Coal Plants, Air Pollution and Anaemia: Evidence from India*

Gaurav Datt[a], Pushkar Maitra[b], Nidhiya Menon[c] and Ranjan Ray[b]

Version: September 2022

Abstract

We examine the impact of pollution from coal–fired power units on the anaemic status of children and women in India. The number of coal units in the district at the time of birth significantly increases the incidence of anaemia in young children as does *in utero* exposure. The number of coal units in the district also adversely affects the anaemic status of women, although the magnitude of impact is smaller than that for young children. The impacts are driven by the increase in $PM_{2.5}$ pollution generated by coal-fired units. Our evidence points to anaemia as a significant health cost of coal-fired power generation in rapidly growing economies that use coal as a major source of fuel to meet increasing energy demands.

Key Words: Anaemia, Coal Units, $PM_{2.5}$, Children, Women, India

JEL Codes: I15, Q32, Q53, O12

Disclosure Statement: *The authors report there are no competing interests to declare*

Word Count: 8,582

---

* We thank the Editor, two anonymous referees, Teevrat Garg, Olexiy Kyrychenko, Subha Mani, Nitya Mittal, Dean Spears, Rahul Tongia, conference participants at AHES, PAA, EAERE, AMES, CMES, NEUDC, Delhi School of Economics and the Econometric Society Winter School, Monash Environmental Economics Workshop and INSEE Conference, seminar participants at the Harvard Social Demography Seminar, the Center for Health Economics at Monash University, and the Institute of Economic Growth, Delhi for their comments and suggestions. We thank Teevrat Garg for sharing the data on air pollutants and Sagnik Dey and Sourangsu Chowdhury for providing us with the data on $PM_{2.5}$. The usual disclaimer applies. Data and code will be made available to bona fide researchers on request.

[a]: Center for Development Economics and Sustainability, Monash University, Australia.
[b]: Department of Economics, Monash University Australia.
[c]: Department of Economics, Brandeis University, USA.

1. Introduction

The detrimental impact of air pollution on various measures of health, especially in early life, is now fairly well documented (see Currie *et al.*, 2009, Jayachandran, 2009, Coneus and Spiess, 2012, Graff Zivin and Neidell, 2013, Currie *et al.*, 2014, Greenstone and Hanna, 2014, Luechinger, 2014, Greenstone and Jack, 2015, Bharadwaj, *et al.*, 2016, Bharadwaj, *et al.*, 2017, Anderson, 2020). While some of the early evidence on the harm to human health from air pollution came from developed countries, recent literature using data from developing countries paints an even bleaker picture. This is particularly true of India. According to the WHO global air pollution data base, of the 20 most polluted cities in the world over the period 2008–2017, 13 are in India. Pollution is widespread and growing as is its health and economic burden. ICMR (2020) estimates that almost 18 percent of deaths in India in 2019 can be attributed to pollution, and premature deaths and morbidity due to pollution resulted in lost output equivalent to 1.36 percent of India's GDP.

We investigate the effect of air pollution on anaemia, a health outcome that is relatively under–researched. This is despite the relevance of this problem for India where anaemia prevalence rates are high. We go a step further and document the chain from coal–fired power plants to air pollution to anaemia. The two sub–links in this chain (coal plants → ambient air pollution and ambient air pollution → anaemia) have been separately studied in the literature, as has the link between coal and more general measures of health. A distinguishing feature of this paper is its joint investigation of the two sub–links in the chain in a relevant developing country context.

Our focus is on iron–deficiency (ID) anaemia, which is the most common form of anaemia. An ID anaemic (henceforth anaemic) individual does not have enough haemoglobin, a protein that enables red blood cells to carry oxygen to other cells in the body. Consequences of anaemia include lack of energy, impaired growth, and a weakened immune system, which decreases the body's ability to prevent infections. Further, anaemia retards cognitive abilities and performance. In children, this often leads to attention deficiency, delays in reading ability, and poor school performance (Maluccio *et al*., 2009, Currie and Vogl, 2013, Chong *et al., 2016*); impairment of physical and cognitive development (Grantham-McGregor and Ani, 2000); and child mortality (Scott *et al*., 2014). Among adults, evidence establishes ID anaemia to be a leading cause of maternal mortality (Rush, 2000) and premature deaths in pregnant women (Khaskheli *et al.,* 2016); and can have significant adverse effects on productivity (Thomas *et al*., 2004). Through its impact on health and educational outcomes, anaemia plays a role in the inter–generational transmission of poverty (Bobonis *et al*., 2006). Macroeconomic estimates suggest that the average impact of ID anaemia, through physical and cognitive channels, could be as large as 4 percent of GDP (Horton and Ross, 2003).

The medical/epidemiological literature has documented the connection between air pollution and anaemia. Among children in particular, such pollution reduces the formation of red blood cells by inhibiting enzymes. Further, air pollutants destroy cell membranes and interfere with cell metabolism thus reducing cell survival (Nikolic *et al*., 2008). Among the elderly, there is evidence that $PM_{2.5}$ reduces levels of haemoglobin and increases levels of c–reactive proteins (markers of inflammation) that retard the absorption of iron by the body

(Honda *et al*., 2017, Elbarbary *et al*., 2019).

The other sub–link between coal–based power generation and ambient air pollution is also established. Lancet (2018) documents that global population weighted $PM_{2.5}$ exposure has increased by 11.2 percent, and that coal is responsible for more than 90 percent of $PM_{2.5}$ emissions, much more than its corresponding contributions to sulphur dioxide (SO_2) and nitrogen dioxide (NO_2) emissions. As $PM_{2.5}$ is the main pollutant generated by coal, we focus on it in our analysis, though we also conduct secondary checks with SO_2 and NO_2.

In sum, while the link from coal–fired power to anaemia mediated through air pollution appears plausible, to the best of our knowledge, there has been no prior investigation that simultaneously evaluates these sub–links in a developing country context. This is our main contribution.

The relevance of India to this topic cannot be overemphasized. Anaemia, among children and adults, is a major health problem. According to the World Bank, the global anaemia rates in 2016 for children under 5 years and for women ages 15–49 were 42 percent and 33 percent, respectively. The corresponding figures for India were substantially higher at 59 percent and 54 percent, underlining the significantly greater severity of this problem in the country. Anaemia rates for India have exceeded the world average since 1900.

Growing levels of ambient air pollution are a major policy concern in India, and the contribution of coal–fired power generation to such pollution has been large. As the country has grown economically, demand for energy has risen. Like other developing countries, India has addressed the need for energy largely by building additional capacity in coal–fired power generation, adding new units (new turbines/generation unit) to existing power plants that use coal as the primary fuel (henceforth coal plant units or coal units). The growth of coal–based power has out–run the growth of any other form of electricity generation; coal is now the primary fuel for almost 75 percent all electricity generated in India (Ali, 2008). However, coal units in India often do not meet emission standards or other guidelines that are commonplace in the developed world. This has exacerbated the health, economic and environmental consequences of relying on this mode of energy generation. For instance, Cropper *et al*. (2020) estimates that coal plants that existed as of 2018 in India were responsible for approximately 78,000 premature deaths (annually) associated with $PM_{2.5}$, and the true negative externalities are likely to be higher when morbidity and damage to environmental systems and agriculture are taken into account.

To analyse the detrimental effects of coal–fired power generation and the associated pollution on anaemia, we assembled a unique data set combining information on the anaemic status of women and children from the Fourth National Family Health Survey (NFHS–4) conducted in 2015–16, district and time wise information on $PM_{2.5}$, and the district–wise location of energy units in India demarcated by primary fuel source. We leverage the spatial variation in the location of coal units at a point in time, and construct measures of change in the spatial variation in coal units overtime, to examine the effects of coal induced air pollution on anaemia while simultaneously controlling for a wide range of other covariates.

Our main finding is that coal units exert harmful consequences on anaemia. For children aged 0–5, the number of coal units in the district at the time of birth is associated with a significant increase in the likelihood of being anaemic.[1] These results are stronger for girls, those born to less educated mothers, and those residing in rural areas. We conduct specification and falsification tests, and control for a range of factors that could affect the relationship (these include rainfall, temperature, district prosperity as measured by nightlights, and a wide range of child, mother and household attributes including diet) to demonstrate that the results may be plausibly attributed to coal. Analysing periods that precede birth, we find that exposure *in utero* matters. For women aged 18–49 years, our results show that the number of coal units in the district at the time of the survey significantly increases their likelihood of being anaemic. Taken together, our findings underline that the expedient generation of coal–fired power is occurring at the significant expense of exacerbating anaemia among children and women.

The harmful health consequences of coal on at–risk populations are a relatively well researched topic in the developed world in the epidemiology literature (Honda *et al*., 2017, Amster and Levy, 2019). In developing countries, the focus has been on the adverse effects of coal–fired power units on general health indicators, morbidity, and mortality (Gupta and Spears, 2017, Gao *et al*., 2018, Morales-Ancajima *et al*., 2019, Barrows *et al*., 2019, Vyas, 2019, ICMR, 2020). We add anaemia as an additional health concern. Our results are consistent with the epidemiological evidence on the association between ambient pollution and anaemia in India (Mehta *et al.,* 2020). However, we go significantly further in understanding this relationship. First, we link $PM_{2.5}$ to emissions from coal plant units.

Second, we control for a larger set of observable covariates and for unobservables that may vary by district, month of birth, and year of birth (and their interactions). Third, we investigate differences in the timing of exposure (*in utero*, at birth, post–birth), and fourth, we consider impacts on adult women. Finally, we conduct a range of falsification tests and robustness checks including those related to selection that may arise due to differential sorting and migration.

In addition to providing evidence on the link from coal plant units to air pollution to anaemia, our paper adds to the literature on the determinants of anaemia. Previous literature in this area has focused on the role of poor nutrition (Bhattacharya *et al.*, 2004, Thomas *et al.*, 2004, Krämer *et al.*, 2021), malaria (Sachs and Malaney, 2002), and hygiene and sanitation including open defecation (Coffey *et al.*, 2017). We add an additional contributing factor: air pollution from coal fired power plants.

2. Data

Our data on anaemia prevalence and individual and household variables of interest come from the NFHS–4, conducted in 2015–2016. It is a large nationally representative (one-time) household survey with detailed information on individual and household characteristics for women aged 15–49, children aged 0–5 years, and men aged 15–54 years.

The location of power units in India as of 2016 is obtained from the Central Electricity Authority (Ministry of Power) of India (GOI, 2018). The data on $PM_{2.5}$ at the district–month–year level from 2001 to 2015 are obtained from satellite measurement estimates generated from aerosol optical depth information collected using techniques developed in Dey *et al*., 2012. We describe our data in detail below.

2.1. Measures of Anaemia and Control Variables from NFHS–4

Our outcome variable is anaemia, defined as altitude adjusted haemoglobin (HBA) level below a certain threshold.[2] Blood specimens were collected after receiving consent from eligible children and women, where a drop of blood was obtained from a finger prick (or heel prick for children 6–11 months) and analysis was conducted on-site with a portable haemoglobin analyser (more details are in IIPS, 2015). Anaemia in children six months through 60 months of age is defined as HBA below 11.0 g/dl. While anaemia has similar debilitating impacts on adult women, the threshold is slightly higher: an adult woman is anaemic if HBA is below 12.0 g/dl.[3] Table 1 presents the average levels of HBA and anaemia prevalence rates for the different sub–groups that we consider: children aged 0–5 years and women aged 18–49. Close to 3 out of 5 children aged 0–5 are anaemic. The corresponding rate is 54 percent for women. We consider moderately or severely anaemic (HBA < 9.9 g/dl for both children and women) as a separate category to document effects that are likely to be especially dire.[4]

[Table 1 near here]

Panels A and B of Figure 1 presents heat maps of anaemia rates for children and women respectively across India's districts. The darker shaded districts represent those with higher anaemia rates. There is a similarity between the two spatial maps of anaemia: areas of high (low) female anaemia are also those with high (low) rates of child anaemia. Districts with red dots denote those with at least one coal unit. Patterns in these figures suggest that coal units may exert a systematic influence on anaemia in children and women.[5]

[Figure 1 near here]

2.2. Measures of Coal and Other Power Units

Over the last three decades, India has experienced rapid economic growth which has contributed to a substantial increase in electricity demand rising from 388 terra watt hours (TWh) in 2001 to 989 TWh in 2015, a jump of 150 percent (Ali, 2018). Given its cost–effectiveness, the country has built coal–fired power units to meet this demand. The data on power units and capacity obtained from the Central Electricity Authority provide the year of commission of all units generating power (from 1922–2016), the type of these units (thermal, hydro, or nuclear), the main source of fuel for these units (coal, diesel, oil, gas, lignite, or naphtha), the capacity of each unit, and the district of location of each unit. We denote thermal units that use coal as their primary fuel source as coal units. Figure 2 shows that while total capacity from hydro powered units exhibits a modest increase from 2000 onwards; total capacity generated from coal units has increased exponentially, especially in the last 10 years. Figure 3 presents the number of new units (Panel A) and the corresponding new capacity (Panel B) commissioned in each year between 2000 and 2016. Since 2006, there has been a heavy reliance on coal.[6] Approximately 50.2 percent of children 0–5 years have at least one coal unit in their district at the time of birth, and 56.6 percent of women 18–49 years have at least one coal unit in their district at the time of the survey.

[Figure 2 near here]

[Figure 3 near here]

Coal–fired power units burn coal to heat water that produces steam, which, in turn, drives the turbines that generate electricity. These units therefore require water and hence are typically situated near sources of water and coal deposits. Much of the coal used in generating power in India is domestic coal that is of low quality; it has high ash and moisture content but low sulphur and calorific content. Consequently, Indian coal generates relatively low levels of emissions of sulphur dioxide (SO_2), but high levels of nitrogen dioxide (NO_2) and fine particulate matter ($PM_{2.5}$). See Barrows *et al.* (2019).

2.3. *Measure of $PM_{2.5}$*

Panel C of Figure 1 depicts the concentration of average $PM_{2.5}$ across the districts of India. The intensity of $PM_{2.5}$ pollution is higher in districts with at least one coal power plant. Pair–wise correlation coefficients affirm a positive and significant relationship between the number of coal units and the natural log of $PM_{2.5}$ in the district at the time of birth (correlation coefficient = 0.17, p–value < 0.05), and between anaemia in children aged 0–5 years and the natural log of $PM_{2.5}$ (correlation coefficient = 0.08, p–value < 0.05).

2.4. *Summary Statistics*

Tables A1 and A2 (available in Supplementary Materials) present summary statistics of demographic variables in our sample of children and women, along with corresponding statistics for the power and pollution–specific variables. Focusing on the key variables, the average age of children in the sample is about 2.5 years, and 52 percent are male. The vast majority (79 percent) of these children were delivered in public sector hospitals. The average age of women in our sample is about 34 years, and average height (a measure of their health endowment) is approximately 152 centimetres. Over 36 percent of the women in these data are uneducated, and only 15 percent have completed secondary school or higher. About a third of the women's sample consumes meat or fish on a regular basis (daily or weekly), while the proportions reporting regular consumption of eggs, green vegetables, fruits, beans, and yoghurt are higher.

Table A1 and Table A2 also report summary statistics of the power plant unit and pollutant variables. The average number of coal units in the district in the month and year of birth of the child is about three. While the number of hydro units are comparable, there are significantly fewer nuclear units. The estimates for mean $PM_{2.5}$ ranges from 30–37 micrograms per cubic meter of air, about three times the World Health Organization's target $PM_{2.5}$ guideline level.[7] As noted in Barrows *et al.* (2019), the level of SO2 in India is low on average.[8]

The statistics for power plant units in the women's sample in Table A2 indicate that the average number of coal units in the district in the month and year of survey is about 3.5. As in the children's sample, the number of hydro units in the district is higher than the number of nuclear units. The summary statistics on pollutant measures in the women sample are comparable to those in the children sample.

Table A1 also reports statistics for two variables that measure total exposure to coal units after birth for children aged 0–5. The first is cumulative exposure to coal units since birth, which conditions on the timing of when new coal units were established in the district after the child's birth (there may be multiple), and how many total years of exposure since birth the child has had to these new coal units. Given variation in the years when units were established in a district since a child's birth, similar–aged children in different districts may have different years of total exposure since birth. The average child has had about five additional coal unit years of exposure since birth. The second variable measuring exposure after birth is a cumulative count of new coal units established in the district after the child's birth year. The mean for this variable is about 1.5.

3. Estimation Framework

Using the initial date of operation of each coal unit, we match the total number of coal units in the district at the time of birth (month and year of birth) of the child to each child aged 0–5 years in the NFHS–4 data. For the *in-utero* estimations, we match the total number of coal units in the district to the three trimester windows of time for each child. Given the low rates of spatial mobility in India (Munshi and Rosenzweig, 2009) and because we consider relatively very young children, we assume that the district of birth is the current district of residence of the child.[9] For the women's sample, we match the number of coal units in the district at the time of the survey to each woman aged 18–49 years.

The regression specification used to examine the effect of exposure to coal units at the time of birth (month and year of birth) on a child's anaemic status is:[10]

$$H_{chmyd} = \beta_0 + \beta_1 Num\ Coal^{b}_{myd} + \gamma X_{chd} + \delta Z_{yd} + M_m + Y_y + \left(M_m \times Y_y\right) + D_d + \varepsilon_{chmyd} \qquad (1)$$

Here H_{chmyd} is the health status of child c in household h, born in month m in year y in district d. We use either HBA levels or anaemic status as the measure of health.

$Num\ Coal_{myd}^{b}$ is the total number of coal units in month m and year of birth y in district of birth d. X_{chd} includes a set of child, mother, household, and district characteristics. The child characteristics include gender, birth order, whether twin birth, whether the child was nursed, recent illnesses (diarrhoea, cough), place of delivery, whether C–section birth, and whether a bed net is used. The mother–specific controls include mother's height, haemoglobin level and anaemic status, whether the mother works, mother's age at first birth, age at marriage, highest level of completed schooling, total number of children born, number of children less than 5 years old in the household, and whether the household regularly consumed meat, fish, eggs, vegetables, fruits, yoghurt, and beans.[11] Household–specific characteristics include religion and caste identifiers, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, access to electricity, type of toilet facility, and the household's primary source of drinking water.[12] Z_{yd} is a set of district specific time–varying controls given by natural log of night–lights in year of birth y and district of birth d.[13] We also condition on whether the mother took iron supplements when she was pregnant; this variable is not used in all models as the information was collected only for the last birth (thus reducing sample size). Further, the regressions control for month of birth (M_m) and year of birth (Y_y) fixed effects, their interactions $(M_m \times Y_y)$, and district fixed effects (D_d). The district fixed effects absorb a range of unobserved district level characteristics that might render the placement of power units across districts endogenous.[14] The presence of district fixed effects and year fixed effects in equation (1) means that we are estimating the effect of exogenous deviations from district and year trends in the number of coal units, similar to the framework in Deschenes and Greenstone (2007), which denotes such deviations as 'presumably random' from their long-run averages. Finally, ε_{chmyd} is the idiosyncratic error term. In additional specifications, we focus on the moderate or severely anaemic category as the dependent variable. Standard errors are clustered at the district level.

The key explanatory variable in equation (1) is $Num\ Coal_{myd}^{b}$, the total number of coal plant units in the district at the time of birth. Given that exposure to coal plant units may be correlated over time, it is difficult to determine the specific months of child development that are of most relevance. Our primary focus is on exposure at birth because *a priori*, this is the point in time when environmental risks for the health and development of children are thought to have the greatest impacts (Currie and Vogl, 2013).

While our specification resembles a two-way fixed effects difference-in-difference (TWFEDD) model as in equation (2) of Goodman-Bacon (2021) (see also Chaisemartin and D'Haultefoeuille, 2020), our key explanatory variable is not from a panel dataset. We use a single cross-section of data. What gives our specification a panel-type representation is that we exploit the variation in the time of birth (year and month) of each child in our sample and link it to the number of coal units in the district at that time. We constructed this variable using data provided by India's Central Electricity Authority which notes the district of location and the month and year of establishment of each unit.

To analyse the impact of coal units on the anaemic status of women we employ the following variant of equation (1):

$$H_{ihmyd} = \alpha_0 + \alpha_1 Num\ Coal^{s}_{myd} + \varphi X_{ihd} + \pi Z_{yd} + M_m + Y_y + \left(M_m \times Y_y\right) + S_s + \epsilon_{ihmyd} \tag{2}$$

Here H_{ihmyd} is either HBA levels or anaemic status of woman i in household h in month m in year y of survey in district of residence d. $Num\ Coal^{s}_{myd}$ is the number of coal units in month m in year y of survey in district d; X_{ihd} includes a set of individual and household characteristics that are the same as those in the child regressions (these women are the mothers of those children).[15] Z_{yd}, M_m, Y_y are defined analogously as in equation (1) except that they refer to the month and year of the survey, and the woman's district of residence. S_s are state fixed effects. Finally, ϵ_{ihmyd} is the error term. Standard errors are clustered at the district level.

Why should the presence of coal units in the district at the time of birth increase the likelihood of anaemia? We argue that the intermediate link is provided by ambient air pollution: coal units in the district increase the concentration of $PM_{2.5}$, which, in turn, adversely affects anaemia. To examine whether this is indeed the mechanism, we proceed in two steps. First, we examine the relationship between the number of coal units in a district in a particular month and year and the corresponding district level average $PM_{2.5}$ in that month and year. Results from this regression are presented in Table 2 (and discussed in Section 4.1). They support the contention that emissions from coal units contribute to $PM_{2.5}$.

[Table 2 near here]

Second, we undertake the following mediation analysis to investigate whether anaemia is sensitive to the number of coal units at the time of birth conditional on $PM_{2.5}$:

$$\begin{aligned} H_{chmyd} = {} & \beta_0 + \beta_1 Num\ Coal_{myd}^{b} + \beta_2 \log(PM_{2.5})_{myd} + \gamma X_{chd} + \delta Z_{yd} \\ & + M_m + Y_y + \left(M_m \times Y_y\right) + D_d + \varepsilon_{chmyd} \end{aligned} \tag{3}$$

Equation (3) is an extended version of equation (1) in that it includes $\log(PM_{2.5})_{myd}$. If anaemia is due to pollution, the impact of coal units at the time of birth should decline in magnitude or become insignificant when we include $\log(PM_{2.5})_{myd}$ as a control.

A few caveats to our approach are worth noting. First, our analysis assumes that each new unit opened has a homogeneous effect, which might not be the case. For example, units may use different coal types or have different abatement technologies. We are not privy to these data and thus cannot control for them. Second, conditional on their number, we assume that coal units have uniform effects in the district without accounting for the fact that individuals residing closer to the unit may be differentially affected as compared to those residing further away.[16] We undertake additional tests at finer geographical levels to demonstrate that this is indeed the case, but our main results are at the district level. Third, we implicitly assume that there are no spill overs or diffusion of air pollution across districts. The second and third assumptions impart a conservative bias to our results in that our estimates give lower bound effects. We present heterogeneity and robustness tests as well as checks to rule out selection arising from migration and sorting.

4. Results

4.1. Coal Units and PM$_{2.5}$

Table 2 reports results on the impact of coal units on PM$_{2.5}$ net of district, year, and month of birth fixed effects. As the first column shows, the number of coal units in a district in any month and year has a significant positive effect on PM$_{2.5}$ levels in that district at that time. Each additional coal unit increases PM$_{2.5}$ concentration by 1.9 percent. Columns 2 and 3 of Table 2 present results of falsification tests: an increase in the number of hydro units and nuclear units in a district is not associated with increases in PM$_{2.5}$. This confirms for our sample the well–documented significant contribution of coal–fired power generation to higher particulate matter pollution.[17]

4.2. Coal Units and Anaemia in Children

Table 3 presents our main results. Each cell in this table is a coefficient from a separate regression. The first set of regressions in column 1 are those corresponding to our main specification in equation (1). The dependent variable is HBA of children aged 0–5. The results in row 1 of column 1 indicate that an additional coal unit in the district at the time of birth is associated with a statistically significant reduction in HBA by almost 0.06 g/dl. Given that the average level of HBA in this population is 10.6 g/dl, this represents a 0.6 percent decline relative to the mean. In column 2, the dependent variable is the anaemic status of a child. The presence of an additional coal unit in the district at the time of birth is associated with a statistically significant 1.1 percentage point increase in the likelihood of the child being anaemic. Given that on average a child in this age group is 59 percent likely to be anaemic, an additional coal unit in the district increases this likelihood by 1.7 percent.

[Table 3 near here]

Columns 3–5 of Table 3 present results corresponding to alternative specifications of equation (1). Columns 3 shows that the presence of coal units has a significant impact on the likelihood of the child being moderately or severely anaemic; the magnitude is close to that in column 2. Controlling for whether the mother took iron supplements when pregnant (column 4) increases the coal unit effect. We note however that the sample size in this regression is smaller than that in the baseline regression (column 2) as iron supplementation is asked only for the last birth.[18] Finally, conditioning on average temperature and rainfall in the district (column 5) results in an effect that is similar in magnitude to that in column 2.[19]

The remaining rows of Table 3 may be interpreted as falsification tests where instead of coal units, we sequentially include the number of hydro units or the number of nuclear units in the district in the month and year of birth. The results show that these alternative sources of power generation have few discernible impacts on the likelihood of being anaemic. We conclude that the impact on anaemia is primarily driven by coal.

We also conduct this analysis at a finer geographical level. Using the geocodes of the DHS cluster (PSU), the geocodes of the coal plant units, and the date each unit started operating, we re-estimate equation (1) but this time include the number of coal units 0–50 km from the PSU, number of coal units 50–100 km from the PSU, and the number of coal units 100–200 km from the PSU, as the relevant explanatory variables.[20] The results are presented in Table A4 and show that an increase in the number of coal units closer to the cluster of residence has a stronger effect on the likelihood that the child is anaemic (column 2). The effects are similar but weaker when the outcome variable is HBA.

4.3. Coal Units and Anaemia in Adult Women

Turning to the effects of coal units on the anaemic status of adult women, the first row of Table 4 presents the main results following the framework in equation (2). The number of coal units at the time of the survey has a weaker effect on the HBA level for women (column 1) in the baseline sample (the coefficient has the expected sign and a *t*–statistic > 1). The presence of an additional coal unit however significantly increases the likelihood that the woman is anaemic by 0.2 percentage points in column 2. Similarly, once we control for rainfall and temperature, an additional coal unit increases the likelihood that the woman is anaemic by 0.3 percentage points (column 4). The result in column 3 suggests that coal units have little measurable impact on the likelihood of women being moderately or severely anaemic. As with children, the detrimental effect on anaemia arises due to coal units and not units of other types.[21]

[Table 4 near here]

Table A5 presents results when we replace the number of coal units in the district at the time of survey by three variables that capture proximity to the nearest coal units from the PSU (number of coal units 0–50 km from the PSU, number of coal units 50–100 km from the PSU, and the number of coal units 100–200 km from the PSU). The results show that the strongest effects are for the number of coal plant units 0–50 km from the PSU at the time of the survey, and there is a negative correlation between distance from the PSU and the magnitude of the harmful impact of coal on women's health.

4.4. Heterogeneity and Robustness Checks

We examine heterogeneity by splitting the sample of children along different dimensions: demographics and socio-economic and health status (gender, mother's work status and mother's health), and heterogeneity by region of residence. The following results in Table A6 are worth noting. The negative effect of coal at the time of birth on anaemic status is statistically significant for girls and not for boys; the overall adverse effect of coal-fired pollution on child anaemia is driven by the sample of children whose mothers do not work outside the home and children of mothers who are relatively short (and so likely to be in poorer health). In terms of geography, the significant association between coal and the anaemic status of children is evident for residents in the Gangetic Plain, rural residents, and in states with reserves of coal.[22]

What about population sorting? If the population that settles in districts with coal units has systematically different characteristics as compared to populations in other districts (for instance, has lower socio–economic status, or is more lacking in access to health infrastructure and facilities), then our results on children's and women's anaemic status could be on account of this fact rather than emissions from coal units. Note that the inclusion of district fixed effects should largely control for such factors. However, to allay concerns further, we restrict our sample of children to those whose parents have resided in the current district of residence for greater than the median number of years (eight years). These results are reported in Table A7 and show that in general, the results remain the same: additional coal units in the district at the time of birth significantly increase the prevalence of anaemia (and decrease HBA levels).

Selection may also result from the fact that a proportion of conceptions may be lost due to pollution (Xue *et al.*, 2021). Note however that if weaker foetuses are lost, which is likely to be the case, then our sample of children is positively selected. This results in a conservative bias in our results. Finally, we condition on mother fixed effects as another robustness check for our results. We find that the coefficient on coal units in the district has the same sign as that reported in Table 3 but is imprecisely estimated.[23]

Sorting of plants to more disadvantaged districts could potentially exaggerate the effects of air pollution from coal. Recall that the major expansion in the use of coal as a major source of electricity generation started post 2000, and the main approach adopted was to add additional units to existing coal power plants. We compute the average pre–2000 prosperity of districts (measured by the average night lights over the period 1992–2000) for districts with and without coal plants in 2000. Districts with coal plants in 2000 were more prosperous relative to districts without coal plants in 2000 (p-value = 0.000). Thus, there does not appear to be a sorting of coal plants to more disadvantaged locations.

Are the impacts on anaemic status of young children confounded by the possible impacts of coal on nutrition in general? To investigate this, we re–analyse the baseline specifications for anaemia and HBA on the sample of children who are stunted, that is, have height for age z–scores that are less than 2 (2 standard deviations below the mean). Low height for age z–scores as indicators of stunting are a widely used measure of undernourishment and long run health. Re–estimating on the sample of children who are stunted reveals a coefficient estimate for coal on anaemic status of 0.02, which is comparable to the impact of coal in column 2 of Table 3. Similarly, re–estimating the model for HBA on the sample of stunted children reveals a coefficient estimate for coal equal to –0.07, which is close to the baseline estimate in column 1 (first row) of Table 3. We conclude that the detrimental impacts of coal on child anaemic status do not only reflect nutritional challenges that this population may face, and are over and above the effects of nutritional deficiencies in our sample of children.

Finally, instead of using the number of coal plants in the district at the time of birth as the relevant explanatory variable of interest, we instead consider the capacity of coal plants in the district at the time of birth. We present the corresponding results in Table A8.[24] An increase in capacity of coal units in the district at the time of birth leads to a statistically significant reduction in HBA (column 1), and an imprecisely estimated increase in the likelihood of the child being anaemic.

Heterogeneity in Treatment

Our regressions control for month and year fixed effects, their interactions and district fixed effects. However, it is possible that the treatment effects vary over time. To control for time-varying district-level unobservables, we re-estimate equation (1) while including interactions of district fixed effects and year of birth fixed effects. The results, which are presented in Table A9, are in general stronger.

In light of recent developments in the broadly related literature, we check for the possibility of negative weights (Chaisemartin and D'Haultefoeuille, 2020). We aggregate the dependent variable in equation (1) over all children born in a district in a particular month and then computed the associated weights. Where the dependent variable is aggregate HBA, only 3 percent of weights are negative while 97 percent are positive. Furthermore, the negative weights sum to only -0.0005 while the sum of the positive weights is 1.0005. Where the dependent variable is aggregate anaemic status, only 3 percent of weights are negative while the remaining are positive. Again, the sum of the negative weights is exceedingly small (-0.00047), while the sum of the positive weights is 1.0004746. We conclude that potential bias in our results arising from negative weights is relatively small.

Cumulative Exposure

In an environment of increasing investment in coal, cumulative exposure after birth may also be important. Table 5 reports results that allow for cumulative exposure (post birth) measured in two ways: (*a*) the total number of coal unit years a child is exposed to after birth (columns 1 and 2), or (*b*) the total number of new coal units established after birth (columns 3 and 4).[25] Results in columns 2 and 4 additionally control for the number of coal units in the district at the time of birth.

[Table 5 near here]

Consistent with our main results, cumulative exposure measured in either of the two ways has a significant effect on child anaemia. However, conditional on the number of coal units at the time of birth, cumulative exposure after birth is not significant. This indicates that the adverse effect of coal is primarily realized by the number of units present at the time of birth, and subsequent additional exposure does not appear to induce further measurable increases.[26]

In Utero Impacts

The results thus far have investigated the effects of coal units at or since birth. However, could some of the adverse effects be traced to *in utero* exposure as well? The foetal origins hypothesis (see Almond and Currie, 2011) postulates that the *in utero* period critically determines a variety of human capital outcomes (mortality, susceptibility to disease, even education and earnings) in both the short and the long term. We investigate *in utero* effects by conditioning on measures of the number of coal units present in each of the three trimesters prior to birth. The results presented in Table 6 affirm the existence of *in utero* effects. The number of coal units in the district in all three trimesters significantly increases the likelihood that the child is anaemic.

[Table 6 near here]

As an additional falsification test, we created two more exposure windows: number of coal plant units in the district six months before pregnancy and three months before pregnancy. The regression results are presented in columns 4 and 5 of Table 6. As expected, the number of coal plant units in the district before the mother is pregnant does not have any effect on the likelihood of the child being anaemic.

In interpreting the effects of exposure to coal units before birth, at birth, or after birth, we choose not to be prescriptive about the primacy of effects at any point in time. Taken together, our results suggest that exposure around the time of birth is what matters most. However, given that exposure to coal plant units is likely to be correlated over time, it is difficult to identify the precise time window when these realized effects have the strongest impacts.

4.5. Ambient Air Pollution as the Intermediate Link Between Coal Units and Anaemia

We have already presented evidence that coal–fired units significantly contribute to higher $PM_{2.5}$ levels in Table 2. We now explore the effects of coal units at the time of birth on anaemia conditioning on $PM_{2.5}$. Table 7 presents the results. Column 1 reports the expected statistically significant positive impact of $PM_{2.5}$ on anaemic status. With the inclusion of the number of coal units in this specification (column 2), $PM_{2.5}$ continues to exert a significant positive effect on child anaemic status while the impact of the number of coal units is no longer significant. It is also smaller in magnitude compared to results in Table 3, falling from 0.011 (column 2 of Table 3) to 0.008 (column 2 of Table 7), a 27.3 percent decline. This suggests that the adverse anaemia impact of coal is due to particulate matter pollution. The results remain unchanged when we control for NO_2 and SO_2, the other common air pollutants in India. As evident in columns 3 and 4 of Table 7, the estimated parameter for $PM_{2.5}$ is relatively unchanged while NO_2 and SO_2 are insignificant.[27]

[Table 7 near here]

In view of the large adverse effects of $PM_{2.5}$ on the health of women and children that this study finds, it is useful to undertake a willingness to pay (WTP) to curtail $PM_{2.5}$ exercise. We follow the ‘dose function’ framework in Moretti and Neidell (2011) to implement a simple back–of–the–envelope calculation. Using data from Horton and Ross (2003) and Menon *et al.* (2016), the slope estimate of 0.014 in column 4 of Table 7 translates into a WTP of approximately \$2.19 (Rupees 148.7) per month per child to reduce $PM_{2.5}$ by 1 percent (in 2016 US Dollars).[28]

5. Conclusion

Assembling a unique dataset from diverse sources, we examine the following important and policy relevant question: what is the effect of coal unit induced pollution on the anaemic status of children and women in India? We combine data on haemoglobin assessments conducted by the National Family Health Survey (2015–16) with time– and district–referenced data on coal–fired power generation from the Central Electricity Authority of India and satellite data on $PM_{2.5}$ pollution to examine this question. We have two principal findings. First, controlling for a comprehensive set of child, mother, household, and district characteristics, an increase in the number of coal plant units in the month and year of birth leads to a worsening of the anaemic status of young children and prime–age women. Second, we identify that the effects of coal units on anaemia are mediated through particulate matter pollution as the primary channel.

Overall, our research adds to the growing evidence on the health costs of a fossil fuel-based energy policy to which damages from the harmful effects on anaemia ought to be added. Evidence suggests that the health consequences are significant. Factoring in these added costs related to anaemia strengthens the case for a progressive shift to cleaner sources of energy. While our focus is on health outcomes, given the significant distributional and welfare consequences of coal-based air pollution, our results are of considerable wider developmental and policy interest in rapidly growing emerging economies around the world.

Endnotes

---

[1] An individual is characterized as anaemic if the altitude adjusted haemoglobin (HBA) level is less than a critical threshold. Details are presented in Section 2.

[2] Haemoglobin levels are adjusted for altitude in enumeration areas that are above 1,000 meters and for smoking status, if known, using the CDC formulas (CDC, 1989), updated in 1998).

[3] The threshold for pregnant women is HBA below 11.0 g/dl. Less than 3% of women in the sample are pregnant. In additional checks, we categorize women as anaemic depending on their pregnancy status or exclude the small proportion that is pregnant. Our results remain unaltered.

[4] Mild anaemia is HBA of 10.0–10.9 g/dl for children and pregnant women, and HBA of 10.0–11.9 for non–pregnant women. Among anaemic children, the majority are either moderately or severely anaemic. Anaemic women are more likely to be mildly anaemic.

[5] We cannot rule out that the patterns are driven by the hereditary nature of anaemia, with mother–to–child transmission of iron deficiency. All regression models for children control for mother's anaemic status and mother's health, as measured by her height.

[6] Figure A1 depicts the geographical spread in commissioned units separated by fuel type (Panels A, B and C for coal, hydro and nuclear, respectively). in 2000 and 2016. There are many more districts with at least one coal unit in 2016, compared to 2000. Figure A2 presents the change in the number of coal units by district during 2000–2016. As of 2016, 61% of the coal plants (not units) are based in the 7 states with the highest coal reserves (Jharkhand, Odisha, Chhattisgarh, West Bengal, Madhya Pradesh, Maharashtra, and Telangana).

[7] The World Health Organization's target $PM_{2.5}$ guideline is an annual mean of 10 micrograms per cubic meter of air (WHO, 2005). Estimates for NO_2 are also relatively large in India.

[8] SO_2 appears to be zero in Panel A of Table 2 because of rounding to three decimal digits.

[9] We run additional tests that control for migration. See Section 4.4.

[10] This specification follows the design in equation (4) of Deschenes and Greenstone (2007), equation (1) in Barreca (2012), equation (1) in Garg *et al*. (2020), and equation (1) in Geruso and Spears (2018).

[11] While the NFHS–4 data contain information on whether children in the household consume these items, we are unable to use these variables because of the large number of missing values.

[12] In additional regressions, we include the proportion of households in the cluster that report engaging in open defecation in place of household's toilet facility. Our main results remain unchanged.

[13] Night-time lights (or nightlights) are increasingly used as a measure of local economic development in the absence of data on income at lower administrative levels (Asher *et al*., 2021).

[14] As noted above, coal units are likely to be established in regions of the country that have easy access to water and are naturally endowed in coal deposits (61% of coal plants are in the states with significant coal reserves). Since the location of water sources and coal deposits is geographically pre–determined, it is unlikely that there exist factors that are simultaneously correlated with our outcome of interest and the presence of coal plant units. Other studies that leverage geographical exogeneity in the location of deposits/reserves include von der Goltz and Barnwal (2019).

[15] Here month m and year y refer to the month and year when a woman was interviewed. We have some variation in y as the NFHS-4 survey began in early 2015 and continued into end 2016.

[16] District fixed effects control for regional parameters that may differ such as district size. When we interact coal units with district size measured in square kilometres, the coefficient on this interaction term is negative indicating that the net impact coal units is lower in larger districts.

[17] We condition on $\log(PM_{2.5})$ to be consistent with Barrows *et al*. (2019). Table A3 presents corresponding results when we use non-transformed $PM_{2.5}$. These results are similar.

[18] We re-estimated the specification in column 2 using the reduced sample in column 4. The coal unit effect is identical to that in column 4 (1.9 percentage points); this indicates that the difference in the impact of coal units between columns (2) and (4) is due to differences in sample size.

[19] Rainfall and temperature may influence how pollution impacts health (Graff Zivin and Neidell, 2013)). However, including these variables reduces the sample size as there are missing values. Therefore, our baseline specification does not include the weather variables.

[20] The geo-codes of the DHS clusters are displaced to preserve anonymity. The distance from the PSU to the

---

nearest coal unit could, therefore, be measured with error. We believe that our approach of accounting for distances within specific ranges minimizes this error.

[21] In Table A9, we further control for changes in unobserved district-level characteristics that may potentially be correlated with both health and the number of coal units. This is done by including month of birth fixed effects, year of birth fixed effects, district fixed effects, and the interactions of district and year of birth fixed effects. In general, the results become stronger.

[22] Given that our main results are stronger for children, in this Section and the next (Sections 4.4 and 4.5), we restrict our analysis to the sample of children aged 0–5.

[23] We believe the smaller sample size reduces precision. Also, in this framework, key mother level variables that are time–invariant (like height, education, or mother's haemoglobin level) cannot be identified.

[24] Capacity is likely to be measured with error since recording numbers for total capacity that are up to four digits in length (maximum capacity in some of the coal plants exceeds 8500 MW) is likely to be especially error prone.

[25] Coal unit years measures the total number of years up until the time of the survey that a unit established after the child's year of birth has been in operation. Where multiple units were established after birth, this variable measures the sum (total) of years of operation for all these units since birth up to the time of the survey. Alternatively, the total number of new coal units conditions on the sum of new coal units established in the district after the year of birth of the child.

[26] This is not surprising as the average child in our sample was born in 2012. Figures 2 and 3 show that by 2012, the meteoric expansion in coal units had already occurred.

[27] As results in Table A10 show, results remain unchanged when we consider non-transformed $PM_{2.5}$.

[28] This estimate of monthly WTP is in a similar ballpark to that estimated for respiratory hospitalizations that result from PM_{10} in Colombia (see Ordoñez, 2020), although the scale there is somewhat different.

References:

Ali, S. (2018). The future of Indian electricity demand: How much, by whom and under what conditions? *Brookings Institution*.

Almond, D., and J. Currie (2011). Killing me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives,* **25**(3), pp 153–172.

Amster, E., and C. L. Levy (2019). Impact of coal-fired power plant emissions on children's health: A systematic review of the epidemiological literature. *International Journal of Environmental Research and Public Health,* **16**, pp 1–11.

Anderson, M. L. (2020). As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality. *Journal of the European Economic Association,* **18**(4), pp 1886–1927.

Asher, S., T. Lunt, R. Matsuura and P. Novosad (2021). Development Research at High Geographic Resolution: An Analysis of Night-Lights, Firms, and Poverty in India Using the SHRUG Open Data Platform. *World Bank Economic Review,* **35**(4), pp 845–871.

Barreca, A. I. (2012). Climate Change, Humidity, and Mortality in the United States. *Journal of Environmental Economics and Management,* **63**, pp 19–34.

Barrows, G., T. Garg and A. Jha (2019). The health costs of coal-fired power plants in India, *IZA DP 12838*

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

Bharadwaj, P., J. Graff Zivin, J. Mullins and M. Neidell (2016). Early-Life Exposure to the Great Smog of 1952 and the Development of Asthma. *American Journal of Respiratory and Critical Care Medicine,* **194**(12), pp 1475–1482.

Bhattacharya, J., J. Currie and S. Haider (2004). Poverty, food insecurity, and nutritional outcomes in children and adults *Journal of Health Economics,* **23**(4), 839–862.

Bobonis, G. J., E. Miguel and C. Puri-Sharma (2006). Anemia and School Participation. *Journal of Human Resources,* **XLI**(4), pp 692–721.

CDC (1989). Current Trends CDC Criteria for Anemia in Children and Childbearing-Aged Women. *Center for Disease Control and Prevention*.

Chaisemartin, C. D. and X. D'Haultefoeuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review,* **110**(9), pp 2964–2996.

Chong, A., I. Cohen, E. Field, E. Nakasone and M. Torero (2016). Iron Deficiency and Schooling Attainment in Peru. *American Economic Journal: Applied Economics,* **8**, pp 222–255.

Coffey, D., M. Geruso, M and D. Spears (2017). Sanitation, Disease Externalities and Anaemia: Evidence from Nepal. *Economic Journal,* **128**(611), pp 1395 - 1432.

Coneus, K., and K. Spiess (2012). Pollution Exposure and Child Health: Evidence for Infants and Toddlers in Germany. *Journal of Health Economics,* **31**, pp 180–196.

Cropper, M., R. Cui, N. Hultman, S. Guttikunda, P. Jawahar, Y. Park, X. Yao and X. P. Song (2020). The Health Impacts of Coal-Fired Power Plants in India and the Co-benefits of GHG Reductions. *American Economic Review, Papers and Proceedings,* **111**, pp 386–390.

Currie, J., M. Neidell and J. F. Schmieder (2009). Air pollution and infant health: Lessons from New Jersey. *Journal of Health Economics,* **28**, pp 688–703.

Currie, J., and T. Vogl (2013). Early-life health and adult circumstance in developing countries. *Annual Review of Economics,* **5**(1), pp 1–36.

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

Deschenes, 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., L. Girolamo, A. V. Donkelaar, S. N. 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.

Elbarbary, M., G. Morgan, Y. Guo and J. Negin (2019). Ambient Air Pollution Association with Anemia Prevalence and Hemoglobil Levels in Chinese Older Adults: Cross-Sectional Study from the WHO

Wave 1 Study on Global Aging and Adult Helath. *Environmental Epodemiology,* **3**, pp 109–111.

Gao, M., G. Beig, S. Song, H. Zhang, J. Hu, Q. Ying, F. Liang, Y. Liu, H. Wang, X. Lu, T. Zhu, G. Carmichael, C. P. Nielsen and M. B. McElroy (2018). The impact of power generation emissions on ambient PM2.5 and human health in China and India. *Environment International*, **121**, pp 250–259.

Geruso, M. and D. Spears (2018). Heat, Humidity, and Infant Mortality in the Developing World. *IZA DP No. 11717*.

GOI (2018). CO2 Baseline Database for the Indian Power Sector. *User Guide, Government of India, Ministry of Powe*r.

Goodman-Bacon, A. (2021). Difference-in-differences with Variation in Treatment Timing, *Journal of Econometrics*, **225**, pp 254–277.

Graff Zivin, J. and M. Neidell (2013). Environment, health, and human capital, *Journal of Economic Literature*, **51**(3), pp 689–730.

Grantham-McGregor, S. and C. Ani (2001). Review of studies on the effect of iron deficiency on cognitive development in children, *Journal of Nutrition*, **131**(2), pp 649S–668S.

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

Greenstone, M. and B. K. Jack (2015). Envirodevonomics: A research agenda for an emerging field, *Journal of Economic Literature*, **53**(1), pp 5–42.

Gupta, A. and D. Spears (2017). Health Externalities of India's Expansion of Coal Plants: Evidence from a National Panel of 40,000 Households, *Journal of Environmental Economics and Management*, **86**, pp 262–276.

Honda, T., V. Pun, J. Manjourides and H. Suh (2017). Anemia prevalence and hemoglobin levels are associated with long term exposure to air pollution in an older population, *Environment International*, **101**, pp 125–132.

Horton, S. and J. Ross (2003). The Economics of Iron Deficiency, *Food Policy*, **28**(1), pp 51–75.

ICMR (2021). Health and economic impact of air pollution in the states of India: the Global Burden of Disease Study 2019, *Lancet Planet Health*, **5**, pp e25–e38.

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

Khaskheli, M. N., S. Baloch, A. Sheeba, S. Baloch and F. K. Khaskheli (2016). Iron deficiency anemia is still a major killer of pregnant women", *Pakistan Journal of Medical Science* **32**(3), 630 - 634.

Krämer, M., S. Kumar and S. Vollmer (2021). Improving Child Health and Cognition: Evidence from a School Based Nutrition Intervention in India, *Review of Economics and Statistics,* **103**(5), pp 818–834.

Lancet (2018). The Lancet Countdown on health and climate change: from 25 years of inaction to a global transformation for public health, *Lancet*, **391**(10120), pp 581–630.

Luechinger, S. (2014). Air pollution and infant mortality: A natural experiment from power plant desulfurization, *Journal of Health Economics*, **37**, pp 219–231.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing and A. D. Stein (2009). The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults, *Economic Journal*, **119**, pp 734–763.

Mehta, U., S. Dey, S. Chowdhury, S. Ghosh, J. E. Hart and A. Kurpad (2020). The Association Between Ambient PM2.5 Exposure and Anemia Outcomes Among Children Under Five Years of Age in India, *Environmental Epodemiology*, **5**(1), pp e125.

Menon, P., C. M. McDonald and S. Chakrabarti (2016). Estimating the cost of delivering direct nutrition interventions at scale: national and subnational level insights from India", *Maternal & Child Nutrition*, **12**, pp 169–185.

Morales-Ancajima, V. C., V. Tapia, B. N. Vu, Y. Liu, D. E. Alarcón-Yaquetto and G. F. Gonzales (2019). Increased Outdoor PM2.5Concentration Is Associated with Moderate/Severe Anemia in Children Aged 6–59 Months in Lima, Peru, *Journal of Environmental and Public Health.* **Article ID 6127845**.

Moretti, E. and M. Neidell (2011). Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles, *Journal of Human Resources*, **46**(1), pp 154–175.

Munshi, K. and M. Rosenzweig (2009). Why is mobility in India so low? Social insurance, Inequality, and Growth, *National Bureau of Economic Research, Working paper w14850*.

Nikolic, M., D. Nikic and A. Stankovic (2008). Effects of Air Pollution on Red Blood Cells in Children, *Polish*

Journal of Environmental Studies, **17**(2), pp 267–271.

Ordoñez, P. J. (2020). Power Plants, Air Pollution, and Health in Colombia, *University of Illinois, Urbana-Champaign*.

Rush, D. (2000). Nutrition and maternal mortality in the developing world, *American Journal of Clinical Nutrition*, **72**(supp), pp 212S–240S.

Sachs, J. and P. Malaney (2002). The economic and social burden of malaria, *Nature*, **415**(6872), pp 680–685.

Scott, S. P., L. P. Chen-Edinboro, L. Caulfield and L. Murray-Kolb (2014). The Impact of Anemia on Child Moratlity: An Updated Review, *Nutrients*, **6**(12), pp 5915–5932.

Thomas, D., E. Frankenberg, J. Friedman, J. P. Habicht, N. Jones, C. McKelvey, G. Pelto, B. Sikoki, J. Smith, C. Sumantri and W. Suriastini (2004). Causal effect of health on labor market outcomes: evidence from a random assignment iron supplementation intervention, *California Center for Population Research, Working Paper*.

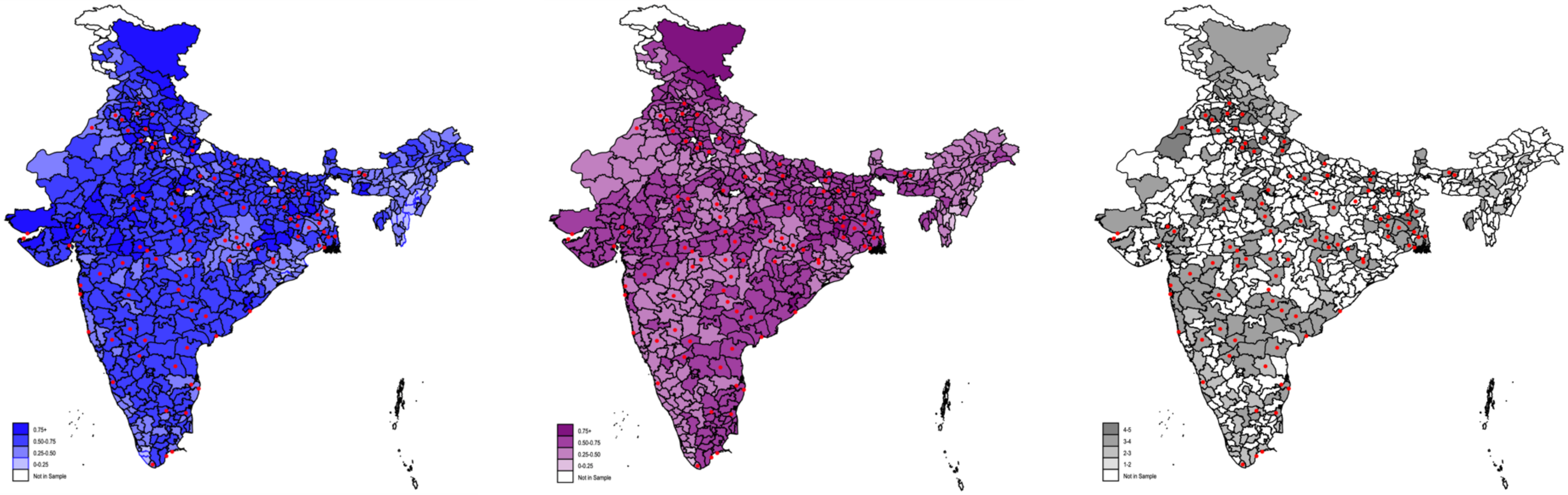
von der Goltz, J. and P. Barnwal (2019). Mines: The local wealth and health effects of mineral mining in developing countries, *Journal of Development Economics*, **139**, pp 1–16.

Vyas, S. (2019). The Child Health Impacts of Coal: Evidence from India's Coal Expansion, *University of Texas at Austin*.

WHO (2005). WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide. Global Update 2005. Summary of Risk Assessment, *World Health Organization*.

Xue, T., T. Guan, G. Geng, Q. Zhang, Y. Zhao and T. Zhu (2021). Estimation of Pregnancy losses Attributable to Exposure to Ambient Fine Particles in South Asia: an Epidemiological Case-Control Study", *Lancet Planetary Health*, **5**, pp e15–e24.

Figure 1. District level Variation in the Proportion Anaemic and $PM_{2.5}$



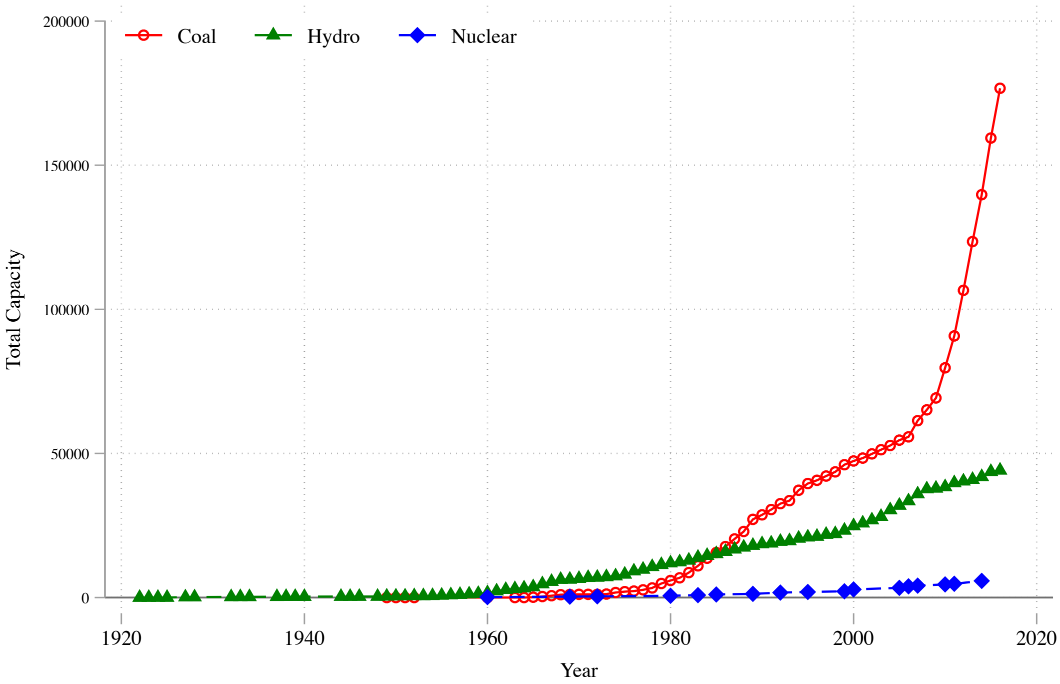
Panel A: Proportion of Children 0–5 who are Anaemic, 2015–16

Panel B: Proportion of Women 18–49 who are Anaemic, 2015–16

Panel C: District level log $PM_{2.5}$

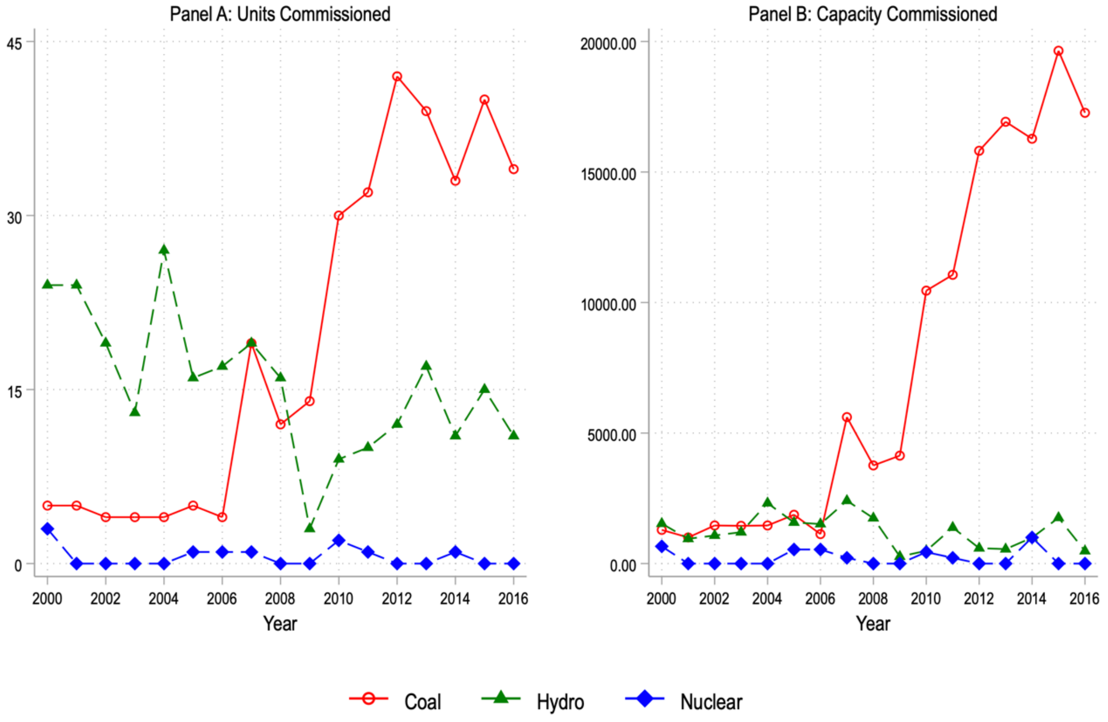
Notes: Authors' calculations. Panel A presents the weighted average of the proportion of children (aged 0–5) in each district categorized as anaemic, that is, altitude adjusted haemoglobin concentration (HBA) < 11.0 g/dl. Category 0–0.25 denotes that up to 25 percent of the children in the district are anaemic, and so on. Districts with red dots have at least one coal plant unit. Panel B presents the weighted average of the proportion of women (aged 18–49) in each district categorized as anaemic, that is, altitude adjusted haemoglobin concentration (HBA) < 12.0 g/dl for non–pregnant women, and HBA < 11.0 g/dl for pregnant women. Category 0–0.25 denotes that up to 25 percent of the women in the district are anaemic, and so on. Districts with red dots have at least one coal plant unit. Panel C presents average values of natural log ($PM_{2.5}$) in each district over the period 2001–2015. Districts with red dots have at least one coal plant unit.

Figure 2. Total Capacity by Year and Fuel Type



Notes: Authors’ calculations. Total capacity (in gigawatts) in each year by fuel type.

Figure 3. Number of New Power Units and Capacity Commissioned by Year and Fuel Type, 2000–2016



Notes: Authors' calculations. Sample restricted to units commissioned between 2000 and 2016. Panel A presents the number of new units (of each fuel type) commissioned in each year between 2000 and 2016. Panel B presents the associated increase in capacity due to these new units.

Table 1. Prevalence of Anaemia

	Children 0–5 (1)	Women 18–49 (2)
Altitude adjusted haemoglobin (HBA) level (g/dl)	10.543	11.628
Anaemic	0.588	0.535
Moderately or Severely Anaemic	0.309	0.137
Sample Size	39, 356	137,227

Notes: Authors' calculations. Anaemia in children six months through 60 months of age is defined as HBA below 11.0 g/dl. Anaemia in women (aged 18–49) is HBA < 12.0 g/dl for non–pregnant women and HBA < 11.0 g/dl for pregnant women. Moderate or severe anaemia uses HBA thresholds of <9.9 g/dl for both children and women.

Table 2. Power Units and $PM_{2.5}$ Concentration

	(1)	(2)	(3)
Number of Coal Units in District	0.0189***		
	(0.004)		
Number of Hydro Units in District		-0.0143	
		(0.016)	
Number of Nuclear Units in District			-0.0626
			(0.053)
Constant	3.4771***	3.5758***	3.5300***
	(0.009)	(0.057)	(0.005)
Sample Size	12,713	12,713	12,713

Notes: Dependent variable in regressions is log($PM_{2.5}$). OLS regression results presented. Number of Coal/Hydro/Nuclear units in the district at each time (month-year) of birth. Regressions control for district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

Table 3. Power Units and Child Health. Children Aged 0–5

	HBA (1)	Anaemic (2)	Moderately or Severely Anaemic (3)	Anaemic (4)	Anaemic (5)
Number Coal Units in District	–0.057***	0.011*	0.010**	0.019***	0.011*
(at time of birth)	(0.018)	(0.006)	(0.004)	(0.006)	(0.006)
Number Hydro Units in District	–0.017	0.000		0.003	–0.003
(at time of birth)	(0.023)	(0.005)		(0.006)	(0.006)
Number Nuclear Units in District	0.039	–0.062		–0.255	–0.071
(at time of birth)	(0.620)	(0.193)		(0.269)	(0.186)
Specification	Baseline	Baseline	Baseline		
Rainfall and Temperature	No	No	No	No	Yes
Iron Supplementation	No	No	No	Yes	No
Sample Size	39, 356	39, 356	39, 356	29, 173	27, 757

Notes: Weighted OLS regression results presented. Child categorized as anaemic (columns 2–5) if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Each cell presents the results from a separate regression. Models include a constant term which is not reported. Sample size is the same across all regressions reported in each cell in a particular column. Child and mother control include those listed in Tables A1 and A2. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

Table 4. Power Units and Women's Health. Women Aged 18–49

	HBA (1)	Anaemic (2)	Moderately or Severely Anaemic (3)	Anaemic (4)
Number Coal Units in District	–0.004	0.002*	-0.000	0.003***
	(0.003)	(0.001)	(0.001)	(0.001)
Number Hydro Units in District	0.001	0.000		–0.001
	(0.004)	(0.001)		(0.001)
Number Nuclear Units in District	0.000	–0.004		–0.005
	(0.015)	(0.004)		(0.004)
Specification	Baseline	Baseline	Baseline	
Rainfall and Temperature	No	No	No	Yes
Sample Size	137,227	137,227	137,227	97,256

Notes: Weighted OLS regression results presented. Sample Restricted to women aged 18 and higher at the time of the survey. A woman is categorized as anaemic if the altitude adjusted haemoglobin count (HBA) is below 12.0 g/dl (if the woman is not pregnant and is below 11.0 g/dl if she is pregnant at the time of the survey. Each cell presents the results from a separate regression. Models include a constant term which is not reported. Dependent variable is the number of units of a specific fuel type at the time of survey. Sample size is the same across all regressions reported in each cell in a particular column. Controls include those listed in Table A2. The regressions control for month and year of survey fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***p < 0.01; **p < 0.05; *p < 0.10.

Table 5. Exposure after Birth and Child Health. Children Aged 0–5

	Anaemic **(1)**	**Anaemic** **(2)**	**Anaemic** **(3)**	**Anaemic** **(4)**
Number Coal Units in District (at time of birth)		0.014*		0.017**
		(0.008)		(0.008)
Cumulative Exposure After Birth	0.002**	–0.002		
	(0.001)	(0.002)		
Number Coal Units in District After Birth			0.009**	–0.008
			(0.004)	(0.008)
Constant	2.118***	0.928*	2.118***	0.910*
	(0.421)	(0.546)	(0.421)	(0.547)
Specification	Baseline	Baseline	Baseline	Baseline
Sample Size	25,178	21,390	25,178	21,390

Notes: Weighted OLS regression results presented. Child categorized as anaemic if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Child and mother control include those listed in Tables A1 and A2. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Table 6. *In utero* Impacts of Coal Units in District. Children Aged 0–5

	Anaemic (1)	**Anaemic (2)**	**Anaemic (3)**	**Anaemic (4)**	**Anaemic (5)**
Average Number of Coal Units in District in Trimester 1	0.015** (0.007)				
Average Number of Coal Units in District in Trimester 2		0.013** (0.006)			
Average Number of Coal Units in District in Trimester 3			0.012** (0.006)		
Average Number of Coal Units in District six months before pregnancy				-0.002 (0.005)	
Average Number of Coal Units in District three months before pregnancy					-0.001 (0.005)
Constant	0.668* (0.404)	0.681* (0.403)	0.700* (0.404)	1.314*** (0.408)	1.284*** (0.403)
Specification	Baseline	Baseline	Baseline	Baseline	Baseline
Sample Size	38,684	38,937	39,179	38,192	38,437

Notes: Weighted OLS regression results presented. Child categorized as anaemic if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Each cell presents the results from a separate regression. Child and mother control include those listed in Tables A1 and A2. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Table 7. Impact of Power Units Conditional on $PM_{2.5}$ Concentration. Children Aged 0–5

	Anaemic (1)	**Anaemic (2)**	**Anaemic (3)**	**Anaemic (4)**
Number of Coal Units in District		0.008		0.008
		(0.007)		(0.008)
Natural log ($PM_{2.5}$)	0.014**	0.013*	0.015**	0.014**
	(0.007)	(0.007)	(0.007)	(0.007)
Natural log (SO_2)			–0.880	–0.576
			(1.851)	(1.835)
Natural log (NO_2)			0.317	0.221
			(0.706)	(0.750)
Rainfall and Temperature	Yes	Yes	Yes	Yes
Iron Supplementation	No	No	No	No
Sample Size	24,315	24,315	24,315	24,315

Notes: Weighted OLS regression results presented. Child categorized as anaemic if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Each cell presents the results from a separate regression. Child and mother control include those listed in Tables A1 and A2. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Coal Plants, Air Pollution and Anaemia: Evidence from India
Supplementary Materials
A1. Descriptive Statistics

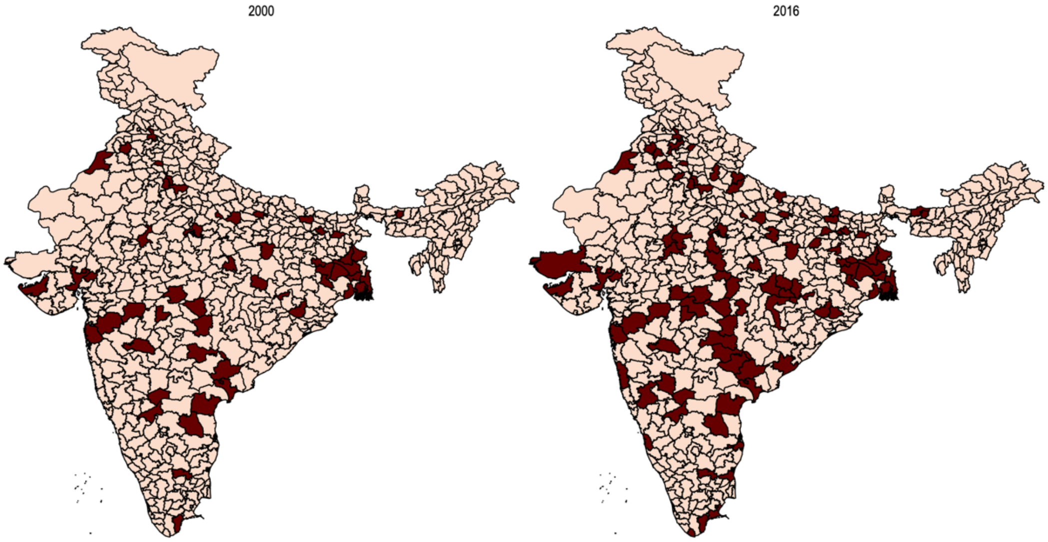
Summary statistics for variables in our sample for children and women are presented in Tables A1 and A2, respectively. We begin by discussing the results in Table A1, focusing on variables beyond the anaemia outcomes as these have been detailed in Table 1 and the other variables discussed in Section 2.4. Average age is about 3 years in the child 0–5 years sample, 52 percent of these children are male, and those who reported illnesses in the last two weeks range from 19–23 percent. Most deliveries occurred in public sector hospitals.

Mother specific parameters reported in Table A1 indicate that average height (measure of genetic health endowment) is about 152 centimetres, and 78 percent of women report taking iron supplements during pregnancy. 58 percent of mothers are anaemic. A very small proportion of mothers' report working outside the home, and age at first marriage and first birth are relatively low. Almost 74 percent of mothers are uneducated and only 24 percent report some or all primary school. The dietary indicators show that about a third of the sample consume meat and fish on a daily or weekly basis, with larger proportions consuming eggs, green vegetables, and fruits. The proportion consuming beans and yoghurt on a regular basis range from 66 to 90 percent. As expected, most households are Hindu, and about 45 percent report belonging to the other backward caste category. Most of these households use unclean sources of fuel for cooking – kerosene, coal, lignite, charcoal, wood, and agricultural crop, and only 19 percent report having a separate space for cooking. Table A1 also reports descriptive measures for the household's wealth index, age of the household head (about 44 years), and proportion of households that is male headed. The proportion that is rural is about 71 percent, and equal proportions report access to sanitary facilities and no facilities. Piped water and ground water are the major sources of drinking water in the child sample. In addition, rates of migration are relatively low as the average years lived in current place of residence is approximately 11 years. Table A1 also reports regional specific variables including measures for nightlights density at the district level, weather conditions, and an indicator for the Southern states where health outcomes for children and women are known to be relatively better.

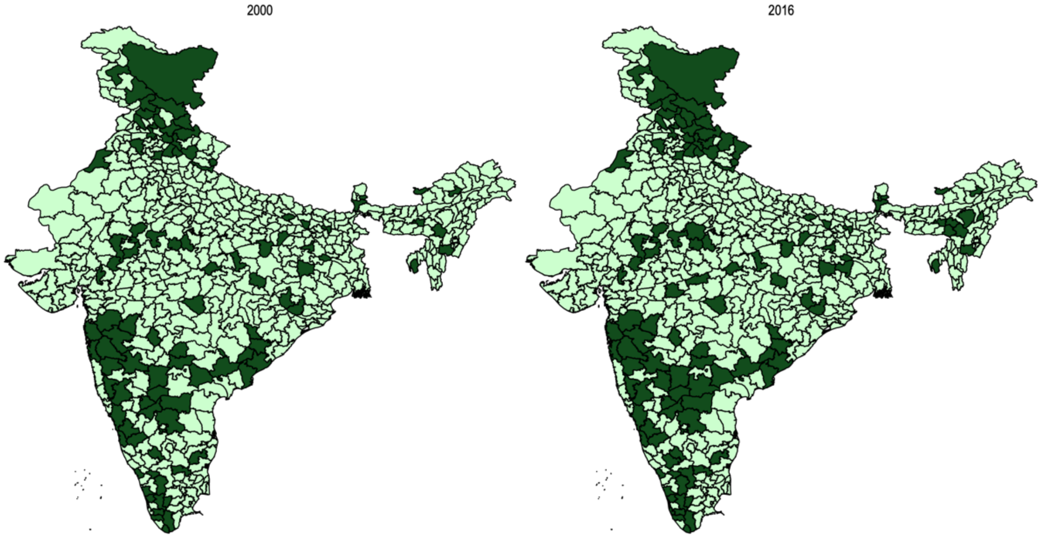
Table A2 reports the summary statistics for women aged 18–49 years and as in the case of Table A1, we focus on variables beyond the anaemia outcomes that we discuss above and those in Section 2.4. The average age of these women is 34 years, and age at first birth and first marriage are similar to those in the child samples. On average, about 36 percent are uneducated, whereas 49 percent have some secondary school or have completed secondary school or higher. The means for the dietary indicators are similar to those reported in Table A1. About 82 percent of these women report living in households that are Hindu, and 58 percent use the relatively more polluting sources of fuel for cooking. Age of the household head is slightly older than in the child samples at 46 years, and about 67 percent of these households are rural. Access to electricity is reported by almost 90 percent of households, and many of the sanitation and access to drinking water measures are similar to those in the child samples. Descriptive statistics for total capacity indicate that most capacity is generated from coal, and coal has experienced a significant rise in number of units from 2000 onwards. The region–specific variables and the weather indicators offer statistics that are broadly in line with those in the children's samples.

Figure A1. Districts with Coal, Hydro and Nuclear Power Units 2000 and 2016

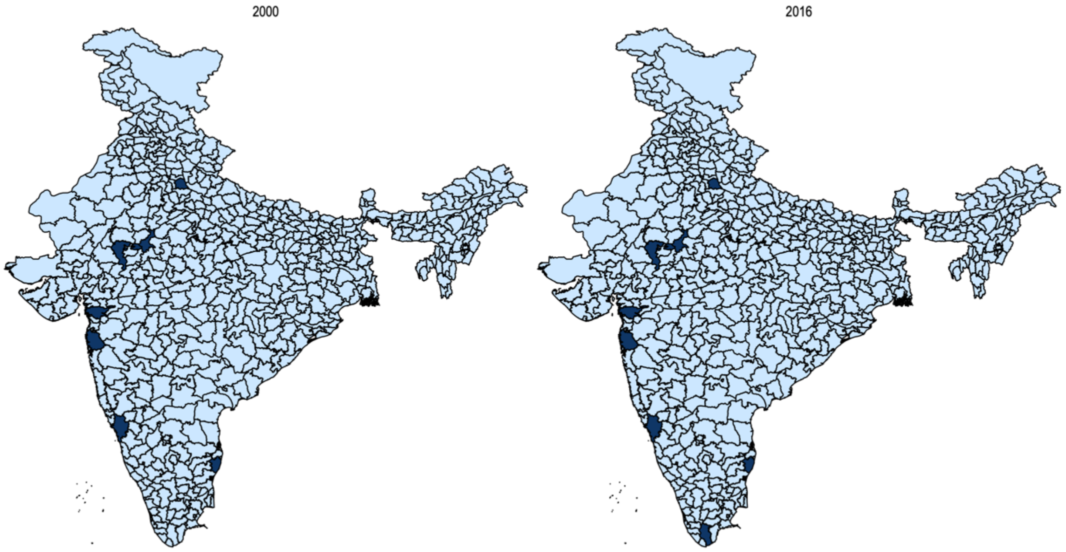
Panel A: Coal Units



Panel B: Hydro Units

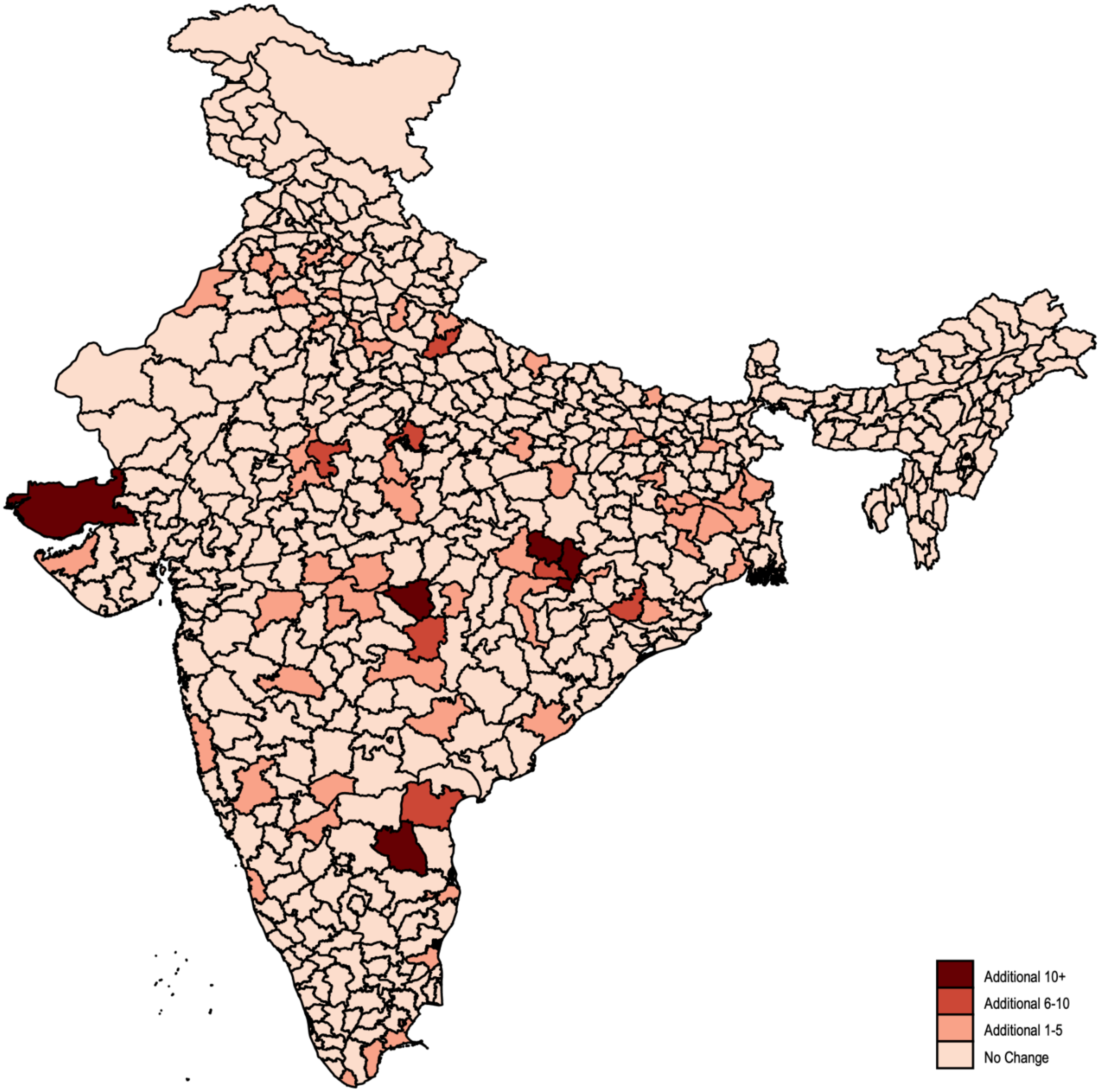


Panel C: Nuclear Units



Notes: Authors’ calculations. Darker shaded districts denote units present. DHS sample does not include Andaman and Nicobar Islands

Figure A2. Change in the Number of Coal Units by District 2000–2016



Notes: Authors' calculations. DHS sample does not include Andaman and Nicobar Islands.

Table A1. Descriptive statistics for children, aged 0–5 years

	Mean (1)	Std. Deviation (2)
Outcomes		
Anaemic	0.588	0.492
Altitude adjusted haemoglobin level (g/dl)	10.543	1.505
Mildly anaemic	0.279	0.449
Moderately or severely anaemic	0.309	0.462
Child–specific		
Age in years	2.566	1.666
Male	0.524	0.499
Birth order	2.060	1.269
Nursed	0.575	0.494
Diarrhoea	0.191	0.635
Cough	0.230	0.668
Delivered at home	0.199	0.399
Delivered in public sector hospital	0.788	0.408
C–section delivery	0.174	0.379
Uses mosquito bed net while sleeping	0.387	0.487
Mother–specific		
Height in centimetres	151.680	6.132
Took iron supplements in pregnancy	0.783	0.412
Anaemic	0.583	0.493
Works outside the home	0.041	0.198
Age at first birth in years	20.951	3.529
Age at first marriage in years	18.660	3.630
Total number of children ever born	2.480	1.365
Number of children below 5 years	1.816	0.899
Not educated	0.738	0.439
Has some or all primary school	0.239	0.427
Has some secondary school	0.022	0.146
Completed secondary school or higher	0.000	0.014
Consumes meat or chicken daily or weekly	0.324	0.468
Consumes fish daily or weekly	0.333	0.471
Consumes eggs daily or weekly	0.409	0.492
Consumes green vegetables daily or weekly	0.853	0.354
Consumes fruits daily or weekly	0.417	0.493
Consumes beans daily or weekly	0.898	0.303
Consumes yoghurt daily or weekly	0.657	0.475
Household–specific		
Hindu	0.785	0.411
Muslim	0.165	0.371
Christian	0.021	0.143
Scheduled caste	0.227	0.419
Scheduled tribe	0.107	0.309

Other backward caste	0.451	0.498
Fuel for cooking: electricity or other	0.006	0.077
Fuel for cooking: liquefied petroleum gas, natural gas, biogas	0.347	0.476
Fuel for cooking: kerosene, coal, lignite, charcoal, wood, straw/shrubs/grass, ag crop, animal dung	0.647	0.478
Food is cooked in a separate building, outdoors, other	0.194	0.395
Wealth index: poorest	0.248	0.432
Wealth index: poorer	0.217	0.412
Wealth index: middle	0.200	0.400
Wealth index: richer	0.184	0.388
Wealth index: richest	0.152	0.359
Age of household head	43.902	15.144
Household head is male	0.881	0.323
Household size	6.430	2.883
House has raw floor	0.433	0.495
House has raw wall	0.217	0.412
House has raw roof	0.102	0.302
Rural	0.714	0.452
Electricity	0.847	0.360
Toilet facility: flush toilet	0.458	0.498
Toilet facility: pit toilet/latrine	0.071	0.256
Toilet facility: no facility/bush/field	0.462	0.499
Toilet facility: other	0.009	0.096
Source of drinking water: piped water	0.418	0.493
Source of drinking water: ground water	0.474	0.499
Source of drinking water: well water	0.075	0.263
Source of drinking water: surface water	0.015	0.123
Source of drinking water: rainwater, tanker truck, other	0.018	0.132
Years lived in place of residence	11.300	10.888
Power Unit and Pollutant–Specific		
Count of coal units in district in month and year of birth	2.968	4.314
Count if hydro units in district in month and year of birth	2.684	4.471
Count of nuclear units in district in month and year of birth	0.172	0.732
Cumulative exposure to coal units after birth (years)	4.691	7.658
Number of new coal units established in district after birth	1.519	2.088
Natural log of $PM_{2.5}$	3.601	0.751
Natural log of NO_2	0.410	0.173
Natural log of SO_2	0.000	0.000
Region–specific		
Natural log of sum of annual nightlights in district	9.984	1.178
Average rainfall in mms.	101.282	153.181
Average temperature in centigrade	25.817	5.196
Southern states	0.188	0.391

Notes: Authors’ calculations. Table reports weighted summary statistics.

Table A2. Descriptive statistics for women, aged 18–49 years

	Mean **(1)**	**Std. Deviation** **(2)**
Outcomes		
Anaemic	0.535	0.499
Altitude adjusted haemoglobin level (g/dl)	11.628	1.630
Mildly anaemic	0.399	0.490
Moderately or severely anaemic	0.137	0.343
Woman–specific		
Age in years	34.241	8.121
Height in centimetres	151.892	5.823
Age at first birth in years	20.291	3.781
Age at first marriage in years	18.067	3.861
Total number of children ever born	2.683	1.495
Number of children below 5 years	0.735	0.956
Not educated	0.362	0.481
Has some or all primary school	0.150	0.357
Has some secondary school	0.338	0.473
Completed secondary school or higher	0.151	0.358
Consumes meat or chicken daily or weekly	0.327	0.469
Consumes fish daily or weekly	0.350	0.477
Consumes eggs daily or weekly	0.417	0.493
Consumes green vegetables daily or weekly	0.860	0.347
Consumes fruits daily or weekly	0.439	0.496
Consumes beans daily or weekly	0.900	0.300
Consumes yoghurt daily or weekly	0.676	0.468
Household–specific		
Hindu	0.815	0.388
Muslim	0.129	0.335
Christian	0.023	0.151
Scheduled caste or tribe	0.312	0.463
Other backward caste	0.458	0.498
Fuel for cooking: electricity or other	0.006	0.079
Fuel for cooking: lpg, natural gas, biogas	0.415	0.493
Fuel for cooking: kerosene, coal, lignite, charcoal, wood straw/shrubs/grass, ag crop, animal dung	0.578	0.494
Food is cooked in a separate building, outdoors, other	0.191	0.393
Wealth index: poorest	0.186	0.389
Wealth index: poorer	0.201	0.401
Wealth index: middle	0.208	0.406
Wealth index: richer	0.209	0.406
Wealth index: richest	0.196	0.397
Age of household head	46.140	12.852
Household head is male	0.872	0.334

Household size	5.617	2.634
House has raw floor	0.357	0.479
House has raw wall	0.190	0.392
House has raw roof	0.083	0.276
Rural	0.672	0.470
Electricity	0.890	0.313
Toilet facility: flush toilet	0.522	0.500
Toilet facility: pit toilet/latrine	0.074	0.263
Toilet facility: no facility/bush/field	0.394	0.489
Toilet facility: other	0.009	0.096
Source of drinking water: piped water	0.479	0.500
Source of drinking water: ground water	0.413	0.492
Source of drinking water: well water	0.077	0.267
Source of drinking water: surface water	0.014	0.118
Source of drinking water: rainwater, tanker truck, other	0.017	0.128
Years lived in place of residence	15.304	11.923
Power Unit and Pollutant–Specific		
Number of coal units in district	3.488	4.952
Number of hydro units in district	2.739	4.846
Number of nuclear units in district	0.163	0.709
Natural log of $PM_{2.5}$	3.570	0.537
Natural log of NO_2	0.413	0.173
Natural log of SO_2	0.006	0.004
Total capacity of coal units in district ($x10^{-3}$) in megawatts	0.734	1.359
Total capacity of hydro units in district ($x10^{-3}$) in megawatts	0.104	0.289
Total capacity of nuclear units in district ($x10^{-3}$) in megawatts	0.007	0.054
Change in the number of coal units in district since 2000	1.545	2.821
Change in the number of hydro units in districts since 2000	0.613	2.215
Change in the number of nuclear units in district since 2000	0.061	0.373
Region–specific		
Natural log of sum of annual nightlights in district	10.202	1.070
Average rainfall in mms.	57.231	95.766
Average temperature in centigrade	28.363	4.283

Notes: Authors' calculations. Table reports weighted summary statistics.

Table A3. Power Units and $PM_{2.5}$ Concentration

	(1)	(2)	(3)
Number of Coal Units in District	0.5711***		
	(0.199)		
Number of Hydro Units in District		-0.5566	
		(0.647)	
Number of Nuclear Units in District			-1.8564
			(1.212)
Constant	43.7108***	47.1354***	45.3032***
	(0.496)	(2.325)	(0.109)
Sample Size	12,713	12,713	12,713

Notes: Dependent variable in regressions is $PM_{2.5}$. OLS regression results presented. Number of Coal/Hydro/Nuclear units in the district at each time (month-year) of birth. Regressions control for district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

Table A4. Power Units and Child Health. Children Aged 0–5

	HBA **(1)**	**Anaemic** **(2)**
Number of Coal Plant Units 0–50 km from PSU	-0.006	0.005**
	(0.005)	(0.002)
Number of Coal Plant Units 50–100 km from PSU	-0.007*	0.004***
	(0.004)	(0.002)
Number of Coal Plant Units 100–200 km from PSU	-0.004**	0.001*
	(0.002)	(0.001)
Specification	Baseline	Baseline
Rainfall and Temperature	No	No
Iron Supplementation	No	No

Notes: Weighted OLS regression results presented. Child categorized as anaemic (columns 2–5) if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Each column presents the results from a separate regression. Models include a constant term which is not reported. Controls include a set of child (gender, birth order, whether the child was nursed, recent illnesses (diarrhoea, cough), place of delivery, whether C–section birth, and whether a bed net was used), mother (mother's height, mother's anaemic status, whether mother works, mother's age at first birth, age at marriage, mother's educational level and diet, total number of children born, and number of children less than 5 years in the household) and household (religion and caste identifiers, SC/ST status, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, presence of electricity, type of toilet facility and primary source of drinking water) characteristics and a set of district specific time–varying controls including natural log of night–lights in year of birth y and district of birth d. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***p < 0.01; **p < 0.05; *p < 0.10.

Table A5. Power Units and Women’s Health. Women Aged 18–49

	HBA (1)	Anaemia (2)
Number of Coal Plant Units 0–50 km from PSU	-0.007***	0.002***
	(0.002)	(0.001)
Number of Coal Plant Units 50–100 km from PSU	-0.005***	0.002***
	(0.002)	(0.001)
Number of Coal Plant Units 100–200 km from PSU	-0.001	0.001**
	(0.001)	(0.000)
Specification	Baseline	Baseline
Rainfall and Temperature	No	No

Notes: Weighted OLS regression results presented. Sample Restricted to women aged 18 and higher at the time of the survey. A woman is categorized as anaemic if the altitude adjusted haemoglobin count (HBA) is below 12.0 g/dl (if the woman is not pregnant and is below 11.0 g/dl if she is pregnant at the time of the survey. Each column presents the results from a separate regression. Models include a constant term which is not reported. Dependent variable is the number of units of a specific fuel type at the time of survey. The regressions control for a set of individual (age, height, age at marriage, age at first birth, education and diet, total number of children born to the woman), household specific (religion and caste identifiers, SC/ST status, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, presence of electricity, type of toilet facility and primary source of drinking water) characteristics. Regressions also include a set of district specific time–varying controls given by natural log of nightlights in year of survey y and district d. The regressions control for month and year of survey fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***p < 0.01; **p < 0.05; *p < 0.10.

Table A6. Heterogeneity in Impacts of Coal Units in District. Anaemic status of Children Aged 0–5

	Boys	**Girls**	**Mother Works outside home**		**Mother's Height**	
			No	**Yes**	**Below median**	**Above median**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
No. of Coal Units in District	0.010	0.011**	0.012**	0.022	0.021***	0.001
	(0.009)	(0.005)	(0.006)	(0.028)	(0.006)	(0.008)
Constant	0.630	0.759	0.815*	-0.503	0.756	0.600
	(0.501)	(0.535)	(0.416)	(1.773)	(0.495)	(0.556)
z	0.057		0.344		-1.964**	
Specification	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Rainfall and Temperature	No	No	No	No	No	No
Iron Supplementation	No	No	No	No	No	No
Sample Size	20,347	19,008	37,798	1,547	19,029	20,327

	South	**Gangetic Plain**	**Other States**	**Rural**	**Urban**	**Coal Reserve**	
						Yes	**No**
	(7)	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**
No. of Coal Units in District	0.008	0.057***	0.004	0.018**	0.001	0.014*	0.007
	(0.009)	(0.009)	(0.006)	(0.008)	(0.009)	(0.008)	(0.007)
Constant	2.721**	-0.519	2.187***	0.754*	0.197	1.033	0.701
	(1.190)	(0.493)	(0.758)	(0.447)	(0.934)	(0.794)	(0.459)
z		3.831***		-1.409		-0.697	
		5.047***					
Specification	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Rainfall and Temperature	No	No	No	No	No	No	No
Iron Supplementation	No	No	No	No	No	No	No
Sample Size	7,687	11,589	20,080	27,627	11,728	11,998	27,358

Notes: Weighted OLS regression results presented. Child categorized as anaemic if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Controls include a set of child (gender, birth order, whether the child was nursed, recent illnesses (diarrhoea, cough), place of delivery, whether C–section birth, and whether a bed net was used), mother (mother's height, mother's anaemic status, whether mother works, mother's age at first birth, age at marriage, mother's educational level and diet, total number of children born, and number of children less than 5 years in the household) and household (religion and caste identifiers, SC/ST status, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, presence of electricity, type of toilet facility and primary source of drinking water) characteristics and a set of district specific time–varying controls including natural log of night–lights in year of birth y and district of birth d. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. z-tests as defined in the text. Two z-tests presented for region specific regressions (Gangetic Plain vs South and Gangetic Plain vs Other states). Standard errors clustered at the district level in parenthesis. Significance: ∗∗∗p < 0.01; ∗∗p < 0.05; ∗p < 0.10.

Table A7. Coal Units and Child Anaemia: Children Aged 0–5 in Households Resident in the District for More than 8 Years

	HBA	**Anaemic**	**Moderately or Severely Anaemic**	**Anaemic**	**Anaemic**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Number coal units in	–0.082***	0.016	0.017**	0.018*	0.021**
District	(0.029)	(0.010)	(0.008)	(0.010)	(0.010)
Specification	Baseline	Baseline	Baseline		
Rainfall and temperature	No	No	No	No	Yes
Iron supplementation	No	No	No	Yes	No
Sample Size	15, 932	15, 932	15, 932	12, 014	11, 136

Notes: Weighted OLS regression results presented. Child categorized as anaemic if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey whose households have been resident in the area for eight or more years. Each cell presents the results from a separate regression. Models include a constant term which is not reported. Controls include a set of child (gender, birth order, whether the child was nursed, recent illnesses (diarrhoea, cough), place of delivery, whether C–section birth, and whether a bed net was used), mother (mother's height, mother's anaemic status, whether mother works, mother's age at first birth, age at marriage, mother's educational level and diet, total number of children born, and number of children less than 5 years in the household) and household (religion and caste identifiers, SC/ST status, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, presence of electricity, type of toilet facility and primary source of drinking water) characteristics and a set of district specific time–varying controls including natural log of night–lights in year of birth y and district of birth d. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***p < 0.01; **p < 0.05; *p < 0.10.

Table A8. Capacity of Power Units and Child Health. Children Aged 0–5

	HBA	Anaemic
	(1)	(2)
Capacity of Coal Units in District at the time of birth	-0.097**	0.012
	(0.039)	(0.013)
Sample Size	17,422	17,422
Specification	Baseline	Baseline
Rainfall and Temperature	No	No
Iron Supplementation	No	No

Notes: Weighted OLS regression results presented. Capacity of Coal Units in District measured at the time (month-year) of birth of the child. Child categorized as anaemic (columns 2–5) if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Models include a constant term which is not reported. Sample size is the same across all regressions reported in each cell in a particular column. Controls include a set of child (gender, birth order, whether the child was nursed, recent illnesses (diarrhoea, cough), place of delivery, whether C–section birth, and whether a bed net was used), mother (mother's height, mother's anaemic status, whether mother works, mother's age at first birth, age at marriage, mother's educational level and diet, total number of children born, and number of children less than 5 years in the household) and household (religion and caste identifiers, SC/ST status, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, presence of electricity, type of toilet facility and primary source of drinking water) characteristics and a set of district specific time–varying controls including natural log of night–lights in year of birth y and district of birth d. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

Table A9. Power Units and Child Health Controlling for Unobserved Time-Varying District Characteristics. Children Aged 0–5

	HBA **(1)**	**Anaemic** **(2)**
Number Coal Units in District	–0.141***	0.031***
	(0.019)	(0.008)
Specification	Baseline	Baseline
Rainfall and Temperature	No	No
Iron Supplementation	No	No
Sample Size	39, 356	39, 356

Notes: Weighted OLS regression results presented. Child categorized as anaemic (columns 2–5) if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Models include a constant term which is not reported. Controls include a set of child (gender, birth order, whether the child was nursed, recent illnesses (diarrhoea, cough), place of delivery, whether C–section birth, and whether a bed net was used), mother (mother's height, mother's anaemic status, whether mother works, mother's age at first birth, age at marriage, mother's educational level and diet, total number of children born, and number of children less than 5 years in the household) and household (religion and caste identifiers, SC/ST status, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, presence of electricity, type of toilet facility and primary source of drinking water) characteristics and a set of district specific time–varying controls including natural log of night–lights in year of birth y and district of birth d. All regressions control for month of birth fixed effects, district fixed effects, year of birth fixed effects, and interactions of district fixed effects and year of birth fixed effects. Standard errors clustered at the district level in parenthesis. Significance: ***p < 0.01; **p < 0.05; *p < 0.10.

Table A10. Impact of Power Units Conditional on $PM_{2.5}$ Concentration. Children Aged 0–5

	Anaemic **(1)**	**Anaemic** **(2)**	**Anaemic** **(3)**	**Anaemic** **(4)**
Number of Coal Units in District		0.008		0.008
		(0.007)		(0.008)
$PM_{2.5}$/1000	0.307**	0.312**	0.343**	0.340**
	(0.143)	(0.139)	(0.150)	(0.145)
SO_2			-1.200	-0.945
			(1.920)	(1.909)
NO_2			0.073	0.018
			(0.352)	(0.367)
Constant	1.347**	1.247**	1.312**	1.247**
	(0.551)	(0.568)	(0.604)	(0.615)
Rainfall and Temperature	Yes	Yes	Yes	Yes
Iron Supplementation	No	No	No	No
Sample Size	24,027	24,315	24,027	24,027

Notes: Weighted OLS regression results presented. Child categorized as anaemic if altitude adjusted haemoglobin count (HBA) is below 11.0 g/dl. Sample restricted to children aged 0–5 at the time of the survey. Controls include a set of child (gender, birth order, whether the child was nursed, recent illnesses (diarrhoea, cough), place of delivery, whether C–section birth, and whether a bed net was used), mother (mother's height, mother's anaemic status, whether mother works, mother's age at first birth, age at marriage, mother's educational level and diet, total number of children born, and number of children less than 5 years in the household) and household (religion and caste identifiers, SC/ST status, type of cooking fuel used, whether the kitchen is located in a separate room of the house, the wealth index, controls for age of the head, gender of the head, household size, type of floor material of the house, wall material of the house, roof material of the house, rural/urban status, presence of electricity, type of toilet facility and primary source of drinking water) characteristics and a set of district specific time–varying controls including natural log of night–lights in year of birth y and district of birth d. All regressions control for month and year of birth fixed effects, their interactions, and district fixed effects. Standard errors clustered at the district level in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.