BROOKINGS INDIA
QUALITY, INDEPENDENCE, IMPACT

Working Paper-06

Early Life Exposure to Outdoor Air Pollution:
Effect on Child Health in India

Prachi Singh

Sagnik Dey

Sourangsu Chowdhury

Kunal Bali

July 2019

Early Life Exposure to Outdoor Air Pollution: Effect on Child Health in India

Prachi Singh[*] Sagnik Dey[†] Sourangsu Chowdhury[†] Kunal Bali[†]

July 2019

[*]Indian Statistical Institute, Delhi & Brookings India
[†]Indian Institute of Technology, Delhi

Early Life Exposure to Outdoor Air Pollution: Effect on Child Health in India

Abstract

This paper examines effect of outdoor air pollution on child health in India by combining satellite PM2.5 data with geo-coded Demographic and Health Survey of India(2016). We use an instrumental variable strategy for identification as local pollution levels may be endogenous due to local household behavioural choices like participation in local fuel wood market, burning crop residue etc which are not observed in survey data. Our identification strategy relies on use of upwind biomass burning events in neighbouring areas to identify the effect of air pollution on child health. Our results indicate that one standard deviation increase in exposure to pollution during first trimester lowers Height-for-age (by 6.7 percent) and Weight-for-age (by 7.8 percent); the effect is prominent for poor people and Northern states of India which have higher pollution levels.

JEL Classification: 012 I15 Q53 Q56

Corresponding author:

Prachi Singh

Associate Fellow, Brookings India

&

PhD Candidate

Economics and Planning Unit

Indian Statistical Institute

New Delhi

Email: prsingh@brookingsindia.org, prachi.singh1987@gmail.com

Effect of Early Life Exposure to Outdoor Air Pollution on Child Health in India [1]

[PRELIMINARY DRAFT]

1 Introduction

Pollution in any form, whether it be air or water, poses an environmental risk to the health of the exposed population. According to WHO global air pollution database, out of the 15 most polluted cities in the world, 14 belong to India. Another recently published report by Health Effects Institute on air pollution in India (2018) reports that air pollution was responsible for 1.1 million deaths in India in 2015. In the absence of effective pollution regulatory policies, air pollution levels have reached alarming levels in various parts of India (Greenpeace, 2017). This warrants a closer look at the air pollution problem from the standpoint of welfare of the younger generation currently being exposed to harmful pollutants with possible long-lasting effect on their health. This article aims at estimating the effect of in-utero exposure to air pollution on child growth indicators, using exogenous changes biomass burning events which contribute to the air pollution.

Recent studies on India which focus on air pollution and child health rely on broad measures of pollution at the city level (Greenstone and Hanna, 2014). In this paper we conduct a pan-India analysis and we rely on rich geo-spatial information on air pollution to study its effect on children's growth indicators. In particular we study the effect of early life exposure to air pollution (as measured by PM 2.5) on children's weight and height measures for children under age five. The rich geo-spatial information on pollution comes from satellite data on aerosol optical depth which has been converted into gridded PM2.5 data (Dey et al.,

[1] We are thankful to Abhiroop Mukhopadhyay & E.Somanathan for their valuable comments. We are also thankful to conference participants at CECFEE, ADEW, EfD Annual Meet, AWEHE (inaugral meet), ACEGD & Brookings India. We are especially grateful for valuable feedback which we got from Randall Ellis, Shiko Maruyama, Sabyasachi Das and Dhritiman Gupta. We thank Athisii Kayina for his immense help with ArcGIS software.

2012; van Donkelaar et al., 2010). We match the gridded PM2.5 data to GPS locations of sampled clusters in Demographic Health Survey (DHS, 2015-16 round for India) to produce rich geo-spatial information about local (75 km radius) pollution levels in the place and time of conception(residence) of a child.

In an empirical exercise which causally links child health to local pollution levels, household income and behavioural choices are omitted variables which make local pollution levels endogenous. We use exogenous changes in biomass burning events like crop-burning and forest fires which are sources of pollution to address the endogeneity problem. In this paper we adopt a instrumental variable strategy where we use neighbouring *upwind* fire-events (that is biomass burning events) as an instrument for local pollution levels. This strategy critically relies on the assumption that exogenous changes in wind direction are not associated with household's income or behavioural choices. Multiple studies (Rangel and Vogl, 2018; Pullabhotla, 2018) have shown that these wind changes impact local pollution levels. The literature linking air pollution to child health has mostly focused on child mortality. In this paper we show that air pollution affects child's growth indicators even if she survives. To the best of our knowledge, this is the first study for India which addresses the endogeneity issues present while studying the link between children's growth indicators with local pollution levels.

Our analysis shows that air pollution negatively affects children's health. Exposure to air pollution during the first trimester decreases both Height-for-age (stunting measure) and Weight-for-age (underweight measure) for children aged below five years. A standard deviation change in PM2.5 is associated with 6.7% decrease in Height-for-age and a 7.8% decrease Weight-for-age measure. The effect is prominent for poorer households, with Northern states being more vulnerable due to high pollution levels in the area. These results are especially important given the link between stunting and other human capital outcomes. Early life stunting leads to irreversible damage, it is associated with shorter adult height, lower cognitive ability, lower educational attainment, reduced adult income, and decreased off spring birth weight (Victor et al., 2008; Mendez and Adair, 1999).

The paper follows the following structure. The next section provides a literature overview of the effect of pollution on child health and highlights the contribution of this paper to the literature. Section 3 describes the various datasets that we use in our analysis. The next section presents the empirical methodology that we follow and it is followed by results in Section 5. Lastly, Section 6 concludes with an estimate of the extent of the problem and discusses current state of policies regarding air pollution in India.

2 Previous Literature

Our work is motivated by the "fetal origins" hypothesis (Douglas and Currie, 2011), which states that the *in-utero* period of a child critically determines mortality outcomes, disease prevalence and future health outcomes, abilities and earnings. Fetal growth, if restricted, can negatively affect future outcomes. The biological link between air pollution and fetal growth has not been documented in the literature, but it is mediated by placental growth which determines supply of oxygen and nutrients to the fetus. Exposure to pollution would affect placental function which can be impacted by inflammation caused by maternal infection. Additionally pollution is known to cause epigenetic changes (interaction between our genes and environment which can cause DNA methylation, which regulates gene expression) which could affect fetal growth as well (Rangel and Vogl, 2018). A recent paper by Chakrabarti et. al (2019) has shown how exposure to biomass burning (which causes pollution) affects respiratory health in adults as well as children. Hence this suggests that mothers can possibly be affected during the pregnancy time due to exposure to pollution which can potentially affect fetal growth as well.

In the economics literature the intrauterine period has been the focus of many studies which have established links between occurrence of early life shocks to multiple outcomes. Early life shocks studied in economics literature include incidence of a) disastrous events (like famines, war, drought); b) nutritional shocks (like introduction of iodised salt, pregnancy during Ramadan) and c) pollution (air or water). Currie and Vogl (2013) provide a review of these early life shocks (a and b) on various outcomes; broadly summarised these shocks negatively affect adult cognition, years of schooling, literacy status, adult height and stunting measures; and increase the likelihood of presence of birth defects, prevalence of heart disease and obesity.

The focus of our study is in-utero exposure to air pollution and Currie et. al (2014) reviews landmark studies which have been conducted in this area. Most of these studies are from developed nations with a few exceptions. Similar to previous studies, a major part of the literature focuses on learning outcomes (test-scores) and earnings which are negatively affected due to in-utero exposure to pollution (Bharadwaj et al., 2013; Isen et al., 2013 & Sanders, 2012). We extend this literature by looking at the link between early life exposure to pollution and stunting and underweight measures.

The strand of literature which is most relevant for our study has mainly looked at the effect of in-utero or early life exposure to air-pollution on infant mortality and birth weight. Few papers in this area have used natural experiments to causally identify the effect of air pollution on infant survival, for example, Chay and Greenstone (2003a and 2003b) use introduction of environmental regulations under Clean Air Act, 1970 and recession in 1981-82 in United States to show that reduction in pollution levels led to reduction in infant mortality. Currie and Walker (2011) show that introduction of congestion-reducing automated toll payment systems in United States (which reduced number of idle vehicles emitting harmful pollutants) reduced pre-mature and low birth-weight births. Currie and Neidell (2005) use spatial and temporal variation in CO levels to analyse the effect of CO levels on infant mortality. Most of the studies in this domain are from developed nations where availability of high resolution pollution data is not a constraint. We focus on a developing nation which has much higher pollution levels in comparison to developed nations. Lack of data on pollution for developing nations has been a major limitation in the past but with availability of rich spatial information on pollution from satellite data we link local exposure to air pollution with child's growth factors.

In a developing country context the paper by Greenstone and Hanna (2014) analyses the effect of water and air pollution regulation policies on infant mortality in India. Another study from a developing nation includes Foster et al. (2009) which uses Mexico's clean industry certification program to study its effect on pollution (we use a similar measure of pollution i.e. satellite data on Aerosol Optical Depth to infer PM2.5 levels) and resulting respiratory related infant deaths. Wildfires and their negative health effects (like increase in infant mortality, reported asthma cases, pre-term births etc) have also been studied in context of Indonesian wildfire of 1997 (Jayachandran, 2009; Rukumnuaykit, 2003; Kunii et al., 2002; Frankenberg et al., 2005 & Barber and James, 2000), California wildfires (Holstius et al.,2012) and Australian wildfires (O'Donnell and Behie, 2015). A few recent papers assess the effect of in-utero exposure to biomass burning events and pollution: Vogl & Rangel (2018), Pullabhotla (2018) and Soo & Pattnayak (2019) study impacts on birth weight, infant mortality and long-term health outcomes like adult height respectively. These papers come closest to our paper as we also explore the link between in-utero exposure to pollution and child health but our sample is much bigger than the Indonesian study; we focus on solving the endogeneity problem in our paper rather than focusing on reduced form effect of biomass burning events on child health and we look at post-natal growth instead of survival. Another recent study on Bangladesh (Goyal & Canning(2017) provides evidence for in-utero exposure to air pollution and increased risk of stunting, underweight and wasting but it doesn't address the endogeneity issues related to local pollution levels.

Our paper also adds to the growing literature of the effect of pollution on child health in India. these effects have been demonstrated by two recent papers on effects of water pollution in India. Brainerd and Menon (2014) have focused on use of fertilisers in India during crop sowing season which increases concentration of harmful chemicals in water. They find that exposure to these pollutants during the month of conception increases infant mortality and reduces Height-for-age and Weight-for-age for children. Do et al. (2018) have shown that regulation targeting industrial pollution in the Ganga River led to reduction in water pollution levels and infant death.

3 Data

3.1 Demographic Data

The demographic data used in this paper is sourced from the Demographic and Health Survey (Round-4 for 2015-16) for India. DHS-IV contains detailed information about birth history of each woman who was interviewed. This survey sampled 601,509 households and interviewed 0.7 million eligible [2] women in the age group 15-49. Further, anthropometric measures of health were collected for 0.22 million children of ages five years and below. The DHS sample is a stratified two-stage sample and the primary sampling units (PSUs or clusters) correspond to villages in rural areas and blocks in urban areas. The DHS-IV comprises of around 28526 clusters with GIS information on almost all clusters[3]. To hide the identity of the village (block in urban areas) all clusters were displaced by five kilometres (two kilometres for urban clusters), with one percent of the clusters being displaced by as much as 10 kilometres. We account for this displacement when we discuss our identification strategy in the next section.

Our focus is on in-utero exposure to pollution for which we need the location and time of conception. To measure in-utero exposure to pollution we use the birth history of every child ever born to a woman. We use the location of the cluster, birth date and pregnancy duration of a child to impute exposure to pollution during the first trimester [4]. We make an important assumption that the place of stay of the mother when the child was in-utero is the same as current residence of a child [5].

We measure impact of air pollution on child health by using anthropometric measures: Height-for-age and Weight-for-age (WHO standard z-scores) for children aged five and below.

[2] Eligible women - married or unmarried women of reproductive ages.

[3] 131 clusters have no GIS information.

[4] We also construct separate measures of exposures to pollution for other trimesters and first three months after birth.

[5] This assumption is a standard assumption which is employed by many papers which have used DHS data for analysis (Brainerd & Menon, 2014). For example, in our sample the mean number of years for which the interviewed family has stayed at the place of residence is around 15 years.

In addition all other demographic and household level variables which are used in our analysis are sourced from the DHS.We provide summary statistics of our analysis sample in Table 1. 52 percent of children in our estimation sample [6] are males with mean age around 29 months (2.5 years old) and the mean birth order of children is 2.2. The birth order is slightly lower for children in South India. The average age at which mothers have children is 24.5 years. Mothers had on an average 6.2 years of education. Three-fourth of our sample consists of rural households and a similar proportion of households report their religion to be Hindu. 37% of our sample belongs to marginalised groups which includes schedule caste and scahedule tribes. The mean household size for our estimation sample is 6.5. 88 percent of the households are headed by a male member and the average age of household head is 44.5 years. 85 percent of the households have an electricity connection, but only 23 percent of the households use piped water as their source of drinking water, 28 percent of our sample uses clean source of cooking fuel like LPG or bio-gas and the mean open defecation rate in a cluster is 43 percent.

The mean Height-for-age (HFA) and Weight-for-age (WFA) z-score for our sample is -1.46 and -1.52 respectively (mean weight-for-height is -0.97 for our sample). Children from South India have much better HFA and WFA as compared to North Indian children. Height-for-age is a measure of stunting and it represents the effect of early life shocks that a child receives. Stunting generally occurs before age two and its effects are largely irreversible. It is associated with an underdeveloped brain, with long-lasting harmful consequences, including diminished mental ability and learning capacity, poor school performance in childhood, reduced earnings and increased risks of nutrition-related chronic diseases such as diabetes, hypertension, and obesity in future. Weight-for-age (underweight measure) reflects body mass relative to chronological age. It is influenced by both the height of the child (height-for-age) and his or her weight (weight-for-height). Deaton and Dreze (2009) advocate the use of Weight-for-age as the health status indicator for children as its a comprehensive measure which captures both stunting and wasting.

[6] Details about estimation sample discussed in pollution section later.

3.2 Pollution Data

In India, ground-based pollution measurement started post 2009 under the National Ambient Air Quality monitoring program maintained by Central Pollution Control Board. The network has slowly expanded to around 90 sites across 35 cities over the years, which leaves majority of India unmonitored [7]. Amongst these cities, only Delhi has greater than 20 monitoring sites while most other cities have a single monitoring site. Furthermore, most of the sites do not have continuous temporal data. We use PM2.5 as our measure of pollution which is a correlate of other pollutants (like NO2, SO2, CO) which are not captured in our analysis. To address the paucity in ground-based pollution data in India, we estimate PM2.5 exposure using satellite data (van Donkelaar et al., 2010). We convert Aerosol Optical Data (AOD) retrieved at 0.5 x 0.5 degree resolution from Multiangle Imaging SpectroRadiometer (MISR) to PM2.5 data (Liu et al, 2004; Kahn and Gaitley 2015; Dey et al. 2010) using a spatially and temporally heterogeneous conversion factor (Dey et al., 2012). The PM2.5 data is further statistically downscaled at 0.1 x 0.1 degree resolution using spline interpolation. The PM2.5 thus obtained is available at monthly frequency at 0.1 * 0.1 degree resolution (10km*10km grid).

We explore the spatial variation in PM2.5 by plotting a heat map in Figure 1. We plot mean annual PM2.5 (average over monthly data for years 2010 to 2016) for each district of India. As the figure shows, the Northern region of the country is severely impacted by high and dangerous levels of pollution, especially the states which lie in Indo-Gangetic plains (Punjab, Haryana, Uttar Pradesh, Bihar) have the highest levels of pollution. On the other hand, the Southern part of the country has much lower levels of pollution as shown by the lighter shades in heat map. The WHO guideline for maintaining safe standards of pollution recommends a threshold of mean annual pollution levels of $10ug/m^3$. Other standards include WHO-IT1 which is $35ug/m^3$, WHO-IT2 which is $25ug/m^3$ and WHO-IT3 which is $15ug/m^3$. The Indian National Ambient Air Quality Standards (NAAQS) sets the threshold at $40ug/m^3$. For our estimation sample we observe in Table 1 that the mean level of PM2.5

[7] India has around 600 ground based monitors to cover the entire country with only 148 monitors which capture PM2.5 for the entire country.

is between $55ug/m^3$ & $60ug/m^3$ during all critical windows of development for children from North India while the corresponding level of pollution exposure is much lower for children from South India, which is around $35ug/m^3$ for all windows.

We use cluster location from DHS data and calculate mean PM2.5 in the 75km radius for each month since the time of conception. We use these monthly pollution measures to construct trimester level pollution exposure by calculating mean PM2.5 for three month periods. Our estimation sample is constrained by the availability of PM2.5 data as we only keep those children in our estimation sample for whom the pollution measure for each month in the first trimester is present. The missing PM2.5 are due to missing satellite retrievals due to cloud covers. In the appendix Table A1, we show that our outcome variables along with our control variables are very similar between the estimation sample and out-sample (with missing PM2.5 information).

We now link exposure to pollution during the first trimester with anthropometric measure (Height-for-age z-score) for children in Figure 2. The descriptive graph is a bin-scatter plot which shows a negative relationship between Height-for-age and exposure to pollution during first trimester. We convert pollution exposure figures to z-scores for ease of exposition. However this correlation may be driven by other factors. We explore this relationship empirically in greater detail in Section 4.

4 Empirical Model

As pointed out above, we seek to investigate whether early life exposure to outdoor pollution during first trimester has an impact on future child health, measured by Height-for-age (z-scores) and Weight-for age (z-scores). Formally, we estimate the following empirical model:

$$H_{icdmt} = \theta_1 PM_{cdmt} + \beta X_{icdmt} + \gamma_c + \delta_t + \lambda_m + \rho^1_{dt} + \rho^2_{mt} + \varepsilon_{icdmt} \quad (1)$$

Our main outcomes of interest (H_{icdmt}) are z-score for Height-for-age (stunting measure) and Weight-for-age (underweight measure) for child i who was conceived in cluster c belonging to district d in month m and year t. The main variable of interest is PM_{cdmt} which captures the standardized PM2.5 (i.e. z-scores) in the 75km radius during first trimester for a child. To be precise, for ease of interpretation, we transform the mean PM2.5 in first trimester into z-scores based on the average and standard deviation of first trimester mean PM2.5 in the estimation sample. We control for confounding factors in the vector X_{icdmt} which includes gender, birth order and age of child, mother's and father's educational status, mother's age at birth, age and gender of household head, dummy for whether household has piped water, has clean cooking source, whether household practices open defecation and the fraction of households who practice open defecation in the cluster (excluding self). Since all children in our sample are aged five or below, we use the assumption that these controls have not changed a lot over time (i.e. from the time of conception to the time when they were surveyed).

Different clusters (villages or blocks) can have different levels of development (health infrastructure) which can affect health of a child hence we include cluster fixed effects in our specification. We also include month and year fixed effects to account for systematic effects related to season and year. We also remove any omitted variables that are related to a district in any particular year as well as any seasonality effect specific to a month of a particular year by including a district year specific fixed effect, ρ^1_{dt} and a month year specific fixed effect, ρ^2_{mt}. The inclusion of these fixed effects means that the variation that remains is the spatial variation in pollution within clusters and temporal variation within a year for a district.

While our estimation exercise removes systematic variation using various fixed effects, endogeneity concerns still remain. These endogeneity concerns arise as the *local* residential area for a household corresponds to the region of economic activity that a household depends on and also affects based on its behavioural decisions. The economic activity of a household determines key inputs (like income) which feed into the production function of health of a child. An example of this can be dependence of a household on nearby forest resources for fuel-wood consumption or for livelihood (if it sells these resources in a market). In this case the choice of use of fuel-wood by household affects the local pollution level in the region. Additionally the forest cover is affected by the demand for forest resources (like fuel-wood) in the market, which in turn affects the pollution level in the area where they are finally consumed. A similar logic holds true for crop residue burning as well, it is a conscious decision taken by a household which impacts local pollution levels and at the same time affects a farmer's income which is a determinant of child health. Thus, local pollution level is endogenous in the region of economic activity of the household.

4.1 Identification

The household behavioural choice of collecting fuel-wood or crop-burning and household income are omitted variables in our specification hence the local pollution variable is endogenous. To solve this endogeneity problem, we use a standardized measure of number of upwind fire-events (more on this below) which take place in the 75 to 100 km radius of the sampled cluster as an instrument. Fire-events are sourced from satellite image that are divided into pixels. Number of fire-events refers to the number of pixels where that atleast one fire-event is located within the pixel. These are biomass burning events that include crop residue burning and forest fires. Further, when fire incidents are recorded then each of them has a confidence value attached (interpreted as probability) which depicts the quality of the observation and therefore, using this we construct a confidence weighted count of fire-events around a cluster. We use only upwind fire-events, that is fire-events from which wind is blowing towards the cluster [8]. Further, following Rangel and Vogl (2018) for ease of interpretation of results, we standardize the events by calculating z-scores for these fire-events occurring in each cluster.

We use such fire-events only in the radius between 75 and 100 km (for ease of exposition we refer to this area as a non-local area) as they impact local mean PM2.5 levels (within 75km radius of a cluster) but are not affected by household behavioural choices. To elaborate further, these fire events belong to a region which is not a part of economic activity area of a household. This essentially removes the effect of dependence on crop-burning or nearby forest resources (or farmlands) for livelihood or fuel-consumption. Thus by capturing fire-

[8] Downwind fire-events refer to events with wind blowing away from the cluster.

events in this non-local area, we ensure that we only capture the part which contributes to the local pollution levels but is not correlated with household behavioural choices. Further exogenous changes in wind direction is unlikely to be correlated with local behavioural choices or economic activity in the area and thus non-local *upwind* fire-events serves as an ideal instrumental variable. Zheng et. al (2019) use a similar instrumental variable (IV) in their paper where they study the impact of air pollution on happiness levels using pollution levels of neighbouring areas as an IV for local pollution levels.

The IV that we use has been explained diagrammatically in Figure 3, where the light grey center denotes the cluster location, the white circle forms the 75 km radius around the cluster and the grey ring represents the area between 75 and 100 km radii around the cluster. Our endogenous variable is the mean PM2.5 variable which is calculated for the white circle (within 75 km) and the probability weighted number of upwind fire-events in the grey ring form the IV (between 75 and 100 km).

4.1.1 Fire-events and Wind Data

Our source of biomass burning events (called fire incidents) is NASA's Fire Information for Resource Management System (FIRMS) data which captures real-time active fire locations across the globe. The FIRMS data that we use is called MODIS (shortform for MODerate Resolution Imaging Spectro radiometer) data and it records fire incidents at pixel level where each pixel is identified by a latitude and longitude reading. Each latitude (and longitude) is centroid of a one kilometre pixel (1 km X 1 km in size). This data records not just the location of a fire but also the brightness (temperature) of fire (in Kelvin units) and date and time when the incident was picked by the Terra satellite. An observation for a fire incident in MODIS data for a latitude and longitude does not necessarily mean that the size of the fire is one square kilometre, but it means that atleast one fire is located within this fire pixel (under good conditions the satellite can detect fires as small as $100\mathtt{m}^2$). The MODIS data is available on a daily basis since November 2000 and NASA reports that the fires captured by this dataset are mostly vegetation fires. NASA data on fire incidents also provides a variable "confidence", which depicts the quality of the observations and it ranges from 0-100

[9]. Following Rangel and Vogl (2018), we use this variable to construct a probability weighted count of fire-events around a cluster (between 75 and 100 km) [10].

We use the cluster GIS information from DHS data and calculate the probability weighted count of fire events which took place between 75 and 100 km radii (non-local exposure) during the first trimester of a child. To ensure respondent confidentiality, all clusters in the DHS data are displaced from their true location. The displacement is done by displacing an urban cluster by two kilometre and a rural cluster by five kilometre with one percent of the rural clusters being displaced by as much as 10 kilometres. The displacement can take place in any direction but the cluster remains within the country boundary, within the same state and district. We take the radius for our analysis to be 75 kilometre which is large enough so that the true location of the cluster and sphere of economic activity of a household is contained within the 75 kilometre radius circle.

Meteorological variables such as wind speed and wind direction are expected to play an important role in modulating the outflow of fire burning residues emitted from a fire event. To account for this, we tag each fire event with the respective wind speed and wind direction. We use ERA-Interim data of u (zonal wind) and v (meridional wind) at 10m from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim dataset at 0.125*0.125 degree resolution.

The wind speed and wind direction was estimated as in equation (2) and (3) respectively (Chowdhury et al., 2017):

$$ws = sqrt[(u)^2 + (v)^2] \tag{2}$$

$$winddirection = [atan(u/v) * (180/pi] + 180 \tag{3}$$

[9]We convert confidence figures into probability figures by dividing them by 100

[10]NASA's FIRMs data can also have some missing values attributable to satellite sensor outage. However major incidents reported for sensor outage happened in years 2001-2003 which precedes our analysis period.

4.1.2 Fire-events and Pollution

India has a substantial amount of land under cultivation(60%) and under forest cover(25%), with majority biomass burning events taking place in these areas. Over the past few decades, Indian agriculture has been marked with expansion of irrigation facilities, adoption of high yield variety seeds and increased mechanisation (like use of combine harvester). A combination of these factors led to adoption of multi-cropping system by farmers which leaves little time in between the harvest of one crop and sowing of another. In this scenario, crop residue burning thus emerged as the quickest and cheapest way to get the farm ready for the next crop. Cereals are the prime contributor to crop burning activity in India, with rice and wheat crop residue burning forming the major chunk of residue burning process (Jain et al, 2014). Two major residue burning seasons are thus related to crop harvest seasons: kharif crop harvest (rice stubble burning) which takes place in the months of October and November; and rabi crop harvest (wheat straw burning) which happens in the months of March to May.

Biomass burning in India is not limited to just crop residue burning, it covers forest fires as well. Forest fires or wildfires are caused by various factors acting in conjunction with each other. These factors include availability of biomass (dry vegetation) and appropriate climatic conditions(high temperature, low pressure, windy conditions). Forest Survey of India lists vulnerable months for each state when forest fires are most likely to happen, which mainly span the high temperature months from March to June. Wildfires happen due to both intentional and unintentional human activity. In North Eastern states and in states along the Eastern Ghats, slash and burn activity is rampant wherein vegetation in forests is cut (slashed) and then burned to clear the piece of land for human use. In a lot of cases unintentional human activities like leaving active cigarette butts behind in open forests lead to forest fires. Other natural factors which cause forest fires include lightening which produces a spark to start a fire in dry vegetation.

Figure 4 provides a linear fit plot between local PM2.5 levels and non-local fire-events (all fire-events - left panel and just upwind fire-events - right panel). A strong positive relationship between the two is evident from this graph and forms the basis for using non-local fire-events as an IV for local pollution levels. In our empirical work though, we use the within cluster variation of these variables.

In Figure 5, we plot the temporal variation in PM2.5 and non-local fire-events. The figure plots mean levels across all sampled clusters in the latest DHS round for India. We look at mean PM2.5, mean count of total fire-events which take place in non-local areas and mean count of total upwind fire-events which take place in non-local areas. As shown in the graph (solid blue line) the winter months (from October to January) have highest pollution levels in comparison to summer months (March to June), with lowest pollution levels recorded in monsoon period (August-September). Corresponding to two harvest seasons we see two peaks in fire-events plots (both all fire-events and upwind fire-events in dashed lines).

In western countries forest fires are mainly responsible for the carbon content release due to biomass burning; however, in case of India (and other South Asian countries) crop residue burning contributes the most to total carbon release. In South Asia, India stands out both in terms of total area burned (4.5 million hectares burned in 2015) and in terms of total carbon content (1.5 million metric tonnes) released due to biomass burning. A raw count of biomass burning events in India shows that roughly both crop residue burning and forest fires contribute equally. However, if we weigh these events based on the population density [11] of the area in which these events occur then crop burning events contribute more to the total biomass burning events (65 percent). This mainly happens because residue burning activities happen in more populated areas as against forest fires which happen in low density areas. Appendix Figure 1 provides the population weighted split between forest fires and crop residue burning in few selected states in India. As can be seen in this graph, with an exception of Punjab, almost all other states are affected by both forest fires and residue burning.

Biomass burning is a major source of pollution as it releases harmful pollutants like

[11] Geo-coded fire events have been projected onto land mask cover for India to categorise each fire event as an event which happens in a forest area vs cropped area. This data is then projected onto density map of India, to get the density of the population in which these events take place.

Carbon Dioxide(CO_2), Carbon Monoxide (CO), Sulphur Oxides and particulate matter (PM) in the atmosphere. The release of harmful pollutants in the atmosphere is captured by aerosol loading [12] in the region. To summarise, fine particulate matter released during biomass burning incidents have long range travel properties and affect not just the local areas but far away regions as well.

Arguments above provide some suggestive evidence about the fact that non-local fire events are associated with local pollution levels. Further evidence on this will be provided when we discuss the first stage of 2SLS regression. However in addition what we require for our IV strategy to work is that our IV should be uncorrelated with other factors which are related to child health. We provide evidence in next section that it is likely to be true.

5 Results

5.1 Pollution and Child Health

We begin by presenting OLS results on effect of mean outdoor pollution in the first trimester on child health outcomes in Table 2. Column 1 and 2 in Table 2 show that pollution exposure during first trimester is negatively correlated to weight-for-age (WFA-Z) and height-for-age (HFA-Z). The OLS estimates are small, a one standard deviation change in local PM2.5 reduces WFA-Z by 0.012 and HFA-Z by 0.011 standard deviation units (not significant). A possible reason behind small coefficients could be the fact that local pollution exposure subsumes the effect of both income and physiological effect of PM2.5 on child health. Since these two effects can affect child health in opposite ways so the OLS estimate we notice is smaller. Also, as described in the previous section, in equation (1) local pollution exposure variable is riddled with endogeneity problem, hence the OLS estimates are biased.

To address the endogeniety problem, we use an Instrumental Variable strategy where upwind fire events in the non-local areas are used as an instrument for local pollution levels.

[12] Aerosol loading is the suspensions of solids and/or liquid particles in the air that we breathe. Dust, smoke, haze are also part of aerosol loading.

These upwind fire-events are assumed to be orthogonal to the income levels, so we are able to capture the pure effect of PM2.5 disentangling it from the income effect. We present the first stage of 2SLS results in Table 3. As hypothesized, we find that the relationship between the endogenous variable - local PM2.5 in 75km radius and non-local upwind fire-events is positive and highly significant. A one standard unit change in number of upwind fire-events leads to a 0.105 standard deviation unit increase in local pollution levels. This is in line with our hypothesis that particulate matter from fire-events far away affect local pollution levels. The first stage rk-LM statistic is 1005 and is much above the Stock & Yoko bias cut off. These results represent that local PM2.5 variation is affected by the seasonality present in biomass burning events happening in non-local adjacent areas.

We test whether our IV meets exclusion restriction by providing some suggestive evidence in Table 4. We regress various characteristics of a household (and its members) on our main IV - upwind fire-events, essentially an insignificant result shows that there is no systematic relationship between our IV and household (and its member's) characteristics. We do this by regressing variables which affect child health on our IV, columns 1 to 5 in Table 4 shows education level of mother, source of water, choice of cooking fuel, open defecation measure and birth order which act as controls in our main specification are not systematically related to the IV. Columns 6 to 12 provides results for other variables (these include wealth class, asset ownership, religion, dummy for minority group, household size and vaccination) which can potentially affect child health and we find that they are also not related to fire-intensity in non-local areas. The only exception is pregnancy duration which is positively correlated with our IV in column 7.

We next move to the second stage results obtained using 2SLS strategy. We find that both WFA-Z and HFA-Z are negatively affected by outdoor pollution experienced in-utero during the first trimester. Columns 3 and 4 in Table 2 present our 2SLS results using upwind fire events in non-local area as an IV. We find that a standard deviation unit change in mean PM2.5 during first trimester leads to a decrease in WFA-Z(HFA-Z) by -0.102(-0.115) standard deviation units which translates into a 6.7 percent decrease in WFA-Z and 7.8 percent decrease in HFA-Z.

Additional results

We provide results on our full model in Appendix Table A2. We find that being a male child, or being born later (higher birth order) is associated with lower HFA-Z and WFA-Z. Similar to previous findings in the literature, we find that child growth indicators are positively associated with mother's education level and also age at which mother gives birth. Source of water being pipedwater seems to have no affect on child health while use of clean cooking fuel is associated with better child health outcomes. Household's open defecation practice is negatively associated with stunting and underweight measures. Finally, an older household head perhaps contributes to better child care and hence is associated positively with child health outcomes, while gender of the household head being male only affects stunting measure.

We now focus on other time windows of critical development, that is second, third trimester and the post-natal period of first three months after birth. Table 5 summarises our results, we find that in-utero exposure to outdoor pollution which is experienced by the mother (and her foetus) for second, third trimester and post-natal period[13] has no impact on Height-for-age, but some negative effect is present for Weight-for-age corresponding to exposure in second trimester.

5.2 Robustness Checks

5.2.1 Extended Controls

a) Local weather conditions

In this section we provide multiple robustness checks for our results. Local weather condition like rainfall can play an important role as rainfall makes the ash and other pollutant particles settle on the ground thereby reducing pollution levels. Temperature also plays an important role in pollution dynamics. We control for both local temperature and rainfall in columns 1 and 2 in Table 6. The number of observations is slightly smaller than before due

[13]The number of observations differ depending upon availability of pollution data for all months for the window of analysis.

to missing rainfall and temperature information for some clusters. Our original results still hold and the magnitude of the effect is slightly larger after accounting for weather controls.

b) Gestational period, Household Size, Caste

The gestational period or pregnancy duration is also an important determinant of intrauterine growth of a foetus which affects future child health. We additionally control for household size and minority status of a household (being schedule caste or schedule tribe) to see if extended controls affect our original results. We find in column 3 and 4 (in Table 6) that our estimates remain unchanged. Duration of gestational period is positively associated with child growth indicators while children belonging to minority group have worse health outcomes. Household size seems to have no effect on child growth indicators.

5.2.2 Sensitivity Analysis

The analysis uptil now used upwind fire-events happening in 75 to 100 km radius as the IV for local mean PM2.5 in the 75 km radius around the cluster location. We now provide results for alternate radii specifications to test the sensitivity of our model. In Table 7, columns 1 and 2, the IV being used is the probability weighted total number of upwind fire-events in 50 to 100 km radius (compressing the white inner circle in Figure 5). In columns 3 and 4, the IV being used is the probability weighted total number of upwind fire-events in 50 to 75 km radius for local mean PM2.5 in the 50 km radius (compressing the donut in Figure 5). Reducing the local pollution radius to 50 kms leads to a significant drop in total number of observations as PM2.5 information is missing for a lot of observations. However we still find that our results are of similar magnitude (they are slightly smaller for HFA-Z analysis) and still remain significant. The HFA-Z result in Table 7 column 2 is significant at 10 percent level while in column 4 it is marginally significant at 10 percent level (p-value = 0.109). Lastly in columns 5 and 6, we drop the observations corresponding to the state of Punjab. This has been done to ensure that our results are not driven in any way by the state of Punjab which is affected by high levels of pollution corresponding to highest level of recorded fire-events in India[14]. Our results become larger in magnitude and are more significant after dropping the state of Punjab.

5.2.3 Falsification Tests

In Table 8 (column 1 and 2) instead of using upwind fires as an IV we use downwind fires (which lie in opposite octant from that of upwind fires with wind blowing away from cluster location). We find that using downwind fire-events as an IV makes our results insignificant and in case of HFA-Z the insignificant point estimate has the opposite sign. In columns 3 and 4, we provide results on the effect of pollution on child health where the location of a child has been randomly shuffled. This random assignment of location leads to counter-intuitive positive (and insignificant) effect of early life exposure to pollution during first trimester on child health which strengthens our hypothesis that location does matter when it comes to pollution exposure (and in turn affects child health).

5.2.4 Avoidance Behaviour

Do mothers plan conception?

An important threat in our analysis can be avoidance behaviour by mothers, that is if mothers purposely avoid particular months for conception due to their concern about future child health related to seasonal biomass burning activities. We test this by looking at birth history of mothers for the estimation period i.e. years 2010 to 2016. We do this by creating a mother-month-year panel. We create a dummy variable which takes value 1 if a mother successfully conceives in a particular month of an year. We estimate a linear probability model to test whether mother's conception behaviour is systematically linked to non-local fires. We control for mother's education, characteristics of household head along with other household characteristics like source of water, toilet facility, choice of cooking fuel. We introduce the same fixed effects which are present in our initial specification to control for regional and seasonal factors. We present these results in Table 9. In column 1, we present results where we try and see whether there is correlation between three month exposure to

[14] Almost 25% of total fire-events in India take place in Punjab.

non-local upwind fire-events and conception. In column 2, we assess whether exposure to fire-events in the month of conception has any effect on conception behaviour. We find that in both the cases non-local upwind fire-events increase the probability of conception. This suggests that we have some positive selection, as probability of mothers conceiving is more during the time when incidence of fire-events is high. A one standard deviation in non-local fire events in a month increases the probability of conception by 0.015% (corresponding figure for 3 month exposure is 0.03%. Our analysis provides some evidence that mothers do not practice avoidance behaviour.

5.3 Heterogeneity

i) By Background Characteristics

We now provide disaggregated regressions for Height-for-age. We split our estimation into Poor (wealth index lower than 2) and Rich sample (wealth index greater than equal to 3). In Table 10 (Column 1 and 2), we find that the negative effect of pollution is present only for poor households. This can possibly be due to the fact that children in poor households have less access to health care to abate negative effect of pollution on health.

As discussed before Southern India has lower pollution levels in comparison to Northern India. While mean PM2.5 is above $56ug/m^3$ during all critical windows of development in North India, the corresponding figure for South India is as low as $35ug/m^3$. We do sub-sample analysis on observations from Northern and Southern States in columns 3 and 4 and find that most of the effects that we see are limited to North India which have alarmingly high levels of pollution throughout the year.

Finally, we compare children who are born to mothers with different educational attainment. In column 5 and 6 of Table 10, we find the negative effect of pollution on child health is mainly present for mothers who have less (till primary level) or no education. There is negative effect present for educated mothers (secondary or above) as well but it is not significant.

ii) By Child's Age

In Table 11 we test whether the effect of in-utero exposure to pollution persists overtime. We observe that the effect is negative for all age-groups (zero to one year, greater than one but less two years and greater than three years old). Studies have suggested that stunting is irreversible after age of two years, we do find a significant negative persistent effect of early exposure to pollution on stunting outcome for one to two year old children. Although the effect is smaller but it continues to be present for children older than 3 years as well.

6 Conclusion

Outdoor pollution in India breaches safe standards in many areas. We link outdoor pollution to biomass burning which is a significant source of carbonaceous aerosols, it plays a vital role in atmospheric chemistry, air quality, ecosystems, and human health. Our analysis shows that outdoor pollution is affected by neighbouring biomass burning events; this is used to causally infer the effect of outdoor pollution (as measured by PM2.5) on child growth indicators. We find that a z-score increase in PM2.5 levels during first trimester leads to a reduction in Height-for-age (HFA-Z, stunting measure) and Weight-for-age (WFA-Z, underweight measure) by 0.115 and 0.102 standard deviation units respectively. Figure 6 summarises our results graphically, exposure to outdoor pollution during different critical windows of growth of a child is associated with worse child health outcomes. Almost all the estimates are negative with significant effect present for exposure to pollution during first trimester and second trimester (only WFA-Z measure).

The above results establish that exposure to pollution is linked to stunting measure (HFA) in childhood. What impact does this have on the economy? We provide a back-of-an-envelope calculation based on the Galasso et al. (2016) study. This study does a literature review of the effect of stunting on GDP. Stunting affects GDP of a nation via three channels: lower returns to lower education, lower returns to lower height and lower returns to lower cognition. For India, where 66 percent of the workforce was stunted in childhood, this study estimates that a complete elimination of stunting would have increased GDP by 10 percent

[15]. We use a point estimate of probability of being stunted due to outdoor pollution, and find that one standard deviation increase in outdoor pollution leads to a 0.18 percent reduction in GDP.

India needs effective policies regarding regulation and management of outdoor pollution, since the current policies are ineffective. Cross-border policies are needed to tackle the problem of pollution. To curb air pollution, effective management of forest fires is needed; however, the budget allocation for this purpose is really small and remains unused in every financial year. Similarly the government has committed itself to subsidising the use of happy-seeder technology (this is an alternative to combine harvester, it leaves rice residue in form of a mulch on farm which doesn't hamper wheat crop sowing and hence doesn't require burning), however the uptake of this policy remains quite low due to high initial investment in the machine (Gupta and Somnathan, 2016). The National Clean Air Program (2018) is a welcome step in this domain as it plans to extend air quality monitoring network, conduct intensive awareness and monitoring campaigns, create city-specific action plans, among many other initiatives.

[15] This is an average figure for South Asia.

References

Assessment of Air Pollution in Indian Cities 2017. Technical report, Greenpeace.

State of Global Air 2018. Technical report, Health Effects Institute.

Vulnerability of Forest Fire 2017. Technical report, Forest Survey of India.

Acharya, P., Sreekesh, S., and Kulshrestha, U. (2016). Ghg and Aerosol Emission from Fire Pixel during Crop Residue Burning Under Rice and Wheat Cropping Systems in North-West India. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41:753.

Almond, D. and Currie, J. (2011). Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, 25(3):153–72.

Barber, C. and Schweithelm, J. (2000). Trial by Fire: Forest Fires and Forestry Policy in Indonesia's Era of Crisis and Reform. Technical report, World Resources Institute, Washington, DC (EUA). Forest Frontiers Initiative.

Bharadwaj, P., Gibson, M., Zivin, J. G., and Neilson, C. (2017). Gray Matters: Fetal Pollution Exposure and Human Capital Formation. *Journal of the Association of Environmental and Resource Economists*, 4(2):505–542.

Brainerd, E. and Menon, N. (2014). Seasonal Effects of Water Quality: The Hidden Costs of the Green Revolution to Infant and Child Health in India. *Journal of Development Economics*, 107:49–64.

Chakrabarti, S., Khan, M. T., Kishore, A., Roy, D., and Scott, S. P. (2019). Risk of Acute Respiratory Infection from Crop Burning in India: Estimating Disease Burden and Economic Welfare from Satellite and National Health Survey Data for 250 000 persons. *International Journal of Epidemiology*.

Chakravorty, U., Pelli, M., and Risch, A. (2014). Far Away from the Forest? Fuelwood Collection and Time Allocation in Rural India.

Chakravorty, U., Pelli, M., and Risch, A. (2015). Who is Buying? Fuelwood Collection in Rural India.

Chay, K. Y. and Greenstone, M. (2003a). Air Quality, Infant Mortality, and the Clean Air Act of 1970. Technical report, National Bureau of Economic Research.

Chay, K. Y. and Greenstone, M. (2003b). 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.

Chowdhury, S., Dey, S., Tripathi, S. N., Beig, G., Mishra, A. K., and Sharma, S. (2017). "traffic Intervention" Policy Fails to Mitigate Air Pollution in Megacity Delhi. *Environmental Science & Policy*, 74:8–13.

Cook, C. J., Fletcher, J. M., and Forgues, A. (2018). Multigenerational Effects of Early Life Health Shocks. Technical report, National Bureau of Economic Research.

Cummins, J., Agarwal, N., Aiyar, A., Bhattacharjee, A., Gunadi, C., Singhania, D., Taylor, M., Wigton-Jones, E., et al. (2017). Month of Birth and Child Height in 40 Countries. Technical report.

Currie, J. and Neidell, M. (2005). Air Pollution and Infant Health: What Can We Learn from California's Recent Experience? *The Quarterly Journal of Economics*, 120(3):1003–1030.

Currie, J. and Vogl, T. (2013). Early-life Health and Adult Circumstance in Developing Countries. *The Annual Review of Resource Economics*, 5(1):1–36.

Currie, J. and Walker, R. (2011). Traffic Congestion and Infant Health: Evidence from E-Z Pass. *American Economic Journal: Applied Economics*, 3(1):65–90.

Currie, J., Zivin, J. G., Mullins, J., and Neidell, M. (2014). What Do We Know about Short-and Long-Term Effects of Early-Life Exposure to Pollution? *The Annual Review of Resource Economics*, 6(1):217–247.

Deaton, A. and Drèze, J. (2009). Food and Nutrition in India: Facts and Interpretations. *Economic and Political Weekly*, pages 42–65.

Dey, S. and Di Girolamo, L. (2010). A Climatology of Aerosol Optical and Microphysical Properties over the Indian Subcontinent from 9 Years (2000–2008) of Multiangle Imaging Spectroradiometer (MISR) Data. *Journal of Geophysical Research: Atmospheres*, 115(D15).

Dey, S., Girolamo, L., Van Donkelaar, A., Tripathi, S., Gupta, T., and Mohan, M. (2012). Decadal Exposure to Fine Particulate Matters (PM 2.5) in The Indian Subcontinent Using Remote Sensing Data. *Remote Sens Environ*, 127:153–161.

Do, Q.-T., Joshi, S., and Stolper, S. (2018). Can Environmental Policy Reduce Infant Mortality? Evidence from the Ganga Pollution Cases. *Journal of Development Economics*, 133:306–325.

Douglass, R. L. (2008). Quantification of the Health Impacts Associated with Fine Particulate Matter due to Wildfires.

Fletcher, J. M., Green, J. C., and Neidell, M. J. (2010). Long Term Effects of Childhood Asthma on Adult Health. *Journal of Health Economics*, 29(3):377–387.

Foster, A., Gutierrez, E., and Kumar, N. (2009). Voluntary compliance, Pollution Levels, and Infant Mortality in Mexico. *American Economic Review*, 99(2):191–97.

Foster, A. D. and Rosenzweig, M. R. (2003). Economic Growth and the Rise of Forests. *The Quarterly Journal of Economics*, 118(2):601–637.

Frankenberg, E., McKee, D., and Thomas, D. (2005). Health Consequences of Forest Fires in Indonesia. *Demography*, 42(1):109–129.

Galasso, E., Wagstaff, A., Naudeau, S., and Shekar, M. (2016). The Economic Costs of Stunting and How to Reduce Them. *Policy Research Note World Bank, Washington, DC.*

Goyal, N. and Canning, D. (2017). Exposure to Ambient Fine Particulate Air Pollution In-Utero as a Risk Factor for Child Stunting in Bangladesh. *International Journal of Environmental Research and Public Health*, 15(1):22.

Greenstone, M. and Hanna, R. (2014). Environmental Regulations, Air and Water Pollution, and Infant Mortality in India. *American Economic Review*, 104(10):3038–72.

Gupta, P. K., Sahai, S., Singh, N., Dixit, C., Singh, D., Sharma, C., Tiwari, M., Gupta, R. K., and Garg, S. (2004). Residue Burning in Rice–Wheat Cropping System: Causes and Implications. *Current Science*, pages 1713–1717.

Gupta, R. (2014). Low-Hanging Fruit In Black Carbon Mitigation: Crop Residue Burning In South Asia. *Climate Change Economics*, 5(04):1450012.

Gupta, S., Agarwal, R., and Mittal, S. K. (2016). Respiratory Health Concerns in Children at Some Strategic Locations from High PM Levels during Crop Residue Burning Episodes. *Atmospheric Environment*, 137:127–134.

He, G., Fan, M., and Zhou, M. (2016). The Effect of Air Pollution on Mortality in China: Evidence from The 2008 Beijing Olympic Games. *Journal of Environmental Economics and Management*, 79:18–39.

Holstius, D. M., Reid, C. E., Jesdale, B. M., and Morello-Frosch, R. (2012). Birth Weight following Pregnancy during The 2003 Southern California Wildfires. *Environmental Health Perspectives*, 120(9):1340.

Isen, A., Rossin-Slater, M., and 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.

Jain, N., Bhatia, A., and Pathak, H. (2014). Emission of Air Pollutants from Crop Residue Burning in India. *Aerosol and Air Quality Research*, 14(1):422–430.

Jayachandran, S. (2009). Air Quality and Early-Life Mortality Evidence from Indonesia's Wildfires. *Journal of Human Resources*, 44(4):916–954.

Kahn, R. A. and Gaitley, B. J. (2015). An Aanalysis of Global Aerosol Type as Retrieved by MISR. *Journal of Geophysical Research: Atmospheres*, 120(9):4248–4281.

Kaskaoutis, D., Kumar, S., Sharma, D., Singh, R. P., Kharol, S., Sharma, M., Singh, A., Singh, S., Singh, A., and Singh, D. (2014). Effects of Crop Residue Burning on Aerosol Properties, Plume Characteristics, and Long-Range Transport Over Northern India. *Journal of Geophysical Research: Atmospheres*, 119(9):5424–5444.

Kunii, O., Kanagawa, S., Yajima, I., Hisamatsu, Y., Yamamura, S., Amagai, T., and Ismail, I. T. S. (2002). The 1997 Haze Disaster in Indonesia: Its Air Quality and Health Effects. *Archives of Environmental Health: An International Journal*, 57(1):16–22.

Lai, W., Li, Y., Tian, X., and Li, S. (2017). Air Pollution and Cognitive Function: Evidence from Crop Production Cycles. *Available at SSRN 3039935*.

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.

O'Donnell, M. and Behie, A. (2015). Effects of Wildfire Disaster Exposure on Male Birth Weight in an Australian Population. *Evolution, Medicine, and Public Health*, 2015(1):344–354.

Prinz, D., Chernew, M., Cutler, D., and Frakt, A. (2018). Health and Economic Activity Over the Lifecycle: Literature Review. Technical report, National Bureau of Economic Research.

Pullabhotla, H. (2018). Fires, Wind, and Smoke: Air Pollution and Infant Mortality; Job Market Paper.

Rangel, M. A. and Vogl, T. S. (2018). Agricultural Fires and Health at Birth. *Review of Economics and Statistics*, (0).

Rukumnuaykit, P. Crises and Child Health Outcomes: The Impacts of Economic and Drought/Smoke Crises on Infant Mortality and Birth Weight in Indonesia. *Economics Department, year=2003, publisher=Citeseer.*

Sahu, S. K., Zhang, H., Guo, H., Hu, J., Ying, Q., and Kota, S. H. (2019). Health Risk Associated with Potential Source Regions of PM 2.5 in Indian Cities. *Air Quality, Atmosphere & Health*, pages 1–14.

Sanders, N. J. (2012). What Doesn't Kill You Makes You Weaker Prenatal Pollution Exposure and Eucational Outcomes. *Journal of Human Resources*, 47(3):826–850.

Sharma, A. R., Kharol, S. K., Badarinath, K., and Singh, D. (2010). Impact of Agriculture Crop Residue Burning on Atmospheric Aerosol Loading–A Study over Punjab State, India. *Annales Geophysicae*, 28(2).

Simeonova, E., Currie, J., Nilsson, P., and Walker, R. (2018). Congestion Pricing, Air Pollution and Children's Health. Technical report, National Bureau of Economic Research.

Singh, C. P. and Panigrahy, S. (2011). Characterisation of Residue Burning from Agricultural System in India Using Space Bbased Observations. *Journal of the Indian Society of Remote Sensing*, 39(3):423.

Soo, J.-S. T. and Pattanayak, S. K. (2019). Seeking Natural Capital Projects: Forest Fires, Haze and Early-Life Exposure in Indonesia. *Proceedings of the National Academy of Sciences.*

Srivastava, P. and Garg, A. (2013). Forest Fires in India: Regional and Temporal Analyses. *Journal of Tropical Forest Science*, pages 228–239.

Vadrevu, K. P., Lasko, K., Giglio, L., and Justice, C. (2015). Vegetation Fires, Absorbing

Aerosols and Smoke Plume Characteristics in Diverse Biomass Burning Regions of Asia. *Environmental Research Letters*, 10(10):105003.

van Donkelaar, A., Martin, R., Verduzco, C., Brauer, M., Kahn, R., Levy, R., and Villeneuve, P. (2010). A Hybrid Approach for Predicting PM 2.5 Exposure: van Donkelaar et al. respond. *Environmental Health Perspectives*, 118(10):A426–A426.

Van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P. J. (2010). Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. *Environmental Health Perspectives*, 118(6):847.

Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L., Sachdev, H. S., Maternal, Group, C. U. S., et al. (2008). Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital. *The Lancet*, 371(9609):340–357.

Vijayakumar, K., Safai, P., Devara, P., Rao, S. V. B., and Jayasankar, C. (2016). Effects of Agriculture Crop Residue Burning on Aerosol Properties and Long-Range Transport over Northern India: A Study Using Satellite Data and Model Simulations. *Atmospheric Research*, 178:155–163.

Zheng, S., Wang, J., Sun, C., Zhang, X., and Kahn, M. E. (2019). Air Pollution Lowers Chinese Urbanites' Expressed Happiness on Social Media. *Nature Human Behaviour*.

Tables and Figures

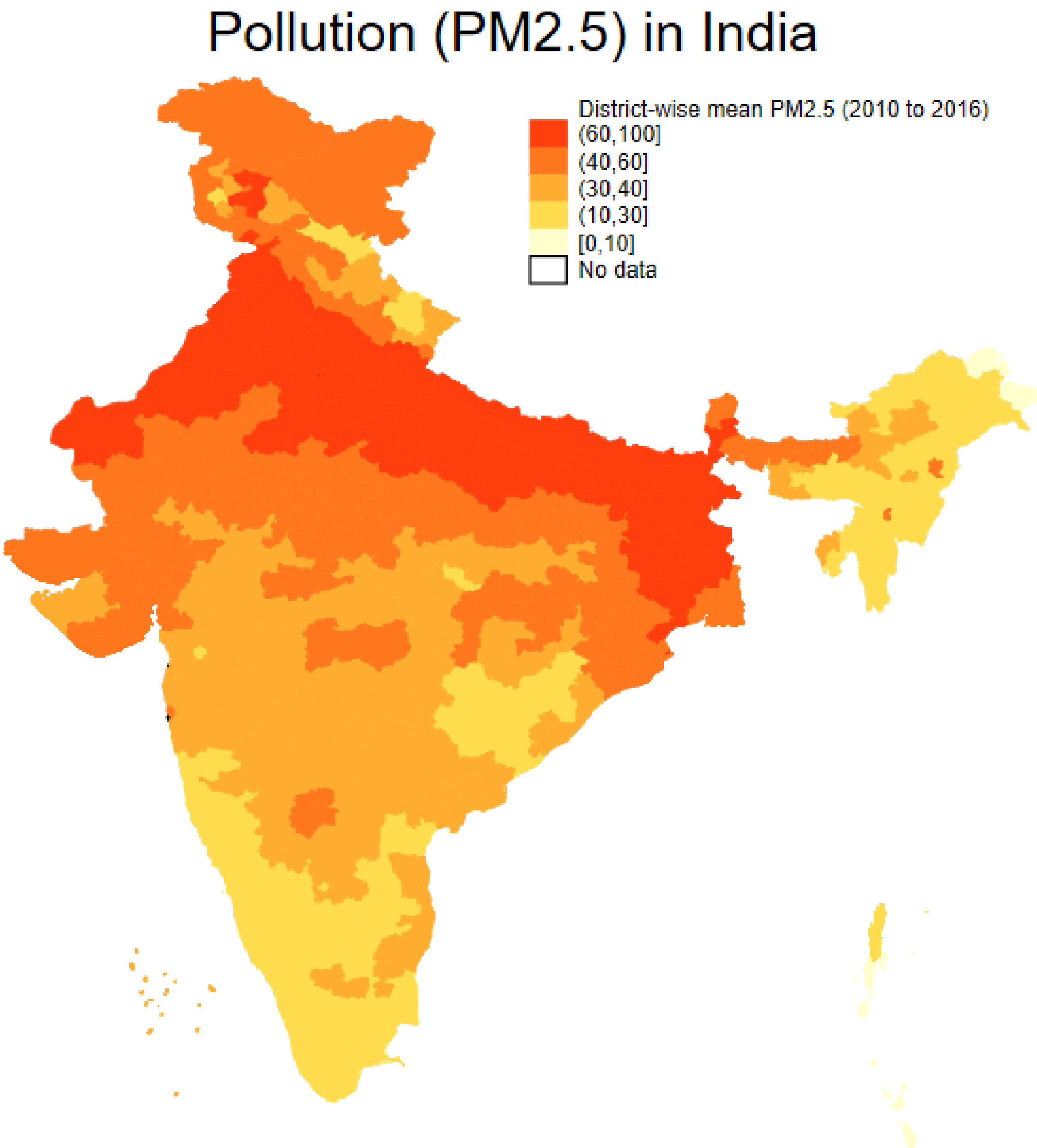

Figure 1: Spatial variation in Pollution: Mean PM2.5 in districts of India (2010 to 2016)

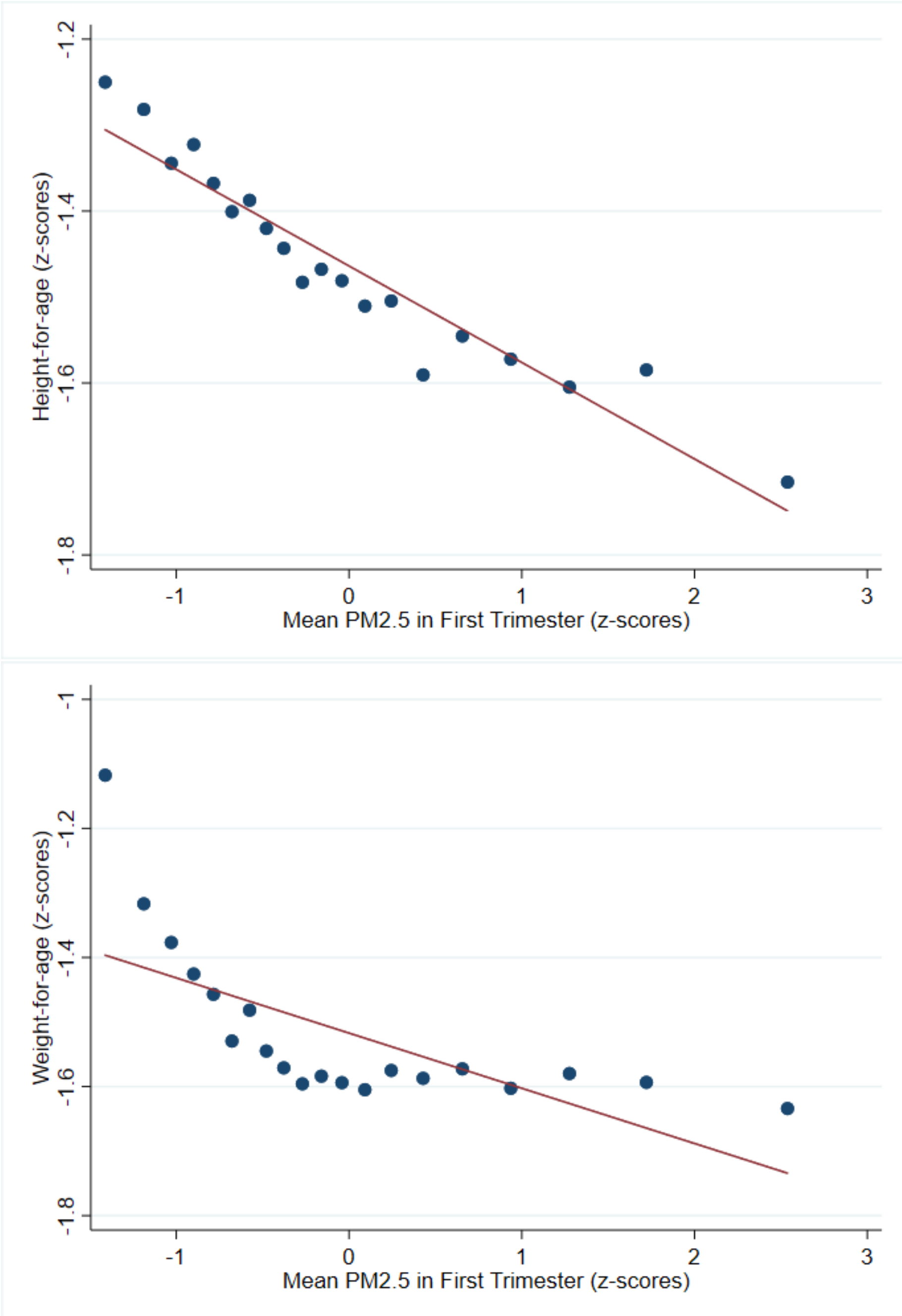

Figure 2: Binscatter plot for relationship between Height-for-age & Weight-for-age (z-scores) and Mean PM2.5 in first trimester (z-scores).

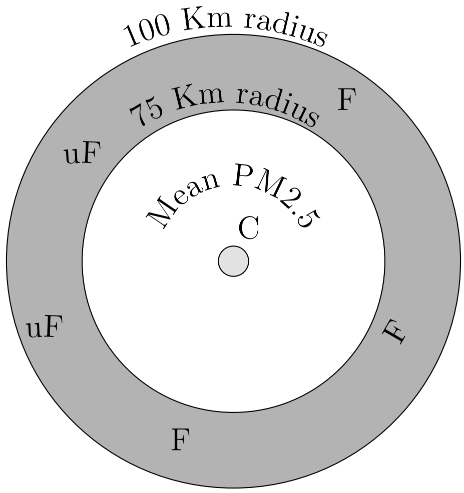

Figure 3: Identification Strategy: Center (smallest grey circle) represents the cluster location, White circle corresponds to 75km radius circle around the cluster location, grey ring area (donut shape) corresponds to area between two circles (75 and 100 Km radii circles) with cluster location as the center. Mean pollution level is calculated for the white circle, we call this *local* pollution level for cluster C.
Local pollution level is instrumented using *upwind non-local* biomass burning events which take place in the grey ring area (only uF). Probability weighted counts used everywhere.

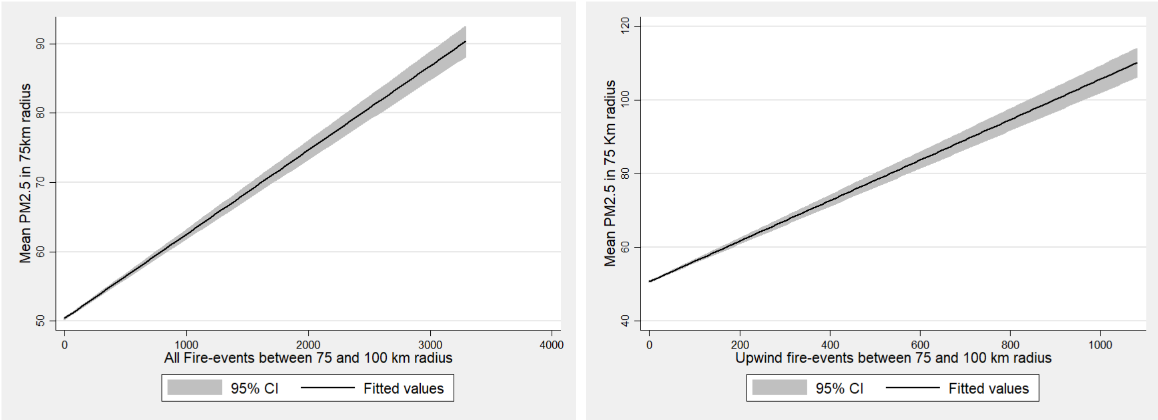

Figure 4: Linear fit plot between Mean PM2.5 (in 75 km radius), Total number of fire-events between 75 and 100 km radius(Non-local fire-events) & Total number of upwind fire-events between 75 and 100 km radius (Non-local upwind fire-events). Unit of observation is a child, shaded area is 95% confidence interval.

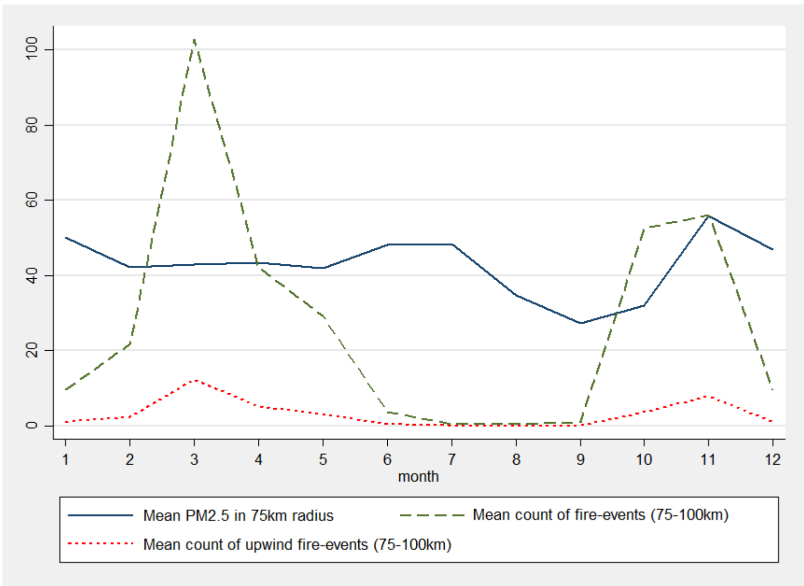

Figure 5: Mean PM2.5, Mean count of all fire-events (in 75-100 km radius) & Mean count of upwind fire-events (in 75-100 km radius) for each month in every year from 2010 to 2016. Figure represents mean over all sampled clusters belonging to all states of India.

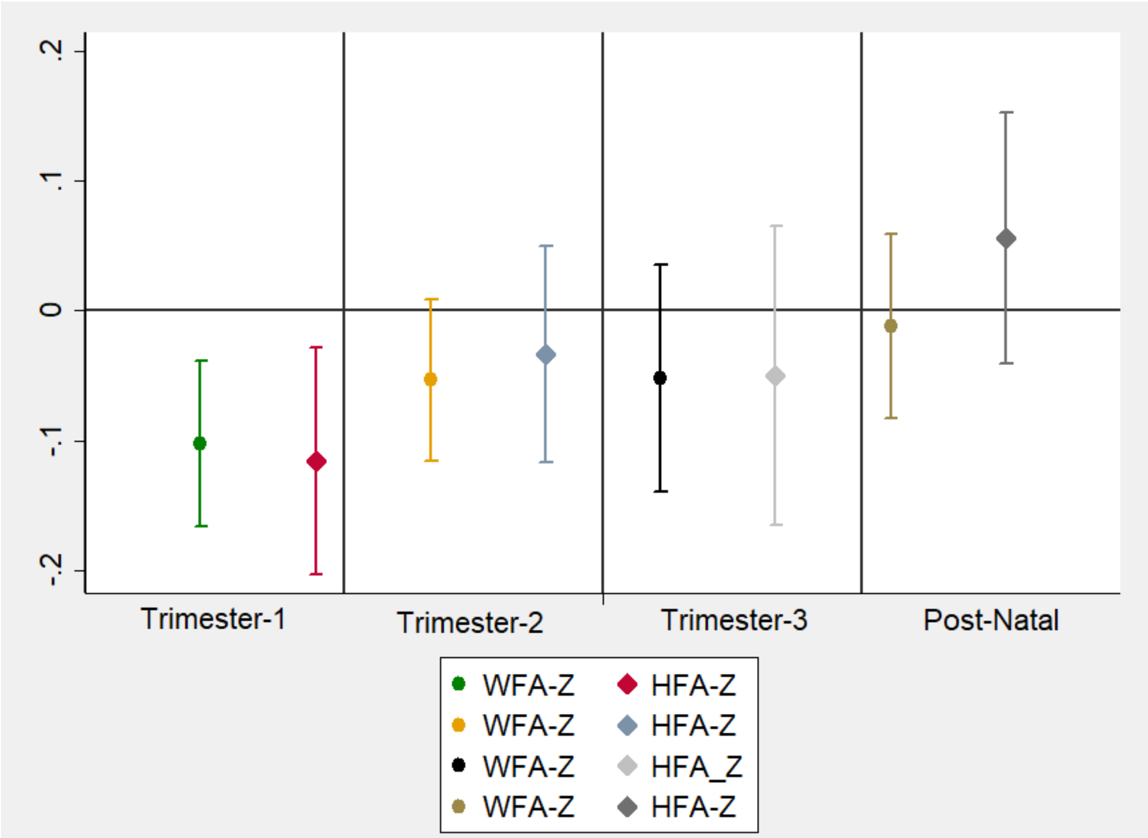

Figure 6: Coefficient of 2SLS regression of outcomes(HFA-Z and WFA-Z) on outdoor air pollution for different critical windows of development of a child. Vertical lines represent 95 percent confidence intervals.

Table 1: Summary Statistics

	All India		**NorthIndia**		**SouthIndia**	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Outcomes						
Height-for-age z score	-1.46	1.68	-1.52	1.68	-1.27	1.66
Weight-for-age z score	-1.52	1.22	-1.54	1.22	-1.44	1.22
Weight-for-height z score	-0.97	1.38	-0.96	1.38	-1.01	1.41
Child characteristics						
Dummy for male child	0.52	0.50	0.52	0.50	0.51	0.50
Birth-order	2.26	1.46	2.38	1.54	1.88	1.05
Childage in months	29.00	16.55	29.02	16.58	28.94	16.46
Pregnancy duration	9.02	0.48	9.00	0.46	9.10	0.52
Mother's characteristics						
Mother's age at birth	24.50	4.91	24.67	5.00	23.91	4.56
Mother's number of education years	6.26	5.16	5.81	5.18	7.81	4.77
Household characteristics						
Rural	0.76	0.43	0.77	0.42	0.70	0.46
Dummy for head of household being a male	0.88	0.33	0.88	0.33	0.89	0.32
Age of head of household	44.59	15.18	44.51	15.21	44.88	15.09
Dummy for source of water: Pipedwater	0.23	0.42	0.23	0.42	0.24	0.43
Dummy for using clean cooking fuel	0.28	0.45	0.26	0.44	0.37	0.48
Dummy for household practicing open defecation (OD)	0.42	0.49	0.41	0.49	0.44	0.50
Fraction of HHs practicing OD in a village	0.43	0.35	0.43	0.35	0.45	0.33
Has electricity connection	0.85	0.36	0.82	0.38	0.94	0.23
Religion = Hindu	0.73	0.44	0.70	0.46	0.83	0.38
Caste = SC or ST	0.37	0.48	0.37	0.44	0.39	0.49
Household size	6.57	2.87	6.75	2.93	5.99	2.57
Mean pollution in 75 km radius						
1st Trimester	54.06	31.87	59.33	32.90	35.96	19.07
2nd Trimester	52.23	29.84	56.98	30.90	35.90	18.02
3rd Trimester	53.30	32.40	58.52	33.68	35.32	18.47
Post-natal (3 months after birth)	53.32	35.38	58.46	36.89	35.44	21.46
Observations	1,81,361		1,40,476		40,885	

Table 2: Instrumental variable regression of outcomes on PM2.5 (z-score)

	OLS		IV: Upwind fire-events between 75 to 100 kms radius	
	(1)	(2)	(3)	(4)
	WFA-Z	**HFA-Z**	**WFA-Z**	**HFA-Z**
Trimester-1: Mean PM2.5 in 75km radius (z-score)	-0.0120**	-0.0116	-0.103***	-0.116***
	(0.005)	(0.007)	(0.032)	(0.044)
Mean of Dependent Variable	-1.52	-1.46	-1.52	-1.46
Includes Child, Mother and Household characteristics	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes
Observations	179816	179816	179816	179816

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual OLS or 2SLS regression of HFA-Z or WFA-Z on weighted mean PM2.5 in first trimester (z-score).Regressions include other controls - gender, birth order and age of child, mother's years of education, mother's age at birth and its square, age and gender of household head, dummy for whether household has pipedwater, has clean cooking source and whether household practices open defecation.

Table 3: First-stage regression

	Mean PM2.5 in 75km radius (z-score)
IV: Number of upwind fire events between 75 and 100km radius (z-score)	0.105***
	(0.003)
First Stage F-stat	863
rk LM statistic	1005
Anderson Rubin wald statistic (p-value)	0.0016
Stock & Yoko critical values:	
10 %	16.38
25 %	5.53
Observations	179816
Includes other controls from 2nd stage	Yes
Includes FEs for Month, Year & Cluster	Yes
Includes FEs for Month*Year & District*Year	Yes

Note: Each coefficient corresponds to an individual FIRST stage 2SLS regression of HFA-Z or WFA-Z on variables mentioned in the first column. Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. PM2.5 & Fire-events variables have all been converted into z-scores. Regressions include controls which are same as those mentioned in Table 2 notes.

Table 4: IV Validity

Identifying Instruments	(1) Mother's Education	(2) Source of water is Pipedwater	(3) Uses clean cooking fuel	(4) Fraction of HHs who OD in the cluster	(5) Birth Order	(6) Asset Ownership
Upwind fire events between 75 and 100km radius in 1st Trimester (Z)	0.016	-0.000	0.000	-0.000	0.002	-0.002
	(0.012)	(0.001)	(0.001)	(0.000)	(0.003)	(0.003)
Observations	179816	179816	179816	179816	179816	179816
Includes Child, Mother & HH characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes	Yes	Yes

Identifying Instruments	(7) Pregnancy Duration	(8) Religion is Hindu	(9) Caste is SC or ST	(10) Poor	(11) Vaccination	(12) Household Size
Upwind fire events between 75 and 100km radius in 1st Trimester (Z)	0.002*	0.000	0.0009	0.001	0.0007	0.001
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.006)
Observations	179816	179816	179816	179816	179816	179816
Includes Child, Mother & HH characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Fire-events have been converted to z-scores.

Table 5: Instrumental variable effects:

Impact of weighted PM2.5 in 2nd Trimester to Post-natal period

	(1)	(2)
	WFA-Z	**HFA-Z**
Trimester-2: Mean PM2.5 in 75km radius (Z)	-0.05*	-0.03
	(0.03)	(0.04)
Observations	184183	184183
Trimester-3: Mean PM2.5 in 75km radius (Z)	-0.05	-0.05
	(0.04)	(0.05)
Observations	172917	172917
Post-natal: Mean PM2.5 in 75km radius (Z)	-0.01	0.05
	(0.03)	(0.04)
Observations	190717	190717
Includes Child, Mother & HH characteristics	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual 2SLS regression of HFA-Z or WFA-Z on weighted mean PM2.5 in first trimester (z-score). Regressions include controls which are same as those mentioned in Table 2 notes.

Table 6: Robustness Checks: Extended controls

	IV: Upwind fire-events b/w 75 to 100 km radius			
	(1)	(2)	(3)	(4)
	WFA-Z	HFA-Z	WFA-Z	HFA-Z
Trimester-1: Mean PM2.5 in 75km radius (Z)	-0.111***	-0.125***	-0.103***	-0.116***
	(0.0328)	(0.0453)	(0.0323)	(0.0448)
Mean rainfall in 75km radius	-0.0463**	-0.0492		
	(0.0217)	(0.0300)		
Mean temperature in 75 km radius	0.00336**	0.00348		
	(0.00151)	(0.00213)		
Gestational period			0.141***	0.172***
			(0.0252)	(0.0347)
Household Size			0.00549***	0.00304
			(0.00143)	(0.00194)
Caste is SC or ST			-0.131***	-0.153***
			(0.00877)	(0.0118)
Observations	179459	179459	178718	178718
Includes Child, Mother & HH characteristics	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual 2SLS regression of HFA-Z or WFA-Z on weighted mean PM2.5 in first trimester (z-score). Regressions include controls which are same as those mentioned in Table 2 notes.

Table 7: Sensitivity Analysis

	IV: Upwind Fire events b/w 50 and 100 Km radius		IV: Upwind Fire events b/w 50 and 75 Km radius		IV: Upwind Fire events b/w 75 and 100 Km radius Dropping Punjab	
	(1)	(2)	(3)	(4)	(5)	(6)
	WFA-Z	**HFA-Z**	**WFA-Z**	**HFA-Z**	**WFA-Z**	**HFA-Z**
Trimester-1: Mean PM2.5 in 50km radius (Z)	-0.101***	-0.0867*	-0.111***	-0.0770		
	(0.0331)	(0.0461)	(0.0349)	(0.0481)		
Trimester-1: Mean PM2.5 in 75km radius (Z)					-0.0910***	-0.121***
					(0.0285)	(0.0413)
Observations	164462	164462	164462	164462	175141	175141
Includes Child, Mother & HH characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual 2SLS regression of HFA-Z or WFA-Z on weighted mean PM2.5 in first trimester (z-score). Regressions include controls which are same as those mentioned in Table 2 notes.

Table 8: Falsification Tests

	IV: Downwind Fire-events in 75 to 100km radius		IV: Upwind Fire-events in 75 to 100 km radius (shuffled location)	
	(1)	(2)	(3)	(4)
	WFA-Z	**HFA-Z**	**WFA-Z**	**HFA-Z**
Trimester-1: Mean PM2.5 in 75km radius (Z)	-0.008	0.003	0.033	0.054
	(0.03)	(0.04)	(0.033)	(0.048)
Observations	179816	179816	179539	179539
Includes Child, Mother & HH characteristics	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual 2SLS regression of HFA-Z or WFA-Z on weighted mean PM2.5 in first trimester (z-score). Regressions include controls which are same as those mentioned in Table 2 notes.

Table 9: Mother's conception behaviour

	(1)	(2)
	Dummy for successful conception	
3 month exposure to upwind fire-events in 75-100kms	0.000307*** (0.000049)	
Exposure to upwind fire-events in 75-100kms in the month of conception		0.000151*** (0.000042)
Number of Unique Mothers	144833	144833
Includes Child, Mother & HH characteristics	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual OLS regression of log number of conceptions on controls mentioned in the table. Regressions include other controls for mother's education, father's literacy level, characteristics of household head and wealth index of the household.

Table 10: Heteregeneity: By Background Characteristics

	IV regression: Height-for-age Z score					
					Mother's Education	
	Poor	Rich	North	South	Primary or less	Secondary or higher
	(1)	(2)	(3)	(4)	(5)	(6)
Trimester-1: Mean PM2.5 in 75km radius (Z)	-0.139**	-0.0824	-0.120***	-0.0226	-0.176**	-0.0775
	(0.059)	(0.063)	(0.043)	(1.813)	(0.081)	(0.056)
Observations	84458	88932	139656	40159	76965	95642
Includes Child, Mother & HH characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual 2SLS regression of HFA-Z on weighted mean PM2.5 in first trimester (z-score). Regressions include controls which are same as those mentioned in Table 2 notes.

North Indian states: Arunachal Pradesh, Assam, Bihar, Chandigarh, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Madhya Pradesh, Manipur, Meghalaya, Mizoram Nagaland, Delhi, Punjab, Rajasthan, Sikkim, Tripura, Uttar Pradesh and Uttarakhand.

South Indian states: Andhra Pradesh, Karnataka, Kerala, Maharashtra, Chhattisgarh, Odisha, Telangana, West Bengal, Lakshwadeep Islands, Andaman and Nicobar Islands, Dadar and Nagar Haveli, Daman and Diu, Puducherry and Goa.

Table 11: Heteregeneity: By Age

	IV regression: Height-for-age Z score		
	Age of child		
	0 to 1 years	1 to 2 years	3+ years
	(1)	(2)	(3)
Trimester-1: Mean PM2.5 in 75km radius (Z)	-0.0564	-0.719**	-0.110*
	(0.349)	(0.338)	(0.0588)
Observations	30456	28669	100701
Includes Child, Mother & HH characteristics	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Each coefficient corresponds to an individual 2SLS regression of HFA-Z or WFA-Z on weighted mean PM2.5 in first trimester (z-score). Regressions include controls which are same as those mentioned in Table 2 notes.

Appendix

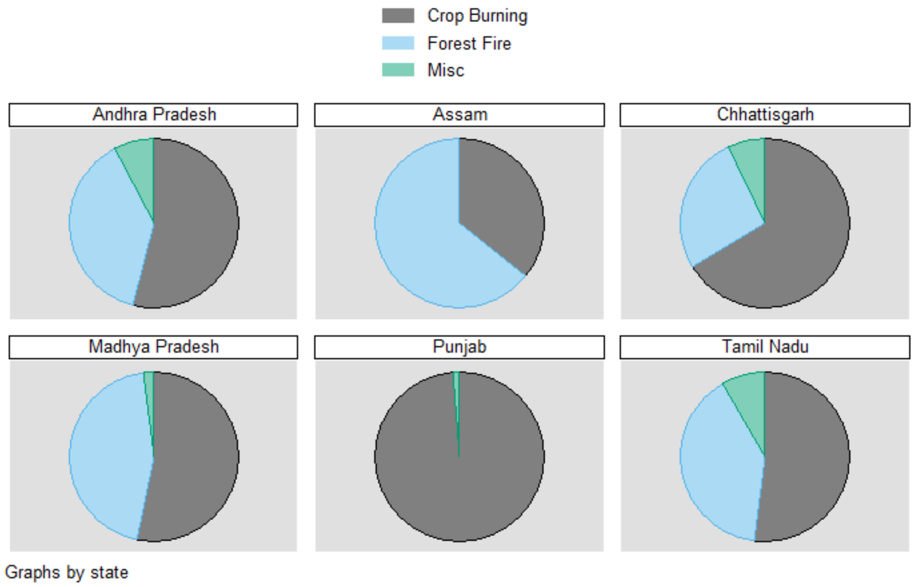

Figure 1: Population weighted split of all biomass burning events which took place from 2010-2016 for select states.

Table A1: Missing PM2.5 analysis

	(1) **HFA-Z**	(2) **Male Child**	(3) **Child's Age**	(4) **Mother's Education**	(5) **Source of water: Pipedwater**	(6) **Uses clean cooking fuel**	(7) **HH Openly Defecates**	(8) **Poor**
Dummy for missing PM2.5 information	-0.003	-0.003	-0.004	-0.017	0.001	-0.002	-0.002	0.002
	(0.014)	(0.004)	(0.004)	(0.032)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	223150	223150	223150	223150	223150	223150	223150	223150
Includes Child, Mother & HH characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month, Year & Cluster	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster.

Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$.

Table A2: Full Model : IV Regression

	(1) WFA-Z	(2) HFA-Z
Trimester-1: Mean PM2.5 in 75km radius (z-score)	-0.103***	-0.116***
	(0.032)	(0.045)
Child is Male	-0.036***	-0.099***
	(0.006)	(0.008)
Birth Order	-0.046***	-0.062***
	(0.003)	(0.004)
Child's Age in Months	-0.082***	-0.106***
	(0.008)	(0.010)
Mother's Education (in years)	0.027***	0.031***
	(0.001)	(0.001)
Household Head's Age	0.002***	0.002***
	(0.000)	(0.000)
Houehold Head is Male	0.011	0.031**
	(0.010)	(0.014)
Mother's Age at Birth	0.048***	0.058***
	(0.005)	(0.007)
Mother's Age at Birth square	-0.001***	-0.001***
	(0.000)	(0.000)
Source of water is Pipedwater	0.016	0.005
	(0.011)	(0.015)
Uses clean cooking fuel	0.134***	0.140***
	(0.010)	(0.013)
Household Defecates in Open (OD)	-0.136***	-0.166***
	(0.009)	(0.013)
Observations	179816	179816
Includes FEs for Month, Year & Cluster	Yes	Yes
Includes FEs for Month*Year & District*Year	Yes	Yes

Note: Standard errors in parentheses are clustered by DHS cluster.
Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$.

QUALITY.

INDEPENDENCE.

IMPACT.

BROOKINGS INDIA