Misra, Shashank; Kulshreshtha, Shobhit

**Working Paper**

# Impact of Air Pollution on Birth Outcomes: Causal Evidence from India

GLO Discussion Paper, No. 1635

**Provided in Cooperation with:**
Global Labor Organization (GLO)

*Suggested Citation:* Misra, Shashank; Kulshreshtha, Shobhit (2025) : Impact of Air Pollution on Birth Outcomes: Causal Evidence from India, GLO Discussion Paper, No. 1635, Global Labor Organization (GLO), Essen

This Version is available at:
https://hdl.handle.net/10419/322534

# Impact of Air Pollution on Birth Outcomes: Causal Evidence from India*

Shashank Misra† Shobhit Kulshreshtha‡

**Abstract**

India consistently ranks among the countries with the highest levels of ambient air pollution worldwide. At the same time, it faces significant challenges in neonatal health, with newborns having low average birth weights and a high incidence of being born within the low birth weight (LBW) and very low birth weight (VLBW) category. Using data from the Indian National Family Health Survey (NFHS), we examine the impact of in-utero exposure to particulate matter on a number of birth weight indicators. We exploit variation in wind direction during the in-utero period to capture quasi-random variation in particulate matter exposure for each child. We find that reducing in-utero $PM_{2.5}$ exposure by one standard deviation would lead to 1.3% increase in average birth weight, a 2.7 percentage point decrease in the incidence of LBW births and a 0.6 percentage point decrease in the incidence of VLBW births respectively. Drawing on estimates from prior studies, we find that the observed improvements in both average birth weight and reductions in LBW incidence from meeting WHO air quality standards could yield substantial long-run economic benefits, potentially amounting to billions of dollars annually in addition to broader gains in child health, cognition, and educational outcomes.

**JEL-codes:** J13, I12, Q53
**Keywords:** birth weight, air pollution, in-utero exposure, India

---

*We would like to thank Valentina Melentyeva, Charlie Moss, participants at the Health Economists Study Group Conference, SEG Seminar Tilburg University and GSS Seminar Tilburg University for their suggestions and comments

†Department of Econometrics and Operations Research, Tilburg University

‡Department of Econometrics and Operations Research, Tilburg University and GLO. Corresponding Author: s.kulshreshtha@tilburguniversity.edu

# 1 Introduction

India has experienced rapid economic growth over the past few decades, accompanied by a significant increase in air pollution levels (Sicard et al., 2023). Air pollution, particularly through particulate matter, has substantial adverse effects on the economy, contributing to the loss of healthy years of life and increasing the burden of disease. In 2019 alone, air pollution is estimated to have cost the Indian economy over $36 billion due to its negative impact on human capital (WHO, 2024). Research has shown that air pollution can have a negative effect on health and cognitive development of children (Balakrishnan & Tsaneva, 2021; Balietti et al., 2022), and a growing body of literature highlights the short- and long-term consequences of fetal exposure to particulate matter on child health (Ai et al., 2023; Palma et al., 2022). However, much of this literature focuses on developed countries, raising concerns that the effects of pollution on child health may differ between developed and developing countries contexts (Arceo et al., 2016). In this study, we examine the effect of in-utero exposure to particulate matter ($PM_{2.5}$) on neonatal health outcomes in India, using birth weight as the primary health outcome.

Studying the impact of air pollution on neonatal outcomes in low- and middle-income countries (LMICs) is essential for two main reasons (Currie et al., 2014). First, pollution levels in these countries are often much higher than in high-income nations. For instance, average $PM_{2.5}$ exposure between 2010 and 2019 was 38 $\mu$g/m$^3$ in India, compared to 10 $\mu$g/m$^3$ in the United States over the period 2010–2019.[1] This higher "dose" of pollution exposure may result in effect sizes that differ substantially from those observed in high-income countries, regardless of the specific outcome being measured. In support of this, Arceo et al. (2016) find that the negative impact of pollution on infant mortality is greater in a developing country such as Mexico than in the United States. Second, baseline maternal health in developing countries is typically poorer, which may cause the effects of air pollution on birth outcomes to be more heterogeneous. On one hand, poorer maternal health may amplify the adverse effects of pollution on neonatal outcomes. On the other hand, it may

[1] These figures are computed using pollution data from the Goddard Earth Sciences Data and Information Services Center (NASA) and the United States Environmental Protection Agency.

obscure pollution's effects, as other health complications during pregnancy could play a more dominant role. This highlights the need for context-specific research in developing countries, where vulnerabilities and exposure patterns differ. In the context of our study, India, despite its recent economic growth, continues to perform poorly in terms of neonatal health. The incidence of adverse birth outcomes in India is high not only relative to industrialized nations, but also compared to other low- and middle-income countries (Marete et al., 2020).

In this study, we use birth weight as the primary indicator of child health. Birth weight is a well-established predictor of short- and medium-term health outcomes (Hummer et al., 2014; McGovern, 2019), and it also has long-term implications, including adult mortality (Risnes et al., 2011). Beyond health, birth weight is closely linked to cognitive development, educational attainment, and future earnings, making it a strong determinant of labor market outcomes (Behrman & Rosenzweig, 2004; Black et al., 2007; Cook & Fletcher, 2015; Royer, 2009). Improving birth weight yields substantial economic benefits through reduced healthcare costs and increased productivity, as documented in both developed and developing countries (Alderman & Behrman, 2006; Almond et al., 2005). By estimating the effect of in-utero exposure to air pollution on birth weight in India, our study offers valuable insights into how early-life environmental conditions can shape human capital development in a low- and middle-income country context.

We use two recent waves of the Indian Demographic and Health Survey (DHS), 2015–16 and 2019–21, to causally estimate the relationship between in-utero exposure to particulate matter ($PM_{2.5}$) and birth weight for children born between 2010 and 2019. A simple OLS regression of birth weight on in-utero pollution exposure may produce biased estimates due to omitted variable bias—such as unobserved local characteristics that affect both pollution levels and birth outcomes—and attenuation bias from measurement error. To address these endogeneity concerns, we exploit quasi-random variation in pollution exposure during gestation, driven by heterogeneity in wind patterns.

Using our instrumental variables (IV) strategy, we find that a one standard deviation increase in $PM_{2.5}$ exposure during pregnancy reduces birth weight by approximately 1.3% relative to the sample mean of 2.8 kilograms—equivalent to a decrease of about 36.4 grams.

In comparison, Pons (2022) find no significant average effect of $PM_{2.5}$ on birth weight in the United States, though they do observe a reduction of 28 grams in the lower tail of the distribution. Similarly, Palma et al. (2022) find that an increase in $PM_{10}$ levels by one standard deviation in Italy reduces birth weight by 0.5%. Our results indicate a larger effect size in the Indian context.

Furthermore, we estimate that reducing average in-utero $PM_{2.5}$ exposure to the WHO-recommended level of 10 $\mu$g/m$^3$ would lead to a 5.4 percentage point decline in the incidence of low birth weight (LBW), and a 1.2 percentage point reduction in the incidence of very low birth weight (VLBW). Back-of-the-envelope calculations suggest that compliance with WHO guidelines could generate economic gains exceeding $1 billion, solely from improvements in neonatal health outcomes. These findings are robust across a range of alternative model specifications and sensitivity tests addressing potential violations of the key assumptions underlying our IV strategy.

We make a significant contribution to the growing literature on the health and economic consequences of air pollution. Numerous studies have provided causal evidence on the adverse effects of air pollution on short-, medium-, and long-run health outcomes (Almond et al., 2009; Chay & Greenstone, 2003; Fan et al., 2023; Neidell, 2004; von Hinke & Sørensen, 2023). Others have highlighted its impacts on cognitive development and labor market outcomes (Balakrishnan & Tsaneva, 2021; Isen et al., 2017; Sanders, 2012). Beyond health, labor, and cognition, research has also explored how pollution affects crime, real estate markets, and time use (Bondy et al., 2020; Chay & Greenstone, 2005; Herrnstadt et al., 2021; Jafarov et al., 2023). As noted by von Hinke and Sørensen (2023), much of the existing literature—particularly within economics—focuses on immediate birth-related health outcomes. However, these studies have predominantly been conducted in industrialized countries due to better data availability, leaving a gap in causal evidence from low- and middle-income countries (Li & Zhang, 2024; Tang et al., 2024).

Using nationally representative Indian data, our study helps fill this gap by providing causal evidence from a middle-income country characterized by high pollution levels and poor neonatal health outcomes. To our knowledge, this is the first paper to causally examine

the relationship between in-utero exposure to air pollution and a range of birth outcomes in India—one of the most polluted countries globally. It is also only the second causal study in India to contribute to the literature on the fetal origins hypothesis, which posits that early-life environmental conditions, including in-utero exposure, have long-term effects on health and human capital formation (Almond & Currie, 2011). The only other Indian study we are aware of, Singh et al. (2019), examines postnatal anthropometric outcomes (e.g., height-for-age, weight-for-age) rather than neonatal birth outcomes.

We also contribute to the emerging body of research analyzing the impacts of pollution on neonatal health in developing countries (Arceo et al., 2016; Bharadwaj & Eberhard, 2008; Jayachandran, 2009; Li & Zhang, 2024). As previously discussed, this line of inquiry is particularly important in low- and middle-income countries where higher pollution exposure and poorer maternal health conditions may lead to different effect sizes and mechanisms compared to high-income contexts. Our study enables meaningful comparisons with estimates from other settings while offering India-specific insights that are crucial for public health policy.

Our findings are especially timely and policy-relevant given the rising national and international attention on air quality in India (Murukutla et al., 2017). By linking our causal estimates to prior literature on the economic costs associated with low birth weight, we also provide back-of-the-envelope calculations that quantify the potential economic benefits of improving in-utero air quality. These estimates offer a compelling case for stronger environmental regulation and public health interventions aimed at mitigating the adverse impacts of air pollution on early-life health.

The rest of the paper is structured as follows, Section 2 outlines the background in India concerning pollution and health dynamics, Section 3 provides a brief description of the data sources used in this study, Section 4 illustrates the identification strategy and model specification, Section 5 reports the results and robustness of our estimates before we use our point estimates to derive the costs of pollution arising from sub-optimal birth outcomes in Section 6 and provide concluding remarks in Section 7.

# 2 Background

We consider the case of India in this study because it exhibits a striking profile in both dimensions of the relationship we are exploring, pollution levels and neonatal health outcomes. In terms of pollution, all Indian states have $PM_{2.5}$ levels exceeding the UN safe limit of 10 $\mu$g/m$^3$, an important measure of air pollutant (Balakrishnan et al., 2019). $PM_{2.5}$ is considered the main pollutant to assess the impact of air pollution on various health indicators (Balietti et al., 2022). In addition, the Central Pollution Control Board of India[2] considers the levels of particulate matter to be the most critical and general indicator of air quality to make policy decisions (Greenstone & Hanna, 2014). Recently, it has been observed that nearly 80% of the Indian population live in regions with annual $PM_{2.5}$ concentration levels of more than 40 $\mu$g/m$^3$, which falls under the severe air pollution level category that can induce health complications according to the World Health Organization (Balakrishnan et al., 2019).

India also performs abysmally when it comes to indicators measuring neo-natal health. Approximately 750 thousand neo-nates die in India every year, i.e., within the first month of birth (Sankar et al., 2016). In addition, 60% of all children deaths under the age of 5 occur within the neo-natal phase (El Arifeen et al., 2017). Birth weight is a critical indicator of neonatal health and India has one of the lowest average birth weight levels not only in the world, but also among low- and middle-income countries (Marete et al., 2020). India continues to have a high prevalence of births that fall into the Low Birth Weight (LBW) category, estimated to be somewhere between 24 to 30 percent. The nation accounts for 40 percent of all LBW births globally (Bhilwar et al., 2016; Sankar et al., 2016). Given India's dual burden of high ambient pollution and suboptimal neonatal health, investigating the impact of in-utero exposure on birth weight is not only relevant, but vital to uncovering the mechanisms through which environmental stressors in-utero shape early-life health trajectories. The findings can also inform policy makers to improve public health and mitigate children health problems due to air pollution in India.

[2]The Central Pollution Control Board (CPCB), established in 1974 under the Water (Prevention and Control of Pollution) Act and later empowered by the Air (Prevention and Control of Pollution) Act of 1981, serves as India's national authority for monitoring and controlling environmental pollution.

Medical literature has reported various mechanisms through which exposure to pollutants, such as fine particulate matter, can impact in-utero fetal development and subsequently, birth weight. Particulate matter exposure can cause pulmonary inflammation among mothers, which can potentially disrupt oxygen supply and nutritional movement to the fetus, resulting in pregnancy related complications and lower birth weight of children born to these mothers (Sun et al., 2016). In addition, studies have argued that pregnant women can face high oxidative stress that is caused by the presence of metals in the particulate matter. An increase in oxidative stress is likely to hinder embryo growth, which can critically impact the development of the child in the early stages of pregnancy (Kannan et al., 2007). Moreover, increased oxidative stress can also cause the formation of DNA adducts within the placenta, impairing the ability of the uterus to support fetal growth (Topinka et al., 1997). As an additional mechanism, particulate matter exposure has also been linked to interference with maternal hormones, with evidence that it can lead to maternal thyroid imbalances, which in turn negatively impact birth-weight outcomes (Blazer et al., 2003; Janssen et al., 2017). These mechanisms highlight the plausible biological pathways through which ambient pollution, particularly $PM_{2.5}$, can adversely influence fetal development and hence cause lower birth weight. Together, these epidemiological patterns and biological mechanisms underscore the urgency of examining the causal impact of ambient $PM_{2.5}$ exposure on birth outcomes in the Indian context, where both environmental and neonatal health vulnerabilities converge.

# 3 Data and Descriptive Statistics

## 3.1 Birth and Demographics Data

### 3.1.1 Birth Weight and Size

We obtain data on birth outcomes from the Demographic Health Survey (DHS) of India which is a nationally representative repeated cross sectional survey covering key metrics related to maternal and child health and nutrition in India. We use the last two waves of the

DHS, which were conducted during the years 2015-2016 and 2019-2021, respectively[3]. The women taking part in the survey who fall between the reproductive age of 15-49 are classified as 'eligible' for questions on the health of their offspring, and are asked detailed questions on births which took place in the last 5 years since the date of the survey. From the two waves of the survey we have information on the birth weight of children born in the period 2010 to 2021.[4] We exclude children born in 2020 and 2021 from our analysis, as we believe that children born during the pandemic period might systematically vary from those born prior to this period. Moreover, there can be other unobservable factors during the pandemic period which may impact child health and the pollution level in the region. Therefore, we restrict our analysis to births between the period 2010 and 2019. The final sample consists of 322,340 children born during this period.

We are able to identify the birth weight of the children using the women's questionnaire. The respondents report birth weight (in grams) of the children born in the last 5 years, either via a written medical card, which was recorded upon the birth of a child, or via recall. In our sample, approximately 60% report the birth weight via the medical card, while the remaining report the birth weight of their children via recall. The data enables us to construct three measures of birth weight, 1) a continuous measure of birth weight, 2) a binary measure of Low Birth Weight (LBW), which is 1 if the child was born with a weight of less than 2500 grams and 0 otherwise, and 3) a binary measure of Very Low Birth Weight (VLBW), which is 1 if the child was born with a weight of less than 1500 grams and 0 otherwise. As an additional outcome, we construct a variable that equals 1 if the size of the child at birth is reported as average or above average, and 0 otherwise.

[3]We are unable to use the previous three waves since they do not hvae information on the exact location of residence of respondents. A precise location of the respondent is crucial since we construct our measure of in-utero pollution exposure by matching district level particulate matter levels over this period

[4]Our data on in-utero variables weather variables (such as wind, pollution and other weather variables) spans from the years 2010-2019, hence, births in the first 9 months of 2010 are not included as in-utero data for these births stretches into 2009.

### 3.1.2 Demographics, and Socio-economic Background

DHS also provides information on various socio-economic and demographic characteristics of the households that are surveyed. We include a number of mother and child characteristics. Specifically, we include mother's anemic level, mother's BMI, age of the mother when child was born, gender of the child, and birth order of the child. As birth weight has been shown to be affected by the socioeconomic background of the mother and other characteristics of the household in the Indian context (Kader & Perera, 2014), we control for these variables in our baseline model. Specifically, we include wealth index to control for socio-economic background.[5] In addition, we include demographic characteristics such as religion and caste of the household. We also include a binary variable for whether the mother lives in an urban area or not. Moreover, the cooking fuel used by the family has been shown to influence the exposure of children to indoor air pollutants (Pope et al., 2010). As a result, we also control for a binary indicator taking the value 1 if the cooking fuel used by the household was dirty and 0 otherwise.[6]

## 3.2 Pollution Data

Air pollution data are taken from the Goddard Earth Sciences Data and Information Services Center (GES DISC) funded by NASA which provides a total surface mass concentration of $PM_{2.5}$ with spatial resolution of $0.5^\circ \times 0.625^\circ$. The data is sourced from the MERRA-2 satellite. For our analysis, we obtained the monthly data for the Indian union for the period 2010 to 2019. For each year and month, we have calculated the average concentration of $PM_{2.5}$ at the district level. To construct in utero exposure to $PM_{2.5}$, we computed the 10-month average concentration including the month of birth and nine months preceding it. We make the assumption that the mother resided in the same district in which the child was born during her pregnancy. We will discuss this assumption in more detail in Section 5.2.

[5]Wealth Index variable is provided on a 5 point scale ranging from 1 to 5 with 1 indicating the lowest wealth level and 5 indicating the highest wealth level

[6]Dirty fuel sources include kerosene, coal, charcoal, wood, straw/shrubs/grass, agricultural crop waste, and animal dung.

## 3.3 Weather Data

### 3.3.1 Wind Direction

To gather information on wind direction—which influences the concentration of $PM_{2.5}$ in a district—we obtain wind data for India from the ERA5 hourly dataset on surface levels, available through the Copernicus Climate Data Store. We use 10 daily observations of wind direction at the surface level to construct district-wise monthly averages of both the share of time the wind blew from each of the four cardinal directions, and the average wind speed, based on the u- and v-components of wind data.[7] The data are available at a resolution of $0.25^\circ \times 0.25^\circ$ (approximately 25 km $\times$ 25 km), enabling the construction of accurate aggregated wind patterns for each district and month.

Following the method used to calculate in-utero $PM_{2.5}$ exposure, we compute the 10-month average of the share of time the wind blew from each of the four cardinal directions during pregnancy and the month of birth of the child. This variable serves as an instrument for in-utero exposure to $PM_{2.5}$ concentration.

### 3.3.2 Temperature, Wind Speed, and Precipitation

In-utero weather conditions are known to influence fetal growth and development (Hong, 2025). Accordingly, we include weather controls such as mean temperature, wind speed, humidity, and precipitation during the in-utero period. Incorporating these controls allows us to account for the atmospheric conditions to which the fetus was exposed throughout gestation. These data are obtained from the same source as the wind direction data—the ERA5 hourly surface-level dataset provided by the Copernicus Climate Data Store—and are constructed at the district-month level using 10 daily observations. In-utero weather variables are calculated by averaging the monthly values over the ten-month period comprising the nine months preceding birth and the month of birth, consistent with the methodology used to construct the pollution and wind direction variables.

[7] Details on how to convert the u- and v-components of wind into wind speed and direction are provided in Appendix A.

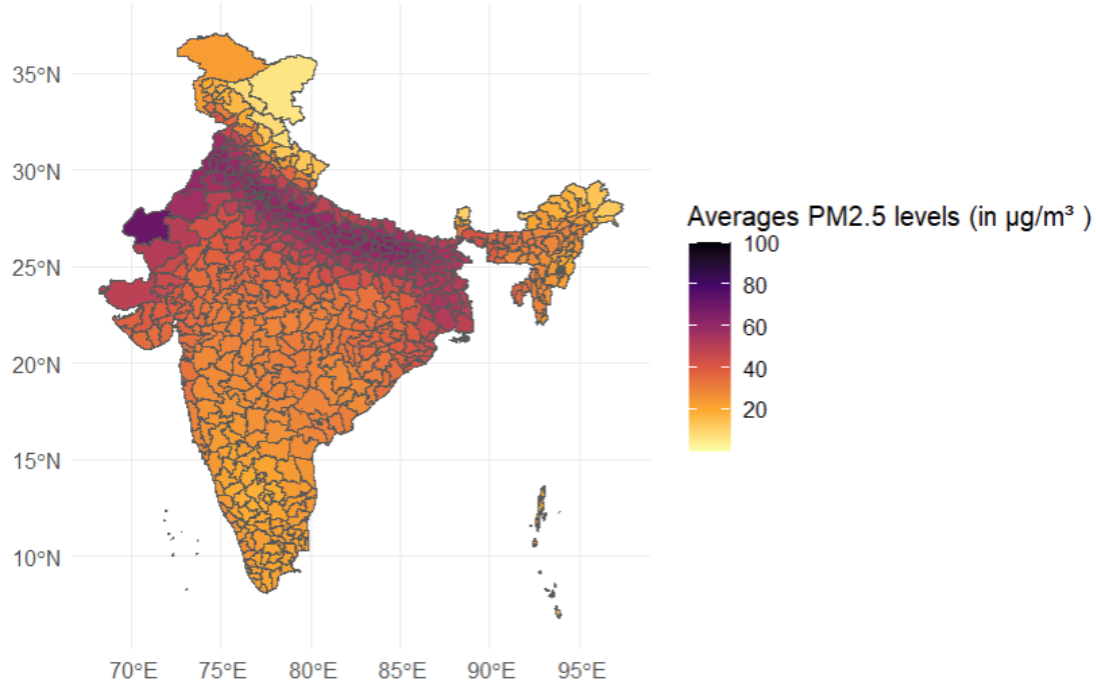
## 3.4 Descriptive Statistics

### 3.4.1 Regional and Temporal Variation: Pollution

Figure 1 illustrates regional patterns in particulate matter trends, based on average $PM_{2.5}$ concentrations from 2010 to 2019. A clear north-south divide emerges, with significantly higher pollution levels in northern states compared to the south. One major source of pollution spikes is crop residue burning in the agricultural plains of North and Northwest India, typically occurring between mid-October and early November to prepare fields for the next harvest (Jain et al., 2014; Jethva et al., 2019). A second pollution wave affects the Indo-Gangetic Plain from November to January, driven by secondary aerosol formation due to increased biofuel use and waste burning, compounded by meteorological conditions that trap pollutants (Kanawade et al., 2020; Saharan et al., 2024; Sen et al., 2017). Additionally, dust storms in the Thar Desert during May to July elevate particulate levels in Northwestern state of Rajasthan and adjacent regions.

In contrast, southern India experiences relatively stable pollution levels year-round, with slight increases between November and April due to dry conditions. More generally, the monsoon period, for which the exact month(s) of incidence varies for each region results in improvements in particulate matter readings. The monsoon period is largely contained between the months of June and September for most regions in India. Figure B1 in the Appendix B presents the seasonal variation in average particulate matter concentration. We observe that temporal patterns are in line with aforementioned descriptions, with the Indo-Gangetic belt experiencing severe particulate matter pollution in the last quarter of the year and the North-Western desert regions encountering pollution spikes in the hot and dry summer months.

Figure 1: Average $PM_{2.5}$ concentrations across Indian districts




### 3.4.2 Regional and Temporal Variation: Wind Direction

Table 1 presents the average within- and between-district variation in wind direction, measured by the standard deviation of the share of wind coming from each direction across districts within a region.[8] We observe considerable variation in wind patterns both across districts and over time within districts, indicating that there is substantial variation that can be used for our identification strategy. The regional and spatial variation in wind patterns implies that children born within the same region–month pair may be differentially exposed to pollution due to exogenous wind patterns, thereby generating plausibly random variation in in-utero $PM_{2.5}$ exposure. Table 1 shows that approximately 15–19% of the variation in wind direction shares arises between districts within the same region. The presence of substantial cross-district variation within the same regions supports our identification strategy, which leverages wind direction as an instrument and includes region-by-month of birth fixed effects.

[8]Following the approach by Deryugina et al. (2019), we divide India into 30 regions using a K-means clustering algorithm based on district centroid coordinates, in order to create clusters of uniform size. Figure C1 displays the resulting clusters.

Table 1: Decomposition of Wind Direction Share Variation: District-Level Averages by Region

| Direction | Overall SD | Between SD | Within SD | Between (%) | Within (%) |
|---|---|---|---|---|---|
| North | 0.1327 | 0.0478 | 0.1197 | 18.72 | 81.28 |
| East | 0.1945 | 0.0603 | 0.1781 | 16.01 | 83.99 |
| South | 0.2527 | 0.0749 | 0.2347 | 15.02 | 84.98 |
| West | 0.1555 | 0.0511 | 0.1429 | 16.04 | 83.96 |

Notes: This table decomposes the standard deviation of district-level wind direction shares into between- and within-district components *averaged across regions*. Between SD reflects cross-sectional variation across districts within each region, while Within SD captures average temporal variation within districts. Percentages indicate the share of total variance attributed to each source.

### 3.4.3 Summary Statistics

Table 2 presents summary statistics for the sample of children used in our baseline analysis. The average birth weight in the sample is approximately 2,800 grams, with 17% of children classified as having low birth weight (LBW) and 1% as very low birth weight (VLBW). Nearly 90% of children were reported to be of average or above-average size at birth. The lower value of birth order (2.07) indicates that the majority of children were either first- or second-born. Approximately 48% of the children in the sample were female. Regarding household characteristics, about 59% of households used dirty cooking fuel, highlighting the importance of controlling for this variable. On average, in-utero exposure to $PM_{2.5}$ was 40 $\mu$g/m$^3$, which exceeds the levels recommended by the World Health Organization (WHO). Other household and demographic characteristics used as covariates are also reported.

Table 3 reports the summary statistics for our $PM2.5$ data with other weather variables such as wind speed, temperature, and precipitation at the month-district level, spanning from the years 2010-2019. We observe a clear pattern of extremely high average $PM_{2.5}$ levels, with a sizable share of all of our district-month pairs having $PM_{2.5}$ levels which fall within the 'severe' category as defined by the WHO.

Table 2: Summary Statistics: DHS Sample

| **Variable** | Mean | SD | Min | Max |
|---|---|---|---|---|
| **Mother and Child Characteristics** | | | | |
| Birth Weight (grams) | 2,825.25 | 582.11 | 500 | 9,000 |
| LBW | 0.167 | 0.373 | 0 | 1 |
| VLBW | 0.010 | 0.100 | 0 | 1 |
| Size is Average or Above | 0.896 | 0.306 | 0 | 1 |
| Mother is Anemic | 0.824 | 0.826 | 0 | 3 |
| Mother's BMI ($\times 100$) | 21.642 | 3.954 | 12.02 | 59.99 |
| Mother's Age at Birth | 25.023 | 4.792 | 11 | 49 |
| Child's Birth Order | 2.071 | 1.205 | 1 | 6 |
| Female Birth | 0.478 | 0.500 | 0 | 1 |
| **Religion** | | | | |
| Hindu | 0.764 | 0.425 | 0 | 1 |
| Muslim | 0.114 | 0.318 | 0 | 1 |
| Christian | 0.079 | 0.269 | 0 | 1 |
| Other Religion | 0.043 | 0.203 | 0 | 1 |
| **Caste** | | | | |
| Scheduled Caste | 0.202 | 0.402 | 0 | 1 |
| Scheduled Tribe | 0.204 | 0.403 | 0 | 1 |
| OBC | 0.408 | 0.491 | 0 | 1 |
| Other Castes | 0.387 | 0.487 | 0 | 1 |
| **Household** | | | | |
| Urban | 0.241 | 0.428 | 0 | 1 |
| Wealth Index | 2.830 | 1.376 | 1 | 5 |
| Dirty Cooking Fuel | 0.589 | 0.492 | 0 | 1 |
| **Pollution Exposure** | | | | |
| $PM_{2.5}$ In-utero ($\mu$g/m$^3$) | 39.837 | 14.678 | 4.739 | 94.203 |
| **Number of Observations** | | 322,340 | | |

Notes: Sample includes children born between 2010–2019 from waves 4 and 5 of DHS. BMI Variable from DHS is scaled by 100 for readability. Household level cooking fuel has over 10 categories; we summarize it as a binary indicator for use of dirty fuel sources: kerosene, coal, charcoal, wood, straw/shrubs/-grass, agricultural crop waste, or animal dung.

Table 3: Summary Statistics: Weather and Pollution Variables

| **Variable** | Mean | SD | Min | Max |
|---|---|---|---|---|
| **Wind Shares** | | | | |
| East Share | 0.228 | 0.219 | 0 | 1 |
| North Share | 0.135 | 0.180 | 0 | 1 |
| South Share | 0.417 | 0.278 | 0 | 1 |
| West Share | 0.220 | 0.205 | 0 | 1 |
| **Weather** | | | | |
| Temperature (K) | 298.001 | 8.724 | 246.784 | 311.945 |
| Wind Speed (m/s) | 3.413 | 1.841 | 0.379 | 11.018 |
| **Pollution** | | | | |
| $PM_{2.5}$ ($\mu$g/m$^3$) | 38.064 | 21.829 | 1.751 | 174.547 |
| **Number of Observations** | 71,160 | | | |
| $PM_{2.5} >$ **35 ($\mu$g/m$^3$)** | 33,051 observations (45%) | | | |

Notes: Summary based on merged monthly panel of 71,160 district-year-month observations (2010–2019). $PM_{2.5}$ values above 35 $\mu$g/m$^3$ are considered above severe by WHO standards. Wind shares represent the proportion of time wind blew from a particular 90-degree bin using 10 daily observations. Summary table reported based on all Indian districts except for Lakshwadeep for which the data is missing.

# 4 Methodology

To estimate the relationship between in-utero pollution exposure and birth outcomes we use the following equation:

$$Y_{idrmt} = \beta \, \overline{PM2.5}_{i(d,m,t)} + \lambda X_i + \nu W_{i(d,m,t)} + \delta_{mr} + \gamma_t + \epsilon_{idrmt} \tag{1}$$

where the dependent variable is the birth outcome of the child $i$, born in district $d$ of region $r$, in month $m$ of year $t$.[9] The key variable of interest is $\overline{PM2.5}$, which measures in-utero exposure to air pollution and is constructed as the average particulate matter ($PM2.5$) concentration in the child's district over the month of birth and the nine preceding months of gestation. The parameter of interest, $\beta$, captures the effect of particulate matter exposure on the selected birth outcome. $X$ and $W$ are a vector of socio-demographic and in-utero weather

[9]The regions as defined in Section 3.4 are construed using the K-means clustering algorithm. We divide India into 30 region clusters which is similar to previous studies which have used this method in the Indian context (Balietti et al., 2022; Jafarov et al., 2023).

controls respectively, as defined in Sections 3.1.2 and 3.3.2. Our specification includes region-by-month-of-birth fixed effects ($\delta_{mr(d)}$), which control for region-specific seasonal confounders that systematically influence birth outcomes as well as year-of-birth fixed effects to control for year-specific shocks that may affect all individuals born within the same cohort.

In a large country like India, seasonal trends vary substantially across regions. Consequently, factors such as cultural factors, labor market dynamic, income and consumption patterns differ not only across months but also across areas of the country. Incorporating region-by-month-of-birth fixed effects allows for greater flexibility in capturing seasonal un-observables that differ across regions, compared to a specification with only a month-of-birth fixed effects, which assume seasonal patterns are uniform nationwide. These fixed effects together imply that our identification comes from variation in pollution exposure among children born within the same region-month pair, while controlling for aggregate nation-wide year-specific confounders.

Despite the inclusion of a comprehensive set of controls and fixed effects, several potential confounders may still bias our estimate of the $\beta$ parameter by simultaneously affecting both pollution exposure and birth outcomes. These include time-varying factors such as changes in industrialization, which can influence pollution levels while also improving birth outcomes through enhanced access to healthcare (Sanders, 2012). In such cases, our estimate of $\beta$ may be biased downward. Moreover, prior health conditions and unobserved behavioral choices during pregnancy may influence relocation decisions. These decisions, in turn, affect pollution exposure while also directly impacting child health, thus introducing endogeneity. In rural areas, practices such as crop residue burning can simultaneously elevate local pollution levels and reduce soil fertility, which negatively affects farm income and potentially child health, thereby confounding the relationship of interest (Singh et al., 2019). In addition, as our analysis relies on satellite-derived pollution data, we are unable to measure individual-level exposure with high precision. While these data enable the construction of district-level estimates of ambient $PM_{2.5}$ concentrations linked to DHS respondents, they are available at a relatively coarse spatial resolution. This limitation introduces potential measurement error in estimating localized pollution exposure, which may result in attenuation bias, leading to estimates of $\beta$ that are biased toward zero.

To address the endogeneity concerns in our main specification, we employ an instrumental variables (IV) strategy following Balietti et al. (2022) and Deryugina et al. (2019). Specifically, we exploit quasi-random variation in in-utero exposure to $PM_{2.5}$ induced by prevailing wind directions during the gestation period. The first-stage equation is specified as follows:

$$\begin{aligned} PM2.5_{idrmt} = &\sum_{r=1}^{30} \rho_1^r \text{Share}^S_{i(d,m,t)} + \sum_{r=1}^{30} \rho_2^r \text{Share}^N_{i(d,m,t)} + \sum_{r=1}^{30} \rho_3^r \text{Share}^E_{i(d,m,t)} \\ &+ \alpha X_{idrmt} + \omega W_{i(dmt)} + \delta_{mr} + \gamma_t + \varepsilon_{idrmt} \end{aligned} \tag{2}$$

We rely on wind patterns to provide exogenous variation in particulate matter levels, driven by the direction of wind flows. In our first-stage equation, we instrument a child's in-utero exposure to particulate matter using the share of time the wind blew from the North, East, and South ($Share^N$, $Share^E$, and $Share^S$, respectively) during the in-utero period in the child's district of residence. Wind from the West serves as the reference category to avoid multicollinearity and to enable meaningful interpretation of the coefficients. We allow the effects of wind direction to vary by region. All other covariates and fixed effects remain as specified in Equation 1.

A potential threat to our identification strategy is the absence of within-region variation in wind direction across districts. However, the regional and spatial variation in wind patterns, as shown in Table 1, suggests that this concern is unlikely to affect our analysis. Our main specification includes region-by-month-of-birth fixed effects, meaning that the estimated impact of in-utero particulate matter exposure is identified from variation among births occurring in the same region and month. If temporal and spatial wind patterns were uniform across all districts within a region, our instrument would fail to capture sufficient exogenous variation in pollution levels attributable to wind trajectories. The observed variation in wind patterns reinforces the validity of our identification strategy by demonstrating that wind-driven differences in pollution exposure are likely to persist even after accounting for the fixed effects embedded in our model.

Another potential threat to our identification strategy arises if the regions are too small, causing the wind instruments to reflect primarily local pollution sources. In such cases, wind patterns may capture emissions from nearby sources that are likely correlated with

unobserved, time-varying determinants of birth outcomes, thereby reintroducing endogeneity. To mitigate this concern, we define regions that are sufficiently large so that wind instruments capture variation in pollution driven by non-local sources—i.e., emissions originating outside the immediate area. Our first-stage specification allows the effects of wind direction to vary only at the regional level, imposing a uniform impact across all districts within a region. This restriction helps ensure that the variation in pollution captured by our instruments reflects broader, regional air flows rather than localized sources. However, defining regions that are too large may violate the monotonicity assumption, as wind from a given direction could have heterogeneous effects across distant districts within the same region. To assess the robustness of our identification, we repeat the analysis using alternative region definitions based on different numbers of clusters.

To causally capture the relationship between $PM_{2.5}$ pollution and birth outcomes based on our instrument, it must hold that wind direction impacts birth outcomes only through its influence on pollution levels. This assumption is plausible, as wind direction itself is a natural and quasi-random meteorological phenomenon that is unlikely to directly affect fetal development or correlate with other determinants of birth outcomes. Based on our first stage fitted values, the second stage specification is the following:

$$Y_{idrmt} = \beta \widehat{\overline{PM2.5}}_{i(d,m,t)} + \lambda X_{idrmt} + \nu W_{i(d,m,t)} + \delta_{mr} + \gamma_t + \epsilon_{idrmt} \tag{3}$$

Equation (3) follows the same structural form as Equation 1, with the key distinction that the main explanatory variable, $\widehat{\overline{PM2.5}}$, represents predicted in-utero $PM_{2.5}$ exposure values obtained from the first-stage regression. We use a Linear Probability Model (LPM) to estimate the relationship between binary outcomes, including LBW and VLBW, and pollution exposure. We choose LPM because it is able to accommodate multiple fixed effects and produces marginal effects that are straightforward to interpret. In addition, LPM with fixed effects tends to provide more accurate probability estimates in cases of rare outcomes, as in our study, compared to alternative binary outcome models with fixed effects, such as fixed effects logit (Timoneda, 2021). In the results section that follows, we present both OLS estimates based on Equation 1 and IV estimates based on Equation 3.

# 5 Results

## 5.1 Baseline Results

Panel A of Table 4 presents OLS results based on Equation 1, while Panel B reports IV estimates from Equation 3 for three outcomes: birth weight (in grams), an indicator for low birth weight (LBW), and an indicator for very low birth weight (VLBW). Using our causal IV estimates, we find that a one standard deviation decrease in $PM_{2.5}$ exposure during the gestational period (15 $\mu$g/m$^3$) is associated with a 1.3% increase in birth weight relative to the mean, and a 2.7 and 0.6 percentage point reduction in the incidence of LBW and VLBW, respectively, holding other factors constant. A one standard deviation change in ambient $PM_{2.5}$ exposure also explains 6.4%, 7.2%, and 6% of the standard deviation in birth weight, LBW, and VLBW, respectively. Compared to existing estimates, we find that a one standard deviation increase in in-utero exposure to $PM_{2.5}$ reduces birth weight by approximately 36.4 grams in India. This effect is substantially larger than the statistically insignificant average impact found for the United States, where only a 28 grams reduction was observed at the lower tail of the birth weight distribution (Pons, 2022) and is higher than the 0.5% reduction in birth weight associated with a one standard deviation increase in $PM_{10}$ exposure in Italy (Palma et al., 2022).

Our OLS estimates are consistently smaller in magnitude than the corresponding IV estimates across all birth outcome indicators. This suggests that the OLS estimates may be biased downward due to unobserved confounders and/or attenuation bias arising from measurement error, as discussed in the previous section. A simple Hausman test confirms that the difference between OLS and IV estimates is statistically significant for the LBW and VLBW outcomes, but not for the continuous birth weight measure. This indicates that endogeneity and measurement error may be more pronounced in the case of binary birth weight indicators. Nevertheless, both the OLS and IV estimates are statistically significant across all birth outcomes. These findings underscore the economically and statistically significant impact of ambient air pollution on birth outcomes and highlight the importance of addressing in-utero environmental exposures as a means of improving early-life health.

Table 4: Effect of $PM_{2.5}$ Exposure on Birth Outcomes: OLS and IV Estimates

| | **Panel A: OLS Estimates** | | |
|---|---|---|---|
| **Dependent variable:** | Birth Weight | LBW | VLBW |
| $PM_{2.5}$ Exposure | -1.759*** | 0.001*** | 0.0001** |
| | (0.389) | (0.0002) | (0.00004) |
| Controls | Yes | Yes | Yes |
| Region×Month of Birth FE | Yes | Yes | Yes |
| Year of Birth FE | Yes | Yes | Yes |
| Mean of Dependent Variable | 2825 | 0.17 | 0.01 |
| Observations | 322,340 | 322,340 | 322,340 |
| Adjusted $R^2$ | 0.053 | 0.016 | 0.002 |
| | **Panel B: IV Estimates** | | |
| **Dependent variable:** | Birth Weight | LBW | VLBW |
| Instrumented $PM_{2.5}$ Exposure | -2.629*** | 0.002*** | 0.0004*** |
| | (0.717) | (0.0004) | (0.0001) |
| First Stage F-statistic | 129 | 129 | 129 |
| Controls | Yes | Yes | Yes |
| Region×Month of Birth FE | Yes | Yes | Yes |
| Year of Birth FE | Yes | Yes | Yes |
| Mean of Dependent Variable | 2825 | 0.17 | 0.01 |
| Observations | 322,340 | 322,340 | 322,340 |

*Notes:* Clustered robust standard errors at the district level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. F-statistics reported are heteroskedasticity-robust F-statistics, based on a joint significance test of excluded instruments, with standard errors clustered at the district level from the first stage.

## 5.2 Robustness Checks

To test the reliability of our findings and demonstrate the robustness of our identification strategy, we conduct a comprehensive set of robustness and sensitivity checks. This section addresses potential methodological concerns and outlines the steps we take to ensure that our results are not driven by alternative explanations or model assumptions.

### 5.2.1 Concerns Related to Instrumental Variable Assumptions

**Monotonicity:** As discussed in Section 4, there is a trade-off in selecting the optimal size of regions over which the wind instruments are allowed to vary. By restricting the impact of wind direction coefficients to be uniform across all districts within a region, we aim to ensure that our instruments primarily capture variation in pollution from non-local

sources—that is, pollution originating from outside the region. However, increasing the size of these regions may risk violating the monotonicity assumption, as the same wind direction could have heterogeneous effects on different districts within a region. Such heterogeneity could compromise the validity of our IV strategy, making it essential to test this assumption.

To address this concern, we conduct several robustness checks. Specifically, we reduce the size of the regions by increasing the number of them and re-estimate the first-stage fitted values of in-utero pollution exposure. This approach, adopted in several prior studies using similar IV strategies, rests on the logic that with smaller regions, it becomes less likely that a given wind direction affects districts in systematically different ways—thereby reducing the likelihood of monotonicity violations. Figures C2, C3, and C4 in the Appendix present the IV coefficient estimates for our three birth outcomes, using alternative specifications with 40, 50, and 60 regions. The results remain robust across these specifications, lending support to the validity of the monotonicity assumption in our setting.

**Hierarchical Clustering:** To test the robustness of our results to the method of region construction, we employ an alternative deterministic clustering algorithm—hierarchical clustering—to form 30 regions. We then re-estimate our main results using this alternative regional specification. As shown in Panel A of Table 5, the coefficients obtained from this approach are similar to those from our baseline model, suggesting that our findings are not sensitive to the specific clustering method used.

**First Stage using District-Month-Year Pairs:** n our baseline IV specification, we use the average wind direction shares over the 10-month in-utero period as instruments for the corresponding average in-utero pollution exposure. However, aggregating wind patterns over such a long duration may smooth out important month-to-month variation, potentially weakening the first-stage relationship. To address this concern, we adopt a more granular approach by predicting monthly pollution levels at the district level using district-month-year-specific wind patterns. These predicted monthly pollution levels are then aggregated over the gestational period for each child, based on their district of residence. The first-stage regression used to estimate monthly pollution is specified as follows:

$$PM2.5_{drmt} = \sum_{r=1}^{30} \rho_1^r \text{Share}^S_{i(d,m,t)} + \sum_{r=1}^{30} \rho_2^r \text{Share}^N_{i(d,m,t)} + \sum_{r=1}^{30} \rho_3^r \text{Share}^E_{i(d,m,t)} + \omega W_{dmt} + \delta_{mr} + \gamma_t + \varepsilon_{drmt} \quad (4)$$

where $PM2.5_{drmt}$ denotes the average concentration of $PM_{2.5}$ in district $d$, belonging to region $r$, in month $m$ of year $t$, and the remaining variables are defined as in Equation 2. This specification allows us to flexibly estimate monthly pollution levels at the district level, which can then be aggregated into individual-level in-utero exposure measures with better temporal precision. Using these first-stage estimates, we present the corresponding second-stage results in Panel b of Table 5. The results remain robust to this alternative method of constructing the exposure estimates.

**Placebo Test:** To ensure that our instruments capture meaningful variation in pollution driven by differential wind directions rather than spurious correlations, we conduct a placebo exercise. Specifically, we randomize wind direction variables across individuals in the sample and re-estimate our baseline model over multiple iterations. Across these placebo regressions, we consistently find null results, indicating that the original estimates are unlikely to be driven by random chance and that the instrumented variation reflects genuine quasi-experimental exposure to pollution.

Table 5: Alternate First Stages

| **Dependent variable:** | Birth Weight | LBW | VLBW |
|---|---|---|---|
| **Panel a: Hierarchal Clustering** | | | |
| $PM_{2.5}$ Exposure | -1.637** | 0.001*** | 0.0003*** |
| | (0.730) | (0.0004) | (0.0001) |
| First Stage F-statistic | 159 | 159 | 159 |
| Adjusted $R^2$ | 0.054 | 0.017 | 0.003 |
| **Panel b: First Stage: District-Month-Year Pairs** | | | |
| $PM_{2.5}$ Exposure | -3.483** | 0.003*** | 0.001*** |
| | (1.064) | (0.001) | (0.0001) |
| First Stage F-statistic | 215 | 215 | 215 |
| Adjusted $R^2$ | 0.054 | 0.018 | 0.003 |
| Observations | 322,340 | 322,340 | 322,340 |

*Notes:* IV estimates. Clustered robust standard errors at the district level are reported in parentheses. All regressions include full controls and fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

### 5.2.2 Alternative Specifications

Maternal health, pregnancy complications, and behavioral or educational characteristics can also influence a child's birth weight. We do not include these variables in our baseline specification, as they may vary over time and could be endogenous—potentially correlated with regional pollution levels. However, to demonstrate the robustness of our estimates to the inclusion of such factors, we control for a set of maternal and birth-specific characteristics in an extended model. These include binary indicators for whether the mother has any chronic health condition, smokes, or consumes alcohol; a categorical variable for maternal education attainment; the number of antenatal checkups attended; a categorical variable for place of delivery; indicators for multiple births (twins), and whether the child was delivered via surgery. We present the results in Panel A of Table 6. While the point estimate is lower than in the baseline specification, the results remain robust to the inclusion of these additional maternal and birth-related controls.

We noted that approximately 60% of our sample reports birth weights of their offspring via a medical card, whereas 40% report birth weights via recall. In order to avoid potential measurement error in our outcomes, we retain only the sample which reported birth weight via a medical card and re-estimate our results. In Panel b of Table 6, we observe that our key results remain similar in terms of magnitude and statistical significance.

Table 6: Alternative Model Specifications

| **Dependent variable:** | Birth Weight | LBW | VLBW |
|---|---|---|---|
| **Panel a: Additional Mother and Birth Controls** | | | |
| $PM_{2.5}$ Exposure | -1.507*** | 0.001*** | 0.0003*** |
| | (0.504) | (0.0003) | (0.0001) |
| Adjusted $R^2$ | 0.073 | 0.041 | 0.013 |
| Observations | 322,340 | 322,340 | 322,340 |
| **Panel b: Card only Sample** | | | |
| $PM_{2.5}$ Exposure | -2.462*** | 0.001** | 0.0001 |
| | (0.797) | (0.0005) | (0.0001) |
| Observations | 183,093 | 183,093 | 183,093 |
| Adjusted $R^2$ | 0.049 | 0.014 | 0.002 |
| Mean of Dependent Variable | 2828 | 0.16 | 0.007 |

*Notes:* Clustered robust standard errors at district level are reported in parentheses. All regressions include full set of controls and fixed effects as the baseline equation. *** p<0.01, ** p<0.05, * p<0.1

### 5.2.3 Sensitivity Checks

We conduct a series of sensitivity checks, we also replicate our analysis after filtering out outliers in birth weight. The results, presented in Tables C1 and C2 in the Appendix, are consistent with our baseline findings. Figures C5, C6, and C7 show the estimates obtained using alternative geographical clusters for standard error computation. The robustness of our results to this alternate clustering approach further reinforces the validity of our findings. In addition, in Figures C8, C9, and C10 in the Appendix, we present the estimates by iteratively excluding births from one State or Union Territory (UT) at a time. This iterative exclusion confirms that our results are not being disproportionately driven by any single region.

# 6 Economic Costs of Pollution

A substantial body of literature highlights the critical role of birth weight in shaping a wide range of long-term child outcomes. In this section, we link our estimates to the findings of some of the studies in this literature. Table C3 in the Appendix provides a brief overview of some of these studies. We focus on reporting estimates from studies which are performed using data from developing countries, preferably in the Indian setting, since their setting and estimates would be the ones which most closely resemble the backdrop of our analysis. This is due to the fact that health shocks at birth in the context of developing countries may be more poignant, as they occur in environments with limited access to medical care, inadequate nutrition, and fewer resources to mitigate their long-term effects, ultimately amplifying their impact on later outcomes (Currie & Vogl, 2013). Hence, similar topical studies from developed countries may not be directly comparable, given the stark differences in baseline health conditions, institutional capacity, and the ability to buffer early-life shocks.

The average in-utero exposure for $PM2.5$ within our sample was 40 $\mu$g/m$^3$. The 2005 WHO Air Quality Guidelines recommend that average annual concentrations of $PM2.5$ should not exceed 10 $\mu$g/m$^3$. Reducing average in-utero $PM2.5$ exposure to this threshold, equivalent to a 30-unit reduction would imply based on our estimates, a 75-gram increase in birth weight on average, and a 5.4 percentage point decline in the incidence of LBW. To

interpret these effects in light of the broader literature on the longer-term consequences of birth weight on child health and cognition, we draw comparisons from the studies summarized in Table C3. Assuming linear transformation of effects reported, a 75g gain translates into approximately 0.32 and 0.19 percentile increases in weight-for-age and height-for-age, respectively (Keshav, 2021). Similarly, based on estimates from Baguet and Dumas (2019), a 75g increase would correspond to a 0.014 standard deviation increase in educational attainment, and according to Kumar et al. (2022), a 0.03–0.04 standard deviation improvement in cognitive test scores. On the LBW margin, Alderman and Behrman (2006) estimate that shifting one child from LBW to normal weight yields a lifetime productivity gain of $510 (Based on 2006 USD, equivalent to $810 in 2025). In the Indian context, where approximately 25 million births occur annually, a 5.4 percentage point reduction in LBW prevalence implies 1.35 million fewer LBW births per year. Applying the estimate from Alderman and Behrman (2006), this corresponds to a potential $1.1 Billion (2025 USD) in aggregate lifetime gains as a result of decline in LBW incidence simply by complying with WHO air quality standards. While figures in this section are useful for contextualizing the potential scale of benefits, they rely on simplifying assumptions and should therefore be interpreted as illustrative rather than precise estimates.

# 7 Conclusion

This paper provides novel causal evidence on the impact of in-utero exposure to air pollution on birth weight outcomes in India, a country characterized by both high ambient air pollution and poor neonatal health indicators. Leveraging exogenous variation in wind direction to instrument for ambient $PM2.5$ exposure during gestation, we document that even modest reductions in particulate matter can yield significant improvements in neonatal health. Specifically, we find that a one standard deviation decrease in PM2.5 leads to a 1.3% increase in average birth weight and a 2.7 percentage point reduction in LBW rates respectively. These findings are robust across a range of sensitivity checks.

Linking our results to previous literature on the consequences of birth weight on a num-

ber of child cognition and health outcomes, our back-of-the-envelope calculations suggest that aligning in-utero $PM2.5$ levels with WHO air quality standards could yield substantial economic benefits, including over one billion dollars annually from reductions in lifetime productivity losses stemming from LBW infants each year, along with broader long-run gains in child cognition and health. While these estimates rely on simplifying assumptions, they underscore the potentially large returns to improving air quality for maternal and child health.

Our research also comes with several limitations. First, we are unable to precisely account for time spent outdoors, pollution avoidance or defensive behaviors, and migration patterns, all of which influence actual ambient exposure during pregnancy. Second, since our data come from a repeated cross-section rather than a longitudinal panel, we are unable to examine the long-term effects of in-utero pollution exposure, or direct mediation effects through birth-weight on these longer run relationships, which remains an important area for future research. Third, and finally, the granularity of our satellite-based pollution data limits our ability to capture micro-level variation in exposure. While our IV strategy addresses concerns of attenuation bias due to measurement error, this coarseness may weaken the first-stage relationship between wind patterns and individual exposure, potentially resulting in noisier estimates than would be obtained with finer-grained pollution data

Overall, our findings highlight the importance of prenatal environmental conditions in shaping early-life health outcomes and suggest that reducing air pollution exposure during pregnancy may be a cost-effective strategy for improving public health in low- and middle-income countries.

# References


Ai, H., Wu, J., & Zhou, Z. (2023). The long-run effects of fetal pm2. 5 exposure on mental health: Evidence from china. *Environmental Science and Pollution Research*, *30*(12), 34158–34173.

Alderman, H., & Behrman, J. R. (2006). Reducing the incidence of low birth weight in low-income countries has substantial economic benefits. *The World Bank Research Observer*, *21*(1), 25–48.

Almond, D., Chay, K. Y., & Lee, D. S. (2005). The costs of low birth weight. *The Quarterly Journal of Economics*, *120*(3), 1031–1083.

Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, *25*(3), 153–172.

Almond, D., Edlund, L., & Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *The Quarterly Journal of Economics*, *124*(4), 1729–1772.

Arceo, E., Hanna, R., & Oliva, P. (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from Mexico City. *The Economic Journal*, *126*(591), 257–280.

Baguet, M., & Dumas, C. (2019). How does birth weight affect health and human capital? a short- and long-term evaluation. *Health Economics*, *28*(5), 597–617.

Balakrishnan, K., Dey, S., Gupta, T., Dhaliwal, R., Brauer, M., Cohen, A. J., Stanaway, J. D., Beig, G., Joshi, T. K., Aggarwal, A. N., et al. (2019). The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: The global burden of disease study 2017. *The Lancet Planetary Health*, *3*(1), e26–e39.

Balakrishnan, U., & Tsaneva, M. (2021). Air pollution and academic performance: Evidence from India. *World Development*, *146*, 105553.

Balietti, A., Datta, S., & Veljanoska, S. (2022). Air pollution and child development in India. *Journal of Environmental Economics and Management*, *113*, 102624.

Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and Statistics*, *86*(2), 586–601.

Bharadwaj, P., & Eberhard, J. (2008). Atmospheric air pollution and birth weight. *Available at SSRN 1197443*.

Bhilwar, M., Upadhyay, R. P., Yadav, K., Kumar, R., Chinnakali, P., Sinha, S., & Kant, S. (2016). Estimating the burden of'weighing less': A systematic review and meta-analysis of low birth-weight in India. *The National Medical Journal of India*, *29*(2), 73.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, *122*(1), 409–439.

Blazer, S., Moreh-Waterman, Y., Miller-Lotan, R., Tamir, A., & Hochberg, Z. (2003). Maternal hypothyroidism may affect fetal growth and neonatal thyroid function. *Obstetrics & Gynecology*, *102*(2), 232–241.

Bondy, M., Roth, S., & Sager, L. (2020). Crime is in the air: The contemporaneous relationship between air pollution and crime. *Journal of the Association of Environmental and Resource Economists*, *7*(3), 555–585.

Chay, K. Y., & Greenstone, M. (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), 1121–1167.

Chay, K. Y., & Greenstone, M. (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy*, *113*(2), 376–424.

Cook, C. J., & Fletcher, J. M. (2015). Understanding heterogeneity in the effects of birth weight on adult cognition and wages. *Journal of Health Economics*, *41*, 107–116.

Currie, J., & Vogl, T. (2013). Early-life health and adult circumstance in developing countries. *Annu. Rev. Econ.*, *5*(1), 1–36.

Currie, J., Zivin, J. G., Mullins, J., & Neidell, M. (2014). What do we know about short-and long-term effects of early-life exposure to pollution? *Annu. Rev. Resour. Econ.*, *6*(1), 217–247.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, *109*(12), 4178–4219.

El Arifeen, S., Masanja, H., & Rahman, A. E. (2017). Child mortality: The challenge for India and the world. *The Lancet*, *390*(10106), 1932–1933.

Fan, M., Jiang, H., & Zhou, M. (2023). Beyond particulate matter: New evidence on the causal effects of air pollution on mortality. *Journal of Health Economics*, *91*, 102799.

Greenstone, M., & Hanna, R. (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review*, *104*(10), 3038–3072.

Herrnstadt, E., Heyes, A., Muehlegger, E., & Saberian, S. (2021). Air pollution and criminal activity: Microgeographic evidence from Chicago. *American Economic Journal: Applied Economics*, *13*(4), 70–100.

Hong, Y. (2025). Heat and humidity on early-life outcomes: Evidence from Mexico. *Journal of Environmental Economics and Management*, *129*, 103082.

Hummer, M., Lehner, T., & Pruckner, G. (2014). Low birth weight and health expenditures from birth to late adolescence. *The European Journal of Health Economics*, *15*, 229–242.

Isen, A., Rossin-Slater, M., & Walker, W. R. (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.

Jafarov, J., Singh, T. P., & Sahoo, S. (2023). *Air pollution and time use: Evidence from India* (tech. rep.). IZA Discussion Papers.

Jain, N., Bhatia, A., Pathak, H., et al. (2014). Emission of air pollutants from crop residue burning in India. *Aerosol and Air Quality Research*, *14*(1), 422–430.

Janssen, B. G., Saenen, N. D., Roels, H. A., Madhloum, N., Gyselaers, W., Lefebvre, W., Penders, J., Vanpoucke, C., Vrijens, K., & Nawrot, T. S. (2017). Fetal thyroid function, birth weight, and in utero exposure to fine particle air pollution: A birth cohort study. *Environmental Health Perspectives*, *125*(4), 699–705.

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

Jethva, H., Torres, O., Field, R. D., Lyapustin, A., Gautam, R., & Kayetha, V. (2019). Connecting crop productivity, residue fires, and air quality over northern India. *Scientific Reports*, *9*(1), 16594.

Kader, M., & Perera, N. K. P. (2014). Socio-economic and nutritional determinants of low birth weight in India. *North American Journal of Medical Sciences*, *6*(7), 302.

Kanawade, V., Srivastava, A., Ram, K., Asmi, E., Vakkari, V., Soni, V., Varaprasad, V., & Sarangi, C. (2020). What caused severe air pollution episode of november 2016 in New Delhi? *Atmospheric Environment*, *222*, 117125.

Kannan, S., Misra, D. P., Dvonch, J. T., & Krishnakumar, A. (2007). Exposures to airborne particulate matter and adverse perinatal outcomes: A biologically plausible mechanistic framework for exploring potential. *Ciencia & Saude Coletiva*, *12*, 1591–1602.

Keshav, V. (2021). Health returns to birth weight: Evidence from developing countries. *European Journal of Social Sciences*, *4*(2), 38–59.

Kumar, S., Kumar, K., Laxminarayan, R., & Nandi, A. (2022). Birth weight and cognitive development during childhood: Evidence from India. *Economic Papers: A Journal of Applied Economics and Policy*, *41*(2), 155–175.

Li, L., & Zhang, X. (2024). The causal impact of fetal exposure to pm2. 5 on birth outcomes: Evidence from rural China. *Economics & Human Biology*, *53*, 101380.

Marete, I., Ekhaguere, O., Bann, C. M., Bucher, S. L., Nyongesa, P., Patel, A. B., Hibberd, P. L., Saleem, S., Goldenberg, R. L., Goudar, S. S., et al. (2020). Regional trends in birth weight in low-and middle-income countries 2013–2018. *Reproductive Health*, *17*, 1–8.

McGovern, M. E. (2019). How much does birth weight matter for child health in developing countries? estimates from siblings and twins. *Health Economics*, *28*(1), 3–22.

Murukutla, N., Negi, N. S., Puri, P., Mullin, S., & Onyon, L. (2017). Online media coverage of air pollution risks and current policies in India: A content analysis. *WHO South-East Asia Journal of Public Health*, *6*(2), 41–50.

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, *23*(6), 1209–1236.

Palma, A., Petrunyk, I., & Vuri, D. (2022). Prenatal air pollution exposure and neonatal health. *Health Economics*, *31*(5), 729–759.

Pons, M. (2022). The impact of air pollution on birthweight: Evidence from grouped quantile regression. *Empirical Economics*, *62*(1), 279–296.

Pope, D. P., Mishra, V., Thompson, L., Siddiqui, A. R., Rehfuess, E. A., Weber, M., & Bruce, N. G. (2010). Risk of low birth weight and stillbirth associated with indoor air pollution from solid fuel use in developing countries. *Epidemiologic reviews*, *32*(1), 70–81.

Risnes, K. R., Vatten, L. J., Baker, J. L., Jameson, K., Sovio, U., Kajantie, E., Osler, M., Morley, R., Jokela, M., Painter, R. C., et al. (2011). Birthweight and mortality in adulthood: A systematic review and meta-analysis. *International Journal of Epidemiology*, *40*(3), 647–661.

Royer, H. (2009). Separated at girth: US twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, *1*(1), 49–85.

Saharan, U. S., Kumar, R., Singh, S., Mandal, T. K., Sateesh, M., Verma, S., & Srivastava, A. (2024). Hotspot driven air pollution during crop residue burning season in the indo-gangetic plain, India. *Environmental Pollution*, *350*, 124013.

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

Sankar, M., Neogi, S., Sharma, J., Chauhan, M., Srivastava, R., Prabhakar, P., Khera, A., Kumar, R., Zodpey, S., & Paul, V. (2016). State of newborn health in India. *Journal of Perinatology*, *36*(3), S3–S8.

Sen, A., Abdelmaksoud, A., Ahammed, Y. N., Banerjee, T., Bhat, M. A., Chatterjee, A., Choudhuri, A. K., Das, T., Dhir, A., Dhyani, P. P., et al. (2017). Variations in particulate matter over indo-gangetic plains and indo-himalayan range during four field campaigns in winter monsoon and summer monsoon: Role of pollution pathways. *Atmospheric Environment*, *154*, 200–224.

Sicard, P., Agathokleous, E., Anenberg, S. C., De Marco, A., Paoletti, E., & Calatayud, V. (2023). Trends in urban air pollution over the last two decades: A global perspective. *Science of The Total Environment*, *858*, 160064.

Singh, P., Dey, S., Chowdhury, S., & Bali, K. (2019). Early life exposure to outdoor air pollution. (06).

Sun, X., Luo, X., Zhao, C., Zhang, B., Tao, J., Yang, Z., Ma, W., & Liu, T. (2016). The associations between birth weight and exposure to fine particulate matter (pm2. 5) and its chemical constituents during pregnancy: A meta-analysis. *Environmental Pollution*, *211*, 38–47.

Tang, Z., Long, X., Wang, K., Berger, K., Zhang, Y., & Mayvaneh, F. (2024). Risk and burden of low birthweight related to maternal PM2.5 exposure in Iran: A national causal inference study. *Ecotoxicology and Environmental Safety*, *288*, 117414.

Timoneda, J. C. (2021). Estimating group fixed effects in panel data with a binary dependent variable: How the LPM outperforms logistic regression in rare events data. *Social Science Research*, *93*, 102486.

Topinka, J., Binkova, B., Mračková, G., Stavkova, Z., Beneš, I., Dejmek, J., Leníček, J., & Šrám, R. (1997). Dna adducts in human placenta as related to air pollution and to GSTM1 genotype. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *390*(1-2), 59–68.

von Hinke, S., & Sørensen, E. N. (2023). The long-term effects of early-life pollution exposure: Evidence from the London smog. *Journal of Health Economics*, *92*, 102827.

WHO. (2024). How is India trying to address air pollution? [Accessed: 2024-12-28]. https://www.worldbank.org/en/country/india/publication/catalyzing-clean-air-in-india

# Appendix

## A. Calculation of Wind Direction and Wind Speed based on U-V wind components

The wind speed ($WS$) and wind direction ($\theta$) can be calculated from the zonal ($u$) and meridional ($v$) wind components using the following formulas:

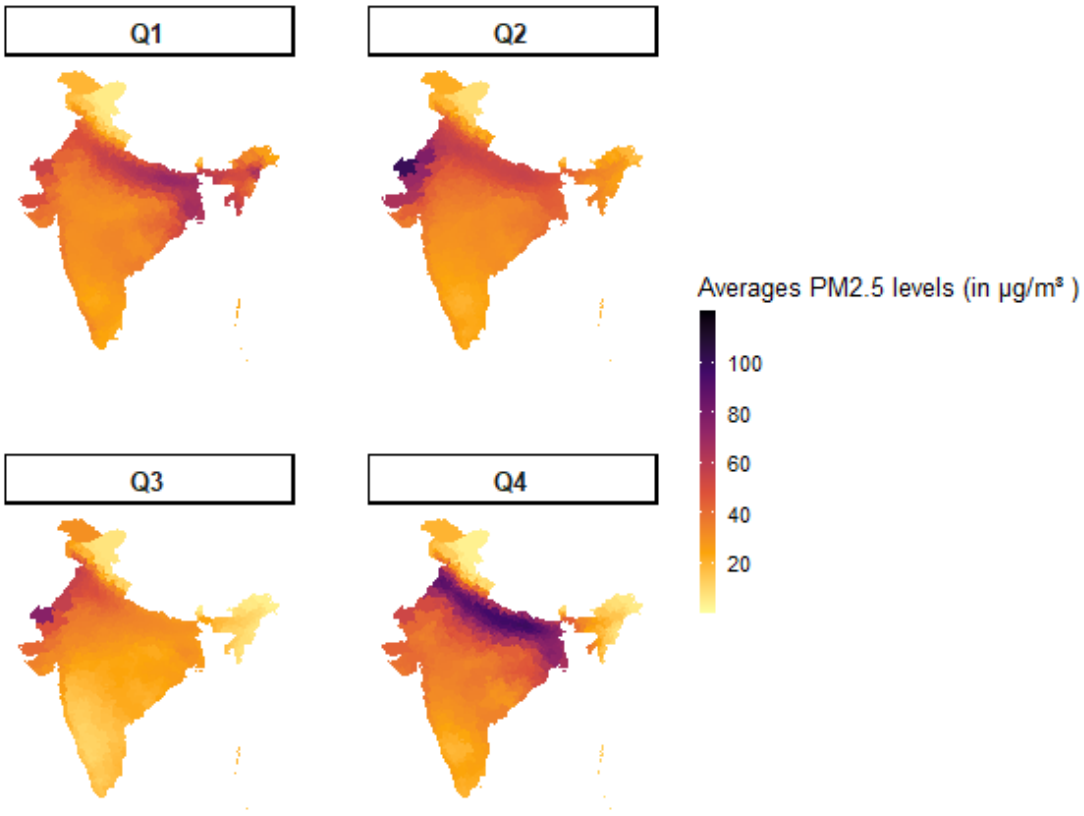
$$WS = \sqrt{u^2 + v^2} \tag{5}$$

$$\theta = \left( \arctan 2(v, u) \cdot \frac{180}{\pi} \right) + 180 \tag{6}$$

- $u$ is the zonal (east-west) component (positive toward the east), $v$ is the meridional (north-south) component (positive toward the north),

## B. Spatial and Temporal Variation of $PM_{2.5}$

Figure B1: Quarter of Year Wise PM2.5



## C: Additional Tables and Figures

Figure C1: 30 clusters based on K-means clustering algorithm

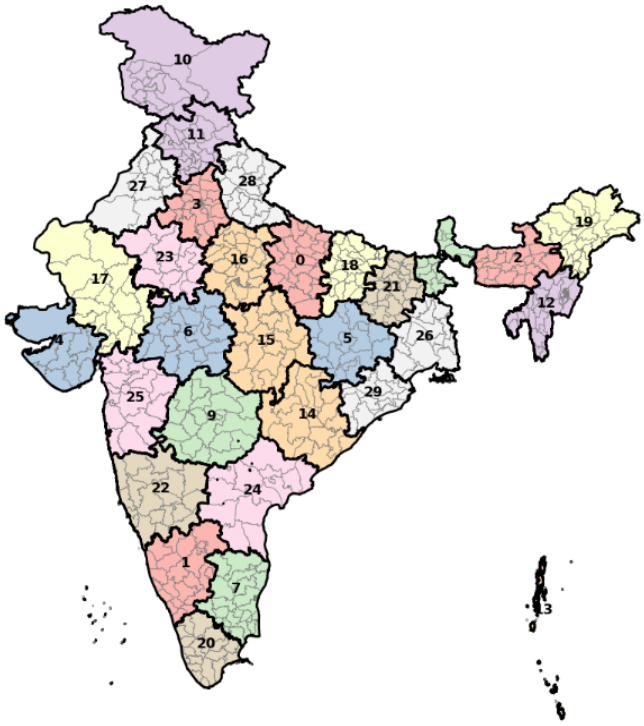



Figure C2: Monotonicity (Continuous Measure of Birth Weight): Alternative Number of Regions for Instrument

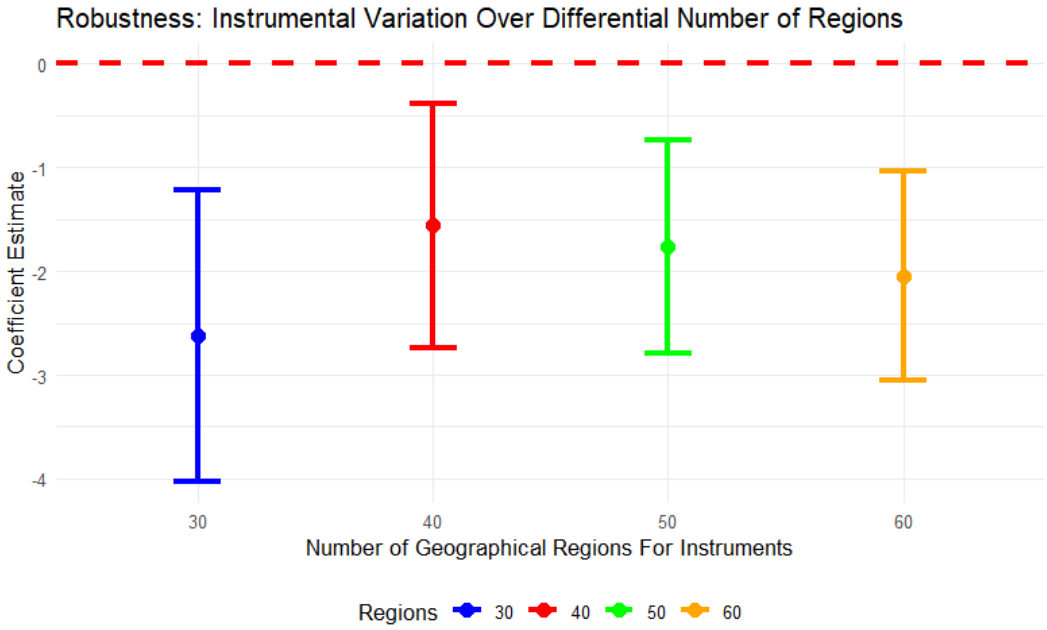


Figure C3: Monotonicity (LBW Binary Measure): Alternative Number of Regions for Instrument

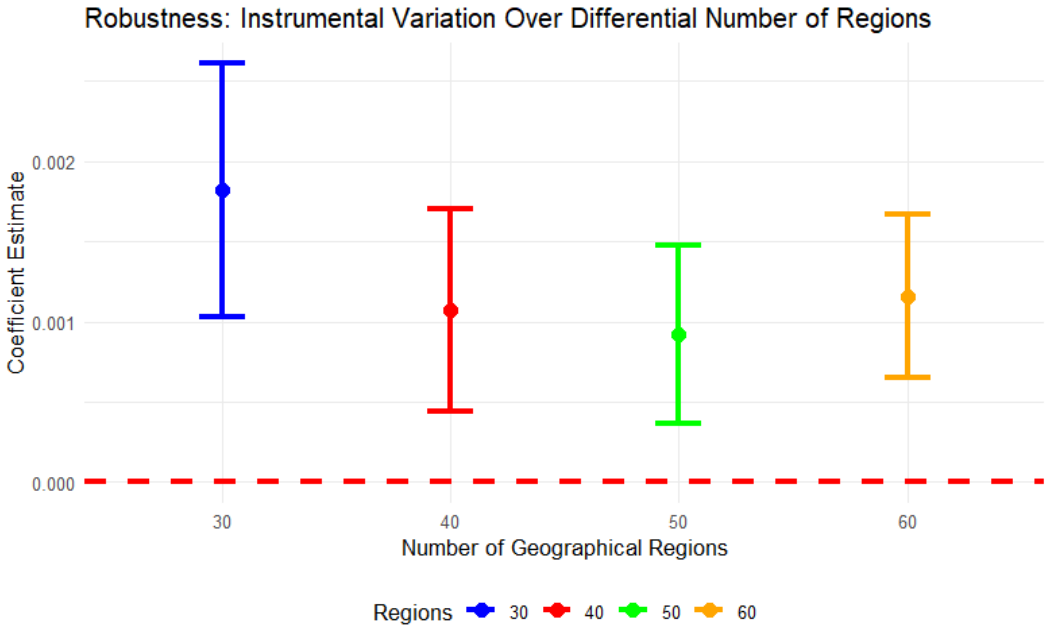



Figure C4: Monotonicity (VLBW Binary Measure): Alternative Number of Regions for Instrument

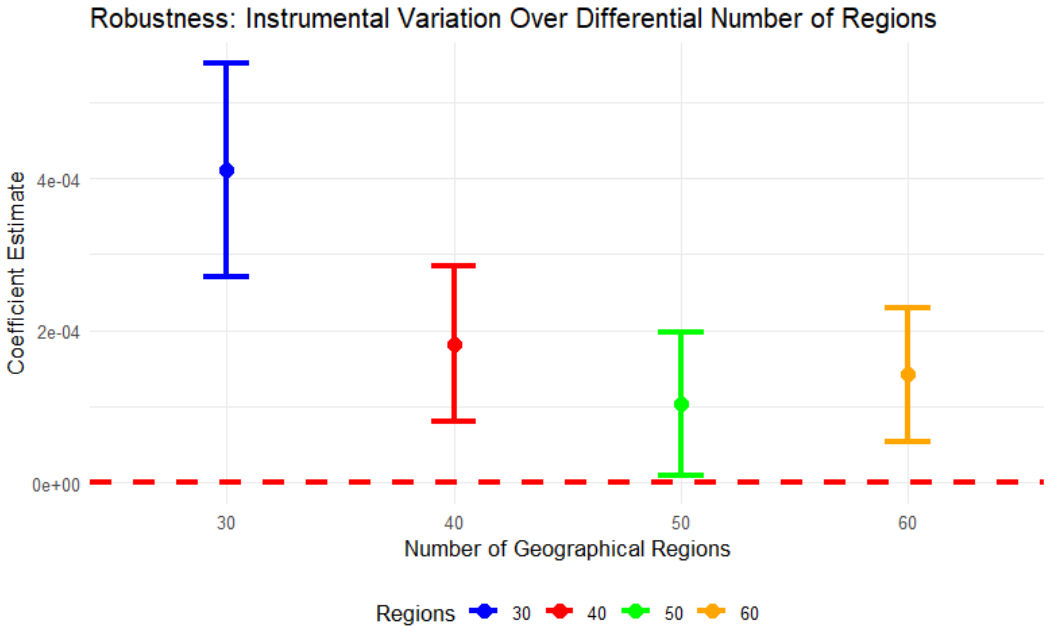


Figure C5: Baseline Results Birth Weight: Alternative Clustering of Standard Errors

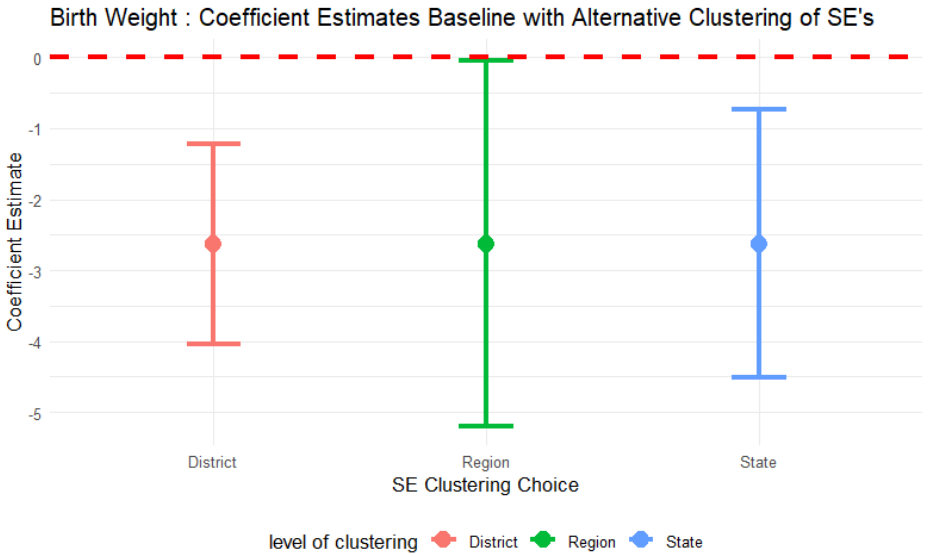



Figure C6: Baseline Results LBW: Alternative Clustering of Standard Errors

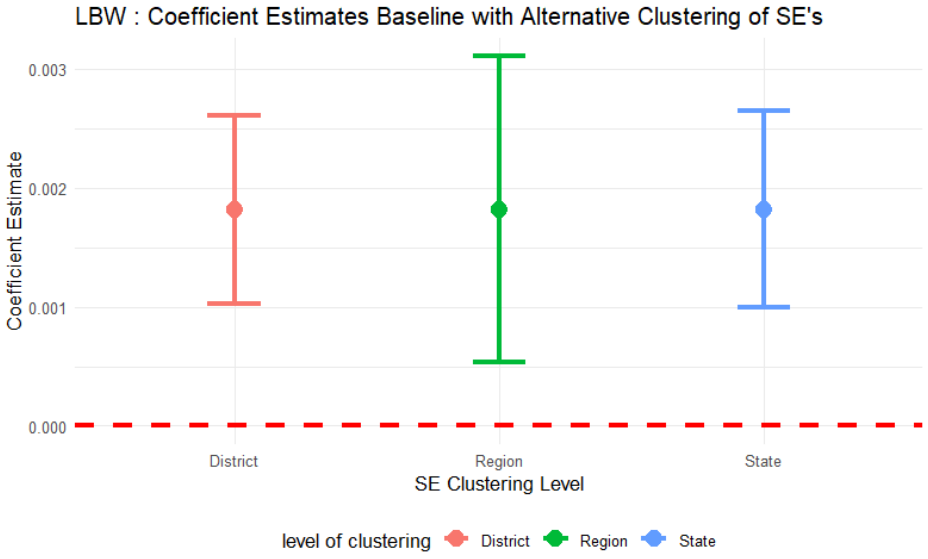


Figure C7: Baseline Results VLBW: Alternative Clustering of Standard Errors

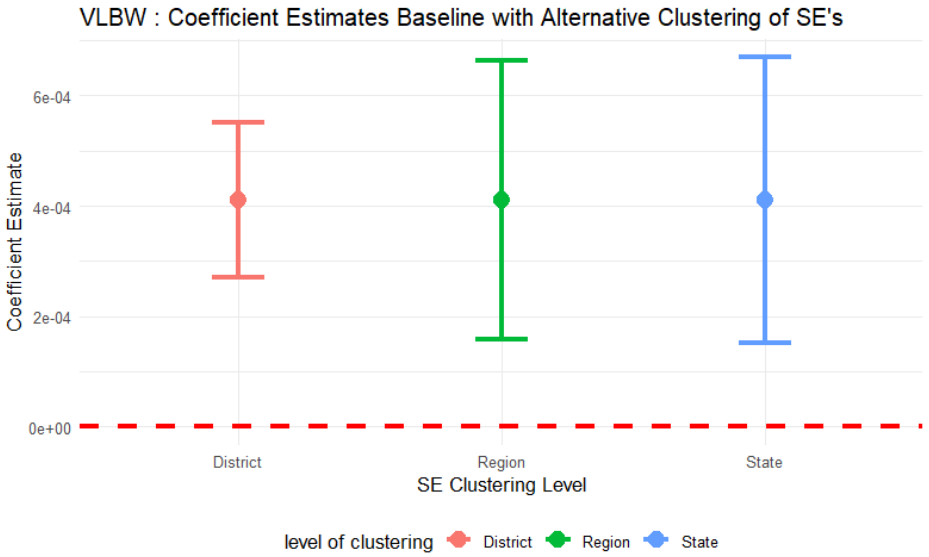



Figure C8: Iteratively dropping States/UTs from the Indian Union one at a time and rerunning baseline first and second stage models for birth weight

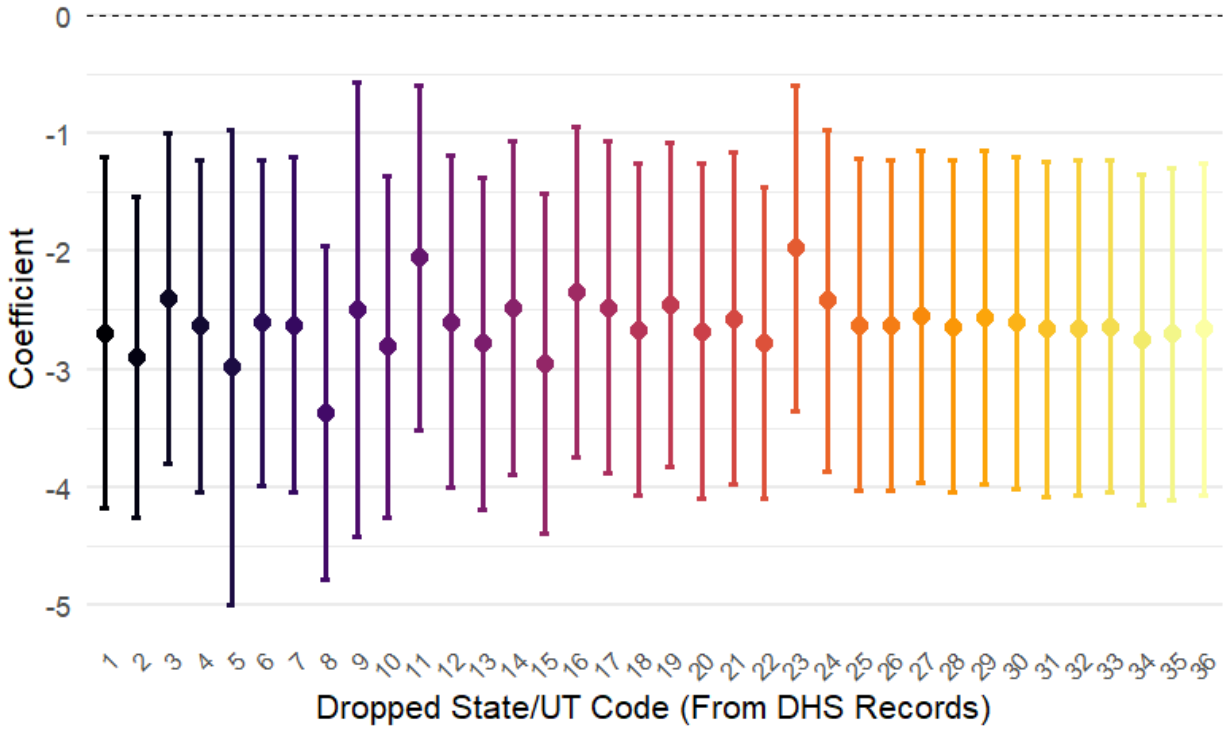


Figure C9: Iteratively dropping States/UTs from the Indian Union one at a time and rerunning baseline first and second stage models for LBW indicator

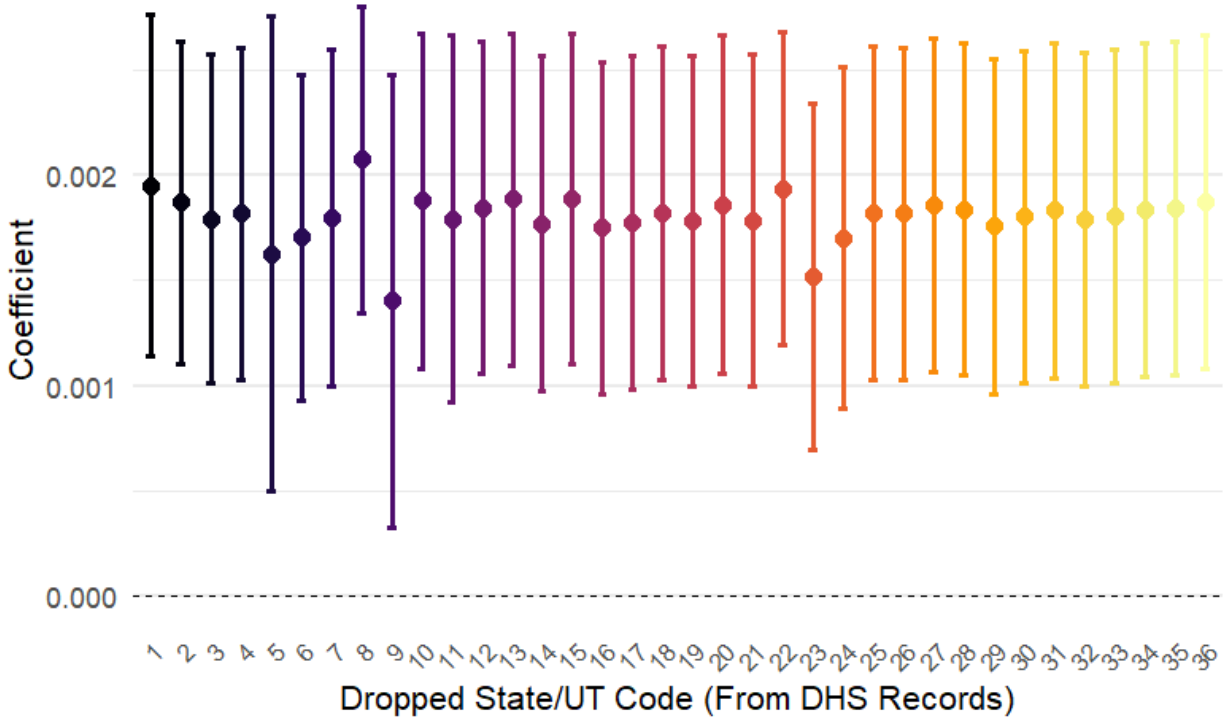


Figure C10: Iteratively dropping States/UTs from the Indian Union one at a time and rerunning baseline first and second stage models for VLBW indicator

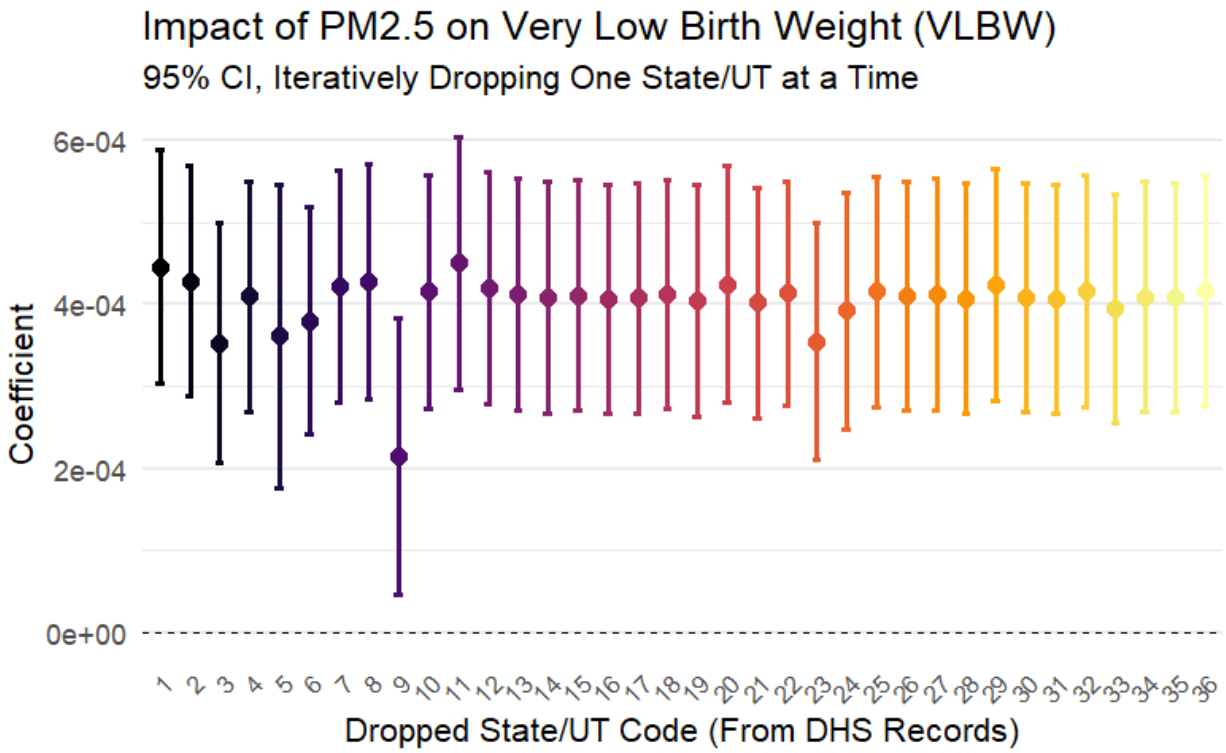


Table C1: Effect of $PM_{2.5}$ Exposure on Birth Outcomes: IV Estimates (1st–99th Percentile Sample)

| **Dependent variable:** | Birth Weight | LBW |
|---|---|---|
| $PM_{2.5}$ Exposure | -2.0079*** | 0.0013*** |
| | (0.6081) | (0.0004) |
| First Stage F-statistic | 127 | 127 |
| Observations | 311,089 | 311,089 |
| Adjusted $R^2$ | 0.0582 | 0.0122 |
| Mean of Dependent Variable | 2838 | 0.15 |

*Notes:* IV estimates. Clustered robust standard errors at the district level are reported in parentheses. All regressions include full controls and fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table C2: Effect of $PM_{2.5}$ Exposure on Birth Outcomes: IV Estimates (Birth Weight 1600g–4000g Sample)

| **Dependent variable:** | Birth Weight | LBW |
|---|---|---|
| $PM_{2.5}$ Exposure | -1.9655*** | 0.0014*** |
| | (0.5557) | (0.0004) |
| First Stage F-statistic | 125 | 125 |
| Observations | 303,235 | 303,235 |
| Adjusted $R^2$ | 0.0549 | 0.0117 |
| Mean of Dependent Variable | 2810 | 0.15 |

*Notes:* IV estimates. Clustered robust standard errors at the district level are reported in parentheses. All regressions include full controls and fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table C3: Birth Weight and Its Long-Run Consequences in Developing Countries

| **Study** | **Country / Data** | **Key Result** |
|---|---|---|
| Alderman and Behrman (2006) | Literature based synthesized evidence from studies in multiple developing countries | Estimated $510 lifetime economic benefit per child moved from LBW to normal weight, primarily via higher productivity |
| Keshav (2021) | 39 longitudinal datasets from developing countries | Birth weight strongly associated with height and weight outcomes; 100g increase linked to 0.43 and 0.25 unit increase in Weight for Age and Height for Age percentiles |
| Baguet and Dumas (2019) | Phillipines, CLHNS Dataset tracking a 1983-84 birth cohort from infancy until age 22 | 100 g increase in birth weight associated with a 0.019 SD increase in years of schooling at age 8 and an increase 0.02 SD increase in IQ at age 8 |
| Kumar et al. (2022) | India, use the Indian Young Lives Surveys (YLS), a survey dataset designed to analyze the impact of early life inequality and poverty on life outcomes | A 10% increase in birth and being born within the LBW threshold leads to a 0.11 SD positive impact and 0.09 SD negative impact on cognitive test scores among children aged 5-8 respectively |
| Sicuri et al. (2011) | Mozambique, rural sample covering over 3000 births | Average cost per LBW infant in the first year following birth for the health system amounts is 4.5 times that of a normal birth weight infant |