# Pollution and Human Capital: Evidence from India*

Magda Tsaneva† and Uttara Balakrishnan‡

January 6, 2020

**Abstract**

Health and well-being during childhood are vital for shaping human capital accumulation. In India, exposure to pollution is increasingly one of the greatest public health challenges facing the country. In this context, we examine the impact of air pollution exposure in India on children's learning outcomes. Using a large scale panel of 5-16 year olds' test scores, we causally estimate the short- and medium-run impacts of air pollution on human capital, using thermal inversions as an instrument for air pollution. We show that high levels of contemporaneous air pollution significantly reduce varying levels of reading outcomes by 0.57 - 1.67 percentage points and math outcomes by 1.37 - 1.43 percentage points, with girls and older children witnessing a larger decline. There is significant inter-state variation with most of the results being driven by the highly polluted northern and eastern states of India. We find that the main mechanism explaining these impacts is a physiological one. During periods of high air pollution, student attendance in school is lower suggesting health effects of pollution. Further, the cumulative effect of pollution in the past two to three years is larger than contemporaneous impacts. Our results suggest that air pollution is a significant threat to human capital in India.

**JEL codes:** O12, O13, I38, J22, J31, R14

**Keywords:** air pollution, human capital, thermal inversions, physiology; India

*We thank ASER India for sharing their data. We also use data from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India, Peru, and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID). The views expressed here are those of the author(s). They are not necessarily those of ASER, Young Lives, the University of Oxford, DFID, or other funders. For comments and suggestions, we would like to thank Kartik Misra. All errors are our own.

†Department of Economics, Clark University; mtsaneva@clark.edu

‡IMPAQ International; uttara.balakrishnan@gmail.com

# 1 Introduction

Does air pollution affect learning outcomes for school aged-children in India? This question is of particular importance given increased global warming and lack of access to clean fuels for most people in India, on the one hand (WHO, 2008), and the steady decline of education productivity in India on the other (ASER, 2018). The effects of air pollution are wide ranging from increased mortality and lower productivity to weaker ecosystems and costly remediation (Maas and Grennfelt (2017)).[1] Worldwide, around 3 million people died prematurely from outdoor air pollution in 2010. For extremely polluted countries like India and China, with large populations, premature deaths due to pollution are expected to be around 20 million by 2060 (OECD, 2016). There is extensive literature documenting the impact of air pollution on a wide range of outcomes including worker productivity; life expectancy; illness; dementia; and child health. However, the evidence on how air pollution impacts human capital, especially in developing countries, is scarce. Further, all the evidence is limited to low and moderately polluted areas like the U.S., Singapore; and Israel (Ebenstein et al. (2016); Zweig et al. (2009), Stafford and Brain (2015); and Chew et al. (2018)). In this context, our paper fills this knowledge gap by estimating the causal impact of air pollution on human capital production for school-aged children in India.

In this paper, we examine the short- and medium-run effects of air pollution on learning outcomes of almost 3.5 million children in primary and secondary schools in India, from 2007 - 2014. The Indian context is important for two main reasons. First, India has one of worst levels of air pollution around the world. Pollution in India has been consistently above the World Health Organization (WHO) recommended levels for the past decade (see Figure 1).[2] Premature deaths due to poor quality are to the tune of 2.1 billion life years (Greenstone and Hanna (2014)). According to the World Bank (2016), the economic cost of air pollution in terms of poor health and higher mortality is almost 8 percent of GDP for India. This is compounded by weak environmental regulations (Duflo, Greenstone, Pande, et al. (2013)) and stiff resistance and lobbying by industry for any move to strengthen them (Bloom et al. (2010)). The second important reason to examine the Indian context, is India's large learning deficit. According to the 2005 ASER report, almost 44 percent of primary school-aged children enrolled in school could not read a basic paragraph and 50 percent could not do simple subtraction (Chakraborty and Jayaraman (2019)). Nationwide productivity of education has declined by 18 percent between 2008 and 2018 (ASER, 2018). While India has been successful in getting children into school (national enrollment rate in 2018 was 96 percent), large sections of children are lacking basic skills. In this context, exposure to high levels of air pollution has the potential to significantly weaken human capital accumulation and future capabilities of these children.

[1]There are two main types of air pollution - ambient air pollution (outdoor pollution) and household (or indoor) air pollution. Outdoor air pollution originates from natural (e.g. dust storms) and anthropogenic sources (e.g. fuel combustion, heat and power generation, industrial facilities, residential cooking). Indoor air pollution refers to pollution generated by household combustion of fuels (caused by burning fuel such as coal, wood or kerosene) using open fires or basic stoves in poorly ventilated spaces. Both indoor and outdoor air pollution can contribute to each other, as air moves from inside buildings to the outside, and vice versa. In this paper we focus on ambient air pollution.

[2]WHO recommended safe air quality guidelines are $10\mu g/m^3$ for maximum yearly average and $25\mu g/m^3$ for maximum 24-hour average.

Estimating the causal impact of air pollution on learning outcomes is not straightforward for two main reasons as highlighted by Hansen-Lewis (2018). First, there is lack of consistent nation-wide data on air pollution. Air pollution in India is monitored by the Central Pollution Control Board (CPCB). As of 2016, air pollution monitors were limited to 342 operating stations in 127 cities across India (Van Donkelaar et al. (2014)). Further, monitors are clustered around bigger cities like Delhi and state capitals. Thus, ground level monitoring of air pollution is not widely and evenly distributed. Second, even if data on pollution was widely available, there is a simultaneity problem in estimating the relationship between pollution and learning outcomes. This is because changes in both, pollution and human capital may simply reflect overall economic conditions. For instance, increased economic growth will increase pollution through higher vehicular and industrial emissions (Hettige et al. (2000); Oates (2006); Myers and Kent (2004); Shafik (1994)), but can also increase learning outcomes through higher incomes (Psacharopoulos (1985); Goldin and Katz (2007); Goldberg and Pavcnik (2007)).

To overcome these concerns, in this study, we examine the impact of air pollution on learning outcomes in an instrumental variables (IV) framework. Specifically, we use thermal inversions to instrument for air quality. Ordinarily, temperature drops with rising altitude. However, a thermal inversion occurs when there is a mass of hot air on top of a mass of cold air. Thermal inversions occur exogenously and trap pollutants. A growing literature uses inversions to instrument for pollution (Arceo et al. (2016); Hicks et al. (2016); Jans et al. (2018); Chen, Oliva, et al. (2017); Sager (2019); and Molina (2016)). Additionally, to overcome the concern of inconsistent and uneven pollution monitoring, we employ satellite data to measure air pollution. Specifically, we use air quality measures from the Moderate Resolution Imaging Spectroradiometer (MODIS) Aerosol Optical Depth (AOD) to proxy for particulate matter (Dey, Di Girolamo, et al. (2012); Hansen-Lewis (2018); Gendron-Carrier et al. (2018)). Van Donkelaar et al. (2014) estimates that a one unit increase in mean AOD, results in an average change in mean PM2.5 by 108 $\mu$g/m$^3$. Data for AOD is available, at fine spatial resolution, for the entire time period of the study, 2007 - 2014. We show that thermal inversions are strongly positively correlated with our measure of pollution, AOD. Further, consistent with the simultaneity bias between human capital and pollution, the IV estimates display more negative impacts on human capital than OLS estimates. Our main outcome of interest is reading and math outcomes for children aged 5-16 years which comes from a large repeated cross-sectional dataset from India, the Annual Status of Education Report (ASER). The ASER survey has been conducted annually from 2005 onwards and has a wide geographic coverage, covering almost all rural districts in India, and surveying over 4,00,000 school-aged children annually. ASER administers learning assessments of basic literacy (reading skills) and numeracy (number recognition and arithmetic skills) to all children aged 5-16 years.

There are four main findings of this study. First, we show that both reading and math outcomes are impacted adversely by air pollution. A 0.01 unit increase in annual AOD (one $\mu$g/m$^3$ in PM2.5) reduces reading outcomes by 0.57 - 1.67 percentage points. Similarly, math outcomes decline by 1.37 - 1.43 percentage points. Even basic skills like probability of reading a letter and probability of reading a word for reading and probability of counting from 10 - 99 for math, are impacted. We corroborate these findings using the Young Lives Survey (YLS) in India - a rich 16-year longitudinal study conducted in Andhra Pradesh. Estimates from the YLS data are larger in magnitude indicating stronger effects. To put these effects into perspective, Shah and Steinberg (2017) analyze the cognitive effects of exposure to drought in utero from India and find that exposure to drought in utero is associated with being 2 percentage points less likely to recognize numbers in childhood.

Second, we examine heterogeneous impacts by state and find almost all the impacts are driven by the north and eastern states of India namely Punjab, Haryana, Uttar Pradesh, Bihar, West Bengal, and Rajasthan. Since the north and eastern states of India are land-locked, poor quality air gets trapped, leading to extremely high particulate matter levels. To compound matters, these states witness high levels of crop stubble burning after the rice and wheat sowing season. We also examine heterogeneous impacts by age and gender. Girls witness a larger decline than boys for both math and reading outcomes in the ASER data. The same pattern is also evident in the YLS data. This is unsurprising since girls have lower reading and math scores to begin with. Worryingly we find that girls in households with a larger number of siblings are worse off. On the other hand, boys in smaller and larger households fare about the same. With respect to age, older children in general, have larger declines due to poor air quality. One plausible explanation for this is that older children in rural areas are more exposed to polluted air, while attending school *and* plausibly while working on or off the farm (Shah and Steinberg, 2019).

Third, there are large cumulative effects of poor air quality on human capital. In general, the effect becomes larger as the duration of exposure to air pollution increases. Specifically, a 0.01 unit cumulative change in AOD (one $\mu g/m^3$ in PM2.5) over the last two years leads to a decline in reading and math outcomes by 0.71 - 1.99 percentage points, while exposure for 3 years leads to a decline in reading and math outcomes by 0.79 - 2.73 percentage points.

Finally, we find that the main mechanism driving the impacts is a physiological one. Changes in learning outcomes could arise both, from pollution directly affecting children's brain development or from changes in health which could impact attendance and drop out rates. We find that attendance on a random day is lower in districts with high levels of current year AOD. We also find that children are less likely to be "on-track" (correct grade for their age) and are more susceptible to falling behind. We do not find any evidence of air pollution impacting drop out behavior. We corroborate the findings using attendance information from several rounds of the National Sample Survey (NSS) data. Additionally, we also find evidence that the physiology of teachers also matters. We explicitly test the impact of AOD on teacher attendance and find that cumulative exposure to pollution leads to marginally higher teacher absences. Thus, pollution is doubly deterimental for children, both through impacts on their own health, and through impacts on the health of their teachers.

We also run a series of robustness checks that ascertain the validity of our main findings. First, we examine whether our results are robust to estimating impacts only for 'on-track' children i.e. children who were the right grade-for-age. Second, we test the robustness of our results using alternate temperature layers to define a thermal inversion. Specifically, instead of using the two pressure levels closest to the ground (1000hPa and 925hPa) to identify inversion episodes, we use the temperature difference between the second and third pressure levels i.e. 925hPa and 850hPa to identify inversion episodes. Third, since the ASER tests are administered between September to November in different villages across the country, we test if exposure in current year, defined as the 12 months prior to the test, impacts learning outcomes. Specifically, we define current year as September - September, as opposed to a calendar year. Fourth, to ensure that the inclusion of state-specific linear time trends are not biasing our results, we estimate impacts omitting these trends. Finally, we test the robustness of our main results by including season specific fixed effects in addition to survey year fixed effects. Specifically, we include indicators for each of the four seasons.

The burden of air pollution is not shared equally. Poor and marginalized communities around the world, including in India, face the highest burden of air pollution (Lipfert (2004); Dasgupta et al. (2003); Rozelle et al. (1997); Duraiappah (1998)).[3] It is also children from these very communities who tend to have the worst learning rates and highest chances of dropouts. High levels of air pollution, thus, impose a double penalty on the poor. First, air pollution leads to worse short-run health outcomes for the poor both, because of higher exposure levels due to agricultural and informal work which happens outdoors, and because of the inability to access proper health care or mitigating equipment like home air purifiers (Houston et al. (2004); Gwynn and Thurston (2001); Osseiran and Chriscaden (2016); Reardon and Vosti (1995)). Second, as shown in this study, air pollution also impacts human capital accumulation of children. The largest impacts in this study come from some of the poorest states in India of Uttar Pradesh, Bihar, and West Bengal. Lack of proper coping strategies, makes poorer children even worse off. The results in this paper imply that impacts on human capital disadvantage the poor even in the long run because by impacting human capital accumulation, pollution impacts long-run capabilities (Sen, 1999) and restricts upward mobility.

This paper contributes to several strands of literature. First, this paper presents the first large-scale evidence on the impact of air pollution on human capital production in India. Existing evidence on the impacts of air pollution in cognitive performance is nascent and comes mostly from low and moderately polluted areas like the U.S., Singapore; and Israel (Ebenstein et al. (2016); Zweig et al. (2009), Stafford and Brain (2015); and Chew et al. (2018)). Cognitive impacts of air pollution, on the scale being witnessed in India have not been examined previously. Importantly, these prior studies examining the short-run effects of pollution assume that the variations in air quality are exogenous. However, as outlined above, this might not necessarily be the case. We use an instrumental variables approach in this paper to address the simultaneity bias arising from human capital and pollution.

Second, we contribute to the broader literature examining the economic impacts of air pollution in various domains including physical and mental health and firm and worker productivity. Studies examining the impact of particulate matter on worker productivity include Graff Zivin and Neidell (2012); Adhvaryu et al. (2014); and Fu et al. (2017). Studies on air pollution and health include: Currie, Zivin, et al. (2014); Borja-Aburto et al. (1998); Loomis et al. (1999); Foster et al. (2009); and Tanaka et al. (2012). Other literature includes impacts on life expectancy (Chay and Greenstone

[3] According to the World Data Lab (2018), until recently India had the largest number of people living in extreme poverty. This co-exists with India having 14 of the 20 most polluted cities in the world.

(2003)); illness (Cohen et al. (2005)); and dementia (Chen, Kwong, et al. (2017), Cacciottolo et al. (2017)). There is also a limited literature examining the impact of air pollution on cognitive outcomes from countries with low to moderate pollution levels like U.S., Singapore, and Israel (Ebenstein et al. (2016); Zweig et al. (2009), Stafford and Brain (2015); and Chew et al. (2018)). Studies examining the impact of air pollution on human capital from developing countries is limited. In one of the few studies in this domain, Zhang et al. (2018) examines the impact of air pollution in China on the cognitive functioning amongst older adults. Our paper adds to this literature by examining the human capital costs of air pollution in India, a developing country, facing a massive challenge with respect to air pollution.

Finally, we add to the literature examining the impacts of air pollution in India. Greenstone and Hanna (2014) examine air pollution regulations in India and find that while air pollution regulations improve air quality, they have only a modest impact on infant mortality. Pullabhotla (2018) finds that fires arising due to agricultural activity in India lead to a higher incidence of neonatal mortality and infant mortality relative to the average mortality rates, resulting in 96,000 additional under-five deaths per year.[4] Hansen-Lewis (2018) examines the impact of air pollution on manufacturing productivity and finds that air pollution substantially lowered productivity among industries with labor intensive technologies, but had little average effect across all industries. Other studies have examined the impact of indoor air pollution in India. Krishnamoorthy et al. (2018) find that indoor air pollution significantly impacts cognitive functioning of adults in their sample from South India. Duflo, Greenstone, and Hanna (2008) find that providing households with improved cooking stoves only had short-run impacts on smoke inhalation. None of these studies examine ambient air quality and its impact on human capital outcomes for children in rural India.

The rest of the paper is organized as follows. Section 2 outlines the context and Section 3, describes the data used in this paper. Section 4 outlines our main methodology and Section 5 discusses the main findings. Section 6 explores potential mechanisms. Finally, in Section 7 we provide concluding remarks.

# 2 Context

## 2.1 Air Pollution and Learning Outcomes

High levels of exposure to air pollution is increasingly becoming an epidemic in both, developed and developing counties. Ambient levels of air pollution in many cities exceed more than 10 to 15 times the clean air guidelines laid down by the WHO (Kilian and Kitazawa (2018)). Major sources of air pollution around the world include, traffic and industrial emissions, agriculture, and burning of fuels for cooking and heating (Craig et al. (2008); Karagulian et al. (2015)). It is estimated that more than 92 percent of people across the world live in areas exceeding the WHO recommended safe air quality guidelines of $10\mu g/m^3$ for maximum yearly average and $25\mu g/m^3$ for maximum 24-hour average.

[4] Previous estimates, based on studies from developing countries, underestimate the true mortality cost of pollution exposure (Lelieveld et al. (2018)).

Air quality in developing countries is much worse, with India and China frequently experiencing annual average PM2.5 levels over $120\mu g/m^3$ (Van Donkelaar et al. (2014)).[5] Globally, the WHO has attributed over 3 million premature deaths to poor air quality in 2012.

While there is a vast literature in economics that studies the relationship between air pollution exposure and human health, similar studies from economics on the effects of pollution on learning outcomes are more limited.[6] In contrast, there are many epidemiological studies that examine the impacts of pollution on human capital. For instance, Porta et al. (2016) and Suades-Gonzalez et al. (2015) examine in-utero exposure to pollution on later life cognitive outcomes. A combination of vehicular emissions, NO2 exposure, and PM2.5 exposure are commonly associated with decreased cognitive ability of children if exposed in-utero. Another study from Taiwan found that SO2 exposure during the second and third trimester (when the brain develops) is associated with impaired gross motor skills among infants (Gou et al. (2016)). Other studies examining in-utero exposure to pollution on cognitive impairment include (Molina, 2016); (Tang et al., 2008); and (Harris et al., 2016). Studies have also examined impacts of air pollution exposure in childhood on cognitive functioning. A longitudinal study from Spain found that exposure to high levels of NO2 impaired the gross motor skills of 5-year old children (Freire et al., 2010). Several other studies from developed countries examine the impact of exposure to air pollution during childhood including Jedrychowski et al. (2015); and Kicinski et al. (2015). There is limited evidence from developing countries apart from two studies in Mexico and China. A study from Quanzhou province in China found that 8-10-year-old children from heavily polluted areas had increased risk of coordination problems (Wang et al., 2009). In another study, children in Mexico city were compared with children living in a rural province on Mexico. Children in Mexico city which is heavily polluted performed worse on multiple tests of IQ and vocabulary (Calderón-Garcidueñas and Torres-Jardón, 2012).

## 2.2 Air Pollution in India

India is one of the fastest growing economies in the world and aims to reach the five-trillion-dollar mark by 2025. However, according to the World Bank, the country lost over 8.5 percent of its GDP in 2013 due to air pollution (World Bank, 2016). Air pollution in India is responsible for 12.5 per cent of all deaths in India (Balakrishnan et al. (2014)). Particulate matter exposure levels in India are more than five times that of the United States (Greenstone and Hanna (2014)). Accounting for almost 17 percent of the world's population, such high exposure levels pose a serious threat to the health of both, current and future generations in India.

[5] In November 2019, particulate levels in India's capital city, Delhi, were more than 20 times the WHO maximum (https://www.bbc.com/news/world-asia-india-50258947).

[6] Poor air quality has been found to lead to a heightened risk of heart disease, stroke, and lung cancer (Dockery and Pope (1996); Chay and Greenstone (2003); Arceo et al. (2016); and Deryugina et al. (2016)). It has been found that particulate matter is small enough to penetrate the thoracic region and form deposits in an individuals airways. It can also lead to inflammation of the airways and worsen respiratory conditions like asthma and bronchitis (Pope 3rd et al. (1995); Nel (2005); Ghio and Devlin (2001); Graff Zivin and Neidell (2012)). In addition to impacts on an individual's respiratory system, according to the WHO, pollution also increases the probability of heart attacks, with the exact causal pathway being unclear (WHO, 2006).

Crop burning, vehicular emissions, dust storms due to construction activity, and industrial emissions are the four most common sources of pollution in India (Guttikunda et al. (2014); Bikkina et al. (2019)). Meteorology over the Indo-Gangetic plains plays a strong role in the observed seasonal cycle of air pollution in cities in this region with the winter time highs (due to high inversion) and the summer time lows (due to rains). Since the north and eastern states of India are land-locked, poor quality air gets trapped, leading to extremely high particulate matter levels. The problem is not restricted to one or two cities in India. Apart from Delhi, dozens of other Indian cities are dealing with severe pollution. For instance, out of the twenty most polluted cities in the world, 14 are in India, according to a 2016 WHO study.

Further, pollution is not merely an urban phenomenon in India. Exposure to PM2.5 is roughly equal across urban and rural India and rural areas accounted for 75 percent of air pollution related deaths in 2015 (Group et al. (2018)). Households in rural areas are also impacted by crop stubble burning (in fact, much of urban air pollution originates in rural areas due to agricultural practices like crop burning).[7] Additionally, households in villages also rely largely on biomass burning for heating and cooking which also leads to unhealthy air quality (Bikkina et al. (2019)).

Air pollution is having a particularly devastating impact on children in India. The noxious air hanging over India's towns and cities kills more than 100,000 children under five every year (Varughese et al. (2009)). This is the highest number of deaths due to pollution in this age bracket across all countries. Further, in the age bracket of 5-14 years, India also saw the deaths of 4,360 children in 2016 (Adair-Rohani (2018)). As highlighted above, children are impacted by pollution both in-utero and after birth. In developing countries, this is compounded by the fact that young children spend more time indoors and inhale smoke generated from wood stoves. Thus children are impacted by both outdoor and indoor air pollution.

Studies examining air pollution and its impact on child outcomes in India are scarce. Moreover, most studies have focused on health and mortality impacts. Greenstone and Hanna (2014) examine air pollution regulations in India and find that while air pollution regulations improve air quality, they have only a modest impact on infant mortality. A recent study has found that fires due to crop burning increased neonatal mortality by 9.2% and infant mortality by 7.7%, resulting in 96,000 additional under-five deaths per year (Pullabhotla, 2018).[8] Other studies examining the impact of indoor air pollution in India include, Krishnamoorthy et al. (2018) who find that indoor air pollution significantly impacts cognitive functioning of adults in their sample from South India. None of these studies examine ambient air quality and its impact on human capital outcomes for children in rural India.

[7]The rice and wheat crop grown in the states of Punjab and Haryana are the major culprits of crop fires. At end of the harvest season in October, farmers burn the left-over crop stubble to prepare their land for the next sowing season. This large scale burning of crop stubble leads to crop fires which leads to pollution in large parts of northern India (Pullabhotla (2018); Venkataraman et al. (2006)).

[8]These mortality costs are nearly twice the number compared to previous estimates (Lelieveld et al. (2018)) that apply exposure-response functions based on studies mostly focused on developed countries to estimate mortality costs of pollution exposure.

# 3 Data

## 3.1 Annual Status of Education Report (ASER)

We use student-level data from the Annual Status of Education Report (ASER) - a household-based survey from most rural districts in India. This annual survey began in 2005 and collects information on reading and arithmetic skills for all school-aged children, irrespective of their schooling status on an annual basis. The survey takes place in the middle of the school year - from end of September to end of November, limiting any spatially systematic seasonality in data collection.

The ASER data is administered only in rural areas and is a repeated cross-section, representative at the district level. The ASER surveyors ask each child, in his or her native language, four potential questions in reading and math. Reading comprehension tests taken during the survey show whether the child can read a letter, a word, a paragraph, or a story. The highest level of reading corresponds to grade 2 curriculum. Mathematics tests show whether the child can recognize numbers from 1 to 9, 10 to 99, can do subtraction, or do division, with the highest level of arithmetic corresponding to grade 3 or grade 4 curriculum, depending on the state.

We use all the ASER rounds currently available in the public domain, 2007 - 2014. In each round of the ASER data, the sample size is large, around 4,00,000 observations, implying over 3 million children for the entire study period. ASER is administered at home, on weekends, and thus includes both, children enrolled in school, out of school children, and children who were never enrolled. This allows us to measure effects on test performance without confounding selection related to school attendance or access to schools. Since the ASER is only administered in rural areas, we are unable to use ASER to get estimated of urban pollution on human capital.

Table 1 presents summary statistics for the study sample from 2007 - 2014.[9] We present summary statistics for four levels of reading and math outcomes and individuals characteristics used as controls in the main empirical specification. With respect to reading, around 91 percent children can read a letter in 2007 and this decline to 85 percent by 2014. There is a general decline in all reading levels fro 2007 to 2014. Further, the harder the reading level, the lower the proportion of children who can complete it. For instance, only 45 (43) percent of children could read Grade 2 text in 2007 (2014). The story is almost identical with respect to math outcomes, with all math outcomes witnessing a monotonic decline from 2007 to 2014. Finally, average age across all rounds is 10 years and roughly, half the sample is male.

## 3.2 Young Lives Survey (YLS)

As a robustness check to our main results, we use a rich longitudinal dataset from India. The Young Lives dataset is a longitudinal panel which surveys two cohorts of children (younger cohort born in 2001-02 and older cohort born in 1994-95). Data was collected from children and their families in 2002, 2006, 2009, and in 2013/14. YLS tracks about 12,000 children in four different countries:

[9] We only include states which have data for all years from 2007 - 2014. Consequently, the states of Jammu & Kashmir and Manipur are excluded from the sample because these states were not included in the 2007 ASER.

India, Ethiopia, Peru, and Vietnam. In India the study has been conducted in the state of Andhra Pradesh. The study includes the districts of Cuddapah, Anantapur, Mahbubnagar, Karimnagar, West Godavari, and Srikakulam and also the capital city of Hyderabad. These districts were chosen in a way to cover the different climactic and geographic variations in the state.

Similar to ASER, the YLS data also collects data on cognitive achievement in all rounds of the survey. Unlike the ASER, however, tests often varied form round to round to better reflect the age of the child and school curriculum in the state. Tests in the YLS were also longer in duration than the ASER, with the math test including 30 questions and the reading test having almost 100 questions. In this paper we use data from the younger cohort consisting of approximately 2011 children since we have data on test scores from at least three survey rounds for them (2007, 2009, 2013/14). The main advantage of using the YLS is that the *same* child is tracked in each survey (as opposed to the ASER). This allows us to account for prior human capital accumulation (Garg et al., 2017). The overall rate of sample attrition is low with only about 4 percent of children lost over a seven-year period. We use cognitive outcomes from the math test and the peabody picture vocabulary test (PPVT).[10] Table 2 presents summary statistics from the YLS data for the three rounds used in this study.

### 3.3 Pollution Data

The Central Pollution Control Board (CPCB) routinely monitors air quality in India through its network of air pollution monitors across the country (Dey, Di Girolamo, et al. (2012)). However, the numbers of monitoring stations are too few for a complete and accurate assessment of regional health risks given the very high spatial and seasonal variability of aerosol loading (Dey & DiGirolamo, 2010). Moreover, most of the CPCB sites are concentrated in the urban areas, leaving the large rural population unchecked.

Thus, we use satellite data to measure air quality which has the advantage of universal coverage. Satellite data can be useful for examining global air quality in the absence of a robust database of in-situ PM2.5 (e.g. Van Donkelaar et al. (2014); Xin et al. (2014)). Following Hansen-Lewis (2018), we proxy air pollution with Aerosol Optical Depth (AOD) using data from Moderate Resolution Imaging Spectroradiometer (MODIS). The MODIS aboard the Terra and Aqua satellites measure AOD twice a day and the data have a spatial resolution of approximately 10 x 10 kilometers. AOD is measured on a log scale of 0 to 5 and measures the fraction of incoming light reflected by the air column before reaching the ground. The resulting estimates have been shown to be good predictors of particulate matter (PM) of different sizes (Chu et al. (2003); Gupta et al. (2006); Kumar et al. (2007)).

As highlighted in Dey, Di Girolamo, et al. (2012), AOD differs from traditional pollution monitoring data since AOD only measures suspended particulates, as opposed to monitoring data which collect data on a variety of pollutants including ozone or sulfur and nitrogen oxides. Prior research

[10] The PPVT is a test for assessing receptive vocabulary. The test requires respondents to select the pictures that best represent the meaning of a series of stimulus words read out by the examiner.

examining the impact of air pollution on economic outcomes in developing countries has used AOD as a proxy for pollution (Foster et al. (2009); Greenstone, Nilekani, et al. (2015); Gendron-Carrier et al. (2018)). Van Donkelaar et al. (2014) examine the extent to which AOD correlates with PM2.5 levels in India and find that there is a strong correlation between ground level PM2.5 and AOD.[11] Environmental and atmospheric science studies have also extensively used AOD as a measure of air quality (see e.g. Guazzotti et al. (2003); Dey and Di Girolamo (2010); Dey, Di Girolamo, et al. (2012); Ten Hoeve and Jacobson (2012)). Using multiple regression analysis, Shaw and Gorai (2018) found that there was a strong correlation between AOD data from the MODIS Aqua satellite and PM2.5 and PM10 concentrations in different parts of India.

Hansen-Lewis (2018) demonstrates that the AOD data are consistent with the main descriptive features of air pollution in India. Across India, AOD exceeds the clear air level of 0.1. Due to inversions and inland accumulation, air pollution is much worse in the northern and eastern states of Punjab, Haryana, Rajasthan, Uttar Pradesh, Bihar, and West Bengal. Figure 2 examines the trend in AOD from 2002 - 2014. In line with rising particulate matter pollution in India, mean AOD in India has risen from around 0.37 $\mu$g/m$^3$ in 2002 to close to 0.45 $\mu$g/m$^3$ in 2014. For the years included in this study i.e. 2007 - 2014, AOD has largely remained between 0.40 - 0.45, with the mean AOD being 0.41 (Table 3).

### 3.4 Thermal Inversions

Normally, at higher altitudes, temperatures are lower. A thermal inversion occurs when this process reverses and a mass of hot air is present above a mass of cold air (Arceo et al., 2016). Broadly inversions are of three types: (a) radiation inversions, which take place at night when the ground and the air in touch with the ground are cooled faster than air layers located higher above, (b) subsidence inversions, which occur from vertical air movements when a layer of cold air descends through a layer of hot air, and (3) marine inversions which take place when air above the sea, which is cooler than the air above land, flows inland and pushes the warm inland air upward. While thermal inversions by themselves do not pose a health risk, when they are accompanied by high levels of air pollution (either vehicular or through industrial emissions) there can be a temporary accumulation of pollutants. Specifically, inversions can lead to a higher concentration of various pollutants including particulate matter (Jacobson and Jacobson (2002); Arceo et al. (2016)) and studies have shown that strong inversions are associated with the worst reported pollution events in history (Malek et al. (2006); Iacobellis et al. (2009); Bailey et al. (2011)).

We identify thermal inversions in India using reanalysis data from the NCEP/NCAR data, which provides air temperatures at a 2.5 x 2.5 degree grid (roughly 250kms by 250kms).[12] The NCEP/NCAR data provides temperature in 17 layers, defined by air pressure in that layer. We use the temperatures for the two pressure levels closest to the ground available in the NCEP/NCAR data (1000hPa and

[11] (Van Donkelaar et al., 2014) show that a one-unit increase in AOD increases PM2.5 by 108$\mu$g/m$^3$. Thus, a 0.01 unit change in AOD is approximately equal to one $\mu$g/m$^3$.

[12] https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.pressure.html

925hPa) to identify inversion episodes. Such reanalysis data has been supported in the literature as generating a consistent best-estimate of weather in a grid-cell and has been used extensively in economics (Garg et al. 2018; Hansen-Leiws, 2018; Schlenker and Roberts, 2009; Schlenker and Lobell, 2010; Auffhammer et al., 2013). The 1000hPa layer temperature corresponds to the surface conditions and 925hPa layer measure conditions at approximately 600m above sea level. We use the temperature differences between these two layers to identify inversion episodes and inversion strength. During normal conditions (inversions), the temperature decreases (increases) with altitude and hence, the temperature difference between the 925hPa and 1000hPa air layer is negative (positive). While the data is collected four times in a day, in our analysis, like Jans et al., (2014) and Molina (2018) we focus on only nighttime inversions. This is because, given the high variation across space and time, nighttime inversions have a stronger correlation with pollution in the first stage of the IV strategy described below. Further, nighttime inversions are less susceptible to endogenous behavior changes since they are less visible than daytime inversions.

To construct the thermal inversion variable, for each day in a year, we create an indicator equal to 1 if the nighttime temperature in the 1000hPa pressure level is higher than the nighttime temperature in the 925hPa pressure level (where nighttime temperature is observed at 12am). We then average this over each district-year. Table 3 shows the mean and standard deviation for our inversion measure. There are about 70 daily nighttime inversions on average, annually across India. Alternatively, approximately 70 days in a year witness a nighttime inversion episode.[13] The variation is large (standard deviation is 76) from 0 nighttime inversions to some places witnessing 264 nighttime inversions in a year.[14]

### 3.5 Other Weather Data

Similar to previous literature, to improve the precision of our estimates, we employ data on temperature, precipitation, wind velocity, and relative humidity. Data on precipitation and temperature are obtained in 0.5 degree grids from the University of Delaware (Willmott and Matsuura (2015)). Data on surface wind velocity and relative humidity was constructed using the ECMWF ERA-Interim and ECMWF ERA-5, respectively (Dee et al., 2011). Controlling for humidity and wind velocity are important since they can impact both, air pollution levels and health. Table 3 reports descriptive statistics of AOD, thermal inversions, and other weather controls. Our main weather controls in all specifications include: mean temperature, precipitation, wind velocity, and relative humidity; and their squares.

[13] Since we measure inversion at only one point in a day i.e. night, the total number of inversions in a year is equivalent to the total days in a year an inversion episode occurred.

[14] There is wide variation in the literature on the annual daily number of inversion episodes. Arceo et al. (2016) report that average number of inversions (combining inversions in a whole day) in a year are approximately 87.36. Fu et al. (2017) report 157 daily inversions in a year in their sample.

# 4 Empirical Methodology

## 4.1 Ordinary Least Squares (OLS)

To examine the effect of air quality on child learning outcomes, we start by first estimating an OLS specification. The main goal of estimating an OLS specification is to demonstrate that OLS estimates of air quality on human capital may not represent causal impacts. Specifically, we estimate the following OLS regression model:

$$Y_{idt} = \beta_0 + \beta_1 AirQuality_{dt} + \gamma X_{idt} + \delta Z_{dt} + \pi_a + \alpha_t + \eta_d + \eta_{s,t} + \epsilon_{idt} \tag{1}$$

where $Y_{idt}$ is the outcome of interest (e.g. reading or math outcome) for student $i$ in district $d$ in survey year $t$ = 2007 - 2014. $AirQuality_{dt}$ is the ambient air quality defined by AOD in district $d$ in survey year $t$.[15] We include district fixed effects, $\eta_d$ to control for time invariant district level heterogeneity. We also include state-specific linear time trends, $\eta_{s,t}$, to control for unobservable factors correlated with education that vary linearly over time within states. Finally, we control for child age fixed effects, $\pi_a$, to account for cohort-specific effects, and year of interview fixed effects, $\alpha_t$, to account for common trends in education. Individual controls, $X_{idt}$, include an indicator for male, household size, and an indicator for mother having attended school. We also include a vector of time varying district level weather controls, $Z_{dt}$, which include controls for precipitation, temperature, wind velocity, relative humidity, and their squares. Finally, standard errors are clustered at the district level to allow for correlation of the error term within a district.

Given the large inequalities between girls and boys in India, we also present heterogeneous effects for girls and boys separately. In addition, in the standard human capital model, children and parents are forward-looking and view schooling as an investment with financial returns. However, children of different ages have different levels of vulnerability to poor air quality due to their differing roles in the household. Older children, both boys and girls are more likely to engage in paid work outside the household, compared to younger children who would perform domestic duties within the household. Exposure to pollution for younger children may come from attending school. For older children, it may be a combination of going to school and from not attending school and instead being exposed to pollution during work on or off the farm. Thus, we study separately the effect of air quality on children between the ages of 6 and 10 and older children ages 11 to 16.

## 4.2 Instrumental Variables (IV) Framework

Variations in pollution levels across time and space are not quasi-random. Additionally, households living in areas with high pollution are different on observable and unobservable characteristics than households who live in cleaner neighborhoods. To address this endogeneity problem, we use thermal inversions as an instrument for air quality to provide causal estimates of the short-term effects of poor

[15] We follow Gendron-Carrier et al. (2018) and do not express the independent variable, AOD, in logarithmic terms. This is because AOD is already a logarithm (see Section 3.3).

air quality on children's human capital. Thermal inversions occur when a mass of hot air situates above a mass of cold air. During inversion episodes, the temperature follows a non-monotonic pattern in altitude. The temperature first increases with altitude up to the inversion layer, and then decreases with altitude. This leads to a sharp deterioration of air quality in the ground level air layer since pollutants are trapped under the inversion layer. Using the NCEP/NCAR data described in Section 3.4, we identify daily inversion episodes for 2007 - 2014. Our main instrument is *number of daily nighttime inversions*. We estimate the following two-stage least squares (2SLS) specification:

$$AirQuality_{dt} = \alpha_0 + \alpha_1 Inversions_{dt} + \gamma X_{idt} + \delta Z_{dt} + \pi_a + \alpha_t + \eta_d + \eta_{s,t} + \epsilon_{dt} \tag{2}$$

$$Y_{idt} = \beta_0 + \beta_1 AirQuality_{dt} + \gamma X_{idt} + \delta Z_{dt} + \pi_a + \alpha_t + \eta_d + \eta_{s,t} + \epsilon_{idt} \tag{3}$$

where Equation 2 and Equation 3 are the first and second-stage equations. $Inversions_{dt}$ is our main instrumental variable which is equal to the number of daily nighttime inversions occurring in district $d$ in year $t$ = 2007 - 2014. All other variables are defined as before.

The relationship between thermal inversions and other meteorological factors might also impact the validity of the instrument. To address this we include controls for ground level temperature, precipitation, relative humidity, and wind velocity and their squares as described above. Controlling for these weather variables is important since inversions have a clear seasonal pattern and temperature may independently affect infant mortality (Deschênes and Greenstone (2007); Arceo et al. (2016)). This is demonstrated in Figure 3, which shows the average number of daily nighttime inversions by month of year for our study period, 2007 - 2014. Inversions are concentrated in the winter months of November - March. However, inversions also occur in the summer months i.e. April and October, when there is relatively little rain. Flexibly controlling for temperature enables us to dientangle the effects of temperature from that of pollution.

# 5 Results

## 5.1 OLS Estimates

We start by first examining the relationship between air quality and human capital in an OLS framework. Specifically, we estimate Equation 1 in Table 4. The OLS estimates imply that AOD has very small to no impact on learning outcomes. In particular, there is no statistically significant impact on reading outcomes. With respect to math outcomes, t7here is a small decline in the ability to do subtraction. That is, a 0.01 increase in AOD (one $\mu$g/m$^3$ in PM2.5) results in a 0.07 percentage point decline in the probability of being able to subtract. These impacts are quite small in absolute terms. This is not unexpected because we expect OLS estimates to be biased upward toward or above zero because of simultaneity bias. Consequently, we next turn to examining the impacts of air quality on learning outcomes in an IV framework.

### 5.2 IV First Stage

Turning next to the IV framework, we first examine the relationship between the number of daily nighttime inversions and our measure of pollution, AOD. This comprises the first stage of our IV strategy. In Figure 4 we present annual mean AOD separately for districts with above and below the median level of annual nighttime inversions. The frequency of high AOD is concentrated in districts witnessing higher than the median number of nighttime inversions (top panel), suggesting a positive relationship between the average annual daily nighttime inversions and AOD.

We also provide corresponding regression analysis in Table 5. Specifically, we present coefficient estimates from Equation 2. In column (1) we include district fixed effects, survey year fixed effects, and weather, and individual level controls. In column (2), we also include state-time linear trends. As suggested by the figure, the number of inversions have a positive and statistically significant effect on average annual AOD. One additional nighttime inversion, on average, results in an increase in AOD by 0.0002 units (roughly $0.02\mu g/m^3$).[16] These are big effects. This implies that a one standard deviation increase in daily nighttime thermal inversions increases AOD by 0.015 units (roughly $1.5\mu g/m^3$ or 3.4%).[17,18]

We test the robustness of the first stage relationship by estimating the first stage relationship in the bottom panel of Table 5 using an alternate measure of inversions, namely 'inversion strength'. Inversion strength is defined as the temperature difference between the ground level air layer and the air layer just above it.[19] We find that using this alternate definition of inversion strength, the first stage is still positive and statistically significant.

The main identifying assumption underlying the IV approach is that the exclusion restriction i.e. thermal inversions impact human capital only through their impact on AOD. This identifying assumption could, be violated if inversion episodes also change children's outdoor activities. We examine if individuals can change their behavior by predicting pollution increases due to nighttime inversions. Even if individuals endogenously move to locations with less pollution, they will not be able to avoid changes in pollution arising from thermal inversions since they will not be able to control or anticipate those changes. Following Hansen-Lewis (2018), in Table 6, we run the first stage regression where *future AOD* is the dependent variable. The trend shows that an additional daily nighttime inversion does not change the trend in AOD next year.

### 5.3 IV Estimates using ASER Data

We next present results from the main instrumental variable specification in Table 7. Results are presented for four levels of reading outcomes: can read a letter, can read a word, can read grade 1 text, can read grade 2 text; and four levels of math outcomes: can count numbers 1 - 9, can

[16] Van Donkelaar et al. (2014) show that a one-unit increase in AOD increases PM2.5 by $108\mu g/m^3$ . This implies a change of 0.0002 units translates to $0.02\mu g/m^3$.

[17] The Angrist-Pischke F-statistics are above the Stock-Yogo 10 percent threshold for weak instruments.

[18] Standard deviation of thermal inversions = 75.79. Thus, a one standard deviation (S.D.) increase in daily nighttime inversions will increase AOD by 75.79*0.0002 = 0.015 units.

[19] In Table 3 we present the mean and standard deviation of the 'inversion strength' variable.

count numbers 10 - 99, can do subtraction, and can do division. All outcome variables are indicator variables taking the value 0/1. Similar to the literature, we present results in terms of a 0.01 unit increase in AOD which equals one $\mu g/m^3$ increase in PM2.5.

On average, both reading and math outcomes decline significantly due to poor air quality. All four reading levels witness a decline, with the decline being smallest for the lowest level of reading. Specifically, for a 0.01 unit increase in AOD (one $\mu g/m^3$ in PM2.5), the probability of reading a letter declines by 0.57 percentage points (significant at the 10 percent level). This is unsurprising because almost 90 percent children at baseline can read letters. The probability of reading level 1 text and the probability of reading level 2 text witnessed the largest declines. For a 0.01 unit increase in AOD (one $\mu g/m^3$ in PM2.5), the probability of reading level one text and the probability of reading level two text declines by 1.67 and 1.40 percentage points, respectively (significant at the 1 percent and 5 percent levels).

With respect to math outcomes, there is no impact on the lowest level of math scores i.e. counting numbers from 1 to 9. Higher levels of math are significantly impacted by poor air quality. The probability of counting numbers from 10 - 99 and the probability of subtraction decline by 1.37 percentage points and 1.43 percentage points, respectively, for a 0.01 unit increase in AOD (one $\mu g/m^3$ in PM2.5).

We also examine heterogeneous effects by age and grade. Regression estimates in Table 8 are presented separately for boys and girls and also separately for children of primary school age (6-10 years) and older children (11-16 years), as they might respond differently to changes in air quality. For younger children, poorer air quality (or a higher AOD) is associated with a lower likelihood of recognizing letters and words. The effect is present for both boys and girls. The estimated effect implies that a 0.01 one unit increase in AOD (one $\mu g/m^3$ in PM2.5), significantly reduces the probability of recognizing letters for boys (girls) by 1.09 (0.94) percentage points. The impact on recognizing words is stronger - for a 0.01 unit increase in AOD, the probability of recognizing words significantly reduces by 1.90 (2.38) percentage points for boys (girls). Young girls also have strong negative impacts on being able to read grade 1 and grade 2 text. Specifically, for girls, a 0.01 unit increase in AOD (one $\mu g/m^3$ in PM2.5) reduces the probability of being able to read grade 1 text and grade 2 text by 2.33 and 1.31 percentage points, respectively. The effect for boys is smaller - the probability of reading grade 1 text reduces by 1.25 percentage points (significant at the 10% level). With respect to math outcomes, we find that the probability of recognizing double digit numbers (10 - 99) declines by 1.63 (2.15) percentage points for boys (girls), for a 0.01 unit increase in AOD (one $\mu g/m^3$ in PM2.5). No other math outcome for young children is impacted significantly. This is unsurprising since more complex math problems like division are taught only from grade 5 onwards, relevant for only the 9-10-year-old children in the sample of 6-10-year-old children. Apart from probability of reading a letter, among younger children, girls witness a larger decline for all other reading and math outcomes.

For older children between the ages of 11 - 16 years, we find stronger effects. An increase in current year AOD by 0.01 units (one $\mu g/m^3$ in PM2.5) reduces all reading outcomes for boys, significant at the 1% level. The strongest effect is on the ability to read grade 2 text which declines by 1.83 percentage points. Similarly, for girls, all reading outcomes decline, with the decline being larger for more advanced reading outcomes. An increase in AOD by 0.01 units (one $\mu$g/m$^3$ in PM2.5) reduces the probability of girls being able to read grade 2 text by 2.48 percentage points. With respect to math outcomes, we once again find a steady decline in all math outcomes, with the more advanced math outcomes witnessing a larger decline. Specifically, an increase in AOD by 0.01 units (one $\mu$g/m$^3$ in PM2.5) reduces the probability of doing division by 1.85 (2.50) percentage points for boys (girls), respectively. For both math and reading outcomes, like for younger children, older girls witness a larger decline.

Putting these effects into perspective - Chakraborty and Jayaraman (2019) find that a one-month exposure to a school feeding program increases the probability that 6-10-year-old children are able to read a word or recognize double-digit numbers by 0.2 percentage points.[20] This suggests that our estimates for younger children are similar to the effect of exposure to more than one year of school feeding. Alternatively, Shah and Steinberg (2017) analyze the cognitive effects of exposure to drought in utero from India and find that exposure to drought in utero is associated with being 2 percentage points less likely to recognize numbers in childhood. Further, Spears and Lamba (2013) find that moving from 0 latrines per capita to 1 latrine per capita in India improves young children's ability to recognize letters or better by 0.72 percentage points. For older children, the effects are larger. They are comparable to those found in other studies. For instance, Barham (2012) estimates the effect on cognitive outcomes of a maternal and child health, family planning, and vaccination program in Bangladesh and finds that early life exposure to the program caused a 0.39 standard deviation in a measure of cognitive functioning when children were 8 to 14 years old.

### 5.4 Heterogeneous Effects by State

As documented in Figure 2 in Hansen-Lewis (2018), the most polluted districts in India i.e. the ones with the highest levels of AOD lie in 6 states. These are: Rajasthan, Punjab, Haryana, Uttar Pradesh, Bihar, and West Bengal. Mean AOD in districts in these states is 0.57 compared to the national average of 0.41. We classify these 6 states as 'polluted states' and examine impacts separately for these states and other states. Intuitively, we would expect that impacts would be stronger in states with her levels of AOD.

Table 9 presents our main specification with the current year AOD variable interacted with whether the district is in a polluted state or not. We present results for all our variables of interest: four levels of reading outcomes and four levels of math outcomes. As expected, there seems to be almost no effects beyond the polluted states. All the impacts of air pollution exposure on learning outcomes are coming from these 6 states. In terms of reading outcomes, the impact in polluted states of air pollution exposure ranges from a decline of 0.89 percentage points for letter recognition to 2.65

[20]The median educational intervention leads to impacts between 0.08 and 0.15 S.D. (McEwan (2015); Garg et al. (2017)). Educational interventions such as school construction programs and other programs from India and other countries have found large effects on learning outcomes: 0.4 S.D. (Akresh et al. (2013)), 0.65 S.D. (Burde and Linden (2013)); 0.5 S.D. (Banerjee et al. (2007)), 0.2 S.D. (Glewwe et al. (2009)), and 0.3 S.D. (Muralidharan et al. (2019)).

percentage points for being able to read Grade 1 text. The effects for the other states are small and insignificant. With respect to math outcomes, for a 0.01 unit change in AOD, in polluted states, outcomes decline by 1.97 to 2.65 percentage points. In other states, impacts are once again small and insignificant. For both reading and math, impacts in polluted states are *larger* than the overall impacts presented in Table 7.

### 5.5 IV Estimates using YLS Data

We check the robustness of our results estimated using the ASER data by examining the impact of air quality on learning outcomes in the Young Lives Survey (YLS). Covering six districts in the state of Andhra Pradesh, the YLS data allows us to estimate the effects of air quality by controlling for individual fixed effects, removing any time invariant individual level characteristics correlated with pollution. We present impacts on the percentage of correct answers since the total number of questions varied in each survey round. We use the following 2SLS specification to examine effects in the YLS data:

$$AirQuality_{dt} = \alpha_0 + \alpha_1 Inversions_{dt} + \gamma X_{idt} + \theta_i + \eta_t + \eta_j + \eta_m + \epsilon_{dt} \tag{4}$$

$$Y_{idt} = \beta_0 + \beta_1 AirQuality_{dt} + \gamma X_{idt} + \theta_i + \eta_t + \eta_j + \eta_m + \epsilon_{idt} \tag{5}$$

where $Y_{idt}$ is the outcome of interest (e.g. PPVT or math test score) for student $i$ in district $d$ in survey year $t$ = 2007, 2009, and in 2013-14. We include individual fixed effects ($\theta_i$); survey round fixed effects ($\eta_t$); day-of-test fixed effects ($\eta_j$); and month of test fixed effects ($\eta_m$). Similar to Garg et al. (2017), standard errors are clusters at the district-week level since test scores in a district in a given testing week could be correlated. Since the YLS data tracks the same child over time, this allows us to account for prior human capital using individual level fixed effects.

First stage results are presented in Table 10. The first stage estimates show a robust relationship between current year AOD and inversions for the state of Andhra Pradesh. Specifically, one additional daily nighttime inversion increases AOD by 0.0002 units. Alternatively, a one standard deviation increase in daily nighttime inversions, AOD rises by 0.0086 units (roughly 0.86$\mu$g/m$^3$ in PM2.5).

Next, we examine the IV results using the YLS individual panel data set in Table 11. Our findings are similar, though qualitatively larger than the estimates using the ASER data. We find that, only reading outcomes for boys are impacted significantly by an increase in AOD. For girls on the other hand, both math and reading outcomes are impacted. We find that for reading scores decline by 5.31 (4.58) percentage points for boys (girls), respectively, for a 0.01 one unit increase in annual AOD (one $\mu$g/m$^3$ in PM2.5). For girls, outcomes in math also decline by 7.29 percentage points.

### 5.6 Robustness Checks

In this section we present five additional robustness checks. First, in Appendix Table A1 we estimate the main results using only the sample of "on-track kids". Shah and Steinberg (2017) define "on-track kids" as those who are in the correct grade for their age. Approximately, 74.6% of children in our sample from 2007 - 2014 are in the correct grade-for-age. To ensure that our results are not driven by poor learning outcomes for children who are not in the correct grade-for-age, we check the robustness of our results with a smaller sample of on-track kids. We find that the main results are smaller but robust. A 0.01 unit increase in AOD (one $\mu$g/m$^3$ in PM2.5), reduces the reading outcomes by 1.19 - 1.43 percentage points and reduces math outcomes by 1.15 - 1.23 percentage points.

Second, we test the robustness of our results using alternate temperature layers to define a thermal inversion. Specifically, instead of using the two pressure levels closest to the ground (1000hPa and 925hPa) to identify inversion episodes (as described in Section 3), we use the temperature difference between the second and third pressure levels i.e. 925hPa and 850hPa to identify inversion episodes. These results are reported in Appendix Table A2.

Third, since the ASER tests are administered between September to November in different villages across the country, we test if exposure in current year, defined as the 12 months prior to the test, impacts learning outcomes. Specifically, we define current year as September-September, as opposed to a calendar year. We report results in Appendix Table A3. Our results remain robust to this alternative specification as well.

Fourth, to ensure that the inclusion of state-specific linear time trends are not biasing our results, we estimate impacts omitting these trends in Table A4. We find that the results are qualitatively similar, which suggest that inversions are uncorrelated with the trends in learning outcomes that are unrelated to pollution. Finally, following Arceo et al. (2016) and Hansen-Lewis (2018), we test the robustness of our main results by including season specific fixed effects in addition to survey year fixed effects. Specifically, we include indicators for each of: January to March, April to June, July to September, and October to December. We present these results in Appendix Table A5.

# 6 Mechanisms

## 6.1 Physiological

Exposure to pollution can lead to physiological changes which can impact health and hence cognition. Pollution can also directly impair brain functioning in addition to direct impacts on health (see e.g. Zhang et al. (2018)). Relatively little is known about the exact mechanisms underlying any health effects (although most pollutants affect the respiratory and cardiovascular systems) especially on children who are more susceptible to the effects of pollution than adults because their bodies are developing and they have higher metabolic rates. Children also typically spend more time outdoors than adults do, increasing their total exposure. Since the ASER data do not contain health outcomes, we test the physiological channel in three ways. First, if pollution has a physiological effect then we would expect stronger effects the more one is exposed to pollution. Consequently, we examine the effects of cumulative exposure to pollution. Specifically, we examine cumulative pollution exposure in the last two years and cumulative exposure in the last three years. Second, we examine if poor air quality impacts the extensive margin. Specifically, we examine effects of air quality on: (a) drop out rates; (b) being on-track; and (c) student attendance.

#### 6.1.1 Cumulative Effects of Pollution

There is a vast body of medical literature that shows that being exposed to poor air quality can lead to a vast array of medical problems including lung cancer, stroke, and heart disease, by impacting blood flow and oxygen circulation in the lungs (Pope 3rd et al. (1995); Mills et al. (2009)). The longer one is exposed to pollution, the higher the likelihood that pollution impacts one's respiratory and/or cardiovascular functioning. To test this channel, we examine the cumulative impact of pollution exposure in the last two years, and pollution exposure in the last three years on current year human capital. This allows us to examine medium-run impacts of air pollution. Following Zhang et al. (2018), we estimate the following equations to study the cumulative impact of air pollution on learning outcomes:

$$AirQuality_{dt} = \alpha_0 + \alpha_1 Inversions_{dt} + \gamma X_{idt} + \delta Z_{dt} + \pi_a + \alpha_t + \eta_d + \eta_{s,t} + \epsilon_{dt} \tag{6}$$

$$Y_{idt} = \beta_0 + \beta_1 \frac{1}{t} \sum_{n=0}^{t-1} AirQuality_{dt-n} + \gamma X_{idt} + \delta Z_{dt} + \pi_a + \alpha_t + \eta_d + \eta_{s,t} + \epsilon_{idt} \tag{7}$$

where the key variable $\frac{1}{t} \sum_{n=0}^{t-1} AirQuality_{dt-n}$ is the mean AOD in the past $t$ years, which measures cumulative exposure. We examine cumulative exposure over the last 2 years and last 3 years, respectively. All other variables are defined as before.

We present the results in Tables 13 and 14. The physiological impacts of past year AOD on current year test scores are large. Cumulative exposure for a longer time horizon, in general, is worse for cognitive development. Specifically, a 0.01 unit cumulative change in AOD (one $\mu$g/m$^3$ in PM2.5) over the last two years leads to a decline in reading and math outcomes by 0.71 - 1.99 percentage points, while exposure for 3 years leads to a decline in reading and math outcomes by 0.79 - 2.73 percentage points. Thus, similar to Zhang et al. (2018), the damage of air pollution on learning outcomes is more sizable when using longer window of exposure measure.

#### 6.1.2 Extensive Margin

We next examine impacts on the extensive margin of schooling. Poor air quality can lead to impacts on cognitive outcomes through the extensive margin of schooling. Changes in learning outcomes observed above could arise both, from pollution directly affecting children's cognition, or from changes on health which impact attendance and drop out rates, and hence learning outcomes. We first begin by examining impacts on drop out rates and an indicator for being "on-track" in Table 12. Neither current year AOD nor cumulative exposure over the last two and three years has a statistically significant impact on the probability of children dropping out school. On the other hand, poor air quality significantly lowers the probability of children being on-track. A 0.01 increase in current year AOD (one $\mu$g/m$^3$ in PM2.5) lowers the probability of being on-track by 1.14 percentage points. Cumulative exposure over two years also significantly reduces the likelihood of being on track by 0.92 percentage points.

Next, we examine if student attendance is adversely impacted by poor air quality, which leads to lower outcomes on math and reading tests. Specifically, children be recovering from short-run illnesses such as increased cough or other breathing problems due to pollution which would increase school absences. Alternatively, certain cities sometimes issue extended vacation periods and/or healthtadvisories due to poor air quality. This will also increase student absences from school. (Currie, Hanushek, et al., 2009) find that high carbon monoxide levels significantly increase school absences in Texas. Other studies which link pollution and absenteeism include (Ransom and Pope III, 1992); (Gilliland et al., 2001); and (Makino, 2000).

Since the student level ASER data does not have information on attendance, we test if an increase in ambient air pollution leads to increased student absences, which then impacts learning outcomes using the school observations conducted by ASER. Following Adukia (2017), we estimate the following 2SLS specification using the school level ASER data:

$$AirQuality_{dt} = \alpha_0 + \alpha_1 Inversions_{dt} + \alpha_t + \eta_s + \eta_{s,t} + \epsilon_{dt} \tag{8}$$

$$Y_{sdt} = \beta_0 + \beta_1 AirQuality_{dt} + \alpha_t + \eta_s + \eta_{s,t} + \epsilon_{sdt} \tag{9}$$

where $Y_{sdt}$ is the proportion of children attending school as a proportion of total enrolled students in grades 1 - 8 in school $s$ in district $d$ at time $t$. We only include school surveys from 2009 - 2014 since variables were defined differently in the 2007 school survey. We also present estimates disaggregated by smaller grades (1 - 5) and higher grades (6 - 8). We include school fixed effects ($\eta_s$), fixed effects for survey year ($\alpha_t$), and state specific linear time trends ($\eta_{s,t}$). Standard errors are clustered at the school level.

Results are presented in Table 15. We find that attendance on the day of the test is lower by 1.11 percentage points, overall. Most of the effect stems from lower attendance in grades 1 - 5 which see a decline in attendance of 1.42 percentage points from a 0.01 one-unit increase in current year AOD. Attendance in grades 6 - 8 also declines by 0.51 percentage points, however, this is not statistically significant. We also examine if cumulative exposure to pollution affects student attendance. Cumulative exposure in the last two years has a large and statistically significant impact on attendance of children in all grades. Children in grades 1 - 5 (grades 6 - 8) are less likely to attend school in the current year by 4.9 (3.13) percentage points if they have been exposed to pollution in the last two years. Cumulative exposure in the last 3 years has smaller, but still statistically significant impacts on current year attendance. These results imply that exposure to pollution has longer term impacts on attendance in the form of health effects. Effects of current year pollution on current year attendance could be driven by both, health effects, and school closures due to heavily polluted weather. However, longer lasting impacts on attendance due to cumulative exposure would most likely stem from health related absences from school.

In Table 16, following, Shah and Steinberg (2017), as a robustness check we estimate the impact of AOD on children's reported "primary activity" using National Sample Survey (NSS) data to corroborate the ASER attendance results. We use data from NSS Rounds 61 (2004-05), 62 (2005-06), 64 (2007-08), 66 (2009-10), and 68 (2011-12). The coefficients represent the effect of current year and cumulative AOD on a dummy variable for whether primary activity is reported as school attendance. The sample is limited to be similar to the ASER - children aged 5-16 years. Both at the extensive and intensive margins, attendance goes down. At the extensive margin, a 0.01 increase in current year AOD (one $\mu g/m^3$ in PM2.5) leads to a 0.69 percentage point decline in children attending school. Effects for cumulative exposure are larger.

Taken together, these findings imply that poor air quality impacts children more on the intensive margin of schooling than on the extensive margin. The impacts on the intensive margin result from lower school attendance due to health-related reasons or from direct impacts on the human brain. These impacts also lead to children falling behind. That is, detrimental impacts on learning due to air pollution lead to a reduced likelihood of children being "on-track" or the correct age for grade. These impacts however, do not result in children dropping out of school altogether.

### 6.2 Teacher Attendance

Teacher quality is an important input into improved learning outcomes. Studies have shown that high teacher quality leads to higher test scores and improved adult life outcomes (Hanushek and Rivkin (2006); Chetty et al. (2014)). If exposure to pollution impacts adult health, then it is possible that teacher absenteeism also increases because of high pollution leading to lower instructional quality.

To test the impact of poor air quality on teacher attendance we use the school ASER survey which records the total number of regular teachers employed and the total number of teachers in attendance during the day of the survey. We estimate the 2SLS specification outlined in Section 6.1.2 in Table 15. The main dependent variable is the proportion of regular teachers attending school on the day of the survey. We find that there is a decline the proportion of teachers attending school on a random day by 1.26 percentage points. However, this effect is not statistically significant.

Turning next to cumulative exposure, we find that cumulative exposure to pollution impacts teacher attendance negatively. Cumulative exposure over the past two years reduces the proportion of teachers attending school in the current year by 5.2 percentage points. Exposure over a three year period has no impact on teacher attendance. Taken together, these findings imply that teacher attendance have a role to play in poor reading and math test scores. This is unsurprising since even adults would be affected by cumulative exposure to pollution and thus have health-related absences.

## 7 Conclusion

Using a large dataset from India, this paper examines the causal impacts of poor air quality on children's learning outcomes. Following recent literature, we use Aerosol Optical Depth (AOD) data from 2007 - 2014, derived from the MODIS Terra satellite to proxy for air quality (Dey, Di Girolamo, et al. (2012); Hansen-Lewis (2018); Gendron-Carrier et al. (2018)). To account for the endogeneity problem when examining the causal relationship between pollution and human capital, we use an instrumental variables (IV) strategy. Specifically, we use thermal inversions to instrument for air quality. Ordinarily, temperature drops with rising altitude. However, a thermal inversion occurs when there is a mass of hot air on top of a mass of cold air. We show that thermal inversions are strongly positively correlated with our measure of pollution, AOD.

We find that poor contemporaneous air quality leads to worse reading and math outcomes for both boys and girls, and children of all ages. Effects are driven primarily by the most polluted states in India i.e. Punjab, Haryana, Uttar Pradesh, Rajasthan, Bihar, and West Bengal. The main mechanism driving these effects is a physiological one. We find that on the extensive margin, student attendance is impacted by both contemporaneous and cumulative exposure to poor air quality. In fact, cumulative exposure over the last two and three years has stronger negative impacts on learning outcomes. Student attendance is also lower due to poor air quality. Additionally, it appears frequent absences leads children to fall behind. That is, we find evidence on a negative impact on being 'on-track' or the correct grade for age. We also find suggestive evidence that teacher absences due to air pollution could also lead to worse learning outcomes for children. This implies that pollution imposes a double penalty on children wherein, not only are they directly impacted due to poor health leading to school absences, they are also impacted indirectly by the impact of pollution on their teachers' health and school attendance.

Our findings have two major implications for understanding the role of environment in impacting population well-being and policy. First, most studies examining the impacts of pollution on well-being, focus on the direct impacts on health. However, to the extent that pollution also affects the human brain, the indirect effect of pollution on social welfare are potentially large. Narrowly focusing on the health effects of air pollution may underestimate the true cost of pollution. Second, adaptation to pollution will require a focus on multiple margins. While short to medium-rum exposure to pollution can be mitigated by the recent proliferation of air purifiers in India and by canceling school on days of elevated air pollution levels. However, longer term adaptation to pollution is a serious concern, made graver by elevated temperatures. Longer-term pollution exposure not only worsens health and increases mortality risk, it also lowers the productivity of both children and adults, and hence may have serious implications for income and future capabilities of the Indian population (Sen, 1999). To mitigate the long-term consequences it is essential to strengthen social protection programs like the National Rural Employment Guarantee Act (NREGA) *and* institute strict environmental policies in place to control emissions from crop burning, vehicular, and industrial emissions.

## References


ADAIR-ROHANI, H. (2018). *Air Pollution and Child Health: Prescribing Clean Air*. World Health Organization.

ADHVARYU, A., N. KALA, and A. NYSHADHAM (2014). "Management and shocks to worker productivity: evidence from air pollution exposure in an Indian garment factory". *Unpublished working paper, University of Michigan*.

ADUKIA, A. (2017). "Sanitation and education". *American Economic Journal: Applied Economics* 9.2, pp. 23–59.

AKRESH, R., D. DE WALQUE, and H. KAZIANGA (2013). *Cash transfers and child schooling: evidence from a randomized evaluation of the role of conditionality*. The World Bank.

ARCEO, E., R. HANNA, and P. OLIVA (2016). "Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City". *The Economic Journal* 126.591, pp. 257–280.

BAILEY, A., T. N. CHASE, J. J. CASSANO, and D. NOONE (2011). "Changing temperature inversion characteristics in the US Southwest and relationships to large-scale atmospheric circulation". *Journal of Applied Meteorology and Climatology* 50.6, pp. 1307–1323.

BALAKRISHNAN, K., A. COHEN, and K. R. SMITH (2014). *Addressing the burden of disease attributable to air pollution in India: the need to integrate across household and ambient air pollution exposures*.

BANERJEE, A. V., S. COLE, E. DUFLO, and L. LINDEN (2007). "Remedying education: Evidence from two randomized experiments in India". *The Quarterly Journal of Economics* 122.3, pp. 1235–1264.

BARHAM, T. (2012). "The Effects of Child Health Interventions on Later Life Cognitive Functioning: Evidence from Matlab in Bangladesh". *American Economic Journal: Applied Economics* 4.1, pp. 245–73.

BIKKINA, S., A. ANDERSSON, E. N. KIRILLOVA, H. HOLMSTRAND, S. TIWARI, A. SRIVASTAVA, D. BISHT, et al. (2019). "Air quality in megacity Delhi affected by countryside biomass burning". *Nature Sustainability* 2.3, p. 200.

BLOOM, D. E., D. CANNING, L. HU, Y. LIU, A. MAHAL, and W. YIP (2010). "The contribution of population health and demographic change to economic growth in China and India". *Journal of Comparative Economics* 38.1, pp. 17–33.

BORJA-ABURTO, V. H., M. CASTILLEJOS, D. R. GOLD, S. BIERZWINSKI, and D. LOOMIS (1998). "Mortality and ambient fine particles in southwest Mexico City, 1993-1995." *Environmental Health Perspectives* 106.12, pp. 849–855.

BURDE, D. and L. L. LINDEN (2013). "Bringing education to Afghan girls: A randomized controlled trial of village-based schools". *American Economic Journal: Applied Economics* 5.3, pp. 27–40.

CACCIOTTOLO, M., X. WANG, I. DRISCOLL, N. WOODWARD, A. SAFFARI, J. REYES, M. L. SERRE, et al. (2017). "Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models". *Translational Psychiatry* 7.1, e1022.

CALDERÓN-GARCIDUEÑAS, L. and R. TORRES-JARDÓN (2012). "Air pollution, socioeconomic status, and children's cognition in megacities: The Mexico City scenario". *Frontiers in Psychology* 3, p. 217.

CHAKRABORTY, T. and R. JAYARAMAN (2019). "School feeding and learning achievement: Evidence from India's midday meal program". *Journal of Development Economics* 139, pp. 249–265.

CHAY, K. Y. and M. GREENSTONE (2003). "The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession". *The Quarterly Journal of Economics* 118.3, pp. 1121–1167.

CHEN, H., J. C. KWONG, R. COPES, P. HYSTAD, A. van DONKELAAR, K. TU, J. R. BROOK, et al. (2017). "Exposure to ambient air pollution and the incidence of dementia: a population-based cohort study". *Environment International* 108, pp. 271–277.

CHEN, S., P. OLIVA, and P. ZHANG (2017). *The effect of air pollution on migration: evidence from China*. Tech. rep. National Bureau of Economic Research.

CHETTY, R., J. N. FRIEDMAN, and J. E. ROCKOFF (2014). "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood". *American Economic Review* 104.9, pp. 2633–79.

CHEW, S. H., W. HUANG, and X. LI (2018). "Does haze cloud decision making? A natural laboratory experiment". *A Natural Laboratory Experiment (August 6, 2018)*.

CHU, D. A., Y. KAUFMAN, G. ZIBORDI, J. CHERN, J. MAO, C. LI, and B. HOLBEN (2003). "Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS)". *Journal of Geophysical Research: Atmospheres* 108.D21.

COHEN, A. J., H. ROSS ANDERSON, B. OSTRO, K. D. PANDEY, M. KRZYZANOWSKI, N. KÜNZLI, K. GUTSCHMIDT, et al. (2005). "The global burden of disease due to outdoor air pollution". *Journal of Toxicology and Environmental Health, Part A* 68.13-14, pp. 1301–1307.

CRAIG, L., J. R. BROOK, Q. CHIOTTI, B. CROES, S. GOWER, A. HEDLEY, D. KREWSKI, et al. (2008). "Air pollution and public health: a guidance document for risk managers". *Journal of Toxicology and Environmental Health, Part A* 71.9-10, pp. 588–698.

CURRIE, J., E. A. HANUSHEK, E. M. KAHN, M. NEIDELL, and S. G. RIVKIN (2009). "Does pollution increase school absences?" *The Review of Economics and Statistics* 91.4, pp. 682–694.

CURRIE, J., J. G. ZIVIN, J. MULLINS, and M. NEIDELL (2014). "What do we know about short-and long-term effects of early-life exposure to pollution?" *Annu. Rev. Resour. Econ.* 6.1, pp. 217–247.

DASGUPTA, S., U. DEICHMANN, C. MEISNER, and D. WHEELER (2003). *The poverty/environment nexus in Cambodia and Lao People's Democratic Republic*. The World Bank.

DERYUGINA, T., G. HEUTEL, N. H. MILLER, D. MOLITOR, and J. REIF (2016). *The mortality and medical costs of air pollution: Evidence from changes in wind direction*. Tech. rep. National Bureau of Economic Research.

DESCHÊNES, O. and M. GREENSTONE (2007). "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather". *American Economic Review* 97.1, pp. 354–385.

DEY, S. and L. DI GIROLAMO (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., L. DI GIROLAMO, A. van DONKELAAR, S. TRIPATHI, T. GUPTA, and M. MOHAN (2012). "Variability of outdoor fine particulate (PM2. 5) concentration in the Indian Subcontinent: A remote sensing approach". *Remote Sensing of Environment* 127, pp. 153–161.

DOCKERY, D. and A. POPE (1996). *Epidemiology of acute health effects: summary of time-series studies.*

DUFLO, E., M. GREENSTONE, and R. HANNA (2008). "Cooking stoves, indoor air pollution and respiratory health in rural Orissa". *Economic and Political Weekly*, pp. 71–76.

DUFLO, E., M. GREENSTONE, R. PANDE, and N. RYAN (2013). "Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India". *The Quarterly Journal of Economics* 128.4, pp. 1499–1545.

DURAIAPPAH, A. K. (1998). "Poverty and environmental degradation: a review and analysis of the nexus". *World Development* 26.12, pp. 2169–2179.

EBENSTEIN, A., V. LAVY, and S. ROTH (2016). "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution". *American Economic Journal: Applied Economics* 8.4, pp. 36–65.

FOSTER, A., E. GUTIERREZ, and N. KUMAR (2009). "Voluntary compliance, pollution levels, and infant mortality in Mexico". *American Economic Review* 99.2, pp. 191–97.

FREIRE, C., R. RAMOS, R. PUERTAS, M.-J. LOPEZ-ESPINOSA, J. JULVEZ, I. AGUILERA, F. CRUZ, et al. (2010). "Association of traffic-related air pollution with cognitive development in children". *Journal of Epidemiology & Community Health* 64.3, pp. 223–228.

FU, Q., Z. MO, D. LYU, L. ZHANG, Z. QIN, Q. TANG, H. YIN, et al. (2017). "Air pollution and outpatient visits for conjunctivitis: A case-crossover study in Hangzhou, China". *Environmental Pollution* 231, pp. 1344–1350.

GARG, T., M. JAGNANI, and V. P. TARAZ (2017). "Human capital costs of climate change: evidence from test scores in India".

GENDRON-CARRIER, N., M. GONZALEZ-NAVARRO, S. POLLONI, and M. A. TURNER (2018). *Subways and urban air pollution*. Tech. rep. National Bureau of Economic Research.

GHIO, A. J. and R. B. DEVLIN (2001). "Inflammatory lung injury after bronchial instillation of air pollution particles". *American Journal of Respiratory and Critical Care Medicine* 164.4, pp. 704–708.

GILLILAND, F. D., K. BERHANE, E. B. RAPPAPORT, D. C. THOMAS, E. AVOL, W. J. GAUDERMAN, S. J. LONDON, et al. (2001). "The effects of ambient air pollution on school absenteeism due to respiratory illnesses". *Epidemiology*, pp. 43–54.

GLEWWE, P., M. KREMER, and S. MOULIN (2009). "Many children left behind? Textbooks and test scores in Kenya". *American Economic Journal: Applied Economics* 1.1, pp. 112–35.

GOLDBERG, P. K. and N. PAVCNIK (2007). "Distributional effects of globalization in developing countries". *Journal of Economic Literature* 45.1, pp. 39–82.

GOLDIN, C. and L. F. KATZ (2007). *Long-run changes in the US wage structure: narrowing, widening, polarizing*. Tech. rep. National Bureau of Economic Research.

GOU, Y.-Y., Y.-C. HSU, H.-R. CHAO, D. E. QUE, L. L. TAYO, C.-H. LIN, S. M. HUANG, et al. (2016). "Pollution characteristics and diurnal variations in polybrominated diphenyl ethers in indoor and outdoor air from vehicle dismantler factories in southern Taiwan". *Aerosol Air Qual. Res* 16, pp. 1931–1941.

GRAFF ZIVIN, J. and M. NEIDELL (2012). "The impact of pollution on worker productivity". *American Economic Review* 102.7, pp. 3652–73.

GREENSTONE, M. and R. HANNA (2014). "Environmental regulations, air and water pollution, and infant mortality in India". *American Economic Review* 104.10, pp. 3038–72.

GREENSTONE, M., J. NILEKANI, R. PANDE, N. RYAN, A. SUDARSHAN, and A. SUGATHAN (2015). "Lower pollution, longer lives: life expectancy gains if India reduced particulate matter pollution". *Economic and Political Weekly* 50.8.

GROUP, G. M. W. et al. (2018). "Burden of disease attributable to major air pollution sources in India". *Special Report* 21.

GUAZZOTTI, S., D. SUESS, K. COFFEE, P. QUINN, T. BATES, A. WISTHALER, A. HANSEL, et al. (2003). "Characterization of carbonaceous aerosols outflow from India and Arabia: Biomass/biofuel burning and fossil fuel combustion". *Journal of Geophysical Research: Atmospheres* 108.D15.

GUPTA, P., S. A. CHRISTOPHER, J. WANG, R. GEHRIG, Y. LEE, and N. KUMAR (2006). "Satellite remote sensing of particulate matter and air quality assessment over global cities". *Atmospheric Environment* 40.30, pp. 5880–5892.

GUTTIKUNDA, S. K., R. GOEL, and P. PANT (2014). "Nature of air pollution, emission sources, and management in the Indian cities". *Atmospheric Environment* 95, pp. 501–510.

GWYNN, R. C. and G. D. THURSTON (2001). "The burden of air pollution: impacts among racial minorities." *Environmental Health Perspectives* 109.suppl 4, pp. 501–506.

HANSEN-LEWIS, J. (2018). "Does Air Pollution Lower Productivity? Evidence from Manufacturing in India".

HANUSHEK, E. A. and S. G. RIVKIN (2006). "Teacher quality". *Handbook of the Economics of Education* 2, pp. 1051–1078.

HARRIS, M. H., D. R. GOLD, S. L. RIFAS-SHIMAN, S. J. MELLY, A. ZANOBETTI, B. A. COULL, J. D. SCHWARTZ, et al. (2016). "Prenatal and childhood traffic-related air pollution exposure and childhood executive function and behavior". *Neurotoxicology and Teratology* 57, pp. 60–70.

HETTIGE, H., M. MANI, and D. WHEELER (2000). "Industrial pollution in economic development: the environmental Kuznets curve revisited". *Journal of Development Economics* 62.2, pp. 445–476.

HICKS, D., P. MARSH, and P. OLIVA (2016). *Air pollution and procyclical mortality: Causal evidence from thermal inversions.* Tech. rep. working paper. Jacobson, M.(2002). Atmospheric Pollution. History, Science fffdfffdfffd.

HOUSTON, D., J. WU, P. ONG, and A. WINER (2004). "Structural disparities of urban traffic in Southern California: implications for vehicle-related air pollution exposure in minority and high-poverty neighborhoods". *Journal of Urban Affairs* 26.5, pp. 565–592.

IACOBELLIS, S. F., J. R. NORRIS, M. KANAMITSU, M. TYREE, and D. R. CAYAN (2009). "Climate variability and California low-level temperature inversions". *California Climate Change Center* 48.

JACOBSON, M. Z. and M. Z. JACOBSON (2002). *Atmospheric pollution: history, science, and regulation.* Cambridge University Press.

JANS, J., P. JOHANSSON, and J. P. NILSSON (2018). "Economic status, air quality, and child health: Evidence from inversion episodes". *Journal of Health Economics* 61, pp. 220–232.

JEDRYCHOWSKI, W. A., F. P. PERERA, D. CAMANN, J. SPENGLER, M. BUTSCHER, E. MROZ, R. MAJEWSKA, et al. (2015). "Prenatal exposure to polycyclic aromatic hydrocarbons and cognitive dysfunction in children". *Environmental Science and Pollution Research* 22.5, pp. 3631–3639.

KARAGULIAN, F., C. A. BELIS, C. F. C. DORA, A. M. PRÜSS-USTÜN, S. BONJOUR, H. ADAIR-ROHANI, and M. AMANN (2015). "Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level". *Atmospheric Environment* 120, pp. 475–483.

KICINSKI, M., G. VERMEIR, N. VAN LAREBEKE, E. DEN HOND, G. SCHOETERS, L. BRUCKERS, I. SIOEN, et al. (2015). "Neurobehavioral performance in adolescents is inversely associated with traffic exposure". *Environment International* 75, pp. 136–143.

KILIAN, J. and M. KITAZAWA (2018). "The Emerging risk of exposure to air pollution on cognitive decline and Alzheimer's disease–evidence from epidemiological and animal studies". *Biomedical Journal* 41.3, pp. 141–162.

KRISHNAMOORTHY, Y., K. S. GOKUL SARVESWARAN, M. SAKTHIVEL, M. G. MAJELLA, and S. G. KUMAR (2018). "Association between indoor air pollution and cognitive impairment among adults in rural Puducherry, South India". *Journal of Neurosciences in Rural Practice* 9.4, p. 529.

KUMAR, N., A. CHU, and A. FOSTER (2007). "An empirical relationship between PM2. 5 and aerosol optical depth in Delhi Metropolitan". *Atmospheric Environment* 41.21, pp. 4492–4503.

LELIEVELD, J., A. HAINES, and A. POZZER (2018). "Age-dependent health risk from ambient air pollution: A modelling and data analysis of childhood mortality in middle-income and low-income countries". *The Lancet Planetary Health* 2.7, e292–e300.

LIPFERT, F. (2004). *Air pollution and poverty: does the sword cut both ways?*

LOOMIS, D., M. CASTILLEJOS, D. R. GOLD, W. MCDONNELL, and V. H. BORJA-ABURTO (1999). "Air pollution and infant mortality in Mexico City". *Epidemiology*, pp. 118–123.

MAAS, R. and P. GRENNFELT (2017). "Towards cleaner air: Scientific assessment report 2016". EMEP Steering Body and Working Group.

MAKINO, K. (2000). "Association of school absence with air pollution in areas around arterial roads". *Journal of Epidemiology* 10.5, pp. 292–299.

MALEK, E., T. DAVIS, R. S. MARTIN, and P. J. SILVA (2006). "Meteorological and environmental aspects of one of the worst national air pollution episodes (January, 2004) in Logan, Cache Valley, Utah, USA". *Atmospheric Research* 79.2, pp. 108–122.

MCEWAN, P. J. (2015). "Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments". *Review of Educational Research* 85.3, pp. 353–394.

MILLS, N. L., K. DONALDSON, P. W. HADOKE, N. A. BOON, W. MACNEE, F. R. CASSEE, T. SANDSTRÖM, et al. (2009). “Adverse cardiovascular effects of air pollution”. *Nature Reviews Cardiology* 6.1, p. 36.

MOLINA, T. (2016). *Pollution, ability, and gender-specific investment responses to shocks.*

MURALIDHARAN, K., A. SINGH, and A. J. GANIMIAN (2019). “Disrupting education? Experimental evidence on technology-aided instruction in India”. *American Economic Review* 109.4, pp. 1426–60.

MYERS, N. and J. KENT (2004). *The new consumers: The influence of affluence on the environment.* Island Press.

NEL, A. (2005). “Air pollution-related illness: effects of particles”. *Science* 308.5723, pp. 804–806.

OATES, W. E. (2006). “An Economic Perspective on Environmental and Resource”. *The RFF Reader in Environmental and Resource Policy.*

OSSEIRAN, N. and K. CHRISCADEN (2016). *Air pollution levels rising in many of the world’s poorest cities.*

POPE 3RD, C., D. V. BATES, and M. E. RAIZENNE (1995). “Health effects of particulate air pollution: time for reassessment?” *Environmental Health Perspectives* 103.5, pp. 472–480.

PORTA, D., S. NARDUZZI, C. BADALONI, S. BUCCI, G. CESARONI, V. COLELLI, M. DAVOLI, et al. (2016). “Air pollution and cognitive development at age 7 in a prospective Italian birth cohort”. *Epidemiology* 27.2, pp. 228–236.

PSACHAROPOULOS, G. (1985). “Returns to education: a further international update and implications”. *Journal of Human Resources*, pp. 583–604.

PULLABHOTLA, H. (2018). “Fires, Wind, and Smoke: Air Pollution and Infant Mortality”. *Job Market Paper. Available here.*

RANSOM, M. R. and C. A. POPE III (1992). “Elementary school absences and PM10 pollution in Utah Valley”. *Environmental Research* 58.1-2, pp. 204–219.

REARDON, T. and S. A. VOSTI (1995). “Links between rural poverty and the environment in developing countries: asset categories and investment poverty”. *World Development* 23.9, pp. 1495–1506.

ROZELLE, S., J. HUANG, and L. ZHANG (1997). “Poverty, population and environmental degradation in China”. *Food Policy* 22.3, pp. 229–251.

SAGER, L. (2019). “Estimating the effect of air pollution on road safety using atmospheric temperature inversions”. *Journal of Environmental Economics and Management* 98, p. 102250.

SEN, A. (1999). *Freedom as development.*

SHAFIK, N. (1994). “Economic development and environmental quality: an econometric analysis”. *Oxford Economic Papers* 46.4, pp. 757–774.

SHAH, M. and B. M. STEINBERG (2017). “Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital”. *Journal of Political Economy* 125.2, pp. 527–561.

– (2019). “Workfare and human capital investment: Evidence from India”. *Journal of Human Resources*, 1117–9201R2.

SHAW, N. and A. GORAI (2018). “Study of aerosol optical depth using satellite data (MODIS Aqua) over Indian Territory and its relation to particulate matter concentration”. *Environment, Development and Sustainability*, pp. 1–15.

SPEARS, D. and S. LAMBA (2013). *Effects of early-life exposure to sanitation on childhood cognitive skills: evidence from India’s total sanitation campaign.* The World Bank.

STAFFORD, E. R. and R. BRAIN (2015). “My mom idles less than your mom! Empowering high school teens to tackle air pollution”. *Solutions J* 6.6, pp. 48–59.

SUADES-GONZALEZ, E., M. GASCON, M. GUXENS, and J. SUNYER (2015). “Air pollution and neuropsychological development: a review of the latest evidence”. *Endocrinology* 156.10, pp. 3473–3482.

TANAKA, S. et al. (2012). *Environmental regulations on air pollution in China and their impact on infant mortality.* Tech. rep. Hiroshima University, Graduate School for International Development.

TANG, D., T.-y. LI, J. J. LIU, Z.-j. ZHOU, T. YUAN, Y.-h. CHEN, V. A. RAUH, et al. (2008). “Effects of prenatal exposure to coal-burning pollutants on children’s development in China”. *Environmental Health Perspectives* 116.5, pp. 674–679.

TEN HOEVE, J. E. and M. Z. JACOBSON (2012). “Worldwide health effects of the Fukushima Daiichi nuclear accident”. *Energy & Environmental Science* 5.9, pp. 8743–8757.

VAN DONKELAAR, A., R. V. MARTIN, M. BRAUER, and B. L. BOYS (2014). “Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter”. *Environmental health perspectives* 123.2, pp. 135–143.

VARUGHESE, G., K. LAKSHMI, A. KUMAR, N. RANA, et al. (2009). “State of environment report: India, 2009.” *State of Environment Report: India, 2009.*

VENKATARAMAN, C., G. HABIB, D. KADAMBA, M. SHRIVASTAVA, J.-F. LEON, B. CROUZILLE, O. BOUCHER, et al. (2006). “Emissions from open biomass burning in India: Integrating the inventory approach with high-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) active-fire and land cover data”. *Global Biogeochemical Cycles* 20.2.

WANG, S., J. ZHANG, X. ZENG, Y. ZENG, S. WANG, and S. CHEN (2009). “Association of traffic-related air pollution with children’s neurobehavioral functions in Quanzhou, China”. *Environmental Health Perspectives* 117.10, pp. 1612–1618.

XIN, J., Q. ZHANG, L. WANG, C. GONG, Y. WANG, Z. LIU, and W. GAO (2014). “The empirical relationship between the PM2. 5 concentration and aerosol optical depth over the background of North China from 2009 to 2011”. *Atmospheric Research* 138, pp. 179–188.

ZHANG, X., X. CHEN, and X. ZHANG (2018). “The impact of exposure to air pollution on cognitive performance”. *Proceedings of the National Academy of Sciences* 115.37, pp. 9193–9197.

ZWEIG, J. S., J. C. HAM, and E. L. AVOL (2009). “Air pollution and academic performance: Evidence from California schools”. *Department of Economics, University of Maryland.*

# Tables & Figures

Table 1: Summary Statistics ASER Data

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|---|
| **Reading outcomes** | | | | | | | | |
| Can read letters (0/1) | 0.91 | 0.90 | 0.92 | 0.91 | 0.89 | 0.87 | 0.86 | 0.85 |
| Can read words (0/1) | 0.76 | 0.76 | 0.76 | 0.76 | 0.73 | 0.69 | 0.68 | 0.68 |
| Can read grade 1 text (0/1) | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| Can read grade 2 text (0/1) | 0.45 | 0.46 | 0.45 | 0.45 | 0.43 | 0.42 | 0.42 | 0.43 |
| **Math outcomes** | | | | | | | | |
| Can count 1 to 9 (0/1) | 0.91 | 0.91 | 0.92 | 0.91 | 0.90 | 0.88 | 0.88 | 0.87 |
| Can count 10 to 99 (0/1) | 0.76 | 0.75 | 0.76 | 0.75 | 0.72 | 0.69 | 0.69 | 0.69 |
| Can do subtration (0/1) | 0.56 | 0.55 | 0.57 | 0.56 | 0.51 | 0.45 | 0.43 | 0.42 |
| Can do division (0/1) | 0.35 | 0.34 | 0.36 | 0.35 | 0.29 | 0.25 | 0.24 | 0.24 |
| **Individual characteristics** | | | | | | | | |
| Age (in years) | 10.09 | 10.24 | 10.34 | 10.29 | 10.33 | 10.34 | 10.36 | 10.39 |
| Male (0/1) | 0.54 | 0.54 | 0.55 | 0.55 | 0.53 | 0.53 | 0.52 | 0.52 |
| Household size | 5.92 | 5.98 | 5.84 | 5.84 | 5.92 | 5.96 | 5.97 | 5.94 |
| Mother attended school (0/1) | 0.45 | 0.43 | 0.46 | 0.51 | 0.49 | 0.48 | 0.49 | 0.50 |
| Number of Students | 515344 | 511647 | 476506 | 449129 | 425591 | 376057 | 357547 | 345600 |

Note: The sample includes all districts in the ASER data for the years 2007 - 2014.

Table 2: Summary Statistics YLS Data

| | Round 1 - 2007 | Round 2 - 2009 | Round 3 - 2013-14 |
|---|---|---|---|
| Currently enrolled (0/1) | 0.94 | 1.00 | 1.00 |
| Math raw score | 9.41 | 12.09 | 12.79 |
| PPVT raw score | 27.23 | 57.56 | 43.25 |
| Female (0/1) | 0.47 | 0.47 | 0.47 |
| Math score (0/1) | 0.67 | 0.42 | 0.46 |
| PPVT score (0/1) | 0.23 | 0.31 | 0.76 |
| Age (in months) | 64.21 | 95.21 | 143.75 |
| Household size | 0.00 | 0.00 | 0.00 |
| Wealth index | 0.45 | 0.50 | 0.57 |
| Consumer durables index | 0.22 | 0.31 | 0.37 |
| Number of Students | 1795 | 1795 | 1795 |

Note: The sample includes data from 6 districts in Andhra Pradesh that were part of the Young Lives sample in Survey Rounds 2007, 2009, and 2013-14. We include only the younger cohort and only those children who were tested thrice in both math and peabody picture vocabulary test (PPVT).

Table 3: Summary of AOD and Weather

| | Mean | S.D. | Min | Max | N |
|---|---|---|---|---|---|
| AOD | 0.42 | 0.16 | 0.06 | 0.86 | 2420 |
| Number of inversions | 69.29 | 75.79 | 0.00 | 264.00 | 2420 |
| Inversion strength | -1.66 | 1.42 | -4.53 | 1.78 | 2420 |
| Wind Velocity (m/s) | 0.76 | 0.42 | 0.02 | 2.91 | 2420 |
| Precipitation (cm/month) | 11.03 | 7.86 | 0.15 | 107.08 | 2420 |
| Temperature (degrees celsius) | 24.84 | 3.80 | -2.20 | 30.08 | 2420 |
| Relative humidity | 64.44 | 5.51 | 49.55 | 79.57 | 2420 |

Note: The unit of observation is the district-year. The sample includes all districts in the ASER data for the years 2007 - 2014. AOD data are from MODIS. Wind velocity data and relative humidity are from ERA-Interim Reanalysis. Precipitation is from the University of Delaware. Inversions are calculated using the NCEP Re-Analysis Data.

Table 4: OLS Estimates - Impact of Current Year Pollution on Human Capital

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD** | -0.0031 | -0.0378 | -0.0095 | 0.0294 | 0.0039 | -0.02 | -0.0738* | -0.0692 |
| | (0.0231) | (0.0313) | (0.0344) | (0.0371) | (0.0224) | (0.0345) | (0.0418) | (0.0429) |
| N | 3000234 | 3000234 | 3000234 | 3000234 | 2989716 | 2989716 | 2989716 | 2989716 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is current year AOD. All specifications include district, year, age fixed effects and state specific linear time trends; contols for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 5: First Stage Relationship between AOD and Inversions

| Current Year AOD | (1) | (2) |
|---|---|---|
| **Number of Inversions** | 0.0002*** | 0.0002*** |
| | (0.00005) | (0.00004) |
| Observations | 3,180,602 | 3,180,602 |
| First Stage F-Statistic | 17.56 | 18.91 |
| Stock Yogo Critical Value | 16.38 | 16.38 |
| **Inversion Strength** | 1.3173*** | 1.3483*** |
| | (0.3218) | (0.3504) |
| Observations | 3,180,602 | 3,180,602 |
| First Stage F-Statistic | 16.76 | 14.81 |
| Stock Yogo Critical Value | 16.38 | 16.38 |
| Weather controls | Yes | Yes |
| District FE | Yes | Yes |
| Year FE | Yes | Yes |
| Individual, Village controls | Yes | Yes |
| State time linear trends | No | Yes |

Note: Weather controls include controls for temperature, rainfall, wind velocity, and humidity and their squares. Individual controls include controls for female, household size, and mother's education. Robust standard errors in parenthesis clustered at the district level. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 6: Effect of Inversions on Future AOD

| Future AOD | (1) | (2) |
|---|---|---|
| **Number of Inversions** | 0.0003 | 0.0004 |
| | (0.0002) | (0.0002) |
| Observations | 3,071,889 | 3,071,889 |
| First Stage F-Statistic | 1.38 | 2.6 |
| Stock Yogo Critical Value | 16.86 | 16.86 |
| Weather controls | Yes | Yes |
| District FE | Yes | Yes |
| Year FE | Yes | Yes |
| Individual controls | Yes | Yes |
| State time linear trends | No | Yes |

Note: Weather controls include controls for temperature, rainfall, wind velocity, and humidity and their squares. Individual controls include controls for female, household size, and mother's education. Robust standard errors in parenthesis clustered at the district level. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 7: IV Estimates - Impact of Current Year Pollution on Human Capital

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD** | -0.5764* | -1.3905** | -1.6740*** | -1.4058** | -0.2336 | -1.3751** | -1.4352** | -1.1577* |
| | (0.3503) | (0.5490) | (0.6324) | (0.6419) | (0.3029) | (0.5580) | (0.6615) | (0.6374) |
| N | 3000234 | 3000234 | 3000234 | 3000234 | 2989716 | 2989716 | 2989716 | 2989716 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is current year AOD. All variables including current year AOD are expressed as %. All specifications include district, year, age fixed effects and state specific linear time trends; contols for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 8: IV Estimates - Heterogeneous Effects by Gender and Age

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Panel A: Boys 6-10** | | | | | | | | |
| Current Year AOD | -1.0910** | -1.8968** | -1.2545* | -0.6112 | -0.5728 | -1.6264** | -0.622 | 0.0576 |
| | (0.5206) | (0.7758) | (0.7051) | (0.5969) | (0.4302) | (0.7495) | (0.6754) | (0.5236) |
| N | 757647 | 757647 | 757647 | 757647 | 754422 | 754422 | 754422 | 754422 |
| **Panel B: Girls 6-10** | | | | | | | | |
| Current Year AOD | -0.9444* | -2.3802*** | -2.3295*** | -1.3097** | -0.5749 | -2.1451** | -1.4912* | -0.3854 |
| | (0.5503) | (0.8803) | (0.8851) | (0.6616) | (0.4829) | (0.8666) | (0.7792) | (0.4932) |
| N | 649757 | 649757 | 649757 | 649757 | 647048 | 647048 | 647048 | 647048 |
| **Panel C: Boys 11-16** | | | | | | | | |
| Current Year AOD | -0.5279*** | -0.9874*** | -1.5856*** | -1.8258** | -0.2493* | -1.0130*** | -1.6657** | -1.8477** |
| | (0.1834) | (0.3747) | (0.5990) | (0.8321) | (0.1479) | (0.3879) | (0.7283) | (0.9415) |
| N | 744881 | 744881 | 744881 | 744881 | 743364 | 743364 | 743364 | 743364 |
| **Panel D: Girls 11-16** | | | | | | | | |
| Current Year AOD | -0.3643 | -1.3184** | -2.4232*** | -2.4765** | -0.2744 | -1.4190** | -2.5270** | -2.5009** |
| | (0.2251) | (0.5178) | (0.8447) | (1.0431) | (0.2099) | (0.5804) | (1.0108) | (1.1645) |
| N | 657151 | 657151 | 657151 | 657151 | 655633 | 655633 | 655633 | 655633 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. All specifications include district, year, and state specific linear time trends; contols for precipitation, temperature, wind velocity, humidity and their squares; and controls for household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 9: IV Estimates - Effects for Polluted States vs. Other States

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD*Polluted States** | -0.8970* | -2.1137** | -2.6545*** | -2.4259*** | -0.3944 | -1.9680** | -2.6519** | -2.2360** |
| | (0.4968) | (0.8680) | (0.9639) | (0.9284) | (0.4016) | (0.8401) | (1.0333) | (0.9717) |
| **Current Year AOD*Other States** | -0.0169 | -0.1283 | 0.0372 | 0.3746 | 0.0458 | -0.3451 | 0.6784 | 0.7155 |
| | (0.3638) | (0.4985) | (0.6544) | (0.3393) | (0.5130) | (0.5130) | (0.6832) | (0.7020) |
| N | 2318084 | 2318084 | 2318084 | 2318084 | 2310640 | 2310640 | 2310640 | 2310640 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 to 2014. All specifications include district, year, age fixed effects and state specific linear time trends; contols for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, and mother's education. PollutedStates is a dummy equal to one for districts in polluted states (Rajasthan, Punjab, Haryana, Uttar Pradesh, Bihar, and West Bengal). Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 10: First Stage Relationship - YLS Data

| AOD | (1) | (2) |
|---|---|---|
| **Number of Inversions** | 0.0002*** | 0.0002*** |
| | (0.00006) | (0.00004) |
| Observations | 4,792 | 4,792 |
| First Stage F-Statistic | 20.31 | 23.31 |
| Stock Yogo Critical Value | 16.86 | 16.86 |
| Weather controls | Yes | Yes |
| Child FE | Yes | Yes |
| Survey Wave FE | Yes | Yes |
| Month and Day of Test FE | No | Yes |

Note: The unit of observation is student-year. Sample includes Young Lives Survey Rounds 2007, 2009, and 2013 for the younger cohort. The independent variable is current year AOD. All specifications include survey wave, child, month of test, and day of test fixed effects; contols for precipitation, temperature, humidity, and wind velocity; and controls for wealth index, consumer durables index, and indicator for debt. Standard errors are in parentheses, clustered by district-week. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 11: IV Estimates using YLS Data

| | PPVT Score | Math Score |
|---|---|---|
| **Panel A: Boys** | | |
| Current Year AOD | -5.3146** | -3.6123 |
| | (2.5502) | (2.6586) |
| N | 2571 | 2540 |
| **Panel B: Girls** | | |
| Current Year AOD | -4.5764** | -7.2856** |
| | (2.0198) | (2.9635) |
| N | 2238 | 2199 |

Note: The unit of observation is student-year. Sample includes Young Lives Survey Rounds 2007, 2009, and 2013 for the younger cohort. The independent variable is current year AOD. All specifications include survey wave, child, month of test, and day of test fixed effects; contols for precipitation, temperature, humidity, and wind velocity; and controls for wealth index, consumer durables index, and indicator for debt. Standard errors are in parentheses, clustered by district-week. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 12: IV Estimates on Extensive Margin

| | Dropped Out | On Track |
|---|---|---|
| **Current Year AOD** | 0.0758 | -1.1383*** |
| | (0.1252) | (0.3902) |
| Observations | 3315347 | 2910649 |
| **Cumulative Exposure Last 2 years** | 0.0252 | -0.9228*** |
| | (0.1169) | (0.3142) |
| Observations | 3314569 | 2909946 |
| **Cumulative Exposure Last 3 years** | -0.0024 | -0.6154 |
| | (0.1599) | (0.03942) |
| Observations | 3684415 | 3230000 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. All specifications include district, year, age fixed effects and state specific linear time trends; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 13: IV Estimates - Cumulative Exposure Last Two Years

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Cumulative Exposure Last 2 Years** | -0.7165** | -1.1881*** | -1.4553*** | -1.3873*** | -0.3542 | -1.0654** | -1.9925*** | -1.6534*** |
| | (0.3026) | (0.4415) | (0.4774) | (0.5174) | (0.2524) | (0.4488) | (0.6159) | (0.5932) |
| N | 3329468 | 3329468 | 3329468 | 3329468 | 3317156 | 3317156 | 3317156 | 3317156 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is cumulative AOD exposure in the last two years. All specifications include district, year, age fixed effects and state specific linear time trends; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 14: IV Estimates - Cumulative Exposure Last Three Years

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Cumulative Exposure Last 3 Years** | -0.7989* | -1.4543** | -1.9954*** | -1.8276** | -0.4526 | -1.6199** | -2.7386*** | -1.4953* |
| | (0.4344) | (0.6273) | (0.7213) | (0.7660) | (0.3614) | (0.6907) | (0.9865) | (0.8682) |
| N | 3328586 | 3328586 | 3328586 | 3328586 | 3316276 | 3316276 | 3316276 | 3316276 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is cumulative AOD exposure in the last three years. All specifications include district, year, age fixed effects and state specific linear time trends; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 15: IV Estimates on Student and Teacher Attendance

| | Attendance All Grades % | Attendance I-V % | Attendance VI-VIII % | Attendance Teachers % |
|---|---|---|---|---|
| **Current Year AOD** | -1.0952*** | -1.4153*** | -0.5121 | -1.2633 |
| | (0.3444) | (0.3589) | (0.3732) | (0.9196) |
| Observations | 33389 | 33022 | 28461 | 14030 |
| Number of Villages | 12368 | 12271 | 10708 | 6238 |
| **Cumulative Exposure Last 2 years** | -4.2210*** | -4.9555*** | -3.1350*** | -3.9874** |
| | (0.9916) | (1.0431) | (1.0750) | (1.8694) |
| Observations | 33361 | 32993 | 28431 | 14012 |
| Number of Villages | 12367 | 12268 | 10705 | 6233 |
| **Cumulative Exposure Last 3 years** | -1.1109** | -1.2081*** | -1.0852* | -5.2946 |
| | (0.4512) | (0.4663) | (0.6083) | (4.0865) |
| Observations | 33332 | 32968 | 28407 | 13986 |
| Number of Villages | 12359 | 12262 | 10699 | 6220 |

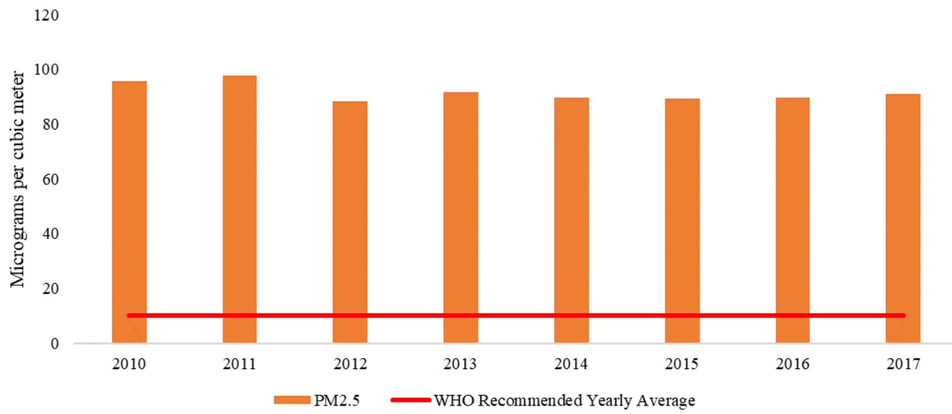
Note: The unit of observation is school-year. Sample includes ASER school level data for 2009 - 2014. The independent variable is current year AOD. All specifications include district, year, and school fixed effects. Standard errors are in parentheses, clustered by school. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table 16: IV Estimates - Attendance using NSS Data

| | Attends School | Number of Days in Previous Week |
|---|---|---|
| **Current Year AOD** | -0.6922* | -4.7477* |
| | (0.4095) | (2.8920) |
| Observations | 232587 | 232587 |
| **Cumulative Exposure Last 2 years** | -1.1618** | -6.8820** |
| | (0.5066) | (3.3663) |
| Observations | 232473 | 232474 |
| **Cumulative Exposure Last 3 years** | -1.1019** | -7.3532** |
| | (0.4978) | (3.4869) |
| Observations | 356310 | 356310 |

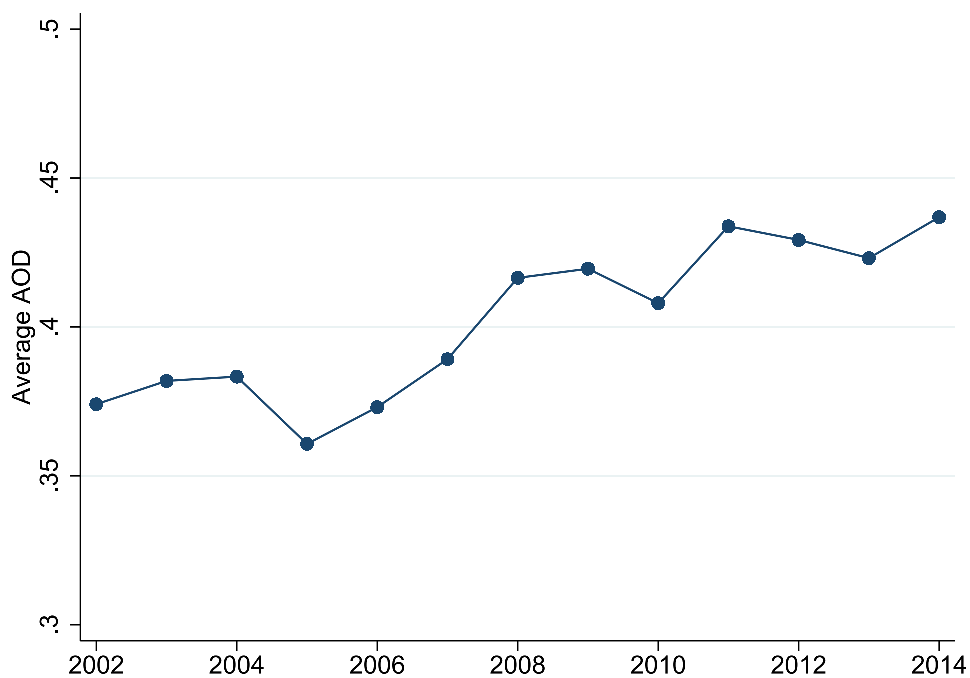
Note: The unit of observation is student-year. Sample includes NSS rounds 61, 62, 64, 66, and 68. All specifications include district, year, age fixed effects; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for household size, gender, caste, and religion . Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Figure 1: PM2.5 Levels in India (in micrograms per cubic meter)



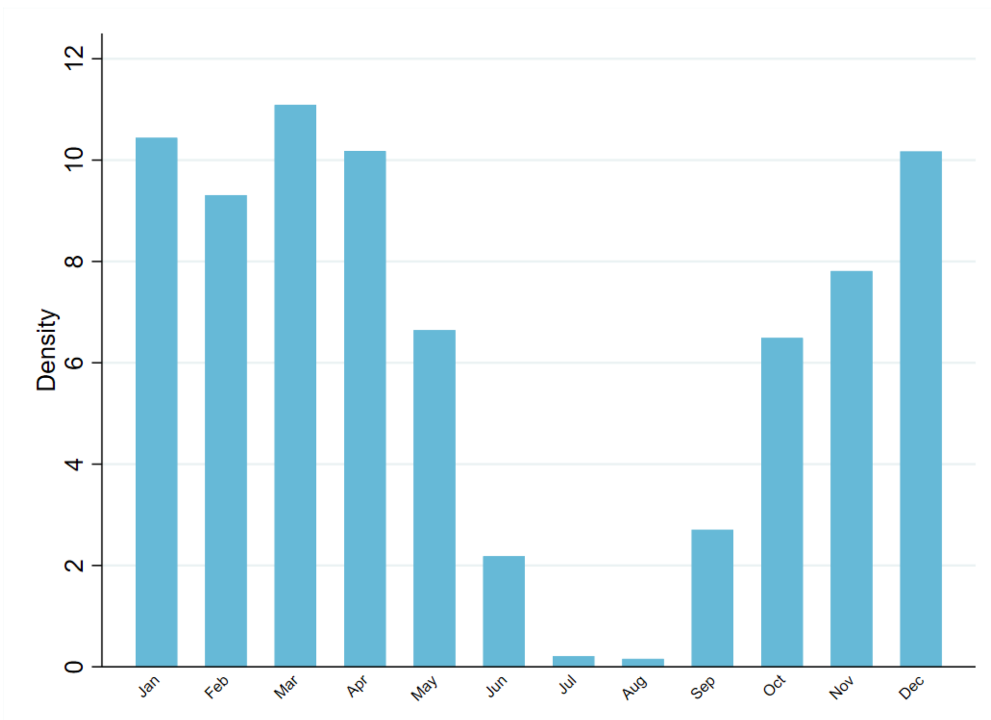

*Note:* This figure plots the trend in PM2.5 levels in micrograms per cubic meter from 2010 - 2014 from the World Bank (2018).

Figure 2: Trend in AOD by Year



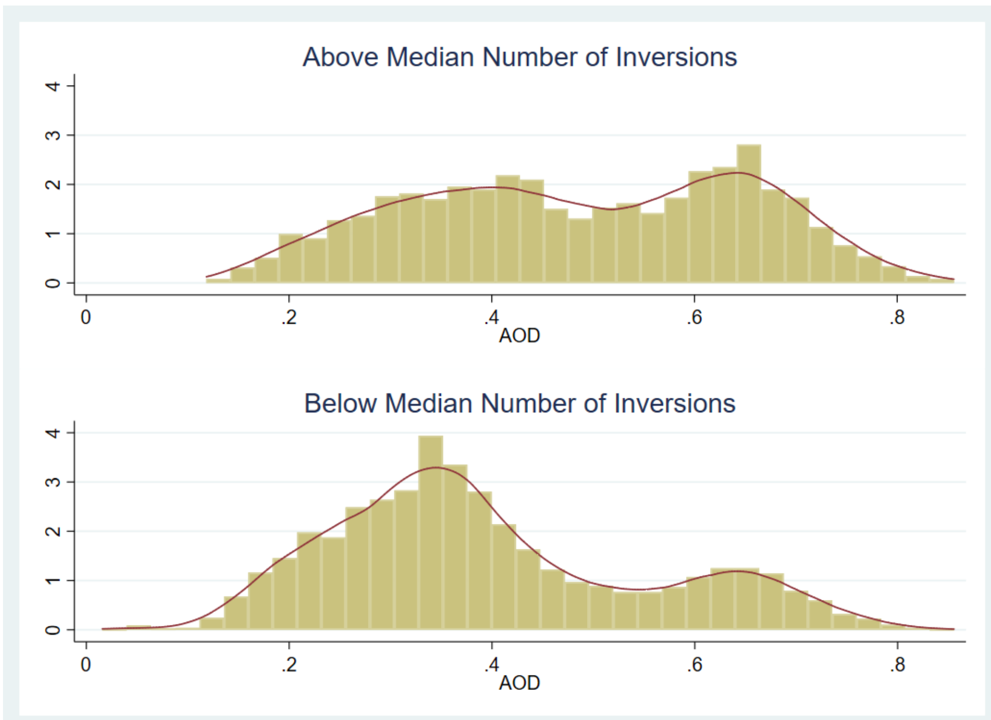

*Note:* This figure plots the trend in AOD from 2002 - 2014 for all of India.

Figure 3: Number of Inversions by Month and Year




*Note:* The unit of observation is the district-month. The sample includes all districts in the ASER data for the years 2007 - 2014.

Figure 4: Distribution of AOD by High and Low Inversion Districts (using Number of Inversions)




*Note:* The unit of observation is a district-year. The sample includes all districts with ASER data from 2007 - 2014. The plot shows a histogram of the district annual mean AOD for (i) district-years above the median number of inversions and (ii) district-years below the median number of inversions.

# A Robustness Checks

Table A1: Robustness - On-Track Kids

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD** | -0.4897 | -1.2478** | -1.4354** | -1.1947* | -0.2906 | -1.1521** | -1.2348* | -0.9339 |
| | (0.3174) | (0.5228) | (0.6026) | (0.6481) | (0.2805) | (0.5288) | (0.6630) | (0.6527) |
| N | 2318084 | 2318084 | 2318084 | 2318084 | 2310640 | 2310640 | 2310640 | 2310640 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 to 2014 and only includes on-track kids i.e. children in the correct grade for age. The independent variable is current year AOD. All specifications include district, year, age fixed effects and state specific linear time trends; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table A2: Robustness - Alternate Inversion Layers

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD** | -0.6806*** | -0.8784*** | -0.6122** | -0.7979** | -0.7463*** | -0.9629*** | -1.1438*** | -0.8599** |
| | (0.2397) | (0.3039) | (0.3099) | (0.3502) | (0.2433) | (0.3385) | (0.4162) | (0.3786) |
| N | 3000234 | 3000234 | 3000234 | 3000234 | 2989716 | 2989716 | 2989716 | 2989716 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is current year AOD. Inversion episodes are defined between the first (1000hPa) and third (875hPa) pressure layers. All specifications include district, year, age fixed effects and state specific linear time trends; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table A3: Robustness - September to September

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD** | -0.5347* | -1.0655** | -1.1392** | -0.6929 | -0.1582 | -0.9640** | -1.4713*** | -1.0885* |
| | (0.2938) | (0.4147) | (0.4619) | (0.4890) | (0.2655) | (0.4324) | (0.5704) | (0.5609) |
| N | 3001098 | 3001098 | 3001098 | 3001098 | 2990582 | 2990582 | 2990582 | 2990582 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is current year AOD. 'Current Year' is defined as the last 12 months (i.e. from September to September) for all variables. All specifications include district, year, age fixed effects and state specific linear time trends; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table A4: Robustness - No State-Time Linear Trends

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD** | -0.8996** | -1.8861*** | -1.9083** | -1.4101** | -0.0511 | -1.3757** | -0.7484 | -0.4185 |
| | (0.4016) | (0.7064) | (0.7535) | (0.7018) | (0.3219) | (0.6506) | (0.7288) | (0.6757) |
| N | 3000234 | 3000234 | 3000234 | 3000234 | 2989716 | 2989716 | 2989716 | 2989716 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is current year AOD. All specifications include district, year, and age fixed effects; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

Table A5: Robustness - Season Fixed Effects

| | Letter | Word | Level 1 | Level 2 | Numbers 1-9 | Numbers 10-99 | Subtraction | Division |
|---|---|---|---|---|---|---|---|---|
| **Current Year AOD** | -0.5745 | -1.3896** | -1.6707*** | -1.6707*** | -0.2335 | -1.3754** | -1.4311** | -1.1599* |
| | (0.3506) | (0.5492) | (0.6325) | (0.6418) | (0.3036) | (0.5583) | (0.6624) | (0.6382) |
| N | 3000234 | 3000234 | 3000234 | 3000234 | 2989716 | 2989716 | 2989716 | 2989716 |

Note: The unit of observation is student-year. Sample includes ASER years from 2007 - 2014. The independent variable is current year AOD. All specifications include season, district, year, age fixed effects and state specific linear time trends; controls for precipitation, temperature, wind velocity, humidity and their squares; and controls for female, household size, and mother's education. Standard errors are in parentheses, clustered by district. *Significant at 10%. **Significant at 5%. ***Significant at 1%.