; He et al. 2020; Lai et al. 2022; Pullabhotla and Souza 2022; Ayesh 2023; Du et al. 2024; Garg et al. 2024). While much of the existing research focuses on the contemporaneous effects of exposure to agricultural fires and a wide range of health and behavioral outcomes (see more detailed discussion in Section 2.2), our paper is among the first that systematically investigate the long-term consequences of in-utero exposure to agricultural fires. One related paper is Carneiro et al. (2024), which examines the effects of in-utero exposure to agricultural fires on students' test scores. Our paper differs in providing more comprehensive evidence on how in-utero agricultural fire exposure leads to health and (non-)cognitive consequences in adolescence and labor market outcomes in adulthood, and sheds light on the potential mechanisms. More broadly, we contribute to the literature that estimates the long-term effects of early-life pollution exposure (Chen et al. 2013; Isen et al. 2017; Ebenstein et al. 2017; Anderson 2020; Barreca et al. 2021). While most of the existing studies focus on the effects of air pollution on urban residents, there is a lack of research that estimates the long-term consequences of pollution exposure for the rural sample, who are more vulnerable to pollution exposure due to income volatility and limited access to health facilities.

Second, we contribute to the literature by empirically examining the mechanisms through which in-utero exposure to air pollution can have long-term consequences (Currie et al. 2014). While there is a vast strand of literature that estimates the long-term consequences of early-life/prenatal pollution exposure (Bharadwaj et al. 2017; Isen et al. 2017; Black et al. 2019; Rosales-Rueda and Triyana 2019; Von Hinke and Sørensen 2023; Ferro et al. 2024; Chen 2025), surprisingly, only a scant amount of literature investigates the potential mechanisms. We add to this broad literature by providing the first empirical evidence on how prenatal exposure to air pollution affects the human capital investment in rural China. Our finding reveals that rural households make reinforcing investments in their children, which suggests an amplification of the effects of in-utero exposure to air pollution. In doing so, we also echo the emerging literature on how early life shocks affect intra-household human capital investment and formation (Yi et al. 2015; Adhvaryu and Nyshadham 2016; Bharadwaj et al. 2018). While the empirical evidence is mixed on how families make compensatory or reinforcing investments in response to early life shocks, our finding suggests that the liquidity constraints and limited access to health insurance may be the reasons that explain why rural households make reinforcing investments in children exposed to prenatal air pollution.[2]

Lastly, we contribute to the literature that investigates the impacts of public health insurance, more specifically, the provision of the New Cooperative Medical Scheme (NCMS) in rural China (Lei and Lin 2009; Wagstaff et al. 2009; Chen and Jin 2012; Cheng et al. 2015; Gruber et al. 2023; Huang and Liu 2023; Wang et al. 2024). While the majority of the literature investigates the potential benefits of health insurance coverage (e.g., increased consumption, education, and reduced mortality), our paper highlights additional benefits of how the coverage of health insurance can mitigate the adverse effects of in-utero air pollution exposure. In doing so, we add to the recent literature that examines how later intervention can mitigate the negative effects of early life shocks (Billings and Schnepel 2018; Duque et al. 2019). Our finding suggests that the provision of health insurance can increase parental investment, which largely mitigates the deleterious effects of in-utero pollution exposure.

[2] Focusing on China, Yi et al. (2015) show that parents act as a net equalizer in which they increase health investment and reduce education investment to children who suffer from adverse early health shocks, while our finding implies that for children who exposed to prenatal air pollution, family reduces both health and education investment, possibly because fatal pollution exposure can have both health and cognitive consequences. Using data from Chile, Bharadwaj et al. (2018) find that parents make compensatory investments regarding initial health. Leveraging a large-scale iodine supplementation program in Tanzania, Adhvaryu and Nyshadham (2016) indicate that children with higher program exposure receive more parental investment, which favors the compensatory mechanism.

2. Background and Literature

2.1 Agricultural fires and air pollution

In rural China, fire has long been a widely used technology for clearing the fields since ancient times. The earliest written record of the use of agricultural fires, found in the "Fundamental Arts for the People's Welfare" (QIMIN Yaoshu in Chinese), dates back to the North Wei dynasty (386-534 AD). It is also commonly believed that the ashes from burning the crop residues can fertilize the soil (Hays et al. 2005), while the heat generated during the burning process can eliminate the hidden pests (Graff Zivin et al. 2020; He et al. 2020; Nian 2023). Following the abolition of the People's Commune in the 1980s, grain production in China increased rapidly (Lin 1992), and the country has since become the largest producer of both grain and straw globally. Wheat, maize, and rice are the primary sources of straw, contributing to over 80% of the total straw production in China.[3]

The rapid increase in grain production has, however, created significant challenges in the disposal and management of crop residues. Due to the benefits of crop burning for agricultural cultivation and production, approximately 31% of crop residues are burned *in situ* (Graff Zivin et al. 2020). However, the burning of these residues generates considerable particulate matter, particularly $PM_{2.5}$, contributing to elevated air pollution (Rangel and Vogl 2019; He et al. 2020; Garg et al. 2024).[4] For example, He et al. (2020) document that 10 additional agricultural fires will lead to a 4.79 μg/m^3 increase in monthly $PM_{2.5}$. Shi et al. (2014) show that in agricultural production areas during the harvest season, the share of fine particulate matter emitted from agricultural fires exceeds more than 50% of the total regional emissions and that pollutant emissions from burning significantly increased the occurrence of regional haze. Despite regulations on crop burning being introduced as early as the 1990s, their enforcement has remained ineffective due to the high costs of monitoring and enforcement (Nian 2023). It was not until 2013 that the Chinese government launched a new round of campaign-style regulation that aimed to comprehensively reduce the number of agricultural fires (Wang et al. 2022; Cao and Ma 2023).

Besides China, the burning of agricultural biomass is a common phenomenon in other developing countries, especially those growing staple foods (e.g., India, Thailand, Vietnam, and the Philippines). It is estimated that every year, after the rice harvest, about 2.5 million farmers in northwestern India burn the remaining straw *in situ* (Keil et al. 2021). And study by Kim Oanh et al. (2018) indicates that the air pollution caused by agricultural fires in Vietnam and the Philippines has already exceeded the pollution caused by forest fires.

Several attributes of agricultural fires make them an intriguing source of air pollution and have drawn growing attention in the literature. First, unlike previous studies that exploit either natural or quasi-experiments as the source of variation to examine the effects of air pollution (Chay and Greenstone 2003; Almond et al. 2009; Sanders 2012; Chen et al. 2013; Isen et al. 2017; Gong et al. 2023), the burning of agricultural biomass is a seasonal and regular activity amid the production of agricultural goods (e.g., wheat, maize, and rice). Moreover, the pollution generated from agricultural fires is relatively lower than traditional sources of industrial pollution (Rangel and Vogl 2019). As numerous studies have indicated that even pollution well below the safety standards can have detrimental effects on human health (Currie et al. 2009; Aizer et al. 2018), it is thus important to understand the effects and magnitudes of such seasonal and (relatively) low-

[3] According to the World Bank data, China's agricultural value added accounted for 31.1 percent of the world's total agricultural value added in 2021. See https://www.gov.cn/xinwen/2022-11/02/content_5723319.htm.

[4] Air pollution from agricultural burning is estimated to involve around 30 million people globally (Landrigan et al. 2018).

level pollution. Second, while the literature provides significant insights into the adverse effects of industrial air pollution (Ebenstein 2012; Greenstone and Hanna 2014; Hanna and Oliva 2015; Anderson 2020; Bombardini and Li 2020; Barreca et al. 2021), much less is known about the potential costs of agricultural air pollution. This is, nevertheless an important issue given that rural populations, often facing pervasive poverty and limited access to public health insurance, may be more vulnerable to the deleterious effects of pollution.

2.2 Fire-induced pollution and related consequences

There is now a burgeoning strand of literature that investigates the broad impacts of fire-induced pollution. Rangel and Vogl (2019) is the first paper in economics that studies the effects of agricultural fires and shows that in-utero exposure to agricultural fires increases infant mortality in Brazil. Carneiro et al. (2024) extend the results from Rangel and Vogl (2019) and show that in-utero exposure to agricultural fires reduces adolescents' cognitive ability. Focusing on China, He et al. (2020) investigate the impacts of short-term exposure to agricultural fires and elderly mortality, while Lai et al. (2022) find that air pollution from straw burning significantly decreases the cognitive function of the elderly. Graff Zivin et al. (2020) studied the effects of agricultural fires that occurred during high-stakes exams on students' performance and find that temporary exposure to agricultural fires significantly decreases students' scores. Using exogenous variation from the construction of rural roads in India, Garg et al. (2024) find that rural roads increase labor exit and result in more crop fires, which in turn lead to an increase in infant mortality in the downwind region. Using detailed data from blood pressure testing, Pullabhotla and Souza (2022) find that the number of upwind fires observed one day before blood pressure testing significantly increases the risk of hypertension. Ayesh (2023) studies the impacts of burning agricultural biomass on crime activities and finds that it increases all types of crimes, particularly violent crimes.

Besides agricultural fires, there are several studies that investigate the impacts of other types of fires (e.g., forest fires). To name a few, Jayachandran (2009) studies the short-term impacts of exposure to Indonesia's forest fires and finds that prenatal exposure to air pollution caused by wildfires increases child mortality. Rosales-Rueda and Triyana (2019) investigate the persistent consequences of the 1997 Indonesian forest fire and find that it leads to significant increases in the presence of stunts and decreases in lung capacity. Borgschulte et al. (2024) and Coulombe and Rao (2025) investigate the impacts of wildfires on labor market outcomes and find that exposure to wildfires decreases local employment growth and reduces quarterly earnings. Du et al. (2024) study the effects of transboundary vegetation fire in Southeast Asian countries on expressed sentiment and find that increases in upwind fire decrease sentiment scores.

2.3 The New Cooperative Medical Scheme (NCMS)

The introduction of the New Cooperative Medical Scheme (NCMS) is a great progress of the health system in rural China (Wang et al. 2024). Before the introduction of NCMS in the 2000s, the vast population in rural China had very limited access to health insurance (either private or public), and was vulnerable to health shocks (Hu et al. 2008; Yip and Hsiao 2008). Typically, more than 90 percent of rural residents had no health insurance throughout the 1990s, with a significant number of households being pushed back into poverty due to unaffordable out-of-pocket payments for health care (Huang and Liu 2023).

The NCMS program was introduced in 2003 and progressively rolled out at the county level. Following guidelines set by the central government, each province is required to select at least two to three pilot counties in the first year of NCMS introduction (Gruber et al. 2023). Over time, additional counties were gradually incorporated, with the goal of achieving nationwide coverage by 2010. In Appendix Figure A1, we show the geographic distribution of the timing of NCMS adoption across different counties. It is evident that, following the vast expansion between 2003 and 2008, nearly all counties have adopted the NCMS. The enrollment rate amounted to 78.6% during the first three years of NCMS expansion (You and Kobayashi 2009). Within six years after its initial launch, the NCMS had expanded to cover over 800 million rural residents, making it the largest health insurance program in modern history (Gruber et al. 2023).

The NCMS program is eligible for only households with local agricultural *Hukou*.[5] Though voluntary, the enrollment rate is particularly high since it is financed by low individual contributions and high government payments, which are shared between local and central governments (Gruber et al. 2023; Huang and Liu 2023).[6] Though being a national policy, the design and implementation of the NCMS program are characterized by great discretion at the local county level. In particular, while all NCMS programs cover inpatient medical care, enrolled counties differ in their coverage for outpatient care. For example, while all counties that adopt the NCMS program offer a 50% subsidy for inpatient expenses, only approximately 80% of counties cover both inpatient and outpatient expenses (Lei and Lin 2009; Wagstaff et al. 2009; Gruber et al. 2023).[7] In addition, the coverage of outpatient care is also heterogeneous across counties. Specifically, within counties that cover outpatient expenses, 25% of counties provide direct payment for outpatient care, while the remaining 75% of counties set up mandatory medical saving accounts that would be used to pay for outpatient care (Burns and Liu 2017). The mandatory saving account is contributed to by both individuals and the government, with the sharing rule determined by local governments (Milcent 2018). Since 2007, more counties have started to incorporate outpatient expenses into the program (Huang and Liu 2023).

There is a modest but emerging strand of literature that investigates the broad impact of the NCMS program (Lei and Lin 2009; Wagstaff et al. 2009; Chen and Jin 2012; Cheng et al. 2015; Gruber et al. 2023; Huang and Liu 2023; Wang et al. 2024). For example, Chen and Jin (2012) show that the introduction of the NCMS program has significantly improved the school enrollment of six-year-olds, while having limited impact on child and maternal mortality. Exploiting a cohort difference-in-difference design, Huang and Liu (2023) document that early-life exposure to NCMS has significantly improved both the health and cognitive outcomes of rural adolescents. Focusing on the elderly, Cheng et al. (2015) find that the NCMS improved daily living activities and cognitive functions, while Gruber et al. (2023) indicate a substantial reduction in elderly mortality. We link the adoption of NCMS with air pollution from agricultural fires and investigate whether the introduction of health insurance can alleviate the long-term effects of pollution exposure.

3. Data

To estimate the effects of agricultural fires on long-term outcomes, we assemble data from multiple sources, including individual surveys that document adolescent outcomes and track cohorts into their adulthood, satellite-derived measures of agricultural potential yield, agricultural fires, air pollution, and other meteorological variables. Additionally, to investigate the mitigating role of rural health insurance, we also collect the timing of NCMS implementation across counties. In what follows, we introduce these data in turn, illustrate how we merge across different datasets, and present summary statistics. To keep the paper concise, we mainly describe the CFPS dataset and our measure of agricultural fires and potential yield in

[5] The *Hukou* system, introduced after the founding of the People's Republic of China, is a household registration system that classifies citizens into two categories: agricultural and non-agricultural *Hukou* holders. This classification plays a crucial role in determining an individual's eligibility for social services and welfare, which are tied to their place of registration. *Hukou* status is inherited from one's parents and is subject to strict government controls, making changes to *Hukou* type or registered location highly restricted.

[6] For instance, the average payment of the NCMS was 246 RMB in 2011 (approximately 35 USD), of which the government financed 84 percent, and households were only required to contribute 39 RMB annually per person.

[7] The remaining 20% counties only cover outpatient services for catastrophic diseases or do not cover outpatient services at all.

this section, and leave detailed descriptions of auxiliary data (e.g., air pollution and meteorological data) in Appendix D.

3.1 The China Family Panel Studies (CFPS)

Our primary data source is from China Family Panel Studies (CFPS), a nationwide survey data implemented by the China Social Science Survey Center of Peking University, which has been conducted biennially as a tracking survey since 2010. It covers 162 counties in 25 provinces in China, representing 94.5% of the country's total population (Xie 2012). The survey is conducted on a household basis for each member of the household, and baseline households are continuously tracked in subsequent surveys.

We rely on CFPS 2010 to explore the effects of in-utero exposure to agricultural fires on adolescent outcomes and provide supplemental evidence on adult outcomes using CFPS 2020. We include only the sample that had a local *Hukou* and were born and resided in the county at age 3 and at the time of the survey, so that the sample would be most likely to be measured with correct exposure intensity and suffer less concerns of endogenous migration. We primarily focus on the effects on the rural sample where the agricultural fire occurs, but for the following empirical exercises, we will also present the corresponding results for the urban sample for either comparison or falsification tests.

Besides birth year, CFPS 2010 additionally provides the birth month of each individual, which we exploit as a source of variation to distinguish the effects of exposure to agricultural fires during different trimesters. Specifically, we denote the last 3 months prior to the birth month as the third trimester, the 3-6 months before birth as the second trimester, and the 6-9 months before birth as the first trimester.[8]

To ensure that our results are not driven by the selection of different birth months, we visualize the distribution of birth months in Figure 1. Though not perfectly balanced, the distribution of birth months is relatively flat and we observe no significant spikes at first glance. This alleviates the potential concerns that parents may strategically choose the timing of birth to avoid pollution exposure.[9] We provide more solid statistical evidence in our subsequent empirical analysis to show that the birth month is not correlated with potential exposure to agricultural fires.

We measure the health and cognitive outcomes of adolescents using the CFPS 2010. Specifically, we measure the health outcomes of adolescents using three variables. The first variable is a categorical measure of general health status, which is self-rated and ranges from 1 to 5 (1 = very good, 2 = good, 3 = fair, 4 = poor, 5 = very poor). We define a dummy variable for not-in-good-health, which takes the value of 1 if the self-rated health status is greater than 3 (i.e., poor or very poor), and 0 otherwise. The second variable is a dummy variable indicating whether the adolescent was hospitalized due to illness in the previous year. The third variable measures the occurrence of respiratory diseases, which are closely related to air pollution.[10] To avoid issues with multiple hypothesis testing, we create an unhealthiness index by first standardizing the three health variables, and then calculating the simple average of their standardized z-scores (Boudreaux et al. 2016; Hoynes et al. 2016). We measure the cognitive abilities of adolescents using two test scores, i.e., a

[8] Since we do not have data on the exact birth date, this definition of trimesters may be measured with error. Nevertheless, as long as the date of birth is randomly distributed, such measurement errors would only lead to an underestimation.

[9] As most of our sampled individuals were born during the 1990s, we believe that the selection of birth month due to air pollution is less plausible. First, information on air quality was relatively scarce back then, and the construction of air quality monitoring stations did not begin until the 2000s, which were measured with significant error due to local discretion (Greenstone et al. 2022). Second, the public awareness of the detrimental effects of air pollution is relatively low (Xie et al. 2023; Barwick et al. 2024).

[10] We use the illness type classification provided in the CFPS to identify whether the adolescent has respiratory diseases. These include upper respiratory tract infections, pneumonia, chronic laryngitis, emphysema, other chronic obstructive pulmonary diseases (including chronic bronchitis), asthma, and other respiratory diseases.

word test score and a math test score. To ensure comparability across different age cohorts, we calculate the age-specific standardized z-scores for both test scores (Huang and Liu 2023).

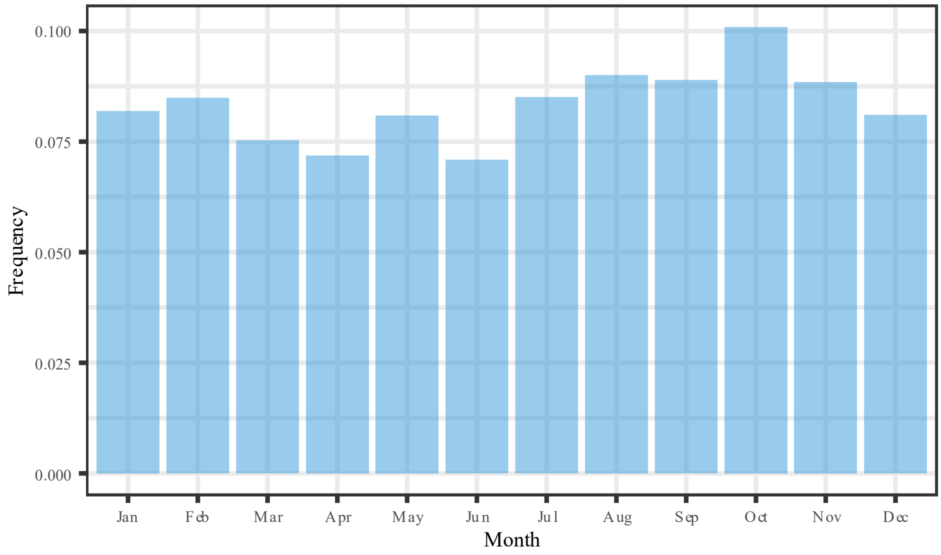


Figure 1 The Distribution of Birth Month

Notes: This figure presents the distribution of birth months of adolescents surveyed in CFPS 2010.

In addition, CFPS 2010 asks adolescents about their attitudes toward a range of statements (e.g., "I decide my own life goals" or "Some children are born lucky"), which we exploit to reflect the personal traits of the respondents. Individuals respond with their level of agreement on each statement, which we re-coded as categorical variables ranging from 1 to 5, with greater values representing a higher degree of agreement. We divide these statements into two broad categories reflecting positive and negative attitudes.[11] We then follow Grönqvist et al. (2020) and use the principal component analysis (PCA) to combine these sub-scores into two general measures of noncognitive ability (reflecting positive and negative attitudes toward life).

Besides current outcomes, CFPS 2010 also provides retrospective data on birth conditions, which enables an investigation of the effect of in-utero exposure to agricultural fires on health outcomes at birth. This serves as a potential channel for the long-term effects of fire-induced air pollution. Specifically, we measure health at birth with three outcome variables. The first is the number of illnesses at age 1. The second is the length of the gestation period (measured in months). The third is the birth weight (measured in 500 grams).

To investigate the channel of intra-household resource allocation, we use two additional measures of health expenditure and education expenses to proxy for the parental investment in adolescents' human capital, in response to the negative health shocks at birth.[12]

We include several control variables to mitigate concerns of omitted variable bias. Specifically, we control for the age and gender of adolescents, parental age, birth month, and education. We also include

[11] Specifically, statements that reflect positive attitudes include (1) "*I pursue my own values instead of following others*", (2) "*I decide my own life goals*", (3) "*Once I start something, I have to finish it no matter what*", and (4) "*I am the kind of person who believes that planning ahead will make things better*". Whereas statements that reflect negative attitudes include (1) "*Some children are born lucky*", (2) "*Don't spend too much time trying, because it will never prove to be useful*", (3) "*Once you make a mistake, it's almost impossible to correct it*", (4) "*The best way to deal with problems is not to think about them*", and (5) "*When bad things are about to happen, they are going to happen no matter how hard you try to stop them*".

[12] The corresponding questionnaires provided in the CFPS dataset are "*Last year, how much was spent in total due to the child's illness? (RMB)*" for health expenses and "*Last year, what's the total education expenses for the child? (RMB)*" for education expenses.

controls for family income, family size, and number of siblings.

To estimate the effects of in-utero exposure to agricultural fires on adulthood outcomes, we track adolescents in CFPS 2010 (aged between 10-15 at the time of survey) to CFPS 2020 using the unique individual ID.[13] We measure the effects of fire exposure on educational attainment and labor market outcomes using three variables. The first variable is the number of years of schooling completed. Since some individuals may not have finished their education by the time of the survey, we normalize this by age to calculate age-specific years of schooling. The second measure is the annual wage. Given that some individuals may not be participating in the labor market, we restrict our sample to those with non-zero earnings. The third measure is a dummy variable indicating whether the individual works in the agricultural sector, which generally requires lower skills compared to work in the manufacturing sector. Panel A of Table A1 provides summary statistics for the above variables.

3.2 Agricultural fires and potential yield

The agricultural fire data used in our paper is sourced from NASA's MODIS aboard the Terra and Aqua satellites, which have been frequently used in recent studies on agricultural fires (Cao and Ma 2023; Nian 2023). These satellites pass over China twice daily, typically occurring between 10 am and 3 pm, and between 9 pm and 2 am China Standard Time. We identify agricultural fires by leveraging land cover data from the China Land Cover Dataset (CLCD), a remotely sensed product providing nationwide land type classifications at 30-meter resolution from 1990 to 2020 (Yang and Huang 2021). We match the fire point data to the land cover raster and define fires as agricultural fires if they occur within cropland pixels. The satellites started to record fire points in November 2000, and we have no available data on fire records prior to this time point. Since the majority of our sampled individuals were born before 2000 (aged between 10-15 in 2010), this data limitation prevents us from directly estimating the effects of the number of fires on adolescent and adult outcomes.

That said, the occurrence of agricultural fire still provides useful information for our research design. Since we identify the effects of fire exposure primarily using variations in the birth month (and thus exposure to fires during different trimesters), it's essential to pin down the month with the highest frequency of agricultural fires. To do so, we calculate the total number of agricultural fires within each county-month cell from 2001 to 2019, and determine the month with the highest frequency of fires for each county (hereafter referred to as "fire month"). Figure 2 presents the distribution of fire months. Unlike the distribution of birth months, the distribution of fire months is uneven. We see three peak months in the figure: March (February), June, and October.[14] The change of fire months across different counties provides sufficient variation for our identification. By comparing the fire month with the birth month of each individual, we are able to pin down the trimester during which the individual is exposed to agricultural fires.

Since data on agricultural fires is not available for our sampled cohorts, we need a valid proxy to produce convincing estimates. To this end, we use the agricultural potential yield calculated using the Global Agro-Ecological Zones (GAEZ) model as a proxy for the occurrence of agricultural fires.[15] The data used

---

[13] The CFPS tracks the same households in each wave of the survey, allowing us to follow approximately 78% of the same individuals from 2010 to 2020. Therefore, sample attrition does not appear to be a major concern. We also provide empirical evidence that attrition is plausibly random and is not correlated with our measure of agricultural fire exposure, as shown in Appendix Table A18.

[14] The three months correspond to the spring peak of straw burning activities in the northeastern region, the summer and autumn peaks in the central and southern regions.

[15] The GAEZ model first estimates the light-temperature production potential for a crop based on temperature and solar radiation, then combines water availability, soil properties, and topography to estimate the light-temperature-water production potential. It simulates the climatic production potential under ideal conditions and, considering factors like agricultural

in this paper is from Liu et al. (2015), who constructed the potential yield raster for China at 1km resolution using different crops.[16,17] The data is generated by combining *cell*-specific land quality attributes with established agronomic models for a given level of water supply and cultivation inputs. Potential yields summarize how detailed geographical attributes translate into productivity. The top three crops used for data construction are wheat, maize, and rice, which are the main contributors to crop residues.

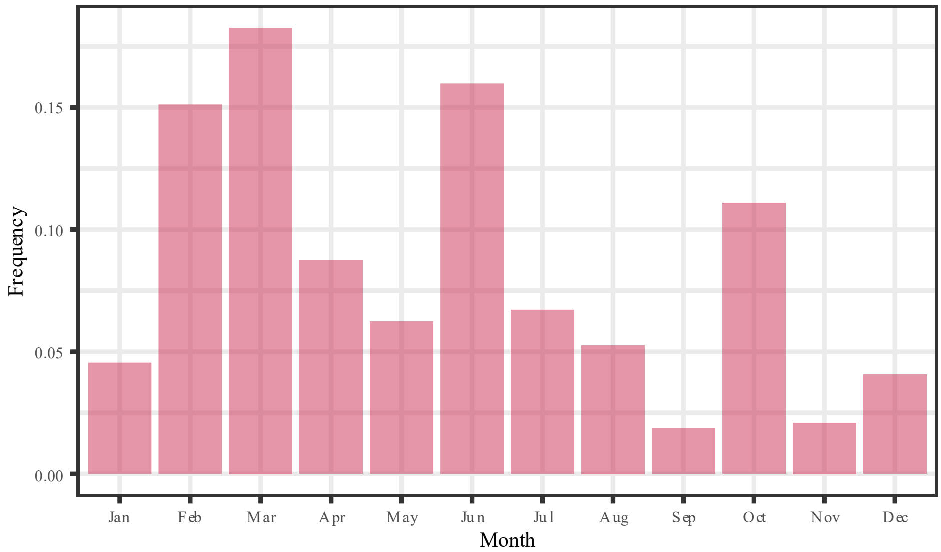


Figure 2 The Distribution of the Fire Month

Notes: This figure presents the distribution of fire months, which is defined as the month with the highest agricultural fire frequency.

Formally, for the potential yield to be a valid proxy for agricultural fire and air pollution, we processed the data with the following steps. We first create a 1km×1km grid covering the entire China's territory, and map the grid with the potential yield raster as well as the county shapefile to determine the relative location of the potential yield grid to the county center. We then leverage the wind direction (described in Appendix D) during each county's fire month and determine whether a specific grid is located in the upwind or downwind direction of a county. In our baseline specification, we use a criterion of 45 degrees to determine whether a grid is located in the upwind, downwind, or non-wind direction.[18] Appendix Figure A2 gives an illustration of how we define the upwind direction. For example, if the absolute difference between the angle of a specific grid to the county center and the wind direction is less than 45 degrees, then the grid is classified as in the upwind direction.

Finally, we calculate the county-level average potential yield for both upwind grids and non-upwind grids. If the potential yield is indeed a valid proxy for agricultural fires and air pollution, we should find strong evidence that the potential yield is positively correlated with both the number of agricultural fires and air pollution. Moreover, we should observe that, while both upwind and non-upwind potential yield correlate with agricultural fire occurrence, only upwind potential yield should have statistically significant predictive power on air pollution. We empirically provide valid support for these hypotheses in the next section.

technology and arable land distribution, calculates the food production potential of each raster using a step-by-step limiting method.

[16] For more details on the data description, see https://www.resdc.cn/DOI/doi.aspx?DOIid=43.

[17] Liu, Xu, and Chen (2015) construct the potential yield data from 1970 to 2010 (in 10-year intervals). To avoid potential endogeneity, we use the potential yield data measured in 1990 and use data measured in alternative years as robustness checks.

[18] Alternative definitions of upwind direction are used as robustness checks.

3.3 NCMS rollout

We manually collect the data on the implementation timing of NCMS across counties from multiple sources, including news and media coverage, government announcements, and other documents. Appendix Figure A1 provides the geographic distribution of the rollout timing of the NCMS policy. To define whether the individual is exposed to the NCMS policy, we follow Huang and Liu (2023) and denote individuals who were less than 5 years old when the NCMS policy was implemented as those exposed to the policy. Appendix Table A2 provides summary statistics and balance tests between cohorts exposed and not exposed to the NCMS policy, finding limited evidence that exposure to the NCMS is based on the selection of individual covariates.[19] To control for other concurrent early-life exposures that may confound the effects of NCMS exposure, we also control for some village-specific characteristics (e.g., accessibility to infrastructure, education, and health facilities) and other village-cohort confounders, e.g., early-life exposure to tap water (Chen et al. 2022; Li and Xiao 2023).

3.4 Summary statistics

Appendix Table A1 presents the summary statistics for the main variables used in our empirical analysis. Panel A provides summary statistics for individual variables, separately for rural and urban samples. As shown in the table, rural adolescents are generally more disadvantaged in terms of health and cognitive outcomes compared with urban adolescents. We see that rural adolescents report that they are more likely to feel in bad health (twice as the urban sample), while reporting (slightly) fewer hospital admissions than the urban adolescents. This may be the result of inadequate access to health facilities. In general, rural adolescents are less likely to have respiratory disease than urban adolescents. This is plausible given that the air quality in urban regions is worse than that in rural areas. We also note that urban adolescents, on average, have better cognitive performance than rural adolescents. Tracking these adolescents ten years later, we find that rural individuals are more likely to complete less education and are more likely to enter the labor market: more than half of the rural sample have already started to work in 2020, while only a quarter of the urban sample have. Finally, the rural sample is more likely to stay in the agricultural sector.

We also find that there is a relative balance in which trimesters individuals were exposed to agricultural fires. Ideally, the probability that an individual was exposed to agricultural fires at a specific trimester is 0.25 if both fire month and birth month are randomly distributed. The summarized mean is close to this probability, and we observe no significant difference between the rural and the urban sample, both of which suggest that the selection of birth month is less likely to occur.

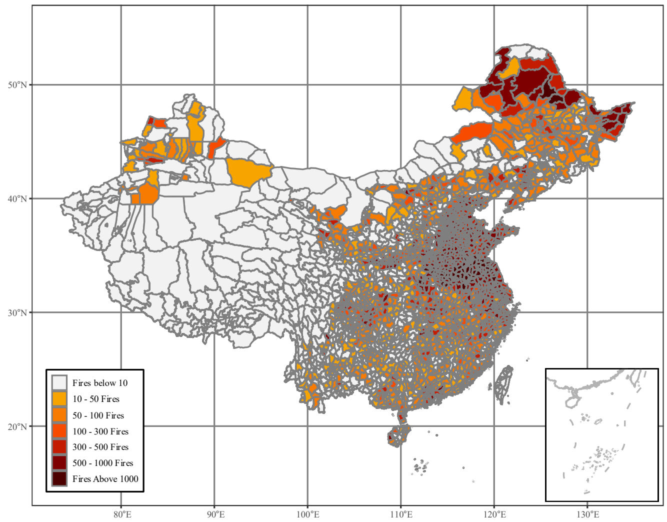
We present summary statistics for county variables in Panel B of Table A1. Except for potential yield variables, all variables are defined at the county-year-month level. Data on agricultural fires is only available from 2001 onward, while data on $PM_{2.5}$ and other meteorological variables are available from 1990 onward. For agricultural potential yield, we report both the upwind and downwind potential yield. Not surprisingly, the two variables are nearly identical regarding the mean and the standard deviation, as the wind direction at the fire month should be (and indeed is) orthogonal to the distribution of potential yield within the county.

4. Validation of Fire Measures

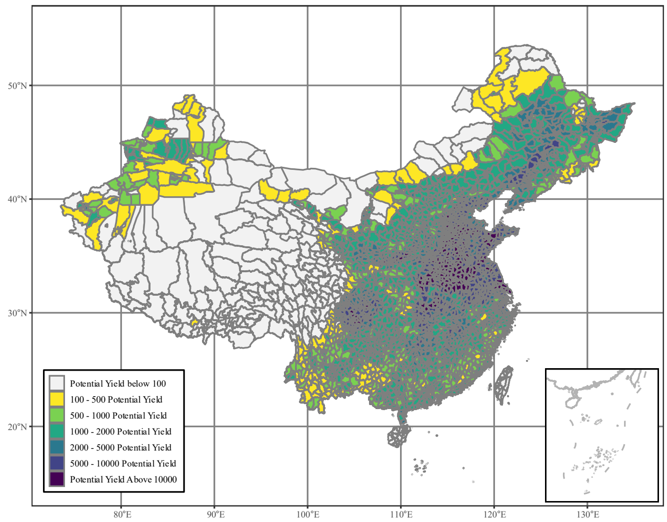
This section provides evidence that the agricultural potential yield is a valid proxy for agricultural fire

[19] Since the timing of the NCMS rollout is plausibly not randomly assigned (Gruber et al. 2023), we also report the results of the balance tests conditional on a set of county characteristics. Conditional on these additional characteristics does not alter our results. See Appendix A for more details.

as well as fire-induced air pollution. To start with, Figure 3 displays the geographic distribution of agricultural fires (Panel A) and potential yield (Panel B). We observe a high correlation between the two variables. This is especially evident in the central region, where most counties are highly suitable for agricultural production and have more burning activities.



Panel A Distribution of agricultural fires



Panel B Distribution of potential yield (kg/ha)

Figure 3 Geographic Distribution of Agricultural Fires and Potential Yield

Notes: This figure presents the geographic distribution of agricultural fires (Panel A) and potential yield (Panel B). Darker colors in Panel A correspond to more agricultural fires, while lighter colors in Panel B correspond to higher potential yield.

To lend further support, Figure 4 presents a binscatter plot showing the correlation between agricultural fires and potential yield. In addition to the spatial correlation documented in Figure 3, we also find strong evidence for a linear relationship between the two variables. Taken together, the above graphical evidence provides preliminary yet persuasive support for the strong correlation between agricultural fire and potential yield.

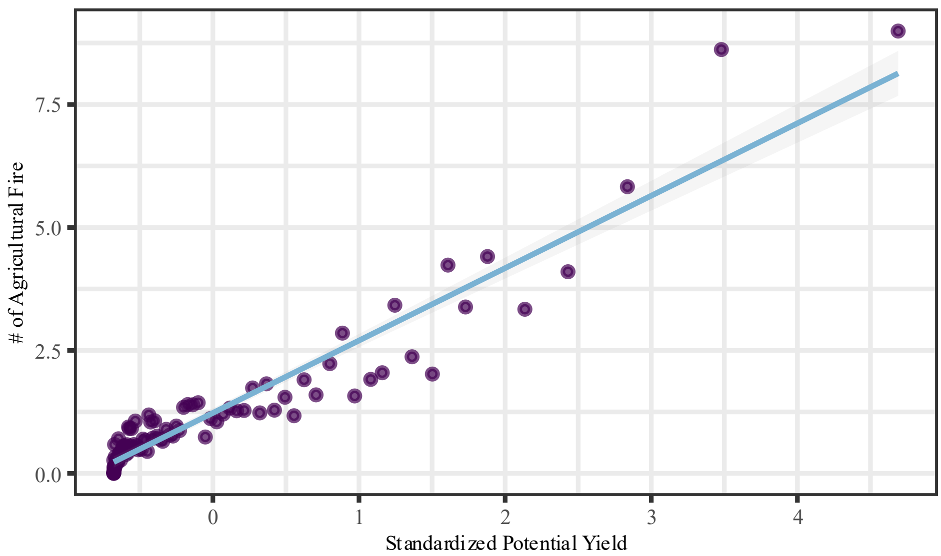


Figure 4 Binscatter of Agricultural Fires versus Potential Yield

Notes: This figure presents the binscatter plot for agricultural fires versus potential yield, with a fitted line colored in blue. The shaded Area is the 95% confidence interval. We standardize the potential yield to have a mean of 0 and a standard deviation of 1.

Formally, we carry out the following regression specification to examine the correlation between agricultural fires and potential yield:

$$fire_{cpmt} = \beta_0 + \beta_1 APY_c + \Gamma \boldsymbol{W_{cpmt}} + \gamma_{pt} + \gamma_{pm} + \gamma_{mt} + \epsilon_{cpmt} \tag{1}$$

Where y_{icmt} is the number of agricultural fires in county c and prefecture p that are observed in month m and year t. APY_c is the agricultural potential yield in county c, which is time-invariant. We standardize the variable to have a mean of 0 and a standard deviation of 1, so that the estimated coefficient, β_1, measures the effects of increasing the agricultural potential yield by one SD on the number of agricultural fires. We control for a set of time-varying meteorological covariates, $\boldsymbol{W_{icmt}}$, which include dew point, sea level pressure, wind speed, temperature, and rainfall. To net out potential confounders and obtain a more precise estimate, we include a detailed set of fixed effects. Specifically, we control for the prefecture-year fixed effects, γ_{pt}, the prefecture-month fixed effects, γ_{pm}, and the prefecture-year fixed effects, γ_{mt}. The inclusion of these fixed effects absorbs substantial time-varying variations at the prefecture level, and allows us to only use within-prefecture variation to identify the effects of agricultural potential yield on the number of agricultural fires.[20] We cluster the standard error at the prefecture level.

Table 1 reports the results estimated using equation (1). We estimate the equation using both least squares and the Poisson Pseudo Maximum Likelihood (PPML) specification. The PPML specification performs well in cases where the dependent variable is non-negative and has possibly many zeros, which is exactly the case with the fire observations.[21] Specifically, the estimated coefficient from column (1) of Table 1 suggests that a one SD increase in potential yield (corresponding to an increase in potential grain output by 2864 kg/ha) is associated with an additional 1.71 agricultural fires per month. This result remains largely unchanged when including meteorological controls in column (2). Considering that on average a county

[20] Due to the time-invariant nature of agricultural potential yield, we cannot control for any county-level fixed effects. We base our fixed effects at the city level as it is the immediate upper administrative division of the county.

[21] In the presence of potential zero values, log-like transformations may not be suitable as the logarithm of zero is undefined, and adding arbitrary constants to address this issue could introduce additional bias (Chen and Roth 2023). The PPML estimation avoids such concerns by directly modeling the non-negative dependent variable in levels, while still allowing for an interpretation in terms of proportional or percentage changes.

experiences 1.18 monthly agricultural fires, our estimates imply that a one SD increase in potential yield would nearly double the number of agricultural fires, representing an increase of approximately 145%. Turning to the results from PPML in columns (3) and (4), we find that a one SD increase in potential yield is associated with an approximately 48.7% increase in agricultural fires.[22] In sum, the results from Table 1 confirm the significant and robust correlation between agricultural potential yield and agricultural fires.[23]

Table 1 The Effects of Agricultural Potential Yield on Agricultural Fires

Dep. Var. # Agri. Fire	(1)	(2)	(3)	(4)
	OLS		PPML	
APY	1.711***	1.722***	0.473***	0.487***
	(0.263)	(0.267)	(0.048)	(0.047)
Observations	660,972	660,972	603,024	603,024
Prefecture-Year FE	Yes	Yes	Yes	Yes
Prefecture-Month FE	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	No	Yes
Dep. Var. Mean	1.181	1.181	1.181	1.181
Adjusted/Pseudo R-squared	0.196	0.199	0.620	0.633

Notes: This table presents the estimated results of the effects of agricultural potential yield on agricultural fires. The observation is at the county-year-month level. The sample period is from 2001 to 2019. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. Standard error is clustered at the prefecture level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

We then explore the relation between agricultural potential yield and air pollution. To do so, we modify our specification in equation (1) to allow for the differential effects of upwind and downwind potential yield. The idea is motivated by previous studies that identify the effects of agricultural fires on air pollution (e.g., Rangel and Vogl 2019; He, Liu, and Zhou 2020). Since pollution within a county is more likely to be caused by upwind fires, we should find strong correlations between upwind potential yield and air pollution, while observing weak correlations between downwind potential yield and air pollution. Specifically, we estimate the following specification:

$$fire_{cpmt} = \beta_0 + \beta_1 UpwindAPY_c + \beta_2 DownwindAPY_c + \Gamma \boldsymbol{W_{icmt}} + \gamma_{pt} + \gamma_{pm} + \gamma_{mt} + \epsilon_{cpmt} \quad (2)$$

Where $UpwindAPY_c$ and $DownwindAPY_c$ represent the average potential yield in the upwind and downwind regions of the county, respectively. We determine the upwind and downwind regions by leveraging the dominant wind direction during the fire month. If the potential yield is indeed a valid measure, then we should estimate a significant result for β_1 and an insignificant result for β_2.

Table 2 presents the corresponding results. In columns (1) and (2), we first confirm that both upwind and downwind potential yield have significant predictive power on the number of agricultural fires. We also show that the coefficients are quantitatively analogous, which suggests that the two variables are balanced in terms of the effects on agricultural fires. Next, in columns (3) and (4), we replace the dependent variable with the monthly $PM_{2.5}$ and find a strong correlation between agricultural potential yield and air pollution. This is not surprising given the significant effects of agricultural fires on $PM_{2.5}$. Finally, in columns (5) and (6), we regress the monthly $PM_{2.5}$ on both upwind and downwind potential yield. Aligning with our expectation, we show that only upwind potential yield has significant predictive power on air pollution, while finding an exactly null effect of downwind potential yield on air pollution. Specifically, the estimated coefficients in Table 2 suggest that a one SD increase in upwind potential yield is associated with a 0.82 increase

[22] Note that the coefficients from OLS are not directly comparable to those from PPML as β_{ols} reflects the absolute change while β_{ppml} only reflects changes relative to the conditional mean of the dependent variable.

[23] To mitigate concerns of omitted variable bias and spurious correlation , in Appendix Table A3, we use non-agricultural fires as a placebo test. Reassuringly, we find no evidence that agricultural potential yield is associated with non-agricultural fires, the estimated coefficients are small in magnitude and insignificant.

in monthly agricultural fire and a 0.56 μg/m^3 increase in monthly $PM_{2.5}$. Put differently, we find that a per 10 points increase in agricultural fires is associated with an increase of monthly $PM_{2.5}$ by 6.83 μg/m^3, which is very similar to the estimates from He, Liu, and Zhou (2020).[24]

Table 2 The Effects of Upwind/Downwind Potential Yield on Agricultural Fires and $PM_{2.5}$

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	# Agri. Fire		$PM_{2.5}$			
Upwind APY	0.822***	0.817***			0.667***	0.564***
	(0.306)	(0.307)			(0.150)	(0.148)
Downwind APY	0.943***	0.959***			-0.015	-0.008
	(0.328)	(0.330)			(0.169)	(0.167)
APY			0.762***	0.630***		
			(0.089)	(0.093)		
Observations	660,972	660,972	1,018,800	1,018,800	1,018,800	1,018,800
Prefecture by Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture by Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	No	Yes	No	Yes
Dep. Var. Mean	1.181	1.181	63.81	63.81	63.81	63.81
Adjusted R-squared	0.196	0.199	0.846	0.847	0.846	0.847

Notes: This table presents the estimated results of the effects of upwind/downwind agricultural potential yield on agricultural fires and $PM_{2.5}$. The observation is at the county-year-month level. The sample period for agricultural fires is from 2001 to 2019, whereas the sample period for $PM_{2.5}$ is from 1980 to 2019. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. Standard error is clustered at the prefecture level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

To ensure that our estimated relationship between upwind/downwind potential yield is indeed driven by the occurrence of agricultural fires, we examine the effects of potential yield on other pollutants. If there are unobservables that drive the correlation between agricultural potential yield and air pollution, then we should find similar effects on other pollutants. But if the correlation is solely driven by the occurrence of agricultural fires, then we should expect to find no effects on other air pollutants. From County Statistical Yearbooks, we derive three common air pollutants (i.e., NO_x, SO_2, and Dust) that are unlikely to be correlated with agricultural fires. We then re-run both equations (1) and (2) to examine whether they are correlated with potential yield. Columns (1) through (6) in Appendix Table A4 report the corresponding results. Consistently, we detect no significant effect for all three air pollutants. Another related concern is the potential effect of CO_2, as recent studies have pointed out that ambient CO_2 can affect agricultural yield (Liu 2025). But since our measurement is the maximum attainable agricultural yield within a specific cell, which is largely determined by soil quality, rainfall, and other natural conditions, changes in CO_2 should be less likely to impact the potential yield (though it affects the actual yield). We show in columns (7) and (8) of Table A4 that CO_2 emissions have no significant correlation with our potential yield measurement.

Taken together, the above analyses offer valid support that the upwind and downwind agricultural potential yields are able to provide valid and exogenous variation to identify the effects of fire-induced pollution on adolescent outcomes. Moreover, the difference in upwind-downwind coefficients allows us to effectively rule out the potential income effects associated with agricultural fires (Rangel and Vogl 2019; He et al. 2020). We will illustrate this point in further detail in the next section.

5. The Impacts of Agricultural Fires

This section outlines our empirical strategy, presents our main findings and robustness, and discusses

[24] Typically, using monthly data in the summer season, they show that per 10 points increase in agricultural fires increases the monthly $PM_{2.5}$ by 4.43-5.03 μg/m^3.

several heterogeneity results. Specifically, we introduce how we build on our previous findings in Section 4 to construct a valid exogenous variation. We then rely on the empirical framework to identify the health effects of in-utero exposure to agricultural fires and explore the robustness and heterogeneity of our findings. Despite health consequences, we also shed light on other outcomes, for instance, the development of cognitive and non-cognitive ability. Finally, we track these cohorts ten years later and explore the effects on education and labor market outcomes.

5.1 Empirical specification

Our empirical strategy exploits three sources of variation. The first is the difference between the fire month and the birth month, which we leverage to identify during which trimester the individual is exposed to agricultural fires. The second is cross-sectional variations from the differences in agricultural potential yield across counties, which we exploit to proxy for fire intensity. The third variation comes from changes in wind direction during the fire month, which we use to conduct upwind-downwind comparison that avoids any confounding factors that are systematically correlated with agricultural production. Specifically, we use the following specification:

$$\begin{aligned} y_{icmt} &= \alpha + \sum_{\tau=1}^{3} \beta_\tau^U UPwind\ APY_c \times \bigcap\{Trimester_i = \tau\} + \sum_{\tau=1}^{3} \beta_\tau^D Downwind\ APY_c \times \bigcap\{Trimester_i = \tau\} \\ &+ \Lambda \boldsymbol{X}_{icmt} + \sum_{\tau=1}^{3} \Gamma_\tau \boldsymbol{W}_{cmt,\{Trimester_i=\tau\}} + \gamma_{mt} + \gamma_c + \epsilon_{icmt} \end{aligned} \tag{3}$$

Where y_{icmt} denotes the outcome of individual i living in county c that was born in month m and year t. $UPwind\ APY_c$ and $Downwind\ APY_c$ are similarly defined as in equation (2). $\bigcap\{\cdot\}$ is an indication function, which denotes during which trimester the individual is exposed to agricultural fire. The trimesters are determined by the relative distance between fire month and birth month. For example, if an individual is born in December and the fire month in her county of residence is November, then we define that she is exposed to agricultural fire during the third trimester. In practice, we first identify the specific months corresponding to the three trimesters based on an individual's birth month, then determine which trimester the fire month falls within. Individuals not exposed to agricultural fires during their whole gestation period are left as the comparison group. $\boldsymbol{X}_{icmt}$ denotes a set of individual covariates, which include gender, age, father's and mother's age, education, family income, family size, and number of siblings. $\boldsymbol{W}_{cmt,\{Trimester_i=\tau\}}$ is individual-birth year-trimester specific controls that account for the confounding effects of weather conditions. Typically, it includes a set of meteorological controls (i.e., temperature, rainfall, humidity, sea level pressure, and dew point), which we average to county-trimester-year level and then match with each individual based on their birth month, birth year, and county of birth. γ_{mt} and γ_c are birth year-month and county fixed effects, respectively. Finally, the standard error is clustered at the county level to account for any unobserved arbitrary correlations within the county.

The parameters of interest are β_τ^Us and β_τ^Ds, which capture the effects of upwind and downwind potential yield (and hence agricultural fires). However, neither of the coefficients has a causal interpretation. Since agricultural fires typically occur shortly after harvesting, this, in turn, could generate substantial income effects that confound the true effects of prenatal agricultural fire exposure (Rangel and Vogl 2019; He et al. 2020). To partial out such income effects, we exploit the fact that only the upwind agricultural potential yield is correlated with air pollution and that the wind direction is plausibly uncorrelated with agricultural production. Therefore, although both β_τ^U and β_τ^D are confounded by income effects, the magnitudes of such effects should be quantitatively the same, and hence a difference between the two coefficients should efficiently remove the confounding income effects.[25] To this end, we focus on the estimation of parameter θ_τ, which is defined as:

$$\theta_\tau = \beta_\tau^U - \beta_\tau^D.$$

Valid identification of θ_τ requires two additional assumptions. The first assumption is that wind direction is orthogonal to potential income effects. To support this assumption, we re-estimate equation (2), with the dependent variable replaced by a set of outcomes that are correlated with agricultural production (e.g., grain output, rural income, agricultural employment, and agricultural GDP).[26] Appendix Table A5 reports the results. We additionally report the estimated results for the differences between upwind and downwind coefficients to examine whether the effects of potential yield are statistically different with respect to wind direction. While we find that both upwind and downwind agricultural potential yields are significantly correlated with all four measures of agricultural production, the estimated coefficients for their differences are small in magnitude and statistically insignificant. This piece of evidence provides support that the upwind-downwind comparison can indeed effectively partial out the confounding income effects.

The second assumption is that there is no selection into the birth month. If households are aware of the negative impacts of fire exposure, then parents may reschedule their timing of reproduction to avoid in-utero pollution exposure *ex-ante*. Such endogenous selections may bias our results by overestimating the true effects of agricultural fire exposure. To ensure that this is not the case, we examine whether agricultural fire impacts the decision on the birth month. To do so, we regress three dummies indicating during which trimesters individuals are exposed to agricultural fires on upwind and downwind potential yields. If households are indeed responding to fire-induced air pollution, then we should find significantly negative correlations in these regressions. Appendix Table A6 reports the corresponding results, based on the regression specification in equation (2).[27] Reassuringly, we find no evidence supporting the presence of birth month selection, with all coefficients being insignificant. In Appendix Table A7, we further examine the effects of agricultural fire intensity on fertility decisions. A recent study by Gao et al. (2024) finds that pollution exposure may distort the fertility decision, therefore, another related concern is the potential fertility selection. We show that this is not the case in the rural sample. Typically, results from column (1) of Table A7 suggest that there are no significant correlations between upwind/downwind agricultural potential yield and the number of children. In columns (2) and (3), we focus on the number of boys and girls. Though there are significantly negative correlations between upwind potential yield and the number of girls, the effects vanished after accounting for the potential income effects. Taken together, the above exercises provide valid support for our identification strategy.

Before we proceed to present our baseline results, several caveats of our research design should be borne in mind. First, the potential yield only varies cross-sectionally but has no variation in the time-series dimension. Thus, it only measures the average intensity of agricultural fires across different counties but cannot account for the variability of fire occurrence across different years. While this limitation prevents us from precisely identifying the treatment effects for cohorts born in different years, it does not essentially affect the estimation for cohorts born in different months, since we mainly exploit the within-year variation (i.e., differences in birth month and fire month) for identification. Since straw burning is more of a seasonal activity, we believe that our measurement still captures important variations of fire intensity.[28] Second, our

[25] For this reason, in our subsequent analyses, we only report the estimated results for the differences between β_τ^U and β_τ^D, a la Rangel and Vogl (2019).

[26] The data is derived from County Statistical Yearbooks, 2000-2019.

[27] As agricultural potential yield is time-invariant at the county level, our regression only controls for birth year by birth month fixed effects and prefecture fixed effects, and clusters the standard errors at the prefecture level.

[28] An alternative way to exploit yearly variations is to leverage monthly wind direction across different years. However, as we show in Appendix C.1 and C.6, the across-year wind direction variation is relatively small and our definition of

indirect measure of agricultural fires only allows us to identify the concurrent effects of in-utero exposure to fire-induced pollution, i.e., pollution that occurred during fire month. Ideally, to fully characterize the effects of in-utero pollution exposure, it would be better to use cumulative pollution or fire exposure. However, this is infeasible in our econometric framework, and our empirical identification only partially captures such effects. Therefore, our estimated effects are better interpreted as a lower bound of the true effects.

5.2 The effects of fire exposure on adolescent outcomes

5.2.1 Health outcomes

Unhealthiness index. Table 3 presents our baseline estimates on the effects of in-utero agricultural fire exposure on adolescent health. The dependent variable is a normalized health index with a greater value representing worse health conditions (see Section 3.1 for variable construction). Column (1) reports the parsimonious specification where we include birth year, birth month, and county of birth fixed effects while controlling for individual characteristics. To control for the confounding effects of prenatal weather conditions, in column (2), we include a set of meteorological controls. Column (3) augments the identification by further controlling for the birth year by month fixed effects. This allows us to effectively partial out any within-year seasonal activity and unobserved time trends. Across different specifications, the estimated coefficients are relatively stable, suggesting less concern about the selection on unobservables (Altonji et al. 2005; Oster 2019).

Regarding coefficients, we find that the effects are more pronounced when individuals are exposed to agricultural fires during the first and third trimesters, which is in line with findings from epidemiological and economic literature that the effects of pollution are more deleterious in the early and later gestation period (Glinianaia et al. 2004; Šrám et al. 2005; Currie and Neidell 2005; Kannan et al. 2006; Currie et al. 2014; Rangel and Vogl 2019).[29,30] Specifically, we estimate that a one SD increase in agricultural potential yield during the first and third trimesters (which approximately corresponds to a 0.8-0.9 unit increase in monthly agricultural fires) is associated with a 1.43 and 1.52 unit increase in the unhealthiness index, respectively. Given that the SD of the outcome is 2.01, this is equivalent to an increase of 0.71 and 0.76 SD. The effects of exposure during the second trimester are smaller in magnitude, which suggests that a one SD increase in potential yield increases the unhealthiness index by 0.40 SD. Given that a one SD increase in potential yield is associated with an increase of $PM_{2.5}$ by 0.76 μg/m^3, our results imply that even a small amount of pollution exposure during the gestation period can lead to substantial long-term consequences.

In columns (4) and (5), we explore the gender heterogeneity in pollution exposure. We find that the effects are mostly concentrated on boys. This also aligns with previous epidemiological literature that male fetuses grow faster and have higher metabolic demands, making them more susceptible and vulnerable to oxidative stress and nutrient deprivation caused by air pollution. Typically, our estimates suggest that the effects of the same amount of agricultural fire exposure can worsen the health outcome of boys by more than 40% compared to girls.

upwind region can mostly account for such fluctuation.

[29] In the first trimester, rapid organogenesis and placental development make the fetus highly vulnerable to structural abnormalities, epigenetic changes, and impaired growth caused by pollution. During the third trimester, rapid fetal growth, brain development, and immune system maturation increase susceptibility to oxidative stress, low birth weight, and preterm birth. The second trimester is relatively less affected as it is a more stable developmental phase. Thus, pollution exposure during the first and third trimesters poses greater risks to fetal health and long-term outcomes.

[30] This evidence also suggests that the persistence of early life exposure is an important channel through which in-utero fire exposure impacts adolescent health outcomes. We discuss this mechanism later in the section.

Our estimates are quantitatively larger than previous findings for several possible reasons. First, since we're focusing on the rural population during an early period (most individuals in our sample were born between 1995 and 2000), the impact of prenatal exposure to air pollution may have been more severe due to insufficient awareness of the hazards of pollution exposure and limited access to adaptation measures. Second, agricultural fire is featured by great seasonality, which could generate substantial air pollution within a relatively short time window. Compared to sustained exposure, short-term acute exposure may have more detrimental effects on fetuses, which is reflected in their later-life health outcomes. Lastly, we note that our estimates are not a direct mapping from early-life exposure to later-life outcomes; instead, they also incorporate parental responses, which could potentially enlarge the damage of air pollution on individuals' health. This point will be discussed in further detail in the next section.

Table 3 The Effects of Agricultural Fires on Adolescent Health

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Unhealthiness Index				
Diff. Upwind-Downwind Trimester 1	1.428***	1.454***	1.433***	2.378***	0.778
	(0.397)	(0.428)	(0.462)	(0.668)	(0.822)
Diff. Upwind-Downwind Trimester 2	0.846**	0.832**	0.813**	1.436**	0.479
	(0.384)	(0.381)	(0.401)	(0.610)	(0.796)
Diff. Upwind-Downwind Trimester 3	1.563***	1.567***	1.524***	2.174***	1.540**
	(0.332)	(0.363)	(0.385)	(0.566)	(0.649)
Observations	1,567	1,567	1,567	746	799
Sample	Full	Full	Full	Boy	Girl
Birth Year FE	Yes	Yes	No	No	No
Birth Month FE	Yes	Yes	No	No	No
Birth Year by Birth Month FE	No	No	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	Yes	Yes	Yes
Dep. Var. SD	2.011	2.011	2.011	2.011	2.011
Adjusted R-squared	0.0304	0.0284	0.0194	0.0538	0.0156

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on adolescent health, using the rural sample. The dependent variable is a normalized health index with a greater value representing worse health conditions. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

In Appendix Table B1, we replicate our baseline results on the urban sample. In contrast to the significant effects we find for the rural sample, we estimate insignificant impacts of agricultural fire exposure on the urban sample. Given that $PM_{2.5}$ generated from straw burning can travel from rural to urban regions (Guo 2021), there are at least two reasons for the insignificant effects on urban adolescents. First, air pollution from agricultural fires has less negative impact on urban residents because (1) the pollution concentration decays as it travels from rural to urban, and (2) the pollution from agricultural fires is less deleterious than industrial air pollution.[31] Second, urban residents have better access to health facilities and can mitigate the negative effects of pollution exposure. Our later investigation suggests that the former seems to be the primary reason.

Health components. Figure 5 visualizes the estimated results for the three health components that we

[31] A recent study by Lee, Wilson, and Hsiang (2025) reveals that pollution from different sources can have different health impacts. Since industrial production is mainly concentrated in the urban sector, and emissions from industrial sources contribute more to air pollution than emissions from agricultural sources (He et al. 2020), it is thus plausible that fire-induced pollution has indistinguishable impacts on health outcomes of urban residents.

used to construct the unhealthiness index, i.e., hospital admission, respiratory disease, and self-rated health status. We find that exposure to agricultural fires in the first trimester increases the probability of hospital admission and feelings of bad health, while exposure during the third trimester significantly increases the probability of having respiratory disease and hospital admission. In Appendix Figure B1, we plot the effects of fire exposure on health components for the urban sample. Again, we find no statistically disenable effects of fire-induced pollution on any of the health components.

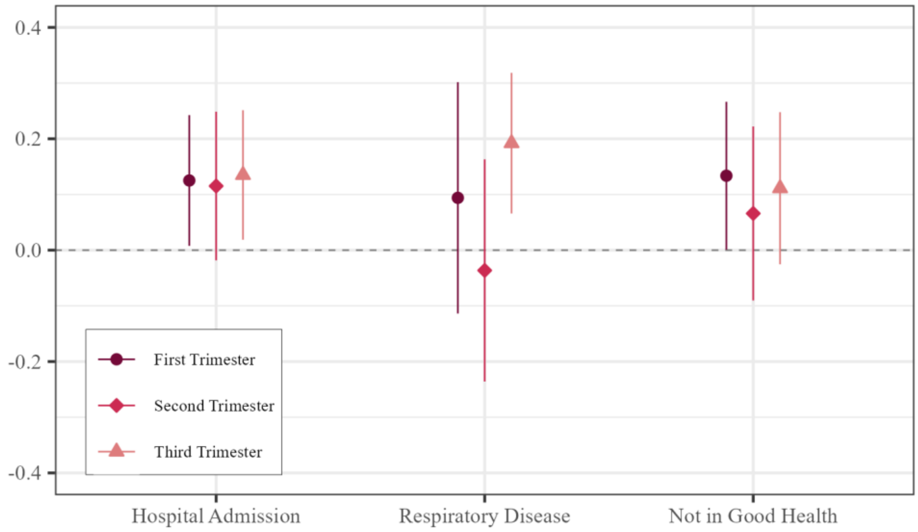


Figure 5 The Effects of Agricultural Fires on Adolescent Outcomes (Health Components)

Notes: This figure visualizes the estimated coefficients of the effects of in-utero agricultural fire exposure during different trimesters on adolescent health outcomes, including hospital admission, respiratory disease, and self-rated status, using the rural sample. All regressions include individual and weather controls. Point estimates and the corresponding 95% confidence intervals are jointly presented.

Heterogeneity. Appendix Table A8 explores the potential heterogeneity of our baseline findings. In columns (1) and (2), we divide our sample by whether the mother has completed middle school education. We show that our estimated effects are mostly driven by individuals whose mothers have lower education levels. In columns (3) and (4), we divide our sample by family income and find that the detrimental effects of pollution exposure are more pronounced for families with lower income. Together, this evidence suggests that agricultural fires have disproportionally stronger effects on more disadvantaged families, and the liquidity constraints seem to be a potential driver of the observed negative outcomes.[32] We will discuss this issue in further detail in the next section when we shed light on the role of parental investment. Finally, in columns (5) and (6), we divide the sample by whether the family is engaged in grain production, which is the major contributor to agricultural fires. We show that the effects are more significant for individuals from households engaged in grain production, which is in line with findings from Fletcher and Noghanibehambari (2024) that these households live closer to the cropland and are more exposed to pollution when the agricultural fire occurs.

5.2.2 Cognitive and non-cognitive outcomes

We then proceed to explore the cognitive and non-cognitive impacts of agricultural fire exposure. Table 4 reports the corresponding results, with Panel A presenting the estimated coefficients for cognitive ability

[32] This finding aligns with previous studies which show that air pollution affect vulnerable families disproportionally (Jans et al. 2018; Suarez Castillo et al. 2025).

while Panel B presents the estimates for non-cognitive ability. In columns (1) and (4), we estimate the effects using a full sample, whereas in columns (2) and (3), as well as in columns (5) and (6), we estimate the gender heterogeneity in pollution exposure. Focusing on the cognitive outcomes, we show that in-utero exposure to agricultural fires significantly reduces word test scores but has an insignificant impact on math test scores. In line with findings from Table 3, we show that the negative effects on word test scores are solely driven by boys, and are more pronounced if the exposure to agricultural fires occurs during the first and third trimesters.

Table 4 The Effects of Agricultural Fires on Cognitive and Non-cognitive Ability

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A Cognitive Ability	Standardized Word Test Score			Standardized Math Test Score		
Diff. Upwind-Downwind Trimester 1	-0.598*	-1.711***	0.557	0.139	-0.185	0.665
	(0.329)	(0.514)	(0.681)	(0.476)	(0.601)	(0.879)
Diff. Upwind-Downwind Trimester 2	-0.0148	-0.722	0.569	0.355	0.697	0.402
	(0.299)	(0.632)	(0.625)	(0.499)	(0.765)	(0.724)
Diff. Upwind-Downwind Trimester 3	-0.256	-1.375***	0.463	0.324	0.372	0.398
	(0.308)	(0.515)	(0.659)	(0.444)	(0.622)	(0.715)
Observations	1,384	659	701	1,393	667	702
Sample	Full	Boy	Girl	Full	Boy	Girl
Adjusted R-squared	0.239	0.222	0.227	0.141	0.052	0.165
Panel B Non-Cognitive Ability	Positive Attitudes			Negative Attitudes		
Diff. Upwind-Downwind Trimester 1	-0.817*	-1.289*	-0.252	1.447**	3.467***	0.617
	(0.538)	(0.735)	(0.939)	(0.559)	(0.910)	(1.122)
Diff. Upwind-Downwind Trimester 2	-0.571	0.362	-0.707	0.0656	2.359**	0.677
	(0.558)	(0.989)	(0.906)	(0.711)	(1.188)	(1.061)
Diff. Upwind-Downwind Trimester 3	-0.844*	-1.204*	-0.150	0.476	2.349**	0.770
	(0.495)	(0.679)	(0.950)	(0.674)	(1.036)	(1.057)
Observations	450	238	211	448	212	236
Sample	Full	Boy	Girl	Full	Boy	Girl
Adjusted R-squared	0.236	0.297	0.169	0.109	0.162	0.128
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on adolescent cognitive and non-cognitive ability, using the rural sample. The dependent variables in Panel A are age-specific standardized word test scores and math test scores. The dependent variables in Panel B are two measures that gauge the positive and negative attitudes of respondents. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

We also document a strong negative impact of agricultural fires on non-cognitive abilities. Specifically, we show that in-utero fire exposure during the first and third trimesters significantly decreases individuals' positive attitudes toward life and increases their negative attitudes in the meantime. Again, these effects are found in the male sample. This is in line with previous studies showing that males are more likely to be severely affected by air pollution exposure (Ebenstein et al. 2016). We also find consistent evidence that the effect is mainly concentrated in the first and third trimesters, aligning with previous findings. For example, using data from Brazil, Carneiro et al. (2024) show that prenatal exposure to agricultural fires, especially during the first and third trimester, significantly lowers students' test scores, and the effects are more pronounced for the male sample. Taken together, our estimates from Table 4 reveal that in-utero exposure to fire-induced air pollution has strong detrimental effects on the development of cognitive and non-cognitive abilities among male adolescents.[33]

[33] We present the estimated results of in-utero fire exposure on cognitive outcomes for the urban sample in Appendix

To streamline our empirics, we refer interested readers to Appendix C for additional evidence that supports the validity and robustness of our baseline findings. Specifically, we show that our results are robust to alternative definitions of the upwind direction, using alternative and more granular fixed effects, alternative clustering adjustment, and accounting for additional confounders and potential measurement errors. We also perform a randomized inference to ensure that our results are not driven by variations that may potentially correlate with our fire intensity measure.

5.3 The effects of fire exposure on adult outcomes

After showing that in-utero fire exposure can lead to worsened health and (non-)cognitive outcomes in adolescence, this subsection proceeds to present more evidence on whether these negative effects penetrate into adulthood and their potential impacts on educational attainment and labor market outcomes. Tracking the cohort in CFPS 2010 to CFPS 2020, Table A17 presents the estimated coefficients on the effects of in-utero agricultural fire exposure on adulthood outcomes. We focus on three outcome variables measured in CFPS 2020: (1) the completed number of years of education, normalized by individuals' age; (2) annual wage, conditioning on entering the labor market; and (3) a dummy variable that denotes whether the individual is employed in the agricultural sector.

Our findings indicate that in-utero fire exposure during the first trimester leads to a significantly shortened year of education completed. The estimated coefficient suggests that first-trimester exposure to agricultural fires is associated with a 0.075 decrease in the completed year of education, and the effects are larger for males and insignificant for females. It is worth noting that, by the time of 2020, not all individuals had completed their education.[34] It is therefore important to normalize the year of education by individuals' age, which allows us to compare individuals born in the same year but in different months. We also show that first- and third-trimester exposure to agricultural fires leads to worsened labor market outcomes. Take the coefficients estimated from third-trimester exposure as an example, our estimates suggest that it would lower the annual wage by 4.3% and increase the probability of being employed in the agricultural sector by more than 34% for the male sample. Again, no significant impacts are detected for the female sample.[35,36] Taken together, our findings suggest that the effects of in-utero exposure to agricultural fires can have persistent deleterious long-term effects. Given the abundant studies that link the development of health, cognitive, and non-cognitive abilities to labor market outcomes (Grönqvist et al. 2020), our empirical exercises thus far provide comprehensive evidence relating pollution exposure to human capital development across different stages.

Table B2, and again find no suggestive evidence that agricultural fires can affect the cognitive performance of urban adolescents. We cannot replicate this exercise for non-cognitive performance as the sample size is too small to run the regression specification.

[34] Our sample size from Table A13 suggests that only half of the rural sample have completed their education and entered the labor market in 2020.

[35] Appendix Table B3 replicates the results of agricultural fire exposure on education attainment in the urban sample, again finding no significant impacts. Unfortunately, due to the small sample size, we cannot recover the effects on labor market outcomes.

[36] In Appendix Table A18, we empirically investigate whether our results could be driven by sample selection by estimating the effect of agricultural fire exposure on a dummy variable indicating whether the individual can be tracked from CFPS 2010 to CFPS 2020. If sample attrition is plausibly random, then there should be no correlations between agricultural fire exposure and attrition probability. Reassuringly, we find no suggestive evidence that non-random sample attrition may affect our results. This further alleviates concerns that sample attrition may significantly affect our estimated effects.

6. Mechanisms

6.1 Early-life outcomes

To shed light on the mechanisms of long-term consequences of agricultural fire exposure, this subsection examines the effects of in-utero exposure on early-life outcomes. As previously documented by Rangel and Vogl (2019) that in-utero exposure to sugarcane fires can increase prenatal mortality and negatively affect health at birth. If our estimated long-term effects are indeed caused by exposure to fire-induced air pollution, then we should find strong negative correlations between in-utero fire exposure and early life outcomes. Table 5 presents the estimated results. Specifically, our evidence suggests that in-utero exposure to agricultural fires is positively correlated with the number of illnesses at age 1, and is negatively associated with gestation month and birth weight, indicating worsened health at birth and in early life. Moreover, the worsened early-life outcomes are exclusively concentrated in the male sample, which (partly) explains why the long-term effects of agricultural fire exposure are primarily driven by males. One particular reason why male individuals are disproportionately affected by agricultural fire is that, biologically, male fetuses tend to be larger than female fetuses, which requires more supply of nutrients and makes them more fragile to prenatal negative shocks, and therefore have less innate health human capital at birth relative to females.

Our results are also comparable to findings from existing studies. For instance, Rangel and Vogl (2019) suggest that doubling exposure to agricultural fire during the third trimester would decrease the gestation period by 4.6 percent, while our estimates suggest that the gestation month would be reduced by 5.9 percent for the male sample (column (5)). Similarly, their results imply that doubling the fire exposure would decrease birth weight by 18.8 percent, while our estimate suggests a corresponding reduction in birth weight by 22.2 percent (column (8)).

Table 5 The Effects of Agricultural Fires on Early Life Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	# Illness at Age 1			Gestation Month			Birth Weight		
Diff. Upwind-Downwind Trimester 1	0.227	0.409*	0.139	-0.327**	-0.672***	-0.151	-0.736*	-1.859***	-0.116
	(0.138)	(0.232)	(0.195)	(0.131)	(0.184)	(0.175)	(0.413)	(0.615)	(0.636)
Diff. Upwind-Downwind Trimester 2	0.130	0.590*	-0.201	-0.245	-0.367*	-0.051	-0.242	-0.326	-0.150
	(0.147)	(0.305)	(0.341)	(0.167)	(0.213)	(0.215)	(0.481)	(0.586)	(0.738)
Diff. Upwind-Downwind Trimester 3	0.191**	0.236	0.245	-0.210	-0.530**	0.106	-0.253	-0.765	0.470
	(0.096)	(0.152)	(0.158)	(0.196)	(0.219)	(0.211)	(0.410)	(0.617)	(0.669)
Observations	2,308	1,178	1,114	2,622	1,331	1,277	2,661	1,349	1,298
Sample	Full	Boy	Girl	Full	Boy	Girl	Full	Boy	Girl
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.142	0.155	0.136	0.216	0.207	0.249	0.137	0.089	0.174

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on early-life outcomes, using the rural sample. In columns (1) to (3), the dependent variable is the number of illnesses at age 1. In columns (4) to (6), the dependent variable is the length of the gestation period, measured in months. In columns (7) to (9), the dependent variable is the individual's birth weight, measured in 500 grams. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

There are two natural concerns pertaining to our findings in Table 5. The first is to what extent survival to birth affects our main estimates, and the second is potential measurement errors in our independent variables due to reduced gestation periods. Theoretically, fetuses with higher health capital are more likely to survive when exposed to air pollution, and therefore the estimated effects are biased downward due to selection into survival. Similarly, as measurement error is more likely to occur when individuals are more susceptible to pollution exposure, the estimated effects are also downwardly biased.

To explore whether exposure to agricultural fires increases infant mortality, we first calculate the number of children who did not survive to the survey year and the mortality rate within the household, where the latter is calculated by dividing the number of children that died by the number of children that survived to the survey year. Our results from Appendix Table A19 show that agricultural fire-induced air pollution has no significant effect on child mortality.[37] However, this result does not preclude the confounding effect of potential unobservables, as we only exploit cross-sectional variation for identification. Thus, the results should be interpreted with great caution. We then examine whether the potential selection affects our main estimates. In Appendix Table A20, we replicate our estimated effects for the unhealthiness index, cognitive, and non-cognitive abilities by additionally controlling for child mortality and gestation length. The results are largely unaffected, suggesting less concern about selection.

Taken together, the results in this section suggest that in-utero exposure to agricultural fires can have persistent detrimental effects on human capital formation and development. Specifically, it negatively affects health outcomes in early life, and leads to worsened health and (non-)cognitive outcomes in adolescence, which further translates into worsened educational attainment and labor market outcomes in adulthood. Undoubtedly, an important mechanism for these outcomes is the transmission of reduced prenatal health capital into later life. However, a largely unexplored mechanism is how pollution triggers intra-household responses and resource allocations (e.g., parental investment), and how these responses contribute to the observed outcomes. The next subsection discusses how parental investment responds to pollution exposure.

6.2 Parental Investment

In this subsection, we examine how parental investments of rural households respond to in-utero exposure to agricultural fires. Theoretically, parents could either make compensatory or reinforcing investments. The compensating investment suggests that parents would devote more resources to their children who are *more* exposed to agricultural fire-induced air pollution, whereas the reinforcing investment suggests that parents would devote more resources to children who are *less* exposed to pollution, as the human capital return of investing in these children is higher. This section brings the theoretical ambiguity to our data and empirically investigates whether rural households make compensatory or reinforcing investments. Specifically, we investigate how parental investments in health and education respond to in-utero pollution exposure.

6.2.1 Health investment

We first examine how families adjust their health investment in response to in-utero pollution exposure. We measure the health investment by the health expenditures on children. Column (1) of Table 6 reports the estimated effects. We find that parents significantly reduced their investment in children's health. Specifically, a one unit increase in fire exposure during the first trimester corresponds to approximately a 15.8% decrease in health expenses.[38] Given that the mean value of health expenses of children in the rural area is 401 RMB (approximately 61.7 USD), our estimates suggest that in-utero exposure to agricultural fires during

[37] Since CFPS 2010 only asks whether the child is alive and does not record in which year the child has passed away, the age of death is thus unknown, so as to whether it is prenatal death or postnatal death. Therefore, our estimated effects are a composite of both the effects on prenatal death and postnatal death.

[38] The magnitude is calculated by (exp (0.126)-1)/0.85, where we divide the coefficient by 0.85 since a one standard deviation increase in potential yield approximately corresponds to a 0.85 increase in the number of agricultural fires.

the first trimester reduces parental health expenses by 63 RMB (corresponding to 9.75 USD).[39]

In columns (2) to (5), we investigate the potential heterogeneity of our results. Specifically, in columns (2) and (3), we divide the sample according to whether the mother has completed at least lower secondary education (corresponding to the low and high education groups). We find that the effects of investment reduction come mainly from households where the mother has lower education. Our estimates suggest that for mothers with lower education, in-utero exposure to agricultural fires reduces health expenses by 119.6 RMB (equivalent to 18.4 USD). This is in line with our results from Appendix Table A8, where we show that the effects of in-utero agricultural fire exposure on adolescent health are more pronounced if the mother has a lower level of education.

Our exploration of the heterogeneity of mothers' education is inspired by the long strand of literature that investigates the effect of maternal education on children's human capital investment (Thomas et al. 1991; Desai and Alva 1998; Brown 2006; Leight and Liu 2020; Lavy et al. 2022). For example, Brown (2006) shows that while parents with higher education invest more in their children, the effect of mother's education is larger than that of father's education. This is in line with traditional Chinese family structure, where mothers exert more effort in raising the child, which makes their education more relevant when making the human capital investment decision.

In columns (4) and (5), we divide the sample according to the median value of family income (corresponding to the low and high income groups). We show that the effects are primarily driven by households with lower income, and the magnitude estimated from column (4) is comparable to the coefficients from column (2). Again, this result aligns with estimates from Table A8, where the effects of fire exposure on health are more significant for households with lower income. These heterogeneities in treatment effects seem to suggest that resource constraint is a plausible driver of the observed outcomes.

Table 6 The Effects of Agricultural Fires on Health Expenses

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Log Health Expenses				
Diff. Upwind-Downwind Trimester 1	-0.126**	-0.226**	-0.162	-0.257*	-0.004
	(0.074)	(0.098)	(0.169)	(0.133)	(0.082)
Diff. Upwind-Downwind Trimester 2	-0.119	-0.191*	-0.071	-0.058	-0.128
	(0.083)	(0.110)	(0.144)	(0.114)	(0.102)
Diff. Upwind-Downwind Trimester 3	-0.057	-0.076	-0.195	-0.215*	0.015
	(0.061)	(0.105)	(0.138)	(0.115)	(0.084)
Observations	1,411	966	391	632	733
Sample	Full	Low Education	High Education	Low Income	High Income
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.144	0.128	0.180	0.175	0.135

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on families' health expenses, using the rural sample. The dependent variable is the logged value of health expenses on children. The observation is at the county-cohort level, with each cohort defined by its birth month. Low or High Education correspond to whether the mother has completed lower secondary education. Low or High Income correspond to whether the family income is below or above the median income. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

In Appendix Table B4, we investigate the effects on health expenses for the urban sample. As urban

[39] Throughout our paper, both RMB and USD are measured on their 2010 monetary values.

residents are less exposed to agricultural fires, they serve as an ideal placebo to examine whether our estimated effects of health expenses reduction are indeed driven by in-utero agricultural fire exposure. Reassuringly, we find neither effects nor heterogeneity of in-utero fire exposure on health expenses for the urban sample.

6.2.2 Education investment

We then examine how parental education investments respond to in-utero agricultural fire exposure. Similarly, we measure the education investments by parents' education expenses on their children. Column (1) of Table 7 reports the estimated coefficients. We find that, in response to in-utero agricultural fire exposure, parents significantly reduced educational investment in their children, and the effects are mostly pronounced for individuals exposed during the third trimester. Specifically, the estimated coefficients suggest that a one unit increase in agricultural fire exposure during the third trimester is associated with an 8.7% decrease in parental education expenses for their children. Given that the mean value of education expenditure is 629 RMB (approximately 96.8 USD), our estimates imply a reduction in education expenditure by 54.7 RMB (corresponding to 8.4 USD).

In columns (2) to (5), we perform the same heterogeneous exercises as in Table 6. We find similar patterns that the reduction in education expenses is more significant for mothers with lower education and families with lower income. Analogously, in Appendix Table B5, we examine the effects of in-utero agricultural fire exposure on education expenditure for the urban sample. Again, no significant effects are found.

Table 7 The Effects of Agricultural Fires on Education Expenses

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Log Education Expenses				
Diff. Upwind-Downwind Trimester 1	-0.030	-0.039	-0.008	-0.047	-0.032
	(0.020)	(0.022)	(0.057)	(0.029)	(0.042)
Diff. Upwind-Downwind Trimester 2	-0.003	-0.001	-0.015	-0.040	0.046
	(0.032)	(0.036)	(0.049)	(0.034)	(0.068)
Diff. Upwind-Downwind Trimester 3	-0.071***	-0.066**	-0.058	-0.078***	-0.011
	(0.021)	(0.025)	(0.047)	(0.027)	(0.057)
Observations	2,049	1,452	571	1,105	881
Sample	Full	Low Education	High Education	Low Income	High Income
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	Yes	Yes	Yes
Adjusted R-squared	0.338	0.331	0.347	0.368	0.310

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on families' education expenses, using the rural sample. The dependent variable is the logged value of education expenses on children. The observation is at the county-cohort level, with each cohort defined by its birth month. Low or High Education correspond to whether the mother has completed lower secondary education. Low or High Income correspond to whether the family income is below or above the median income. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Taken together, our results in this section suggest that in response to negative early life shocks induced by in-utero exposure to agricultural fires, rural households reduce both health and education investment in their affected children. This aligns with the Reinforcement channel in which parents, constrained by limited resources, will invest less in children who have a lower return to human capital investment. More importantly, our results highlight salient unequal effects of early life pollution exposure, even in low-income settings like rural China. These results call for polices that aim to mitigate the negative effects of pollution exposure, especially policies that directly target rural households. In what follows, we examine the mitigating role of an important health insurance coverage in rural China, i.e., the rollout of the New Cooperative Medical Scheme.

7. The Role of Health Insurance

In the last part of our empirical investigation, we examine the role of health insurance in mitigating the adverse effects of pollution exposure during gestation on long-term outcomes. Given that parents reduce their human capital investment in response to in-utero exposure to agricultural fires, a related question is whether the coverage of health insurance can offset such reinforcing mechanisms. Besides, as the long-term persistence of early-life outcomes is another important channel, it is also of great importance to examine whether health insurance coverage can moderate the deleterious effects of pollution exposure on early outcomes. We now present the formal analyses of these questions.

7.1 Response in adolescent outcomes

We first examine the role of health insurance in mitigating the effects of in-utero agricultural fire exposure on adolescent outcomes. Columns (1) and (2) of Table 8 report the estimated results on the unhealthiness index, where we divide our sample by whether the individual is exposed to NCMS before age 5, following Huang and Liu (2023).[40] We show that the adverse health effects of in-utero fire exposure are primarily driven by individuals who are not exposed to NCMS. For individuals who are exposed to NCMS, the estimated coefficients are small in magnitude and are insignificant. To ensure that our results are not driven by other concurrent exposure (e.g., exposure to tap water, electricity, etc.), in columns (3) and (4), we additionally control for whether the individual is exposed to electricity, tap water, road, railway, and natural gas before age 5. Our results remain unaffected after the inclusion of these additional controls. In Appendix Table B6, we re-estimate our regression for the urban sample. As the NCMS only covers rural residents, we should find no mitigating effects for the urban sample. Not surprisingly, we reveal no effect of NCMS exposure on the urban sample.

Table 8 The Effects of Early Exposure to NCMS on Adolescent Outcomes

	(1)	(2)	(3)	(4)
Dep. Var.	Unhealthiness Index			
Diff. Upwind-Downwind Trimester 1	0.714	1.516***	0.434	1.566***
	(1.680)	(0.432)	(1.524)	(0.458)
Diff. Upwind-Downwind Trimester 2	0.0784	1.014**	0.249	1.063**
	(0.882)	(0.448)	(1.074)	(0.444)
Diff. Upwind-Downwind Trimester 3	0.812	1.429***	0.393	1.420***
	(1.505)	(0.396)	(1.396)	(0.389)
Observations	327	1,233	327	1,233
Sample	Exposure to NCMS	Non-Exposure to NCMS	Exposure to NCMS	Non-Exposure to NCMS
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes
Other Exposure	No	No	Yes	Yes
Adjusted R-squared	0.132	0.036	0.137	0.041

Notes: This table presents the estimated results of the role of NCMS coverage in mitigating the effects of in-utero agricultural fire exposure on adolescent health using the rural sample. The dependent variable is a normalized health index with a greater value representing worse health conditions. The observation is at the county-cohort level, with each cohort defined by

[40] Unfortunately, due to our relatively small sample size, we cannot perform estimation for other cognitive and non-cognitive outcomes.

its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Other exposures include controls on whether the individual is exposed to electricity, tap water, road, railway, and natural gas before age 5. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

7.2 Response in parental investment

We then investigate whether the coverage of health insurance can mitigate the negative impacts of in-utero pollution exposure on parental investment. Columns (1) and (2) of Table 9 examine the moderating effects of NCMS on health expenses. We find that for individuals who are exposed to the NCMS, in-utero exposure to agricultural fires does not significantly reduce parental health investment. For individuals who are not exposed to NCMS, fire exposure significantly reduces households' health expenses for their children. As we show in Table 6 that the effects of in-utero fire exposure on health expenditure are mostly driven by mothers with lower education and households with lower income, in columns (3) to (6), we take a one step further and examine whether the rollout of NCMS can mitigate the reduction on health investment for these vulnerable subgroups. Our results confirm this hypothesis and reveal that, compared with individuals whose mother has a lower level of education and was not exposed to NCMS before age 5, individuals exposed to NCMS receive relatively higher health investment, even if their mother has the same lower level of education. The same results hold for individuals who are from households with lower incomes. Moreover, we also notice that the coefficients in columns (3) and (5) are basically positive, though insignificant due to large standard errors. This provides suggestive evidence that the rollout of NCMS may increase parental health investment for individuals who are exposed to agricultural fires, and may turn the reinforcing behavior into compensatory behavior.

Table 9 The Effects of Early Exposure to NCMS on Health Expenses

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Log Health Expenses					
Diff. Upwind-Downwind Trimester 1	-0.068	-0.202*	-0.129	-0.234**	0.095	-0.442***
	(0.088)	(0.115)	(0.245)	(0.118)	(0.215)	(0.150)
Diff. Upwind-Downwind Trimester 2	-0.074	-0.130	0.247	-0.251	0.290	-0.095
	(0.091)	(0.158)	(0.356)	(0.169)	(0.314)	(0.188)
Diff. Upwind-Downwind Trimester 3	-0.099	-0.087	0.109	-0.056	0.189	-0.264
	(0.124)	(0.161)	(0.261)	(0.184)	(0.272)	(0.206)
Observations	636	797	357	568	348	481
Sample	Exposure	Non-Exposure	Exposure & Low Education	Non-Exposure & Low Education	Exposure & Low Income	Non-Exposure & Low Income
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Other Exposure	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.176	0.098	0.087	0.064	0.169	0.102

Notes: This table presents the estimated results of the role of NCMS coverage in mitigating the effects of in-utero agricultural fire exposure on parental health investment using the rural sample. The dependent variable is the logged value of health expenses for children. The observation is at the county-cohort level, with each cohort defined by its birth month. Low or High Education correspond to whether the mother has completed lower secondary education. Low or High Income correspond to whether the family income is below or above the median income. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Other exposures include controls on whether the individual is exposed to electricity, tap water, road, railway, and natural gas before age 5. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

In Table 10, we provide the same estimate for the mitigating effects of NCMS on parental education expenditure. Again, similar patterns emerge. We find that parental education investment does not respond to in-utero agricultural fire exposure for individuals who were exposed to NCMS, and that the mitigating role of NCMS exposure is more pronounced for individuals whose mother has lower education and individuals who are from households with lower income.

7.3 Response in early-life outcomes

We have shown that NCMS exposure can dampen the adverse effects of in-utero agricultural fire exposure on adolescent health outcomes, and one potential mechanism is increased parental investment. A remaining question is whether early life outcomes are also improved due to the rollout of the NCMS. We empirically examine this issue in Table A21, where we compare the estimated coefficients on early-life outcomes for individuals who are exposed and not exposed to NCMS. In contrast to previous findings, we find no discernible differences in treatment effects heterogeneity regarding whether the individual is exposed to the NCMS. This suggests that the rollout of NCMS may not mitigate the adverse effects of agricultural fires on individuals' health at birth. One potential reason for this outcome may be that rural families may not be (sufficiently) aware of the negative effects of agricultural fire exposure during fetal life, so even with health insurance coverage, health outcomes early in life remain unimproved. This calls for further policy improvement to enhance the awareness and dissemination of the harmful effects of potential sources of pollution in rural areas.

Table 10 The Effects of Early Exposure to NCMS on Education Expenses

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Log Education Expenses					
Diff. Upwind-Downwind Trimester 1	0.016	-0.022	0.137	-0.051	0.022	-0.099**
	(0.049)	(0.035)	(0.155)	(0.035	(0.063)	(0.031
Diff. Upwind-Downwind Trimester 2	-0.023	0.057	0.002	0.029	0.017	0.026
	(0.053)	(0.050)	(0.136)	(0.054)	(0.077	(0.076)
Diff. Upwind-Downwind Trimester 3	-0.048	-0.070*	-0.004	-0.110*	-0.044	-0.149***
	(0.049)	(0.040)	(0.147)	(0.031	(0.066)	(0.038
Observations	846	1,151	550	842	490	725
Sample	Exposure	Non-Exposure	Exposure & Low Education	Non-Exposure & Low Education	Exposure & Low Income	Non-Exposure & Low Income
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Other Exposure	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.294	0.408	0.254	0.367	0.322	0.350

Notes: This table presents the estimated results of the role of NCMS coverage on mitigating the effects of in-utero agricultural fire exposure on parental education investment, using the rural sample. The dependent variable is the logged value of education expenses for children. The observation is at the county-cohort level, with each cohort defined by its birth month. Low or High Education correspond to whether the mother has completed lower secondary education. Low or High Income correspond to whether the family income is below or above the median income. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Other exposures include controls on whether the individual is exposed to electricity, tap water, road, railway, and natural gas before age 5. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Although a comprehensive evaluation of the cost-benefit of NCMS is beyond the scope of the paper, we can nevertheless make some rough estimates based on our empirical results. To avoid exaggeration, we only consider the effect of agricultural fire exposure on individuals' earnings, which is monetized and more easily to compare. Our results in Appendix Table A17 imply that exposure to agricultural fires during the third trimester would reduce the wage income by 4.3%. Given the average annual wage of rural individuals in our sample in 72316 RMB in 2020, our estimates suggest a 3110 RMB decline in annual earnings. Given that estimated total cost of insurance coverage per child during early childhood is 2877 RMB in 2014 (Huang and Liu 2023), if we suppose that the NCMS coverage can fully offset the negative effect of agricultural fire exposure on earning, the estimated benefit, deflated to its 2014 value, is then 2762 RMB, which is of similar magnitude to the total cost of NCMS. Since our estimated benefits are likely to understate the overall impact of NCMS coverage (e.g., improved health status, reduced inpatient expenditure and hospital costs), the overall benefits must exceed its total costs.

8. Conclusion

This paper studies the long-term effects of in-utero exposure to agricultural fires. Using a nationally representative household dataset on rural China, we show that in-utero exposure to agricultural fires significantly decreases health outcomes, cognitive and non-cognitive performance in adolescence. The effects are mostly driven by exposure during the first and third trimesters, and are found to be larger in the male sample. Tracking these cohorts into their adulthood, we show that agricultural fire exposure during gestation leads to lower years of education and lower earnings, while increasing the probability of individuals working in low-skill sectors (e.g., agriculture).

Exploring the potential mechanisms, we find that in-utero exposure to agricultural fires significantly worsens early-life health conditions (i.e., more illnesses at age 1, shorter gestation period, and lower birth weight). More importantly, we show that parental investment is another critical driver of the observed effects. Specifically, our evidence suggests that parents reduce their health and education investment in exposed children. The reduction effects are stronger for individuals whose mother has a lower education level and individuals from families with lower income, which suggests that the liquidity constraints may be a potential explanation. Finally, we investigate how the provision of public health insurance can mitigate the adverse effects of pollution exposure. Exploiting the exogenous variation in the implementation of the NCMS program, we show that early-life exposure to NCMS can largely offset the negative effects of agricultural fire exposure. We further show that such mitigation is mainly through improvements in health and education investment, and is more pronounced for more disadvantaged households.

Our findings underscore the critical need for policy interventions to mitigate the lifelong consequences of in-utero environmental shocks. First, given the sensitivity of pollution exposure during the first and third trimesters, region-specific regulations on agricultural burning should prioritize seasonal restrictions aligned with cropping cycles, particularly in areas with high fire density. Complementary measures, such as real-time air quality monitoring and targeted advisories for pregnant women, could reduce fetal exposure during these vulnerable windows. Second, the evidence on parental (dis-)investment, especially among low-income and low-education households, calls for the support of social protection programs. Strengthening the New Cooperative Medical Scheme (NCMS) and embedding liquidity support measures (e.g., conditional cash transfers tied to health and education expenditures) could alleviate financial constraints that perpetuate underinvestment. Importantly, the mitigating role of NCMS highlights the potential for scaling up health insurance to cover pollution-related developmental risks, potentially through pollution-specific insurance or subsidies for vulnerable populations.

Acknowledgement

We thank the editor, Xi Chen, and three anonymous referees for their constructive comments and suggestions. We are also grateful to Dongqin Wang and participants at CES and CHEF for helpful comments. This project is supported by Zhejiang University-International Food Policy Research Institute Center for International Development Study, the National Social Science Fund of China (No. 21&ZD091, No. 23ZDA035), the Humanities and Social Science Fund of the Ministry of Education of China (22JJD790076), and the Agricultural Science and Technology (10-IAED-05-2024; 10-IAED-RY-01-2024).

Conflict of Interest

The authors declare no competing interests.

Data Availability

The primary data used in this study are from the China Family Panel Studies (CFPS). This is publicly available and can be applied through the following website: isss.pku.edu.cn/cfps/. However, the restricted-access dataset enabling matching between pseudocode county code to actual county code is protected under a confidentiality agreement with the Institute of Social Science Survey (ISSS) at Peking University and cannot be shared publicly. Relevant code for data processing and analysis is available upon request.

Reference

Adhvaryu, Achyuta, and Anant Nyshadham. 2016. “Endowments at Birth and Parents’ Investments in Children.” The Economic Journal 126 (593): 781–820.

Aizer, Anna, Janet Currie, Peter Simon, and Patrick Vivier. 2018. “Do Low Levels of Blood Lead Reduce Children’s Future Test Scores?” American Economic Journal: Applied Economics 10 (1): 307–41.

Almond, Douglas, and Janet Currie. 2011. “Killing Me Softly: The Fetal Origins Hypothesis.” Journal of Economic Perspectives 25 (3): 153–72.

Almond, Douglas, Lena Edlund, and Mårten Palme. 2009. “Chernobyl’s Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden.” Quarterly Journal of Economics 124 (4): 1729–72.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.” Journal of Political Economy 113 (1): 151–84.

Anderson, Michael L. 2020. “As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality.” Journal of the European Economic Association 18 (4): 1886–927.

Ayesh, Abubakr. 2023. “Burned Agricultural Biomass, Air Pollution and Crime.” Journal of Environmental Economics and Management 122 (October): 102887.

Barreca, Alan I., Matthew Neidell, and Nicholas J. Sanders. 2021. “Long-Run Pollution Exposure and Mortality: Evidence from the Acid Rain Program.” Journal of Public Economics 200 (August): 104440.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Yongchen Zou. 2024. “From Fog to Smog: The Value of Pollution Information.” American Economic Review 114 (5): 1338–81.

Bharadwaj, Prashant, Juan Pedro Eberhard, and Christopher A. Neilson. 2018. “Health at Birth, Parental Investments, and Academic Outcomes.” Journal of Labor Economics 36 (2): 349–94.

Bharadwaj, Prashant, Matthew Gibson, Joshua Graff Zivin, and Christopher Neilson. 2017. “Gray Matters: Fetal Pollution Exposure and Human Capital Formation.” Journal of the Association of Environmental and Resource Economists 4 (2): 505–42.

Billings, Stephen B., and Kevin T. Schnepel. 2018. “Life after Lead: Effects of Early Interventions for Children Exposed to Lead.” American Economic Journal: Applied Economics 10 (3): 315–44.

Black, Sandra E., Aline Bütikofer, Paul J. Devereux, and Kjell G. Salvanes. 2019. “This Is Only a Test? Long-Run and Intergenerational Impacts of Prenatal Exposure to Radioactive Fallout.” The Review of Economics and Statistics 101 (3): 531–46.

Bombardini, Matilde, and Bingjing Li. 2020. “Trade, Pollution and Mortality in China.” Journal of International Economics 125 (July): 103321.

Borgschulte, Mark, David Molitor, and Eric Yongchen Zou. 2024. “Air Pollution and the Labor Market: Evidence from Wildfire Smoke.” Review of Economics and Statistics 106 (6): 1558–75.

Boudreaux, Michel H., Ezra Golberstein, and Donna D. McAlpine. 2016. “The Long-Term Impacts of Medicaid Exposure in Early Childhood: Evidence from the Program’s Origin.” Journal of Health Economics 45 (January): 161–75.

Brown, Philip H. 2006. “Parental Education and Investment in Children’s Human Capital in Rural China.” Economic Development and Cultural Change 54 (4): 759–89.

Burns, Lawton Robert, and Gordon G Liu. 2017. China’s Healthcare System and Reform.

Cao, Jing, and Rong Ma. 2023. “Mitigating Agricultural Fires with Carrot or Stick? Evidence from China.” Journal of Development Economics 165 (October): 103173.

Carneiro, Juliana, Matthew A. Cole, and Eric Strobl. 2024. “Foetal Exposure to Air Pollution and Students’ Cognitive Performance: Evidence from Agricultural Fires in Brazil*.” Oxford Bulletin of Economics and Statistics 86 (1): 156–86.

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 (3): 1121–67.

Chen, Jiafeng, and Jonathan Roth. 2023. “Logs with Zeros? Some Problems and Solutions.” The Quarterly Journal of Economics, December 14, qjad054.

Chen, Pan. 2025. “Industrialization and Pollution: The Long-Term Impact of Early-Life Exposure on Human Capital Formation.” Journal of Public Economics 241 (January): 105270. https://doi.org/10.1016/j.jpubeco.2024.105270.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. “Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China’s Huai River Policy.” Proceedings of the National Academy of Sciences 110 (32): 12936–41.

Chen, Yuyu, and Ginger Zhe Jin. 2012. “Does Health Insurance Coverage Lead to Better Health and Educational Outcomes? Evidence from Rural China.” Journal of Health Economics 31 (1): 1–14.

Chen, Yvonne Jie, Li Li, and Yun Xiao. 2022. “Early-Life Exposure to Tap Water and the Development of Cognitive Skills.” Journal of Human Resources 57 (6): 2113–49.

Cheng, Lingguo, Hong Liu, Ye Zhang, Ke Shen, and Yi Zeng. 2015. “The Impact of Health Insurance on Health Outcomes and Spending of the Elderly: Evidence from China’s New Cooperative Medical Scheme.” Health Economics 24 (6): 672–91.

Coulombe, Raphaelle G., and Akhil Rao. 2025. “Fires and Local Labor Markets.” Journal of Environmental Economics and Management 130 (March): 103109.

Currie, J., and M. Neidell. 2005. “Air Pollution and Infant Health: What Can We Learn from California’s Recent Experience?” Quarterly Journal of Economics 120 (3): 1003–30.

Currie, Janet, Eric A Hanushek, E. Megan Kahn, Matthew Neidell, and Steven G Rivkin. 2009. “Does Pollution Increase School Absences?” Review of Economics and Statistics 91 (4): 682–94.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. “What Do We Know About Short- and

Long-Term Effects of Early-Life Exposure to Pollution?” Annual Review of Resource Economics 6 (1): 217–47.

Desai, Sonalde, and Soumya Alva. 1998. “Maternal Education and Child Health: Is There a Strong Causal Relationship?” Demography 35 (1): 71–81. https://doi.org/10.2307/3004028.

Du, Rui, Ajkel Mino, Jianghao Wang, and Siqi Zheng. 2024. “Transboundary Vegetation Fire Smoke and Expressed Sentiment: Evidence from Twitter.” Journal of Environmental Economics and Management 124 (March): 102928.

Duque, Valentina, Maria Rosales-Rueda, and Fabio Sanchez Torres. 2019. “How Do Early-Life Shocks Interact with Subsequent Human Capital Investments? Evidence from Administrative Data.” Unpublished.

Ebenstein, Avraham. 2012. “The Consequences of Industrialization: Evidence from Water Pollution and Digestive Cancers in China.” The Review of Economics and Statistics 94 (1): 186–201.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. “New Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China’s Huai River Policy.” Proceedings of the National Academy of Sciences 114 (39): 10384–89.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. “The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution.” American Economic Journal: Applied Economics 8 (4): 36–65.

Ferro, Simone, Alessandro Palma, Chiara Serra, and Massimo Stafoggia. 2024. “Beyond Birth: The Medium-Term Health Impact of Prenatal Exposure to Air Pollution.” Journal of Environmental Economics and Management 127 (September): 103009.

Fletcher, Jason, and Hamid Noghanibehambari. 2024. “The Siren Song of Cicadas: Early-Life Pesticide Exposure and Later-Life Male Mortality.” Journal of Environmental Economics and Management 123 (January): 102903.

Gao, Xuwen, Ran Song, and Christopher Timmins. 2024. “The Fertility Consequences of Air Pollution in China.” Journal of the Association of Environmental and Resource Economists 11 (3): 657–88. https://doi.org/10.1086/726316.

Garg, Teevrat, Maulik Jagnani, and Hemant K. Pullabhotla. 2024. “Rural Roads, Farm Labor Exits, and Crop Fires.” American Economic Journal: Economic Policy 16 (3): 420–50.

Glinianaia, Svetlana V., Judith Rankin, Ruth Bell, Tanja Pless-Mulloli, and Denise Howel. 2004. “Particulate Air Pollution and Fetal Health: A Systematic Review of the Epidemiologic Evidence.” Epidemiology 15 (1): 36–45.

Gong, Yazhen, Shanjun Li, Nicholas J. Sanders, and Guang Shi. 2023. “The Mortality Impact of Fine Particulate Matter in China: Evidence from Trade Shocks.” Journal of Environmental Economics and Management 117 (January): 102759.

Graff Zivin, Joshua, Tong Liu, Yingquan Song, Qu Tang, and Peng Zhang. 2020. “The Unintended Impacts of Agricultural Fires: Human Capital in China.” Journal of Development Economics 147 (November): 102560.

Greenstone, Michael, and Rema Hanna. 2014. “Environmental Regulations, Air and Water Pollution, and Infant Mortality in India.” American Economic Review 104 (10): 3038–72.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2022. “Can Technology Solve the Principal-Agent Problem? Evidence from China’s War on Air Pollution.” American Economic Review: Insights 4 (1): 54–70.

Grönqvist, Hans, J. Peter Nilsson, and Per-Olof Robling. 2020. “Understanding How Low Levels of Early Lead Exposure Affect Children’s Life Trajectories.” Journal of Political Economy 128 (9): 3376–433.

Gruber, Jonathan, Mengyun Lin, and Junjian Yi. 2023. “The Largest Insurance Program in History: Saving One Million Lives per Year in China.” Journal of Public Economics 226 (October): 104999. https://doi.org/10.1016/j.jpubeco.2023.104999.

Guo, Shiqi. 2021. “How Does Straw Burning Affect Urban Air Quality in China?” American Journal of Agricultural Economics 103 (3): 1122–40.

Hanna, Rema, and Paulina Oliva. 2015. “The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City.” Journal of Public Economics 122 (February): 68–79.

Hays, Michael D., Philip M. Fine, Christopher D. Geron, Michael J. Kleeman, and Brian K. Gullett. 2005. “Open Burning of Agricultural Biomass: Physical and Chemical Properties of Particle-Phase Emissions.” Atmospheric Environment 39 (36): 6747–64.

He, Guojun, Tong Liu, and Maigeng Zhou. 2020. “Straw Burning, PM2.5, and Death: Evidence from China.” Journal of Development Economics 145 (June): 102468.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. “Long-Run Impacts of Childhood Access to the Safety Net.” American Economic Review 106 (4): 903–34.

Hu, Shanlian, Shenglan Tang, Yuanli Liu, Yuxin Zhao, Maria-Luisa Escobar, and David De Ferranti. 2008. “Reform of How Health Care Is Paid for in China: Challenges and Opportunities.” The Lancet 372 (9652): 1846–53.

Huang, Wei, and Hong Liu. 2023. “Early Childhood Exposure to Health Insurance and Adolescent Outcomes: Evidence from Rural China.” Journal of Development Economics 160 (January): 102925. https://doi.org/10.1016/j.jdeveco.2022.102925.

Isen, Adam, Maya Rossin-Slater, and W. Reed Walker. 2017. “Every Breath You Take—Every Dollar You’ll Make: The Long-Term Consequences of the Clean Air Act of 1970.” Journal of Political Economy 125 (3): 848–902.

Jans, Jenny, Per Johansson, and J. Peter Nilsson. 2018. “Economic Status, Air Quality, and Child Health: Evidence from Inversion Episodes.” Journal of Health Economics 61 (September): 220–32.

Jayachandran, Seema. 2009. “Air Quality and Early-Life Mortality: Evidence from Indonesia’s Wildfires.” Journal of Human Resources 44 (4): 916–54.

Kannan, Srimathi, Dawn P Misra, J. Timothy Dvonch, and Ambika Krishnakumar. 2006. “Exposures to Airborne Particulate Matter and Adverse Perinatal Outcomes: A Biologically Plausible Mechanistic Framework for Exploring Potential Effect Modification by Nutrition.” Environmental Health Perspectives 114 (11): 1636–42.

Keil, Alwin, P. P. Krishnapriya, Archisman Mitra, et al. 2021. “Changing Agricultural Stubble Burning Practices in the Indo-Gangetic Plains: Is the Happy Seeder a Profitable Alternative?” International Journal of Agricultural Sustainability 19 (2): 128–51.

Kim Oanh, Nguyen Thi, Didin Agustian Permadi, Philip K. Hopke, Kirk R. Smith, Nguyen Phan Dong, and Anh Nguyet Dang. 2018. “Annual Emissions of Air Toxics Emitted from Crop Residue Open Burning in Southeast Asia over the Period of 2010–2015.” Atmospheric Environment 187 (August): 163–73.

Lai, Wangyang, Shanjun Li, Yanan Li, and Xiaohui Tian. 2022. “Air Pollution and Cognitive Functions: Evidence from Straw Burning in China.” American Journal of Agricultural Economics 104 (1): 190–208.

Landrigan, Philip J, Richard Fuller, Nereus J R Acosta, et al. 2018. “The Lancet Commission on Pollution and Health.” The Lancet 391 (10119): 462–512.

Lavy, Victor, Giulia Lotti, and Zizhong Yan. 2022. “Empowering Mothers and Enhancing Early Childhood Investment: Effect on Adults’ Outcomes and Children’s Cognitive and Noncognitive Skills.” Journal of Human Resources 57 (3): 821–67.

Lee, Jaecheol, Andrew Wilson, and Solomon Hsiang. 2025. Empirically Distinguishing Health Impacts of Transboundary and Domestic Air Pollution in Mixture. W33379. National Bureau of Economic Research.

Lei, Xiaoyan, and Wanchuan Lin. 2009. “The New Cooperative Medical Scheme in Rural China: Does More Coverage Mean More Service and Better Health?” Health Economics 18 (S2).

Leight, Jessica, and Elaine M. Liu. 2020. “Maternal Education, Parental Investment, and Noncognitive Characteristics in Rural China.” Economic Development and Cultural Change 69 (1): 213–51.

Li, Li, and Yun Xiao. 2023. “Beyond Boiling: The Effect of in Utero Exposure to Treated Tap Water on Childhood Health.” Journal of Environmental Economics and Management 119 (May): 102814.

Lin, Justin Yifu. 1992. “Rural Reforms and Agricultural Growth in China.” The American Economic Review 82 (1): 34–51.

Liu, Luo, Xinliang Xu, and Xi Chen. 2015. “Assessing the Impact of Urban Expansion on Potential Crop Yield in China during 1990–2010.” Food Security 7 (1): 33–43.

Liu, Ziheng. 2025. “CO2-Driven Crop Comparative Advantage and Planting Decision: Evidence from US Cropland.” Food Policy 130 (January): 102782.

Milcent, Carine. 2018. Healthcare Reform in China: From Violence to Digital Healthcare. Springer.

Nian, Yongwei. 2023. “Incentives, Penalties, and Rural Air Pollution: Evidence from Satellite Data.” Journal of Development Economics 161 (March): 103049.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” Journal of Business & Economic Statistics 37 (2): 187–204.

Pullabhotla, Hemant K., and Mateus Souza. 2022. “Air Pollution from Agricultural Fires Increases Hypertension Risk.” Journal of Environmental Economics and Management 115 (September): 102723. Rangel, Marcos A., and Tom S. Vogl. 2019. “Agricultural Fires and Health at Birth.” The Review of Economics and Statistics 101 (4): 616–30.

Rosales-Rueda, Maria, and Margaret Triyana. 2019. “The Persistent Effects of Early-Life Exposure to Air Pollution: Evidence from the Indonesian Forest Fires.” Journal of Human Resources 54 (4): 1037–80.

Sanders, Nicholas J. 2012. “What Doesn’t Kill You Makes You Weaker: Prenatal Pollution Exposure and Educational Outcomes.” Journal of Human Resources 47 (3): 826–50.

Shi, Tingting, Yongqiang Liu, Libo Zhang, Lu Hao, and Zhiqiu Gao. 2014. “Burning in Agricultural Landscapes: An Emerging Natural and Human Issue in China.” Landscape Ecology 29 (10): 1785–98. Šrám, Radim J., Blanka Binková, Jan Dejmek, and Martin Bobak. 2005. “Ambient Air Pollution and Pregnancy Outcomes: A Review of the Literature.” Environmental Health Perspectives 113 (4): 375–82.

Suarez Castillo, Milena, David Benatia, and Christine Le Thi. 2025. “Air Pollution and Children’s Health Inequalities.” Journal of Environmental Economics and Management 131 (May): 103149.

Thomas, Duncan, John Strauss, and Maria-Helena Henriques. 1991. “How Does Mother’s Education Affect Child Height?” Journal of Human Resources 26 (2): 183.

Von Hinke, Stephanie, and Emil N. Sørensen. 2023. “The Long-Term Effects of Early-Life Pollution Exposure: Evidence from the London Smog.” Journal of Health Economics 92 (December): 102827.

Wagstaff, Adam, Magnus Lindelow, Gao Jun, Xu Ling, and Qian Juncheng. 2009. “Extending Health Insurance to the Rural Population: An Impact Evaluation of China’s New Cooperative Medical Scheme.” Journal of Health Economics 28 (1): 1–19.

Wang, Feng, Min Wang, and Haitao Yin. 2022. “Can Campaign‐style Enforcement Work: When and How? Evidence from Straw Burning Control in China.” Governance 35 (2): 545‐64.

Wang, Zhenggang, Zenan Wu, and Ye Yuan. 2024. “We’ve Got You Covered! The Effect of Public Health Insurance on Rural Entrepreneurship in China.” Journal of Public Economics 235 (July): 105150. Xie. 2012. The User’s Guide of the China Family Panel Studies(2010). Institute of Social Science Survey, Peking University, Beijing.

Xie, Tingting, Ye Yuan, and Hui Zhang. 2023. “Information, Awareness, and Mental Health: Evidence from Air Pollution Disclosure in China.” Journal of Environmental Economics and Management 120 (July): 102827.

Yang, Jie, and Xin Huang. 2021. “The 30 m Annual Land Cover Dataset and Its Dynamics in China from 1990 to 2019.” Earth System Science Data 13 (8): 3907–25.

Yi, Junjian, James J. Heckman, Junsen Zhang, and Gabriella Conti. 2015. “Early Health Shocks, Intra-Household Resource Allocation and Child Outcomes.” The Economic Journal 125 (588): F347–71.

Yip, Winnie, and William C. Hsiao. 2008. “The Chinese Health System At A Crossroads.” Health Affairs 27 (2): 460–68.

You, Xuedan, and Yasuki Kobayashi. 2009. “The New Cooperative Medical Scheme in China.” Health Policy 91 (1): 1–9.

Online Appendix (Not for Publication)

Appendix A: Additional Tables and Figures

Additional Tables

Table A1 Summary Statistics

	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Panel A: Individual variables		Rural Sample			Urban Sample	
Adolescent outcomes						
Unhealthiness index	1846	0.009	2.011	1182	0.006	1.639
Not in good health = 1	1846	0.030	0.172	1182	0.014	0.119
Hospital admission = 1	1846	0.019	0.136	1182	0.022	0.146
Respiratory disease = 1	1846	0.102	0.303	1182	0.173	0.379
Word test score	1762	20.61	7.289	1162	23.36	6.318
Math test score	1773	10.67	4.532	1165	11.91	4.226
Health at birth						
# illness (age 1)	2691	3.058	5.145	2268	3.069	4.569
Gestation month	3065	9.274	0.573	2607	9.363	0.583
Birth weight (500g)	3128	7.105	1.654	2624	6.696	1.201
Adulthood outcomes (CFPS 2020)						
Education year	1504	10.30	2.930	1218	10.40	2.896
Annual wage (per thousand RMB)	830	72.32	34.54	329	36.23	34.99
Work in agricultural = 1	830	0.122	0.327	329	0.083	0.276
Agricultural fire exposure						
Exposed during the first trimester	3128	0.241	0.428	2624	0.229	0.421
Exposed during the second trimester	3128	0.236	0.424	2624	0.220	0.414
Exposed during the third trimester	3128	0.291	0.454	2624	0.312	0.463
Covariates						
Age	3128	10.27	3.414	2624	8.994	3.519
Gender	3128	0.513	0.500	2624	0.530	0.499
Father's age	3118	38.49	6.028	2605	37.09	5.627
Father completed middle school	3128	0.434	0.496	2624	0.755	0.430
Mother's age	3081	36.42	5.703	2590	35.37	5.380
Mother completed middle school	3128	0.295	0.456	2624	0.663	0.473
Family income (per thousand RMB)	2988	25.27	46.58	2489	40.28	57.66
Family size	3128	5.318	1.720	2624	4.672	1.696
# of siblings	3110	2.343	1.291	2615	1.838	1.074
Parental investment						
Health expense (per thousand RMB)	1815	0.401	0.962	1841	0.694	1.449
Education expense (per thousand RMB)	3098	0.629	0.910	2598	1.583	2.109
NCMS exposure						
Exposed to NCMS during 0-5	3128	0.429	0.495	2624	0.406	0.491

Panel B: County variables	Obs	Mean	Std. Dev.
# of agricultural fires	660,972	1.181	10.42
# of non-agricultural fires	660,972	1.460	10.74
Potential yield (kg/ha)	1,046,539	3052	2865
Upwind potential yield (kg/ha)	1,046,539	2957	3013
Downwind potential yield (kg/ha)	1,046,539	2968	3037
$PM_{2.5}$ ($\mu g/m^3$)	1,046,539	63.81	84.30
Dew point	1,046,539	6.672	11.88
Sea level pressure	1,046,539	1016	8.630
Temperature	1,046,539	13.73	10.51
Wind speed	1,046,539	2.333	0.798
Rainfall	1,046,539	2.811	3.068

Notes: This table presents the summary statistics for the main variables that are used in the empirical analysis. Panel A provides summary statistics for the individual sample from CFPS 2010 and CFPS 2020. Variables are separately summarized for both the rural and urban samples. Panel B provides summary statistics for county-level variables. Except for agricultural potential yield variables, all variables are defined at the county-year-month level. The sample period for agricultural fires is from 2001 to 2019, whereas the sample period for $PM_{2.5}$ and other meteorological variables is from 1980 to 2019.

Table A2 Balance Test between NCMS-exposed and Non-exposed Sample

	NCMS-exposed		NCMS-non-exposed		Mean Diff.	
Variables	Obs	Mean	Obs	Mean	Unconditional	Conditional
Age	1517	13.07	349	11.16	1.904***	0.846***
Gender	1517	0.506	349	0.461	0.044	0.073*
Father's age	1513	40.54	348	40.15	0.392	0.0288
Father completed middle school	1517	0.436	349	0.393	0.043	-0.001
Mother's age	1495	38.60	347	37.89	0.715*	0.377
Mother completed middle school	1517	0.281	349	0.272	0.009	-0.008
Family income (per thousand RMB)	1463	24.18	327	24.12	0.056	0.606
Family size	1517	5.115	349	5.372	-0.258	0.144
# of siblings	1510	2.239	346	2.494	-0.255*	0.008

Notes: This table presents the results of summary statistics and balance tests between NCMS-exposed and non-exposed samples. The Unconditional differences perform a simple t-test between exposed and non-exposed individuals. Only the age variable is strongly significant, as we use age and NCMS implementation timing to distinguish whether the individual is exposed to the policy. The conditional differences perform an OLS regression analysis by running each of the individual covariates on the NCMS exposure dummy while conditioning on a set of county-level characteristics, which includes counties' grain output, rural income, agricultural employment, and agricultural GDP.

Table A3 The Effects of Agricultural Potential Yield on Non-agricultural Fires

Dep. Var. # Non-agri. Fire	(1)	(2)	(3)	(4)
	OLS		PPML	
APY	-0.117	-0.099	-0.135	-0.127
	(0.085)	(0.085)	(0.093)	(0.093)
Observations	660,972	660,972	634,613	634,613
Prefecture-Year FE	Yes	Yes	Yes	Yes
Prefecture-Month FE	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	No	Yes
Dep. Var. Mean	1.460	1.460	1.460	1.460
Adjusted/Pseudo R-squared	0.159	0.161	0.518	0.530

Notes: This table presents the estimated results of the effects of agricultural potential yield on non-agricultural fires. The observation is at the county-year-month level. The sample period is from 2001 to 2019. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. Standard error is clustered at the prefecture level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A4 The Effects of Potential Yield on Agricultural Fires and Other Pollutants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	NO_x		SO_2		Dust		CO_2	
Upwind APY		-0.513		0.063		0.108		3.244
		(0.494)		(0.302)		(0.142)		(2.183)
Downwind APY		0.168		-0.011		0.050		-0.993
		(0.154)		(0.221)		(0.309)		(2.236)
APY	-0.440		-0.015		-0.004		2.916	
	(0.438)		(0.401)		(0.404)		(3.060)	
Observations	53,920	53,920	53,920	53,920	53,920	53,920	53,920	53,920
Prefecture by Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.101	0.101	4.617	4.617	2.570	2.570	22.09	22.09
Adjusted R-squared	0.0747	0.0748	0.122	0.122	0.120	0.120	0.314	0.314

Notes: This table presents the estimated results of the effects of upwind/downwind agricultural potential yield on agricultural fires and air pollutants (NO_x, SO_2, Dust, and CO_2). The observation is at the county-year level. The sample period is from 2000 to 2019. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. Standard error is clustered at the prefecture level.

Table A5 The Differential Effects of Upwind/Downwind Potential Yield on Agricultural Production

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Grain Output		Rural Income		Agri. Employment		Agri. GDP	
Upwind APY		54.891***		0.359***		12.218***		21.237***
		(7.887)		(0.117)		(2.412)		(3.135)
Downwind APY		51.840***		0.377***		11.499***		19.585***
		(8.444)		(0.140)		(2.534)		(3.020)
APY	110.852***		0.889***		22.800***		43.098***	
	(11.984)		(0.127)		(2.949)		(4.031)	
Diff. Upwind-Downwind		3.051		-0.018		0.719		1.652
		(11.397)		(0.221)		(4.038)		(4.774)
Observations	53,920	53,920	53,920	53,920	53,920	53,920	53,920	53,920
Prefecture by Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	224.3	224.3	7.530	7.530	112.9	112.9	113.9	113.9
Adjusted R-squared	0.386	0.387	0.807	0.806	0.490	0.493	0.528	0.527

Notes: This table presents the estimated results of the effects of upwind/downwind agricultural potential yield on a set of variables that are related to agricultural production to examine the income effects. The observation is at the county-year level. The sample period is from 2000 to 2019. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. Standard error is clustered at the prefecture level.

Table A6 Birth Month Selection

	(1)	(2)	(3)
Dep. Var.	Exposed in the 1st Trimester	Exposed in the 2nd Trimester	Exposed in the 3rd Trimester
Upwind APY	-0.006	-0.005	0.003
	(0.040)	(0.056)	(0.042)
Downwind APY	0.117	-0.128	0.028
	(0.083)	(0.094)	(0.051)
Diff. Upwind-Downwind	-0.123	0.123	-0.025
	(0.093)	(0.126)	(0.069)
Observations	1,460	1,460	1,460
Birth Year by Birth Month FE	Yes	Yes	Yes
Prefecture FE	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes
Adjusted R-squared	0.188	0.137	0.205

Notes: This table presents the estimated results of the effects of upwind/downwind agricultural potential yield on the birth month selection. The dependent variables are three dummies indicating during which trimester the individual is exposed to agricultural fires. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the prefecture level.

Table A7 Fertility Selection

	(1)	(2)	(3)
Dep. Var.	# of Children	# of Boys	# of Girls
Upwind APY	-0.452	0.021	-0.473**
	(0.304)	(0.088)	(0.236)
Downwind APY	0.778	0.078	0.700
	(0.778)	(0.246)	(0.564)
Diff. Upwind-Downwind	-1.230	-0.057	-1.173
	(1.018)	(0.316)	(0.749)
Observations	1,452	1,452	1,452
Birth Year by Birth Month FE	Yes	Yes	Yes
Prefecture FE	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes
Adjusted R-squared	0.138	0.148	0.187

Notes: This table presents the estimated results of the effects of upwind/downwind agricultural potential yield on fertility selection. The dependent variable in column (1) is the number of children, whereas in columns (2) and (3), the dependent variables are the number of boys and girls, respectively. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include the father's education and age, the mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. Standard error is clustered at the prefecture level.

Table A8 The Heterogeneous Effects of Agricultural Fires on Health

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Unhealthiness Index					
Diff. Upwind-Downwind Trimester 1	1.137**	0.728	1.896***	4.316	1.757***	0.795
	(0.518)	(1.276)	(0.635)	(2.886)	(0.532)	(1.208)
Diff. Upwind-Downwind Trimester 2	0.516	0.633	1.115	0.402	0.960**	0.188
	(0.418)	(1.308)	(1.081)	(1.637)	(0.428)	(1.387)
Diff. Upwind-Downwind Trimester 3	1.555***	0.366	1.850**	1.517	1.692***	-0.0191
	(0.493)	(1.206)	(0.855)	(1.765)	(0.437)	(1.159)
Observations	1,384	659	701	640	667	702
Sample	Low Education	High Education	Low Income	High Income	Grain Production	Non-grain Production
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.0383	0.0837	0.0673	0.0174	0.0522	0.0381

Notes: This table presents the heterogeneity of our baseline results, using the rural sample. We consider the heterogeneous effects regarding three variables: mother's education level (whether completed middle school), family income level (whether the family income is above the median value), and whether the family is engaged in grain production. The dependent variable is a normalized health index with a greater value representing worse health conditions. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A9 Robustness: Controlling for Additional Time Trends

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Unhealthiness Index		Word Test Score		Positive Attitudes		Negative Attitudes	
Diff. Upwind-Downwind Trimester 1	2.452***	2.452***	-1.710***	-1.710***	-1.317*	-1.318*	3.044***	3.050***
	(0.824)	(0.830)	(0.557)	(0.569)	(0.771)	(0.800)	(0.873)	(0.943)
Diff. Upwind-Downwind Trimester 2	1.002	1.002	-0.540	-0.540	0.537	0.550	0.773	0.775
	(0.728)	(0.734)	(0.671)	(0.683)	(1.117)	(1.189)	(1.343)	(1.456)
Diff. Upwind-Downwind Trimester 3	2.057***	2.057***	-1.296**	-1.296**	-1.254	-1.250	1.511	1.507
	(0.713)	(0.719)	(0.540)	(0.550)	(0.783)	(0.834)	(1.261)	(1.355)
Observations	746	746	659	659	238	238	212	212
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province by Birth Year Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quadratic Provincial Trends	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.257	0.321	0.222	0.186	0.241	0.103	0.110	0.088

Notes: This table presents the estimated results for including provincial birth-year trends, using the rural sample. The dependent variable in columns (1) and (2) is the unhealthiness index, in columns (3) and (4), it's the standardized word test score, in columns (5) and (6), it's the positive measure of non-cognitive ability, while in columns (7) and (8), it's the negative measure of non-cognitive ability. In odd columns, we include linear provincial year trends, whereas in even columns, we additionally control for the quadratic provincial year trends. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A10 Robustness: Controlling for Additional Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Unhealthiness Index		Word Test Score		Positive Attitudes		Negative Attitudes	
Diff. Upwind-Downwind Trimester 1	2.358***	2.550**	-2.010***	-1.959**	-1.321*	-1.853*	3.078***	2.799*
	(0.843)	(1.001)	(0.650)	(0.969)	(0.791)	(1.079)	(0.916)	(1.543)
Diff. Upwind-Downwind Trimester 2	0.987	0.258	-0.479	-0.258	0.549	-0.753	0.756	-2.281
	(0.756)	(1.132)	(0.741)	(1.033)	(1.164)	(1.262)	(1.411)	(2.231)
Diff. Upwind-Downwind Trimester 3	2.111***	2.081**	-1.357**	-1.463*	-1.254	-1.779*	1.496	0.436
	(0.768)	(1.046)	(0.627)	(0.807)	(0.824)	(1.064)	(1.317)	(1.572)
Observations	712	712	659	659	235	235	208	208
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province by Birth Year FE	Yes	No	Yes	No	Yes	No	Yes	No
City by Birth Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.321	0.325	0.175	0.194	0.134	0.204	0.088	0.166

Notes: This table presents the estimated results for including additional fixed effects, using the rural sample. The dependent variable in columns (1) and (2) is the unhealthiness index, in columns (3) and (4), it's the standardized word test score, in columns (5) and (6), it's the positive measure of non-cognitive ability, while in columns (7) and (8), it's the negative measure of non-cognitive ability. In odd columns, we control for province by birth year fixed effects, whereas in even columns, we control for city by birth year fixed effects. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A11 Robustness: Alternative Clustering Adjustments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Unhealthiness Index		Word Test Score		Positive Attitudes		Negative Attitudes	
Diff. Upwind-Downwind Trimester 1	2.378***	2.378***	-1.711***	-1.711***	-1.289	-1.289*	3.467***	3.467***
	(0.716)	(0.705)	(0.569)	(0.555)	(1.015)	(0.734)	(1.194)	(0.821)
Diff. Upwind-Downwind Trimester 2	1.436*	1.436**	-0.722	-0.722	0.362	0.362	2.359	2.359*
	(0.739)	(0.701)	(0.748)	(0.732)	(0.920)	(0.962)	(1.668)	(1.199)
Diff. Upwind-Downwind Trimester 3	2.174***	2.174***	-1.375***	-1.375***	-1.204*	-1.204*	2.349*	2.349**
	(0.701)	(0.639)	(0.533)	(0.579)	(0.708)	(0.664)	(1.304)	(1.024)
Observations	712	712	659	659	235	235	212	212
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.053	0.053	0.297	0.297	0.297	0.297	0.162	0.162

Notes: This table presents the estimated results for the alternative clustering adjustment, using the rural sample. The dependent variable in columns (1) and (2) is the unhealthiness index, in columns (3) and (4), it's the standardized word test score, in columns (5) and (6), it's the positive measure of non-cognitive ability, while in columns (7) and (8), it's the negative measure of non-cognitive ability. In odd columns, we two-way cluster the standard error at the county and province by birth year level, whereas in even columns, we cluster the standard error at the prefecture level. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A12 Robustness: Controlling for Additional Pollution Sources

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Unhealthiness Index		Word Test Score		Positive Attitudes		Negative Attitudes	
Diff. Upwind-Downwind Trimester 1	2.714***	2.714***	-2.010***	-1.984***	-1.336*	-1.348*	3.515***	3.396***
	(0.835	(0.837)	(0.544)	(0.655)	(0.743)	(0.765)	(1.051)	(1.038)
Diff. Upwind-Downwind Trimester 2	1.290*	1.290*	-0.893	-0.609	0.332	0.333	2.555*	2.825**
	(0.699)	(0.697)	(0.657)	(0.763)	(0.984)	(0.981)	(1.287)	(1.252)
Diff. Upwind-Downwind Trimester 3	2.364***	2.364***	-1.410***	-1.194*	-1.221*	-1.207*	2.664**	2.930***
	(0.712)	(0.720)	(0.510)	(0.631)	(0.687)	(0.676)	(1.010)	(1.004)
Observations	737	737	659	659	235	235	210	210
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indoor Air Pollution	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Water Pollution	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.120	0.113	0.239	0.237	0.280	0.274	0.176	0.186

Notes: This table presents the estimated results for controlling for additional pollution sources, using the rural sample. In odd columns, we include controls for whether households use straws as their primary source of fuel. While in even columns, we additionally control for the fertilizer expenses as a proxy for potential exposure to water pollution. The dependent variable in columns (1) and (2) is the unhealthiness index, in columns (3) and (4), it's the standardized word test score, in columns (5) and (6), it's the positive measure of non-cognitive ability, while in columns (7) and (8), it's the negative measure of non-cognitive ability. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A13 Robustness: Controlling for Additional Infrastructure and Facility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Unhealthiness Index		Word Test Score		Positive Attitudes		Negative Attitudes	
Diff. Upwind-Downwind Trimester 1	2.612***	2.675***	-1.617***	-1.636***	-1.289*	-1.289*	3.912***	4.234***
	(0.841)	(0.888)	(0.539)	(0.531)	(0.735)	(0.735)	(1.113)	(1.243)
Diff. Upwind-Downwind Trimester 2	1.177*	1.157*	-0.617	-0.760	0.362	0.362	3.073**	3.393**
	(0.682)	(0.686)	(0.623)	(0.609)	(0.989)	(0.989)	(1.488)	(1.446)
Diff. Upwind-Downwind Trimester 3	2.277***	2.287***	-1.118**	-1.162**	-1.204*	-1.204*	3.308***	3.550***
	(0.724)	(0.744)	(0.520)	(0.526)	(0.679)	(0.679)	(1.198)	(1.165)
Observations	731	731	659	659	234	234	209	209
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health and Educational Facilities	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Infrastructure Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.0741	0.0653	0.253	0.255	0.283	0.269	0.0943	0.0519

Notes: This table presents the estimated results for controlling for additional infrastructure and facility, using the rural sample. In odd columns, we include controls for health and educational facilities (i.e., number of kindergartens and primary schools, and number of hospitals and pharmacies). While in even columns, we additionally control for the infrastructure construction at the village level (i.e., access to electricity, road, and railway). The dependent variable in columns (1) and (2) is the unhealthiness index, in columns (3) and (4), it's the standardized word test score, in columns (5) and (6), it's the positive measure of non-cognitive ability, while in columns (7) and (8), it's the negative measure of non-cognitive ability. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A14 Robustness: Controlling for Contemporaneous Weather and Pollution Conditions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Unhealthiness Index		Word Test Score		Positive Attitudes		Negative Attitudes	
Diff. Upwind-Downwind Trimester 1	2.262***	2.251***	-1.882***	-1.844***	-1.048	-0.838	3.593***	3.497***
	(0.801)	(0.803)	(0.557)	(0.563)	(0.670)	(0.654)	(0.991)	(0.973)
Diff. Upwind-Downwind Trimester 2	1.114	1.102	-0.836	-0.817	0.202	-0.120	2.648**	2.564**
	(0.696)	(0.699)	(0.667)	(0.670)	(0.889)	(0.854)	(1.262)	(1.222)
Diff. Upwind-Downwind Trimester 3	2.201***	2.203***	-1.362**	-1.364**	-1.530**	-1.267*	2.576**	2.342**
	(0.719)	(0.719)	(0.527)	(0.527)	(0.669)	(0.646)	(1.054)	(1.147)
Observations	731	731	659	659	234	234	209	209
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Condition	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pollution	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.0116	0.0109	0.241	0.241	0.374	0.428	0.126	0.121

Notes: This table presents the estimated results for controlling for contemporaneous weather and pollution conditions, using the rural sample. In odd columns, we include controls for contemporaneous weather conditions (i.e., weather controls at the survey month). While in even columns, we additionally control for contemporaneous pollution conditions (i.e., $PM_{2.5}$ at the survey month). The dependent variable in columns (1) and (2) is the unhealthiness index, in columns (3) and (4), it's the standardized word test score, in columns (5) and (6), it's the positive measure of non-cognitive ability, while in columns (7) and (8), it's the negative measure of non-cognitive ability. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A15 Robustness: Addressing Potential Measurement Errors

Dep. Var.	(1) Unhealthiness Index	(2) Word Test Score	(3) Positive Attitudes	(4) Negative Attitudes
Diff. Upwind-Downwind Trimester 1	4.815***	-3.180***	-1.307	5.822***
	(1.381)	(1.013)	(1.492)	(1.889)
Diff. Upwind-Downwind Trimester 2	2.483**	-1.070	-0.574	5.761**
	(1.224)	(1.240)	(1.986)	(2.844)
Diff. Upwind-Downwind Trimester 3	4.404***	-2.073**	-2.622**	5.702**
	(1.208)	(1.038)	(1.194)	(2.110)
Observations	731	659	234	209
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes
Adjusted R-squared	0.0114	0.243	0.243	0.0584

Notes: This table presents the estimated results for accounting for potential measurement errors, using a two-step IV approach. The dependent variable is the unhealthiness index in column (1), the standardized word test score in column (2), the positive measure of non-cognitive ability in column (3), and the negative measure of non-cognitive ability in column (4). The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A16 Robustness: Sample Representativeness

Dep. Var.	(1) Unhealthiness Index	(2) Word Test Score	(3) Positive Attitudes	(4) Negative Attitudes
Diff. Upwind-Downwind Trimester 1	3.492***	-1.331*	-2.314***	3.379***
	(1.106)	(0.710)	(0.775)	(1.121)
Diff. Upwind-Downwind Trimester 2	0.859	-0.522	-1.085	1.903
	(0.928)	(0.748)	(1.017)	(1.458)
Diff. Upwind-Downwind Trimester 3	2.459**	-1.059*	-1.870**	2.659**
	(0.913)	(0.699)	(0.776)	(1.298)
Observations	731	659	234	209
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes
Adjusted R-squared	0.106	0.325	0.364	0.439

Notes: This table presents the weighted least squares estimation results, using the rural sample. We use the national sampling weight provided by CFPS to weight the individuals. The dependent variable is the unhealthiness index in column (1), the standardized word test score in column (2), the positive measure of non-cognitive ability in column (3), and the negative measure of non-cognitive ability in column (4). The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A17 The Effects of Agricultural Fires on Adulthood Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Education Year			Log Wage			Work in Agriculture		
Diff. Upwind-Downwind Trimester 1	-0.075**	-0.133*	-0.032	-0.029**	-0.019	-0.024	0.297**	0.416**	0.373
	(0.036)	(0.076)	(0.054)	(0.014)	(0.083)	(0.028)	(0.128)	(0.208)	(0.226)
Diff. Upwind-Downwind Trimester 2	-0.0200	-0.058	0.019	-0.028**	-0.041**	-0.021	0.337**	0.231	0.207
	(0.041)	(0.079)	(0.070)	(0.014)	(0.017)	(0.032)	(0.162)	(0.190)	(0.192)
Diff. Upwind-Downwind Trimester 3	-0.033	-0.019	-0.067	-0.023*	-0.043**	-0.025	0.217*	0.346*	0.340
	(0.035)	(0.083)	(0.049)	(0.013)	(0.019)	(0.030)	(0.120)	(0.193)	(0.220)
Observations	1,228	614	561	657	323	297	657	323	297
Sample	Full	Boy	Girl	Full	Boy	Girl	Full	Boy	Girl
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.180	0.208	0.145	0.030	0.033	0.198	0.059	0.102	0.040

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on adulthood outcomes, using the rural sample. In columns (1) to (3), the dependent variable is the number of education years completed, normalized by the individual's age. In columns (4) to (6), the dependent variable is the logged value of annual wage, conditioning on whether the individual enters the labor market. In columns (7) to (9), the dependent variable is a dummy that indicates whether the individual works in the agriculture sector. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A18 The Effects of Agricultural Fires on Sample Attrition

	(1)	(2)	(3)
Dep. Var.	Sample being Tracked from 2010 to 2020		
Diff. Upwind-Downwind Trimester 1	0.103	-0.137	0.315
	(0.174)	(0.246)	(0.383)
Diff. Upwind-Downwind Trimester 2	0.008	-0.029	0.048
	(0.194)	(0.362)	(0.352)
Diff. Upwind-Downwind Trimester 3	-0.027	-0.278	-0.178
	(0.182)	(0.317)	(0.329)
Observations	1,459	706	731
Sample	Full	Boy	Girl
Birth Year FE	Yes	Yes	No
Birth Month FE	Yes	Yes	No
Birth Year by Birth Month FE	No	No	Yes
County of Birth FE	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes
Weather Controls	No	Yes	Yes
Adjusted R-squared	0.0304	0.0284	0.0194

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on sample attrition, using the rural sample. The dependent variable is a dummy variable on whether the individual can be tracked from CFPS 2010 to CFPS 2020. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A19 The Effects of Agricultural Fire-Induced Air Pollution on Child Mortality

	(1)	(2)
Dep. Var.	# of Deaths	mortality
Upwind APY	-0.018	-0.006
	(0.033)	(0.011)
Downwind APY	-0.115*	-0.037*
	(0.060)	(0.021)
Diff. Upwind-Downwind	0.097	0.031
	(0.066)	(0.023)
Observations	1,452	1,322
Birth Year by Birth Month FE	Yes	Yes
Prefecture by Birth Year FE	Yes	Yes
Prefecture by Birth Month FE	Yes	Yes
Individual Controls	Yes	Yes
Weather Controls	Yes	Yes
Adjusted R-squared	0.0202	0.0594

Notes: This table presents the estimated results of the effects of upwind/downwind agricultural potential yield on child mortality. The dependent variable in column (1) is the number of deaths, and the dependent variable in column (2) is the mortality rate calculated by dividing the number of deaths by the number of surviving children. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include the father's education and age, the mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. Standard error is clustered at the prefecture level.

Table A20 Main Estimates after Controlling for Potential Selection

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Unhealthiness Index			Standardized Word Test Score		
Diff. Upwind-Downwind Trimester 1	1.781**	2.781***	-0.289	-0.768*	-1.988***	0.564
	(0.875)	(0.875)	(1.395)	(0.400)	(0.551)	(0.818)
Diff. Upwind-Downwind Trimester 2	0.856	1.256*	-0.509	-0.048	-0.784	0.478
	(0.692)	(0.692)	(1.455)	(0.363)	(0.665)	(0.733)
Diff. Upwind-Downwind Trimester 3	1.328***	2.328***	0.868	-0.352	-1.502**	0.512
	(0.350)	(0.750)	(1.227)	(0.376)	(0.518)	(0.750)
Observations	1,561	755	784	1,385	671	690
Sample	Full	Boy	Girl	Full	Boy	Girl
Adjusted R-squared	0.0534	0.0234	0.0517	0.238	0.239	0.200
	(7)	(8)	(9)	(10)	(11)	(12)
Dep. Var.	Non-cognitive Ability (Positive)			Non-cognitive Ability (Negative)		
Diff. Upwind-Downwind Trimester 1	-0.674	-1.319*	-0.052	1.375**	3.244***	0.582
	(0.526)	(0.728)	(0.931)	(0.561)	(1.026)	(0.955)
Diff. Upwind-Downwind Trimester 2	-0.613	0.0097	-0.318	-0.014	2.420*	0.522
	(0.529)	(0.972)	(0.912)	(0.698)	(1.280)	(1.024)
Diff. Upwind-Downwind Trimester 3	-0.798*	-1.323*	-0.382	0.474	2.309**	0.814
	(0.471)	(0.634)	(0.875)	(0.680)	(1.098)	(1.053)
Observations	450	232	211	441	209	229
Sample	Full	Boy	Girl	Full	Boy	Girl
Adjusted R-squared	0.247	0.301	0.187	0.111	0.137	0.165
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on adolescent health, cognitive, and non-cognitive ability by additionally controlling for selection, using the rural sample. The dependent variables are the unhealthiness index in columns (1) to (3), age-specific standardized word test scores in columns (4) to (6), positive measure of cognitive ability in columns (7) to (9), and negative measure of cognitive ability in columns (10) to (12). The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A21 The Effects of Early Exposure to NCMS on Early-life Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	# Illness at Age 1		Gestation Month		Birth Weight	
Diff. Upwind-Downwind Trimester 1	0.157	0.666**	-0.510**	-0.361*	-0.226	-0.967
	(0.209)	(0.266)	(0.228)	(0.194)	(0.617)	(0.641)
Diff. Upwind-Downwind Trimester 2	0.289	0.267	-0.299	-0.273	0.247	-0.099
	(0.344)	(0.341)	(0.245)	(0.201)	(0.646)	(0.776)
Diff. Upwind-Downwind Trimester 3	0.572***	0.302	-0.122	-0.372*	-0.059	-0.284
	(0.199)	(0.208)	(0.332)	(0.201)	(0.628)	(0.632)
Observations	963	1,087	1,089	1,246	1,103	1,262
Sample	Exposure	Non-Exposure	Exposure	Non-Exposure	Exposure	Non-Exposure
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Other Exposure	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.173	0.175	0.225	0.249	0.112	0.116

Notes: This table presents the estimated results of the role of NCMS coverage on mitigating the effects of in-utero agricultural fire exposure on early-life outcomes, using the rural sample. In columns (1) and (2), the dependent variable is the number of illnesses at age 1. In columns (3) and (4), the dependent variable is the length of the gestation period, measured in months. In columns (5) and (6), the dependent variable is the individual's birth weight, measured in 500 grams. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Other exposures include controls on whether the individual is exposed to electricity, tap water, road, railway, and natural gas before age 5. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Additional Figures

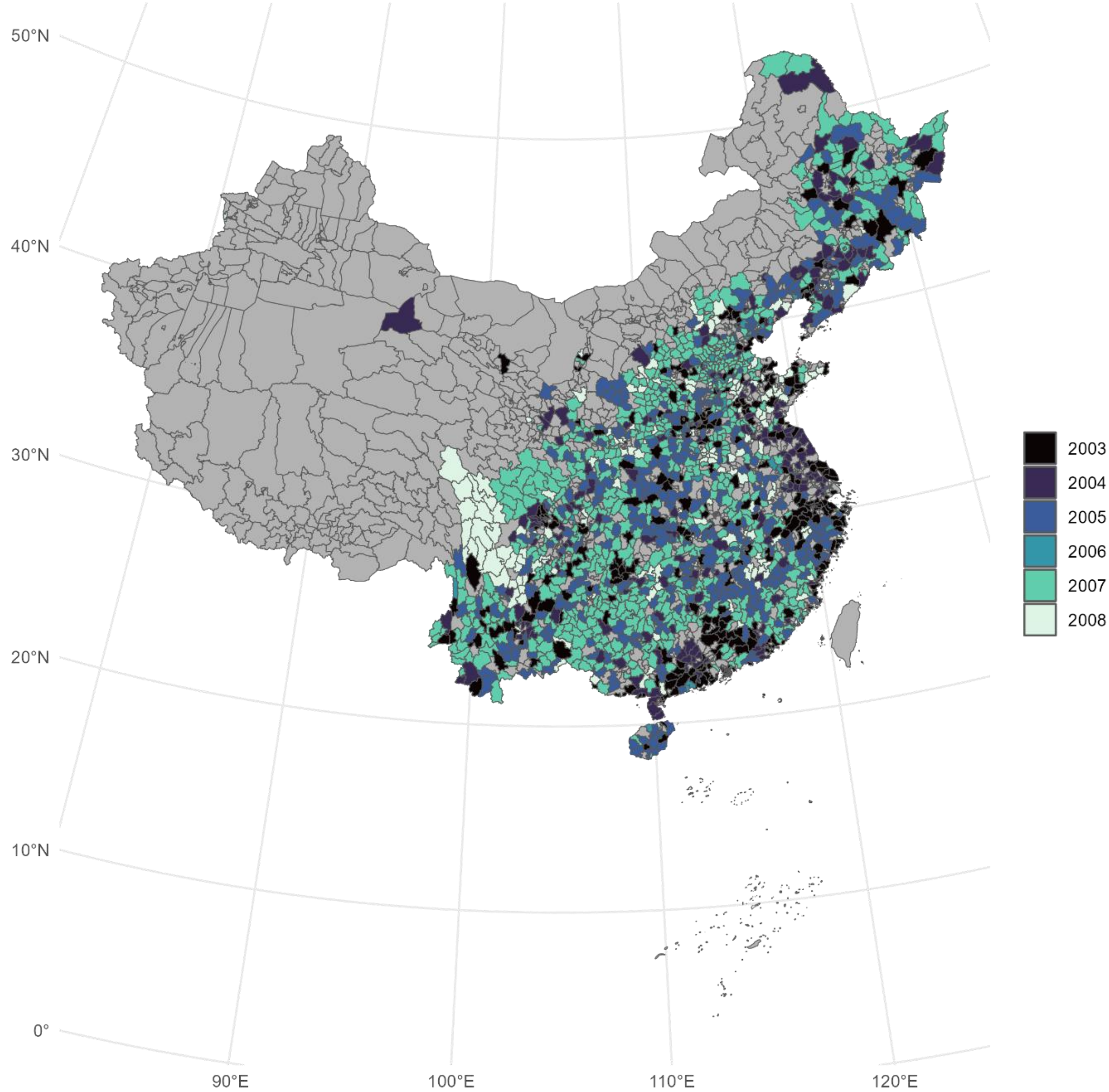


Figure A1 The Geographic Distribution of NCMS Rollout Timing

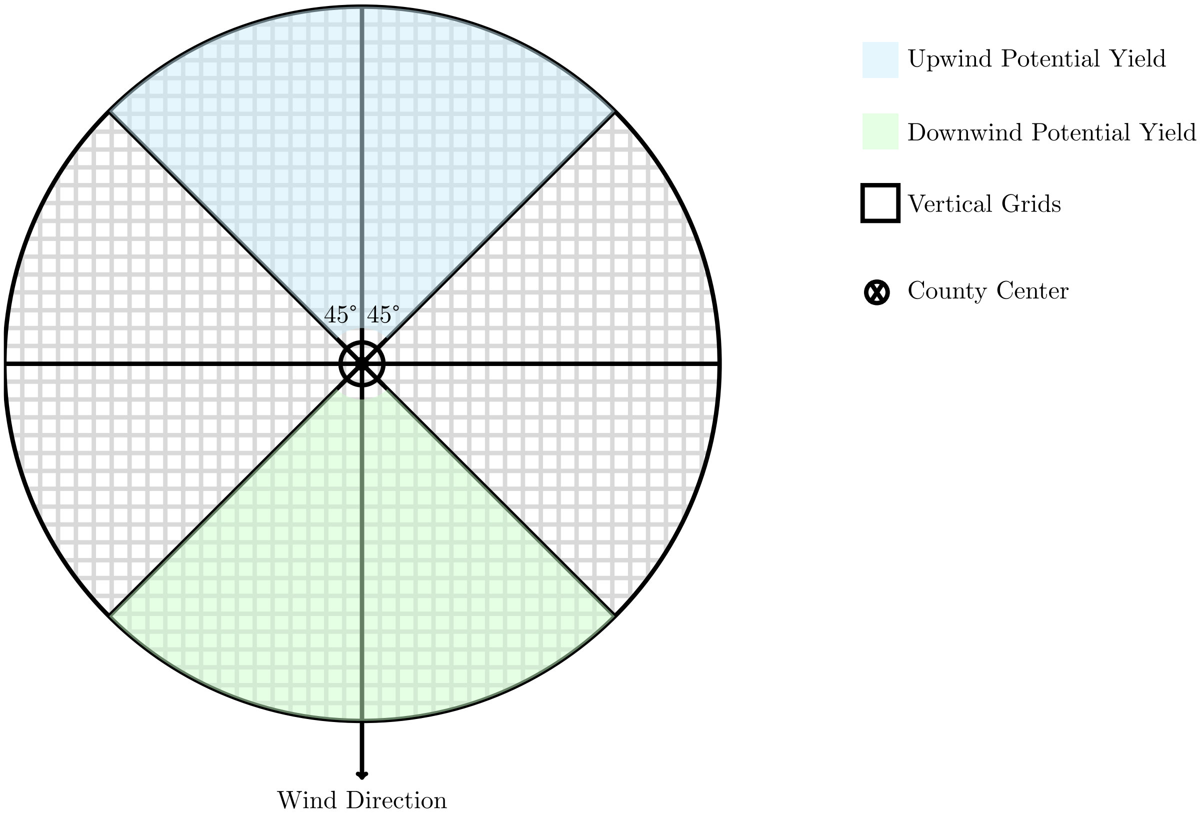


Figure A2 An Illustration of Determining the Upwind Region

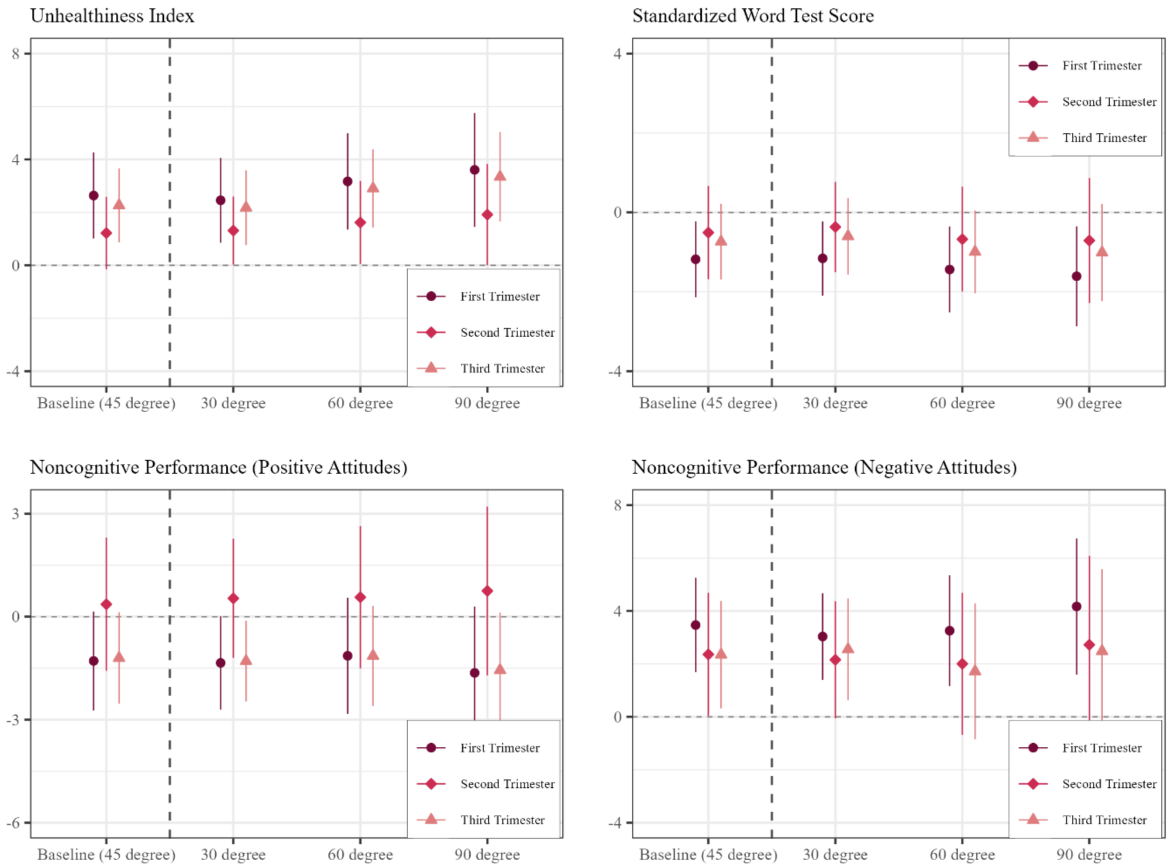


Figure A3 Robustness: Altering the Definition of Upwind Direction

Notes: This figure presents the estimated robustness results of altering the definition of upwind direction. Our baseline results adopt a definition of 45 degrees, and we check for the robustness of our results by using 30, 60, and 90 degrees. The figure is comprised of 4 subplots. The upper left panel depicts the corresponding effects on the unhealthiness index, the upper right panel depicts the effects on the standardized word test score, the lower left panel depicts the effects on the positive non-cognitive performance, and the lower right panel depicts the effects on the negative non-cognitive performance. We focus on the sample of rural male adolescents. Point estimates and the corresponding 95% confidence intervals are jointly presented.

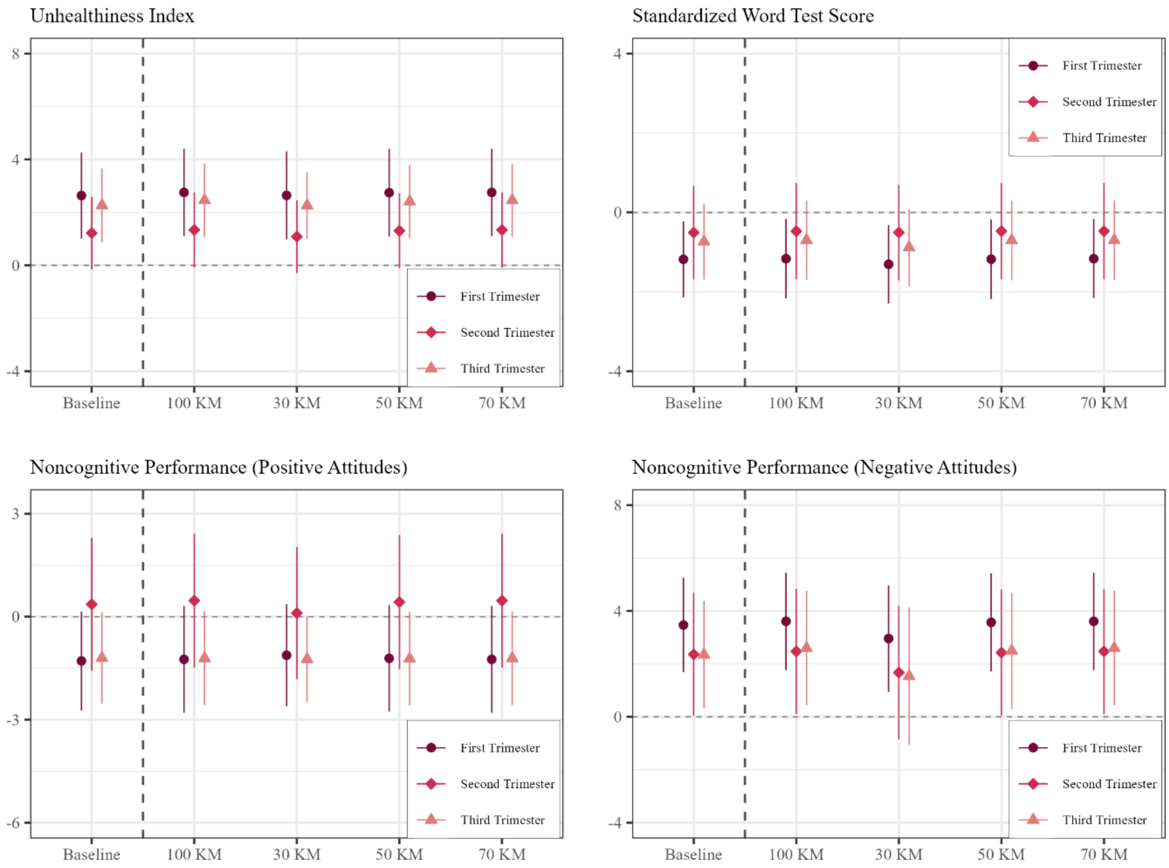
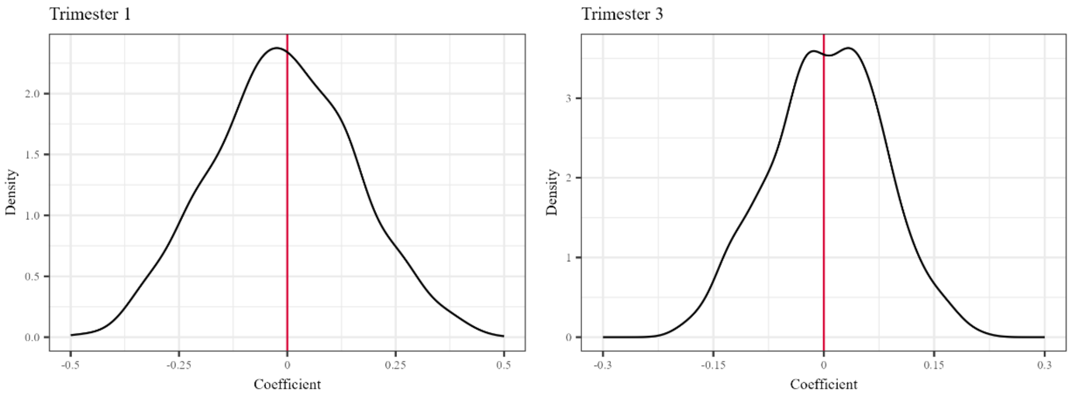
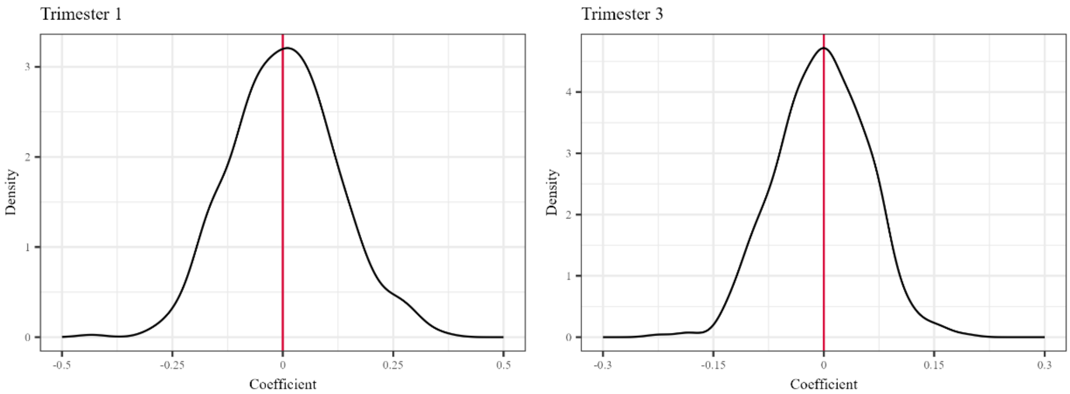


Figure A4 Robustness: Different Choices of Potential Yield Coverage

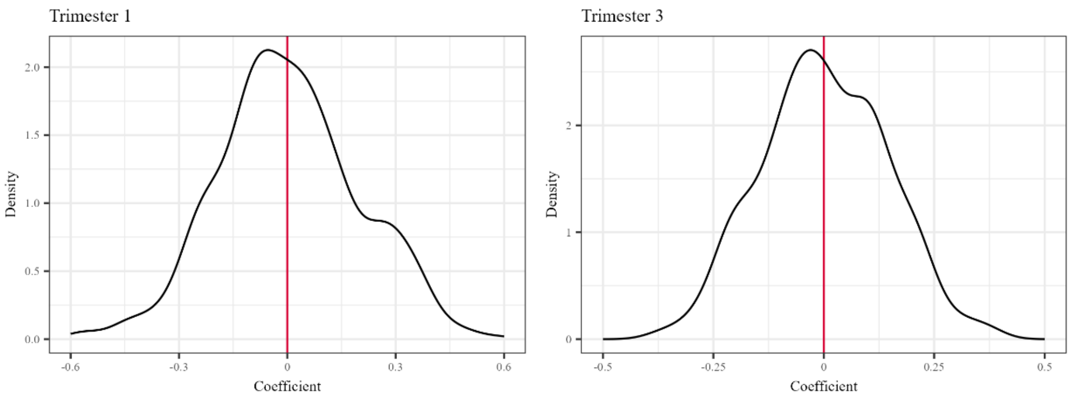
Notes: This figure presents the estimated robustness results of changing the coverage radii of agricultural potential yield. Our baseline results calculate the upwind/downwind potential yield covering the entire county, and we use potential yield grids with radii of 100 KM, 70 KM, 50 KM, and 30 KM to the county center to test the robustness of the results. The figure is comprised of 4 subplots. The upper left panel depicts the corresponding effects on the unhealthiness index, the upper right panel depicts the effects on the standardized word test score, the lower left panel depicts the effects on the positive non-cognitive performance, and the lower right panel depicts the effects on the negative non-cognitive performance. We focus on the sample of rural male adolescents. Point estimates and the corresponding 95% confidence intervals are jointly presented.



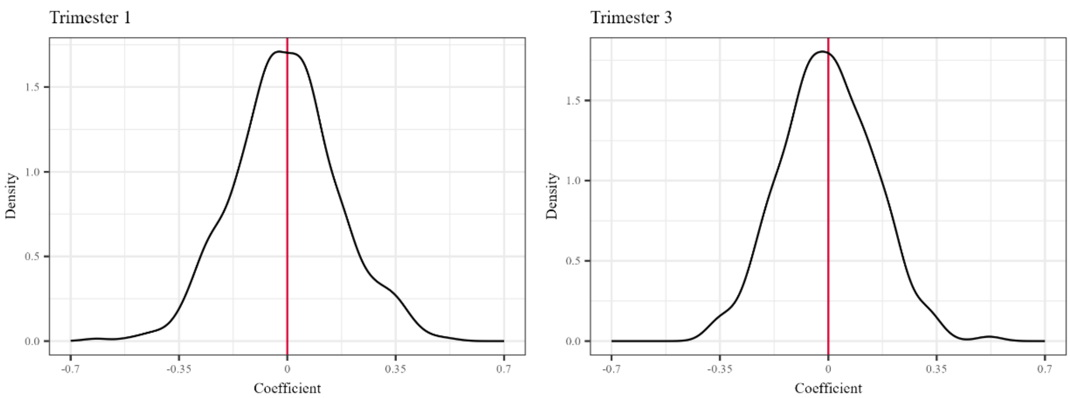
Panel A: Unhealthiness Index



Panel B: Standardized Word Test Score



Panel C: Positive Non-cognitive Performance



Panel D: Negative Non-cognitive Performance

Figure A5 Robustness: The Distribution of Placebo Coefficients

Notes: This figure presents the distribution of placebo coefficients on four of our main outcome variables, focusing on the effects of fire exposure during the first and third trimesters. In Panel A, we plot the distribution of coefficients from regressing the unhealthiness index on placebo upwind/downwind potential yield. Whereas in Panels B, C, and D, the dependent variables are the standardized word test score, the positive non-cognitive performance, and the negative non-cognitive performance. The sample is rural male adolescents.

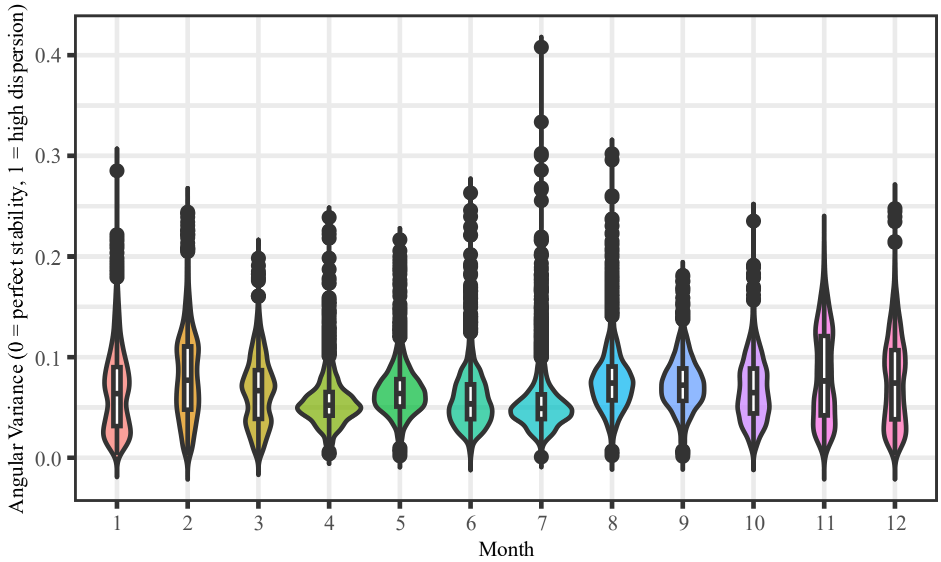


Figure A6 Distribution of Angular Variance of Monthly Wind Directions

Notes: This figure presents the distributions (box plot and violin plot) of angular variance of monthly wind directions.

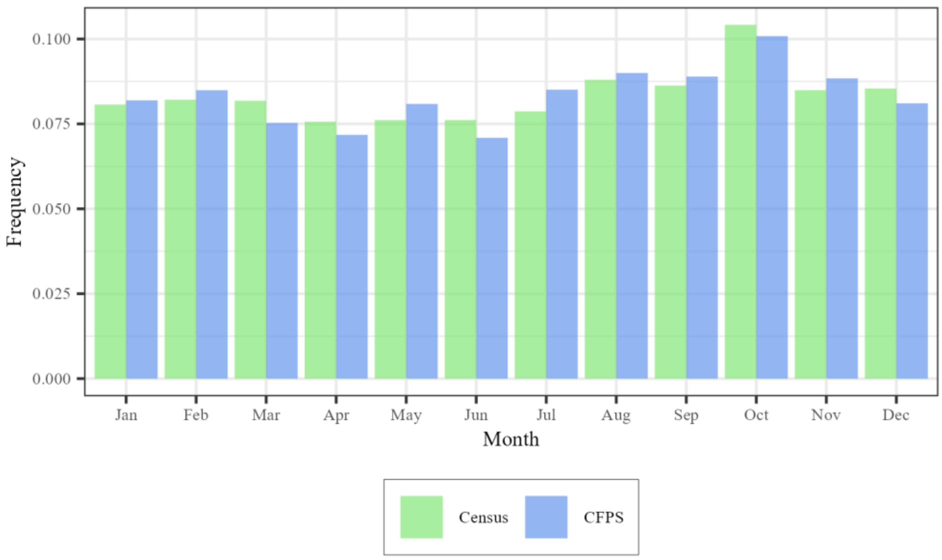


Figure A7 Comparison between Birth Month Distributions

Notes: This figure presents the distributions of birth month from the 2010 Population Census and 2010 CFPS.

Appendix B: Additional Results on the Urban Sample

Table B1 The Effects of Agricultural Fires on Adolescents' Health (Urban Sample)

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Unhealthiness Index				
Diff. Upwind-Downwind Trimester 1	-0.344	-0.305	-0.355	-0.170	-0.929
	(0.347)	(0.362)	(0.409)	(0.709)	(0.576)
Diff. Upwind-Downwind Trimester 2	-0.0521	-0.0341	-0.0870	0.0452	-0.435
	(0.356)	(0.363)	(0.387)	(0.782)	(0.681)
Diff. Upwind-Downwind Trimester 3	0.264	0.255	0.141	0.0504	0.290
	(0.329)	(0.336)	(0.375)	(0.688)	(0.676)
Observations	1,837	1,837	1,834	1,096	945
Sample	Full	Full	Full	Boy	Girl
Birth Year FE	Yes	Yes	No	No	No
Birth Month FE	Yes	Yes	No	No	No
Birth Year by Birth Month FE	No	No	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	Yes	Yes	Yes
Dep. Var. SD	1.638	1.638	1.638	1.638	1.638
Adjusted R-squared	0.440	0.441	0.440	0.432	0.449

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on adolescent health, using the urban sample. The dependent variable is a normalized health index with a greater value representing worse health conditions. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table B2 The Effects of Agricultural Fires on Cognitive Test Scores (Urban Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Standardized Word Test Score			Standardized Math Test Score		
Diff. Upwind-Downwind Trimester 1	-0.417*	-0.473	-0.298	0.00588	0.0688	0.0799
	(0.246)	(0.711)	(0.274)	(0.248)	(0.520)	(0.326)
Diff. Upwind-Downwind Trimester 2	-0.488	-0.599	-0.509	-0.491	-0.943	-0.222
	(0.309)	(0.812)	(0.348)	(0.310)	(0.688)	(0.371)
Diff. Upwind-Downwind Trimester 3	-0.139	-0.0927	0.467	-0.236	-0.113	-0.182
	(0.237)	(0.780)	(0.320)	(0.254)	(0.543)	(0.423)
Observations	816	408	368	818	409	369
Sample	Full	Boy	Girl	Full	Boy	Girl
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.232	0.217	0.228	0.202	0.157	0.232

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on adolescent cognitive ability, using the urban sample. The dependent variables are age-specific standardized word test scores and math test scores. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table B3 The Effects of Agricultural Fires on Adulthood Outcomes (Urban Sample)

	(1)	(2)	(3)
Dep. Var.	Education Year		
Diff. Upwind-Downwind Trimester 1	-0.0207	0.0246	0.0219
	(0.0332)	(0.0527)	(0.0377)
Diff. Upwind-Downwind Trimester 2	-0.0116	0.0546	-0.0247
	(0.0287)	(0.0537)	(0.0335)
Diff. Upwind-Downwind Trimester 3	0.00531	0.0493	0.00612
	(0.0311)	(0.0499)	(0.0281)
Observations	994	470	472
Sample	Full	Boy	Girl
Birth Year by Birth Month FE	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes
Adjusted R-squared	0.377	0.367	0.479

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on adulthood outcomes, using the urban sample. The dependent variable is the number of education years completed, normalized by the individual's age. The observation is at the county-cohort level, with each cohort defined by its birth month. All regressions include both individual and weather controls. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table B4 The Effects of Agricultural Fires on Health Expenses (Urban Sample)

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Log Health Expenses				
Diff. Upwind-Downwind Trimester 1	0.008	0.032	-0.049	0.043	0.068
	(0.008)	(0.026)	(0.092)	(0.104)	(0.141)
Diff. Upwind-Downwind Trimester 2	0.006	0.035	-0.015	-0.034	0.085
	(0.008)	(0.023)	(0.080)	(0.123)	(0.116)
Diff. Upwind-Downwind Trimester 3	0.008	0.018	-0.050	0.065	0.088
	(0.008)	(0.024)	(0.077)	(0.131)	(0.121)
Observations	1,834	623	1,185	928	885
Sample	Full	Low Education	High Education	Low Income	High Income
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.140	0.087	0.175	0.211	0.125

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on families' health expenses, using the urban sample. The dependent variable is the logged value of health expenses for children. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table B5 The Effects of Agricultural Fires on Education Expenses (Urban Sample)

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Log Education Expenses				
Diff. Upwind-Downwind Trimester 1	-0.030	0.013	-0.043	0.006	-0.024
	(0.032)	(0.075)	(0.038)	(0.060)	(0.048)
Diff. Upwind-Downwind Trimester 2	-0.023	0.037	-0.037	-0.063	0.040
	(0.033)	(0.067)	(0.040)	(0.057)	(0.044)
Diff. Upwind-Downwind Trimester 3	-0.008	0.025	-0.019	-0.027	0.021
	(0.028)	(0.072)	(0.038)	(0.052)	(0.041)
Observations	1,662	538	1,092	823	811
Sample	Full	Low Education	High Education	Low Income	High Income
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	Yes	Yes	Yes
Adjusted R-squared	0.402	0.342	0.357	0.368	0.336

Notes: This table presents the estimated results of the effects of in-utero agricultural fire exposure on families' education expenses, using the urban sample. The dependent variable is the logged value of education expenses for children. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table B6 The Effects of Early Exposure to NCMS on Adolescent Outcomes (Urban Sample)

	(1)	(2)	(3)	(4)
Dep. Var.	Unhealthiness Index			
Diff. Upwind-Downwind Trimester 1	-0.618	-0.114	-0.668	-0.213
	(0.728)	(0.526)	(0.740)	(0.549)
Diff. Upwind-Downwind Trimester 2	-0.573	0.546	-0.615	0.498
	(0.566)	(0.612)	(0.553)	(0.607)
Diff. Upwind-Downwind Trimester 3	-0.229	0.329	-0.273	0.286
	(0.735)	(0.478)	(0.734)	(0.492)
Observations	828	985	828	985
Sample	Exposure to NCMS	Non-Exposure to NCMS	Exposure to NCMS	Non-Exposure to NCMS
Birth Year by Birth Month FE	Yes	Yes	Yes	Yes
County of Birth FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes
Other Exposure	No	No	Yes	Yes
Adjusted R-squared	0.321	0.448	0.321	0.449

Notes: This table presents the estimated results of the role of NCMS coverage on mitigating the effects of in-utero agricultural fire exposure on adolescent health, using the urban sample. The dependent variable is a normalized health index with a greater value representing worse health conditions. The observation is at the county-cohort level, with each cohort defined by its birth month. Individual controls include gender, age, father's education and age, mother's education and age, family income, family size, and number of siblings. Meteorological controls include dew point, sea level pressure, temperature, wind speed, and rainfall. All weather controls are constructed by averaging the weather conditions during different trimesters. Other exposures include controls on whether the individual is exposed to electricity, tap water, road, railway, and natural gas before age 5. Standard error is clustered at the county level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

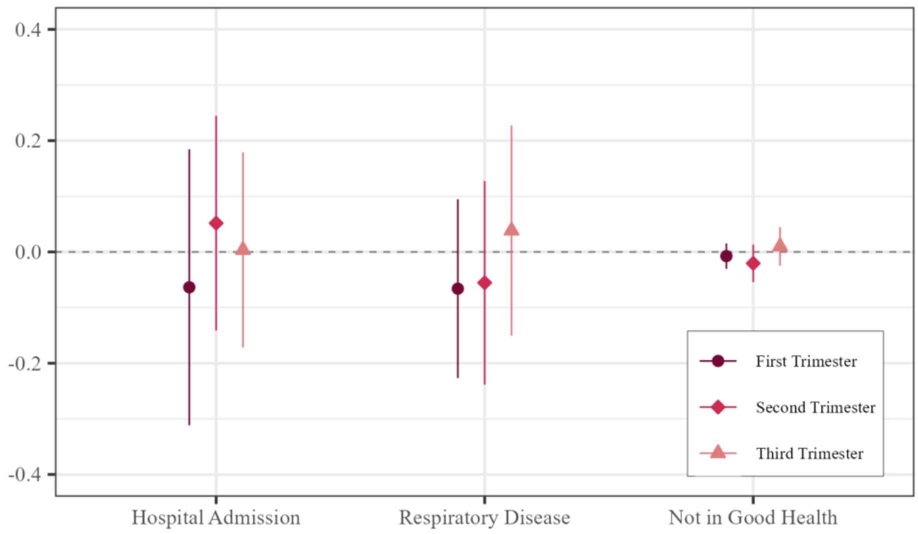


Figure B1 The Effects of Agricultural Fires on Adolescent Health Components (Urban Sample)

Notes: This figure visualizes the estimated coefficients of the effects of in-utero agricultural fire exposure during different trimesters on adolescent health outcomes, including hospital admission, respiratory disease, and self-rated status, using the urban sample. Point estimates and the corresponding 95% confidence intervals are jointly presented.

Appendix C: Robustness Checks

This section briefly discusses the robustness of our baseline estimates. Since our main analyses above suggest that the effects are primarily driven by the male sample, we focus on the male sample for our following robustness checks to avoid redundancy.

C.1 Alternative definition of upwind direction and potential yield coverage

In the baseline regressions, we use the 45-degree criterion to define upwind agricultural potential yields, i.e., whether the angle between the direction of the prevailing wind and the direction of the particular potential yield grid to the county center is less than 45 degrees. To mitigate concerns that our results are driven by the specific choice of wind direction. In Appendix Figure A3, we re-estimate the baseline effects of agricultural fire exposure on health and (non-)cognitive outcomes by changing the definition of upwind direction. Specifically, we consider alternative definitions of 30, 60, and 90 degrees. The estimated results in Figure A3 suggest that our results are robust to alternative definitions of upwind directions.

We then examine whether our estimated results are sensitive to different choices of potential yield coverage radii. In our baseline construction, we calculate the upwind/downwind potential yield covering the entire county. To avoid potential measurement error and examine the sensitivity of our results to different choices of potential yield coverage, we choose alternative radii of 100KM, 70KM, 50KM, and 30KM of coverage. Appendix Figure A4 presents the corresponding results for these alternative specifications. Again, we find that our estimates are stable across different choices of potential yield coverage.

C.2 Controlling for additional trends and fixed effects

Our baseline specification includes birth year by birth month fixed effects to account for the confounding effect of unobserved aggregate time-varying shocks specific to individuals born in different years. However, these unobserved shocks (e.g., shocks to agricultural productivity, natural disasters, or extreme temperatures) may also vary across regions, resulting in the observed association between agricultural fire exposure and adolescent outcomes. We address such concerns in Appendix Table A9 and Table A10. Specifically, in Table A9, we re-estimate our baseline results by additionally including the linear and quadratic province-by-birth-year trends. Whereas in Table A10, we augment our baseline specification by controlling for the province by birth year fixed effects and the city by birth year fixed effects. Throughout different specifications, our results remain quantitatively the same and significant at the conventional level. The inclusion of these granular time trends and fixed effects helps to rule out the possibility of local time-varying confounders and lends us additional credit that our estimated effects are not driven by unobserved shocks. Reassuringly, the estimated coefficients of fire exposure on health and (non-)cognitive measures remain significant, and the magnitudes are similar to our baseline results.

C.3 Alternative clustering adjustments

We examine the robustness of our baseline results to alternative clustering adjustments in Appendix Table A11. Specifically, in odd columns, we augment our baseline clustering (in which the standard error is clustered at the county level to account for serial and cross-sectional correlations within the same county) by using a two-way clustering that cluster the standard error at both the county and the province by birth year level to additionally account for any arbitrary correlations subject to each province within the same birth year. In even columns, we further cluster the standard error at the prefecture level to account for inter-county interactions within the same prefecture. This allows for a more robust inference of our estimated results, which prevents the possibility that our statistical inference may be confounded by unobserved correlations between different counties. Reassuringly, most of our baseline findings remain statistically significant under alternative clustering adjustments, suggesting that our results are less likely to be driven by spurious statistical correlations.

C.4 Accounting for additional confounders

Our baseline estimates reveal a significant correlation between in-utero agricultural fire exposure and adolescent outcomes. However, the presence of several potential confounders may prevent us from convincingly establishing the causal relationship. Specifically, there are two main challenges to our identification. First, despite air pollution from agricultural fires, the potential yield may correlate with other sources of rural pollution. Apart from *in situ* burning, another use of straw residues is for household fuel, which is associated with indoor air pollution. Besides, since higher potential yield is correlated with higher grain output, which may be associated with the intensive use of fertilizer and pesticide, leading to potential water pollution. Second, the agricultural potential yield may be correlated with other factors that simultaneously affect adolescent outcomes. For example, higher potential yield may be correlated with higher agricultural income, which in turn may result in differential access to infrastructure and other facilities.

First, to ensure that our results are not driven by pollution from other sources (e.g., indoor air pollution or fertilizer-induced water pollution), in Appendix Table A12, we include additional controls on whether households use straws as their primary fuel source and the household expenses on fertilizer. The former accounts for the potential channel of indoor air pollution, while the latter accounts for the potential confounding of water pollution. We find that the inclusion of these additional controls for pollution exposure merely affects our estimated effects, suggesting that we can be less concerned that our potential yield measures may be correlated with other sources of rural pollution.

Next, we examine whether our estimated effects are driven by the differential access to infrastructures and other facilities. Specifically, we consider facilities that are related to education and health (e.g., the number of kindergartens, the number of primary schools, the number of hospitals, and pharmacies in the village). We also include controls on whether villages have access to electricity, roads, and railways, to serve as proxies for Infrastructure construction. The corresponding results are reported in Appendix Table A13. We show that the estimated coefficients are stable in both magnitude and significance level after the inclusion of these additional controls.

Lastly, since our measured health and (non-)cognitive outcomes may also be affected by contemporaneous environmental factors (e.g., meteorological conditions and temporal pollutants), we control for a set of contemporaneous weather and pollution conditions to ensure that our estimates are not confounded by these later-life environmental factors. Typically, in addition to the set of early-life weather conditions we have included throughout our empirical estimations, we examine the robustness of our baseline results in Appendix Table A14 by controlling for weather conditions and pollution levels (i.e., $PM_{2.5}$) at the exact survey month. The results remain significant, and we find no large changes in the magnitude of the coefficients. Taken together, the above exercises suggest that our results are less likely to be driven by omitted variables, which again provides support for the validity of our research design that leverages exogenous sources from wind directions for identification.

C.5 Randomized inference

As a further validation, we provide the randomized inference to show that our estimated effects are not the result of any arbitrary idiosyncratic variations. To do so, we first simulate placebo upwind/downwind agricultural potential yield, which are drawn from the same distribution and have the same mean and standard deviation as the original data. We then follow equation (3) and interact the placebo upwind/downwind potential yield with dummies indicating during which trimester the individual is exposed to agricultural fires. For each random draw of potential yield, we re-estimate equation (3) and record the corresponding coefficients. The process is repeated 500 times. As the effect of in-utero agricultural fire exposure on adolescent outcomes is primarily driven by exposure during the first and third trimesters, we mainly focus on examining the extent to which the placebo coefficients for fire exposure in these two trimesters can replicate our baseline estimates. Appendix Figure A5 plots the corresponding distribution of placebo coefficients on four of our main outcome variables. We find that the placebo coefficients are centered around zero and are small in magnitude, and are far away from the true coefficients. This piece of evidence suggests that our estimated effects are plausibly unlikely to be accounted for by arbitrary idiosyncratic variations.

C.6 Measurement error

One potential challenge to our empirical design is the presence of measurement error in the right-hand side of equation (3). Specifically, there are three potential sources of measurement errors. The first is the measurement of trimesters during which the individual is exposed to fire, the second is the measurement of fire intensity, and the third is the measurement of wind directions. We discuss these potential measurement errors in turn.

Consider first the measurement of trimesters. A key information to pin down the trimester of fire exposure is the birth month. However, birth month, especially in rural China, may be recorded according to either the Gregorian Calendar or the Lunar Calendar. Since the difference between the two calendars could be more than one month, this may significantly bias the allocation of trimesters. To ensure this is not the case, we combine information on the exact year and month in which the survey was conducted, which allows us to calculate the age of individuals on a monthly basis. We can then compare the calculated age in months with the actual age in months recorded in the dataset. If the month of birth is recorded using the Lunar Calendar (or a combination of calendars), then there could be substantial differences between our calculated age in months and the actual age in months reported in the dataset. Reassuringly, we find only slight differences between the two measures; less than 1% of the survey sample have inconsistent ages in months. This alleviates our concerns that there may be substantial errors in the measurement of birth month and thus the allocation of trimesters.

Consider next the measurement of fire intensity. Since data on agricultural fire before 2000 is unavailable, we exploit agricultural potential yield as an indirect measure of fire intensity. Nevertheless, there may be concerns that this measure may have measurement error and may be unable to provide sufficient variation to identify the effect of prenatal agricultural fire exposure. We adopt an alternative strategy, which leverage additional variation of agricultural fires, to ensure that our results are robust to potential measurement errors. Our solution is similar to a two-step IV estimation in which we first use equation (1) to separately estimate the impact of upwind/downwind potential yield on the number of agricultural fires, and obtain the corresponding fitted values (which we refer to as $\widehat{Upwind\ Fire}$ and $\widehat{Downwind\ Fire}$). Since the sample period does not match between years in the county sample and birth years in CFPS 2010, we average the fitted value to the county level and then match it to the survey sample. We then replace the explanatory variable $Upwind\ APY$ and $Downwind\ APY$ in equation (3) with $\widehat{Upwind\ Fire}$ and $\widehat{Downwind\ Fire}$, and estimate this modified specification. The results are presented in Appendix Table A15. The estimated results remain significant, and we see that the magnitude are larger than the baseline estimates. This is plausible as classic measurement errors on the RHS leads to a lower bound estimate. We note that, however, the significance of our results should be interpreted with caution as the two-step IV can leads to misleading estimates of the standard errors.

Consider last the measurement of wind direction. Since we only exploit monthly variations for identification, fluctuations in monthly wind direction across different years may cast potential measurement errors to our design. We believe this is not a big concern as wind pattern in China is relatively stable across years. To provide empirical support to our argument, we leverage a similarity metric, the Angular Variance, to examine whether the distribution of wind direction in a given month is similar across different years. The key idea of this metric is to treat each wind direction as a vector on a circle, average all the vectors, and then see how long the resulting average vector is. A long average vector would imply all directions are similar; whereas a short one means they cancel each other out and thus the directions are more dispersed. Specifically, the Angular Variance is calculated with the following steps:

First, for a given wind direction observation θ_{cmt} in county c, month m and year t, we calculate its x and y coordinates on the unit circle, i.e., $x_{cmt} = cos(\theta_{cmt})$ and $y_{cmt} = sin(\theta_{cmt})$. Next, for each county-month pair, we aggregate the average x and y coordinates across years to get $x_{cm} = \frac{1}{T}\sum_t x_{cmt}$ and $y_{cm} = \frac{1}{T}\sum_t y_{cmt}$, where $t \in \{1990, \cdots ,2010\}$. We can then calculate the length of the mean resultant vector (R_{cm}), with $R_{cm} = \sqrt{x_{cm}^2 + y_{cm}^2}$. It is clear that the value of R_{cm} always lies between 0 and 1. The closer R_{cm} is to 1, the more concentrated the wind direction; the closer it is to 0, the more dispersed the wind direction. Finally, the angular variance is then defined as $Var(\theta_{cm}) = 1 - R_{cm}$. Opposite to R_{cm}, a larger value of $Var(\theta_{cm})$ would imply higher dispersion in wind directions across different years.

To examine the extent of wind direction dispersion, in Appendix Figure A6, we visualize the distribution of angular variance for all months, using box and violin plots. We see that for most of the months, the distribution of angular variance is centered well below 0.1. While there are some fluctuations and outliers for several months, the overall pattern suggests that the monthly wind directions are less dispersed.

Besides measurement error, a related concern is that the statistical power may be low, considering our relatively small sample size. We show that this is not a concern by directly calculating the statistical power of our regressions. Reassuringly, most of our specifications have a statistical power larger than 0.95, even for the Positive Attitude outcome, which has the smallest sample size. The empirical statistical power is larger than the suggested threshold of 0.8, implying that insufficient statistical power is less of a concern.

C.7 Sample representativeness

Due to our small sample size, another concern is the sample representativeness. Sample representativeness could either bias our estimation or reduce the external validity of our findings, if not both. We provide two pieces of evidence to show that our sample is representative. First, we show in Appendix Figure A7 that the distribution of birth month in our sample is similar to that from the 2010 Population Census, which covers information on more than 2 million individuals. This further corroborates that sample selection is of less concern. Second, we show that our results are robust to an alternative weighting estimation that exploits the national sampling weights provided in the CFPS dataset. The corresponding results are presented in Appendix Table A16, where we find that the estimated coefficients are even larger and more significant. This lends additional confidence that our sample is indeed representative.

Appendix D: Additional Data Description

D.1 Air pollution and other meteorological data

We measure air pollution using $PM_{2.5}$, the major pollutant emitted from agricultural fires. We obtain ground-level $PM_{2.5}$ data from the NASA Socioeconomic Data and Applications Center (SEDAC).[41] The dataset combines AOD retrievals from multiple satellite algorithms and exploits the GEOS-Chem chemical transport model to relate the total column measure of aerosol to near-surface $PM_{2.5}$ concentration. Calibration is performed using Geographically Weighted Regression (GWR) to produce the final products. We aggregate the raw raster data of $PM_{2.5}$ concentrations, originally captured at a 1km resolution, to the county level and compute the year-month average of $PM_{2.5}$ concentration for each county.

We obtain meteorological data from the fifth-generation European Center for Medium-Range Weather Forecasts reanalysis dataset (ECMWF ERA-5). The ERA-5 dataset provides hourly, daily, and monthly atmospheric conditions at a resolution of 0.1 degrees (which is approximately 11km). We download a sequence of monthly weather conditions, including temperature, precipitation, humidity, sea level pressure, and wind speed. We collapse the weather data to the county-year-month level.

The weather data serves two purposes. On the one hand, it allows for a more precise estimation of the effects of agricultural fires, as well as the potential yield on air pollution. More importantly, it allows us to control for the confounding effects of other in-utero weather conditions. For instance, a large strand of literature has shown that in-utero or early-life exposure to exogenous weather shocks (e.g., extreme heat, rainfall, drought, flood, etc.) could have both short-term impacts on birth outcomes and enduring effects on individuals' life trajectories. To account for the confounding effects of these weather conditions, we control for the in-utero exposure to weather shocks by calculating the weather conditions experienced during each trimester.

[41] See more details from https://sedac.ciesin.columbia.edu/data/set/sdei-global-annual-gwr-pm2-5-modis-misr-seawifs-aod-v4-gl-03.