Gender Disparities in the Prevalence of Undernutrition in India: The Unexplored Effects of Drinking Contaminated Water[1]

Khushboo Aggarwal[a] and Rashmi Barua[b]

August 2022

Abstract

Stunting, a manifestation of chronic malnutrition, is widespread in India. This coupled with biased preferences of parents towards their eldest sons has led to stunting and underweight among girls that grows sharply with increasing birth order. We study the impact of an environmental water pollutant on child growth outcomes in arsenic contaminated regions of India. Using a large nationally representative household survey (NFHS-4) and exploiting variation in soil textures across districts as an instrument for arsenic, we find that arsenic exposure beyond the safe threshold level is negatively associated with Height-for-age and Weight-for-age. Negative effects are larger for girls who are born at higher birth orders relative to the eldest. This, we argue, suggests that the lack of adequate nutrition and health care during early childhood can make girls more vulnerable to external environmental hazards due to their lower immunity and under developed bodies.

[1] We are grateful to seminar participants at the NEUDC 2021 and the 3rd annual workshop of women in the economy, ISI Delhi. We thank Nidhiya Menon for her feedback and suggestions. All errors are our own.
[a]Rishihood University, Email: Khushboojnu14@gmail.com
[b]Corresponding author. Centre for International Trade and Development (CITD), Jawaharlal Nehru University (JNU), Email: rashmibarua@mail.jnu.ac.in

1. Introduction

Across the globe, one out of every four children under the age of five suffer from severe stunting (UNICEF 2017). More than 30 percent of world's stunted children live in India. Child stunting which is associated with chronic malnutrition, has long lasting effects on health and overall development of a child. Stunted children fall sick more often, are more likely to have learning difficulties, under perform in school and have reduced future earnings (Glewwe and Miguel, 2008; Barker et al., 1993; Case and Paxson, 2008).

Stunting, as measured by low Height-for-Age z-scores (HAZ), is caused by long-term insufficient nutrient-intake and frequent infections. Studies in India suggest that height disadvantage among children increases with steeper birth order gradient, particularly among girls. This height disadvantage materializes at second birth order and increases thereon with increasing birth order (third and higher). This can be explained by biased preferences of parents towards their eldest sons which in turn affects their fertility decision and resource allocation across children (Jayachandran and Pande, 2017; Jayachandra and Kuziemko, 2011). Unequal intra-household allocation of health inputs, made available on the basis of gender and birth order, is thus a major determinant of nutritional status of children.

Numerous studies have investigated the relation between gender and child growth indicators (height and weight) as determined by their respective share in households' available resources. However, in addition to adequate nutrition, safe drinking water acts an indispensable input to child health. Across the world, more than 2,000 children under the age of five die every day from gastrointestinal diseases. Out of these deaths, 90% are attributed to unsafe water consumption (UNICEF, 2013). All things held constant, the effect of drinking contaminated water on child health outcomes should not differ by gender. However, in the presence of gender bias, girls might be more likely than boys to be adversely affected by environmental pollutants in drinking water. Lack of adequate nutrition and health care during their early childhood can make girls more vulnerable to external environmental hazards due to their lower immunity and under developed bodies. To the best of our knowledge, no study has addressed the role that gender plays in the relation between child health and access to safe drinking water.

We investigate the impact of exposure to arsenic contaminated groundwater on child health outcomes in India. Overconsumption of arsenic can lead to fatal health outcomes such as kidney and heart failure, mental illnesses, cancer, skin-related diseases, and adverse pregnancy outcomes.[2] Children are more susceptible to arsenic because of their lower immunity levels and relatively higher proportion of body water compared to adults. In a longitudinal study conducted in rural Bangladesh with 1,505 mother-infant pairs, Gardner et. al. (2013) find an inverse association between arsenic exposure and children's growth outcomes with significantly larger effects among girls.[3] They find that nutritional deficiencies act as a primary factor for larger effects among girls from low SES households.

While there is ample epidemiological evidence that arsenic affects child growth outcomes (Watanabe et al. 2007; Minamoto et al. 2005; Rahman et. al. 2009), the mechanisms by which arsenic may affect growth in early life are unclear. Some studies suggest that arsenic interferes with the distribution and function of micronutrients while others argue that arsenic exposure is associated with increased risk of anemia (Gardner et. al. 2013; Heck et. al. 2008). There is also strong evidence that arsenic crosses the placenta and adversely impacts health in utero and later in life (Rahman et al. 2009; Kile et al. 2016).

In this paper, we argue that in the presence of gender bias, girls may be more likely than boys to be adversely impacted by drinking arsenic contaminated water. This is because nutritional deficiencies, and shorter duration of breastfeeding, might exacerbate the adverse impact of environmental exposure to arsenic on health outcomes. While arsenic is known to readily cross the placenta, exclusive breastfeeding protects infants against arsenic (Fängstrom et al. 2008).[4]Thus, if girls are less likely to be breastfed or given adequate nutrition in childhood, the adverse health effects of arsenic exposure can be more severe among girls.

Using geographical variation in arsenic concentration in water, we estimate the association between arsenic levels and child health outcomes (HAZ and WAZ scores) in India using data from

[2]Arsenic poisoning, or Arsenicosis, is a chronic illness resulting from drinking water with high levels of arsenic over a period of time.

[3] Exposure was based on urinary concentrations of arsenic in urine samples collected from pregnant women and their children at ages 1.5 and 5 years

[4]Consistent with this, following an arsenic awareness campaign in Bangladesh, Keskin, Shastry and Willis (2017) find that mothers were more likely to exclusively breast-feed infants and for longer. These babies had lower mortality rates and fewer episodes of diarrhea during childhood.

the 2015-16 round of the National Family Health Survey (NFHS-4). But relying on regional variation in groundwater arsenic levels is problematic due to the correlation between concentration levels of arsenic in groundwater and economic activity of region. For instance, agriculturally dominant regions in India have higher levels of arsenic contamination in groundwater. This is primarily due to overexploitation of groundwater, since naturally occurring arsenic dissolves out of rock formation when groundwater level drops significantly (Madajewicz et al. 2007). To overcome this identification challenge, we use an instrumental variable framework in our analysis.

We use the variation in fraction of clayey soil textures across districts within a state to instrument for arsenic levels in groundwater to measure its impact on child health. Finer soils such as clay have relatively higher particle density and are less porous than coarse sandy soil which increases the concentration of contaminated water (Brammer & Ravenscroft, 2009; Madajewicz et. al. 2007).

Instrumental variable estimates indicate that exposure to arsenic in groundwater has negative and significant impact on HAZ and WAZ among children less than five years of age, regardless of gender. To test if the effects are larger among girls due to a nutritional disadvantage, following Jayachandra and Kuziemko (2011), we study the effect of birth order on the association between arsenic and health outcomes. We find that a one standard deviation increase in arsenic levels in groundwater leads to a reduction in HAZ and WAZ by 0.63 and 0.53 standard deviations for later born girl child, respectively, relative to a male child born at first birth order.[5]These findings are robust to the inclusion of district level controls for health infrastructure, weather, other water contaminants, sex ratio, literacy and income.

Existing studies finds that children born at higher birth orders have a higher probability of being from a large size family (Behrman and Taubman, 1986; Spears, Coffey, Behrman, 2019). We explore this further by including in a regression both the sibling size effect and birth order effects. The results are robust, even after accounting for the endogeneity of sibling size. Finally, we

[5]Jayachandran and Pande (2017) attribute the disadvantage of being a later born daughter in India to two effects. First, girls who are born at higher birth order have older siblings with an increased likelihood of having an older brother. This would lead to a "sibling rivalry effect" with a larger share of the household resources being spent on the boy child. The second mechanism is fertility stopping behavior related to the disadvantage associated with being a later born girl in a family with no boys. Parents with only daughters would be keen on having a son, irrespective of their desired family size. Hence, birth of late parity daughters' acts as a negative income shock and thus limited income will be spent on them.

conduct several falsification and robustness checks to confirm that we are measuring the causal effect of drinking contaminated water on child health.

Our findings contribute to the under studied link between gender, environmental pollutants and child growth measures. To the best of our knowledge, this is the first study to explore the role of gender in the relation between environmental pollutants and child health outcomes.

The remainder of the paper is structured as follows: Section 2 reviews the existing literature. In section 3 we provide a detailed description of the dataset followed by the empirical framework presented in section 4. In section 5 we report the primary findings of our study including heterogenous effects and robustness checks. Lastly, in section 6 we give concluding remarks and policy implications of our analysis.

2. Relevant Literature

Our paper is related to the literature that studies the effect of gender discrimination, measured by unequal parental investment in childhood feeding, health care, and nutrition, on child health by birth order. Although such difference might prevail in both developed and developing countries, but the magnitude is quite significant for developing countries (Lundberg 2005; Chung and Das Gupta 2007). For instance, in Ghana, Garg and Morduch (1998) find that higher birth order children experience more stunting and are more likely to be underweight as compared to their elder sibling, particularly if the elder child is a son, suggesting parental differences in resource allocation among sons and daughters.

Jayachandran and Pande (2017) examine variation in provision of pre-natal and post-natal health inputs across birth order gradient. Their findings suggest that parents allocate more prenatal inputs during a pregnancy when they do not have any sons. Surprisingly, the authors find a reverse pattern for post-natal inputs such as vaccination and duration of breastfeeding, when the elder child is a girl. Jayachandra and Kuziemko (2011) show that mothers, with no sons or fewer sons, who want to conceive again would limit their breastfeeding duration for new born daughter. The authors argue that lower rate of breastfeeding for girls increases their vulnerability to water related contaminants and thus, in turn increases their mortality rate.

Some studies explains the differential pattern of investment in child rearing and health inputs due to difference in resources available with parents and their preferences (Behrman, Pollak and Taubman 1986; Becker and Tomes 1976). Son preference in India can be explained by a combination of economic, religious and sociocultural factors such as patrilineality and patrilocality associated with the Hindu Kinship system (Dyson and Moore 1983). Moreover, inheritance rights are in favor of sons and religious rites in Hinduism, including death rituals, are conducted only by the male heir (Arnold, Choe and Roy 1998). Bardhan (1974) finds that the neglect of girl child in Northern regions of India could be attributed to the lower participation of females in agricultural activities that leads to their lower economic value.

Other studies show that women's nutritional status may be worse off due to lack of access to formal healthcare and differential child care practices (DeRose et al. 2000; Marcoux 2002). Thus, the evidence of gender bias in health and nutrition is inconclusive.

Our paper is also related to the literature on the effect of environmental pollutants on health outcomes of children. Epidemiological studies have established that early-life environmental exposure plays a role in growth outcomes (Walker et. al. 2007). Evidence also supports that the effect of air pollution on respiratory health among children differs by gender. However, it is unclear whether the differential effects are due to gender bias in nutritional intakes and health investment, sex specific physiological differences or an interplay of both (Clougherty, 2009). In economics, most studies have focused on the negative health outcomes of air pollution (Arceo-Gomez et al. 2012). Foster et al. (2009) evaluate the impact of clean industry certification program on pollution and consequentially on respiratory diseases among infants in Mexico. Goyal and Canning (2017) find negative impact of air pollution on in-utero health and other child growth indicators in Bangladesh.

A handful of papers have looked at the effect of drinking contaminated water on child health in developing countries. Kile et al. (2016) show that mothers who drank arsenic contaminated water during pregnancy were more likely to give birth to low-weight infants. Greenstone and Hanna (2014) study the relation between environmental regulations (air & water) and infant mortality in India. They find that regulations related to water pollution have no effect on infant mortality rates. Do et al. (2018) show that curtailment of industrial pollution in the River Ganges led to lower incidences of infant mortality in India. Brainerd and Menon (2014) study the impact of harmful chemicals released in water via fertilizer use on infant mortality and child health outcomes and find that exposure to fertilizers during pregnancy has a negative impact on child health outcomes.

3. Data and Data Source

Our data comes from the Demographic and Health Survey (National Family Health Survey, NFHS-4, 2015-16), administered by the Ministry of Health and Family Welfare (MoHFW), Government of India (GoI). NFHS is a nationally representative dataset that comprises of 1,11,667 children who belong to the age group of 0 to 5. The survey provides information on key demographics, health, nutrition and related emerging issues in India. It is the only dataset that provides information on anthropometry measures such as height and weight of children in the age group of 0-5 years using z-scores calculated in accordance with WHO guidelines.

To assess the impact of water pollution on child health, we use two measures of child health. First, we study Height-for-Age (HAZ) for children in the age group of 0 to 5 years. HAZ is a commonly used yardstick to measure stunting or nutritional status of children (Deaton and Dreze 2009). It is a cumulative measure of nutritional dearth from birth or conception onwards and is the best aggregate measure of malnutrition among children that is correlated with outcomes at later stages of life. Stunting is linked to underdeveloped brains, lower retention and reduced learning ability that adversely affects productivity and earning capacity of an individual.

Apart from stunting, we also study the effect of arsenic contamination on underweight measured by Weight-for-Age z-scores (WAZ). Underweight is a symptom of acute malnutrition and is a dire consequence of inadequate intake of food or high incidence of infectious diseases such as diarrhea. Stunting and underweight are aspects of malnutrition that are closely linked to each other. Presence of both stunting and underweight in a child intensifies the risk of mortality (Briend et al. 1986; Waterlow 1974).

Figure 1 plots the HAZ scores by birth order among boys and girls. It is clear from the figure that HAZ among girls decreases with increasing birth order. In particular, the percentage of girls who are moderately or severely stunted increases with birth order[6]. For instance, at first birth order approximately 10 percent of girls suffers from severe stunting which at later birth order increases

[6] Moderate stunting refers to HAZ that lie between -1 to -2 while severe stunting implies a HAZ of less than -3.

to 12 percent and 15 percent for 2^{nd} and 3^{rd}+ birth order, respectively. Similar pattern is visible for boys. Similarly, figure 2 shows that the percentage of girls with moderate to severe underweight increases with increasing birth order. Birth order effects on stunting and underweight reflects the poor nutritional status among girls and boys particularly at higher birth order.

The average HAZ and WAZ for our sample is -1.66 and -1.64, respectively. The NFHS data also includes a host of individual, household and family background characteristics. The summary statistics of the variables that are included in our analysis are shown in Table 1. The data is gender balanced with girls comprising 48 percent of the sample with an average age of 27 months. While 34 percent of the sample consists of children at first birth order, 29 percent are at second birth order and 37 percent of children are at higher than second birth order. 37 percent of mother's are uneducated while only 9 percent have post-secondary education. The average age of mothers in the sample is 27 years. More than three fourth of our sample comprises of rural households with 31 % scheduled caste (SC) and scheduled tribes (ST) and 50 % belong to other backward classes.

3.2 District level Control variables

Data for rainfall is provided by the Indian Meteorological Department (IMD) at district level in India, with a mean value of 76.7 mms. District level sex ratio and literacy data is from the 2011 Census of India. The average sex ratio and literacy rate in our estimation sample is 925 and 68 percent, respectively. To control for district level gross domestic product, we use data on Monthly Per Capita Expenditure (MPCE) from 68^{th} round of NSSO (National Sample Survey Office) as a proxy for district level GDP.

Data for the level of arsenic and iron in groundwater is provided by the Central Ground Water Board. Following the WHO guidelines, the Bureau of Indian Standards (BIS) has notified a standard of 50 $\mu gL-1$ (microgram per liter) for arsenic in drinking water. The level of arsenic in groundwater is aggregated at the district level from block level data. The appendix provides a map of arsenic affected regions of India. We restrict the analysis to only those states where the presence of arsenic is measured beyond the threshold limit in at least one district in that state. The final dataset comprises of more than 85,000 children under the age of five, across 261 districts from 9 arsenic affected states, where 105 districts are arsenic affected and 156 are non-arsenic affected districts. These 9 states are Punjab, Uttar Pradesh, Chhattisgarh and Haryana in the North; Assam, West Bengal, Jharkhand and Bihar in the East and North-East; Karnataka in the South. As shown in Table 1, the average level of arsenic is 94 microgram per liter across districts in India, remarkably higher than the threshold limit.

The data on soil texture is obtained from Harmonised World Soil Database (HWSD) which was established in July 2008 by the Food and Agricultural Organisation (FAO) and International Institute for Applied System Analysis (IIASA). HWSD is global soil database framed within a Geographic Information System (GIS) and contains updated information on world soil resources. It provides data on various attributes of soil including texture and composition. As reported in Table 1, the average clayey soil across arsenic affected states is approximately 28 percent.

4. Empirical Model

We start by investigating whether exposure to arsenic has an impact on growth of children under the age of 5. The following OLS regression is estimated separately for boys and girls:

$$\boldsymbol{Y_{ids} = \alpha_1 Ars_{ds} + \alpha_2 X_{ids} + D_{ds} + S + e_{ids}} \mathbf{(1)}$$

We are interested in measuring the effect of arsenic on two outcome variables: height-for-age (HAZ) and weight-for-height (WAZ) of child *i* in district *d* of state *s* as given in equation (1). The main explanatory variable is Ars_{ds} which indicates the concentration level of arsenic in groundwater in district d and state s. $\boldsymbol{X_{ids}}$represents vector of controls for individual level characteristics (gender, age and age square), mother characteristics (mother's education and age), family background characteristics and socio-economic characteristics (religion, caste, family size[7], wealth index[8] and place of residence). We also control for district level controls ($\boldsymbol{D_{ds}}$) for rainfall, presence of other contaminants (iron),[9] per capita consumption expenditure, sex ratio, number of public health facilities and literacy. Finally, we include state fixed effect in our regression analysis. Heteroskedasticity robust standard errors are clustered at the PSU (Primary Sampling Unit) level[10].

[7] Children born at higher birth orders have a higher probability of being from a large size family. Moreover, family size and resource allocated to each child are highly correlated, which might in turn could affect the health outcomes of children (Kugler and Kumar 2017; Booth and Kee 2005).

[8] The NFHS reports a five category wealth variable that ranges from poorest to richest.

[9] Additional robustness checks were conducted controlling for fluorides and nitrates (tables available upon request). The sample size shrinks drastically due to the non-availability of data on these variables for several districts making the estimates less precise.

[10] PSUs (Primary sampling unit) are unique and smallest working unit in NFHS-4 survey. It has well defined and identifiable boundaries and represents either a village (rural) or census enumeration block (urban). Our findings are robust to clustering at the district level instead of PSU.

Estimating the effects of arsenic on nutritional outcomes in equation (1), using regional variation in arsenic levels, is problematic since the intensity of economic activities in a region may be correlated with arsenic concentration levels. In areas with high economic activity, overexploitation of groundwater is a major cause of arsenic contamination since naturally occurring arsenic dissolves out of rock formations when groundwater levels drop significantly (Madajewicz et al., 2007). Hence, to overcome the potential endogeneity of arsenic levels, we use an instrumental variable approach.

4.1 Instrumental Variable Approach

A variety of natural geochemical processes play a vital role in the release, transport, and distribution of arsenic in groundwater. One of the important determinants of arsenic released in groundwater is the age of groundwater, which, in turn is related to soil permeability. Finer soils have relatively more particle density and lower porosity levels, and, as a result, their permeability level is relatively lower than loamy soil[11] which facilitates arsenic concentration in groundwater (Mac Arthur et al. 2001; Madajewicz et al. 2007).

Herath et. al (2016) find that in the Ganges–Meghna–Brahmaputra basin of India and Bangladesh, aquifers covered by finer sediments (clay) contain greater concentrations of arsenic in groundwater, whereas arsenic concentrations are significantly lower in aquifers with permeable sandy materials at the surface. Because arsenic concentration is higher in clayey relative to coarse soil, we exploit the variation in soil texture across districts within a state to instrument for ground water arsenic contamination.

The first stage equation is given by:

$$\boldsymbol{Ars_{ds} = \beta_1 Soil_{ds} + \beta_2 X_{ids} + D_{ds} + S + \epsilon_{ids}} \ \textbf{(2)}$$

We instrument arsenic soil contamination using $Soil_{ds}$ i.e. the percentage of clayey soil in district *d*. Rest of the specification is same as in equation (1) above. The main identifying assumption is

[11]Loamy soil consists of a higher proportion of sandy and silty soil relative to clayey soil.

that soil texture fractions affect education outcomes only through the impact on the level of arsenic in groundwater.[12]

A plausible threat to the identification assumption is that income might be affected by pattern of cultivation which is determined by soil texture. For instance, in India, water intensive crops (rice) are cultivated in areas with clayey soil due to its water retention capacity unlike sandy soil. As we show in the results, our regressions are robust to the inclusion of two separate measures of district level income proxy variables (rainfall and per capital consumption expenditure). Further, based on recent work by Carranza (2014) who finds that exogenous variations in soil texture can explain sex ratios in India, we include sex ratio at the district level in all regressions. As shown later, we also conduct several falsification checks and robustness tests to negate channels other than arsenic exposure.

An additional threat to our identification strategy exists if clayey soil varies with other weather, geographic or demographic factors and these might in turn affect economic outcomes. In the appendix, we provide evidence of no correlation between proportion of clayey soil and several district level indicators of weather (rainfall and temperature), other contaminants (iron, nitrate, nitrogen, phosphorous, potassium and fluoride), economic and demographic factors (monthly per capita expenditure, rice to wheat production, literacy, sex ratio, male and female employment in agriculture), conditional on state fixed effects.[13]There is significant difference by soil permeability in iron, as districts with higher iron also have higher proportion of clayey soil. However, this would be against finding a negative impact of arsenic on health outcomes and if anything, underestimate our findings as groundwater with a high iron concentration is associated with a decreased risk of childhood anaemia. There is also a positive correlation between rainfall and clayey soil. Though there is no direct effect of rainfall on soil permeability levels as both are exogenous in nature, but both can combinedly determine the level of groundwater and presence of contaminated metals in groundwater.[14]

[12]Note that while groundwater arsenic levels could also rise through increased use of fertilizers, the literature suggests that use of fertilizers does not alter the physical properties of soil (Carranza 2014). Unlike commercial crops like rice and wheat, arsenic-based pesticides are applied in specific crops such as fruit trees, potatoes, vegetables and berries. Use of such pesticides might alter some properties of superficial soil (upper most layer of soil), but not the subterranean soil used in our analysis.
[13] Results are reported in Table A.2.2 in the Appendix.
[14] If the amount of rainfall is less than the soil can absorb, it will infiltrate; there will be no run-off or no discharge of water in the ground. But if rainfall is more than the absorption capacity of soil (defined by soil permeability level), there will be more discharge.

To check if health effects of arsenic exposure vary by gender and birth order, we also estimate the following OLS and first stage equations, respectively:

$$\boldsymbol{Y_{ids} = a_1 Ars_{ds} + a_2 girl_{ids} + a_3 2ndchild_{ids} + a_4 3rd^{+}child_{ids} + a_5(Ars_{ds} * girl_{ids} * 2ndchild_{ids}) + a_6(Ars_{ds} * girl_{ids} * 3rd^{+}child_{ids}) + a_7(Ars_{ds} * 2ndchild_{ids}) + a_8(Ars_{ds} * 3rd^{+}child_{ids}) + a_9(Ars_{ds} * girl_{ids}) + a_{10}(2ndchild * girl_{ids}) + a_{11}(3rd^{+}child_{ids} * girl_{ids}) + a_{12}X_{ids} + D_{ds} + S + e_{ids}} \quad \textbf{(3)}$$

$$\boldsymbol{Ars_{ds} = \pi_1 Soil_{ds} + \pi_2 girl_{ids} + \pi_3 2ndchild_{ids} + \pi_4 3rd^{+}child_{ids} + \pi_5(Soil_{ds} * girl_{ids} * 2ndchild_{ids}) + \pi_6(Soil_{ds} * girl_{ids} * 3rd^{+}child_{ids}) + \pi_7(soil_{ds} * 2ndchild_{ids}) + \pi_8(soil_{ds} * 3rd^{+}child_{ids}) + \pi_9(Soil_{ds} * girl_{ids}) + \pi_{10}(2ndchild * girl_{ids}) + \pi_{11}(3rd^{+}child_{ids} * girl_{ids}) + \pi_{12}X_{ids} + D_{ds} + S + \epsilon_{ids}} \textbf{(4)}$$

Where, $\boldsymbol{2ndchild}$ is an indicator for a child i whose birth order is 2. Similarly, $3rd^{+}$child indicates whether the child born is at 3rd or higher birth orders. Children born at first birth order are taken as the base category in our analysis. Here, the main coefficient of interest to be estimated is a_5 and a_6 which are associated with the three way interaction ($\boldsymbol{Ars_{ds} * girl_{ids} * 2ndchild_{ids}}$) and ($\boldsymbol{Ars_{ds} * girl_{ids} * 3rd^{+}child_{ids}}$) respectively. X_{ids} accounts for individual, maternal and family background characteristics as explained earlier. All regressions include district level controls (D_{ds}) as before and state fixed effect (S). Heteroskedasticity robust standard errors are clustered at the PSU level.

4. Results

4.1 Arsenic and child health by gender

We first show results for OLS estimates using equation (1). Column 1 and Column 2 (Table 2) shows OLS estimates of the effect of arsenic on HAZ and WAZ, respectively. For HAZ scores, OLS estimates are insignificant. OLS Estimates for WAZ shows that one standard deviation (SD) increase in arsenic is associated with a 0.03 SD increase in weight-for-age.

In the remaining columns we check whether the impact of arsenic on stunting and underweight varies by gender. Coefficients are statistically insignificant for HAZ scores and there are no differences by gender in WAZ scores.

To overcome the issue of endogeneity, we use an instrumental variable approach, where variation in soil texture across districts within a state is used as an instrument for arsenic levels in groundwater. The first stage regression results in Table 3 shows a positive and statistically significant relationship between arsenic and soil texture (clayey soil). The F-statistic (63.6) suggests that soil texture is a strong instrument for arsenic levels. Here we show the first stage results for only the simple specification in equation 2 above (corresponding IV estimates shown in the next table 4). All other specifications yield similar first stage results and large F statistics.

The IV results for HAZ, shown in table 4, indicate that the OLS is severely downward biased. A one standard deviation increases in arsenic leads to decrease in height-for-age by 0.54 SD units. Column 2 of Table 4 indicates that higher level of arsenic exposure is negatively associated with WAZ (0.53 SD).

We further analyze whether the effect of arsenic on child growth outcomes varies by gender. As is evident from the remaining columns of table 4, there is no difference by gender in the effect of arsenic contamination on HAZ scores though girls have a higher coefficient for WAZ compared to boys.

While the IV results show that arsenic has an adverse effect on stunting and underweight as measured by lower HAZ and WAZ scores, the simple gender segregated regressions indicate that girls are not much worse off than boys. All children, regardless of gender, have worse growth outcomes associated with arsenic found in the groundwater. To examine the channels, we study if the effect of arsenic on height-for-age and weight-for-height varies by household wealth.

4.3 Heterogenous effect of household wealth

As discussed in the introduction, arsenic may impact child growth via several channels such as, by affecting the distribution and function of micronutrients in the body, via nutritional deficiencies in childhood, duration of breastfeeding and/or in-utero exposure due to arsenic contaminated groundwater consumption during pregnancy. A priori, children belonging to poorer and low socio-economic status households should exhibit detrimental effects of arsenic exposure since they are more likely to suffer from nutritional deficiencies.

In the next table (Table 5), we study the relation between arsenic and health outcomes separately for poor and rich households. To define high/low wealth, we use the wealth index variable in the

NFHS which codes households into five groups, namely, poorest, poorer, middle, richer and richest. We code as *low wealth* households belonging to the first two categories (poorest and poorer). The remaining three categories are defined as *high wealth*. The results are robust to excluding middle income households from the analysis though the reduction in sample size affects the precision of estimates for the high wealth group.

The results clearly show that among children belonging to low income households, a one SD increase in arsenic is associated with 0.64 SD decrease in height-for-age with statistically insignificant effects for the high wealth group. On the other hand, there is no differential effect of wealth status on WAZ scores, both groups have comparable estimates. Thus, arsenic exposure leads to low height for age, and the effects are more pronounced among children from low income households. These results capture the effect of arsenic on health among children with poor nutritional intake, however, it does not necessarily capture the effect of gender bias induced breastfeeding behavior or nutritional biases. Though we cannot directly test the interlinkages between arsenic exposure and gender bias in nutritional intakes, we can rely on a well-established birth order literature to test the hypothesis.

4.4 Arsenic and child health across gender and birth order

We study the interaction between arsenic exposure, gender and birth order in Table 6. We show OLS results in columns 1 and 4 and the remaining columns show the preferred IV specification using soil quality as an exogenous instrument for arsenic levels. OLS results show no differences in the health effects of arsenic exposure by birth order and gender. On the other hand, IV results in Table 6 for the triple interaction terms (arsenic*girl*birth order) suggest that girls in arsenic affected regions have higher height disadvantage than boys, and the effects are magnified for later born girls relative to the eldest. In column 3 with all control variables included, a one standard deviation unit change in arsenic leads to decrease in height-for-age (stunting) for third (or later) born girls by 0.63 standard deviation units. Significance of our estimate for third (or later) born girls indicates that arsenic induced stunting in girls increases with steeper birth gradient. These estimates are robust to the inclusion of various district level controls. We find similar IV results for weight-for-age as shown in column 6. IV estimates on WAZ indicates that a one standard deviation increase in arsenic leads to a decrease in weight-for-age (underweight) for second and third (or later) born girls by 0.42 and 0.56 standard deviations, respectively.

When the concentration of arsenic in groundwater increases, later born girls (born at higher birth order) experience more height and weight disadvantage relative to their older sibling (lower birth order), particularly if the elder sibling is male. Some studies attribute the birth order effect to the *sibling rivalry effect* i.e. having an older brother limits the availability of essential nutrients along with other health inputs to later born daughters in the family (Fledderjohann et al., 2014; Victoria et al. 1987).

5. Robustness and Falsification Tests

5.1 Arsenic, birth order and sibling size

In a recent paper, Coffey and Spears (2021) show that later born children born in India have an advantage in terms of neo natal mortality. They find that a large disadvantage to high sibling size co-exists with a large advantage to later birth order emphasizing the endogeneity of sibling size for estimating birth order effects. To account for this potential bias in our estimates, we control for the number of siblings under the age of five in the household in the main regressions. Further, to overcome the issue of endogeneity of sibling size, we use the gender of first child as an instrument for sibling size. Having a girl as first child is positively associated with fertility, particularly in the presence of son preference, as parents will continue to have more children until desired number of boys are born in a family (Pande and Astone, 2007). Further, gender of first child is exogenously determined and should affect child health outcomes only through fertility (Kugler and Kumar 2017).

In this specification, we run the regressions separately by gender[15]. We estimate the following OLS and first stage regressions separately by gender:

$$\boldsymbol{Y_{ids} = b_1 Ars_{ds} + b_2 2ndchild_{ids} + b_3 3rd^{+}child_{ids} + b_4(Ars_{ds} * 2ndchild_{ids}) + b_5(Ars_{ds} * 3rd^{+}child_{ids}) + b_6\, sib\_size_{ids} + b_7 X'_{ids} + D_{ds} + S + e_{ids}} \qquad \textbf{(5)}$$

$$\boldsymbol{Ars_{ids} = \lambda_1 soil_{ds} + \lambda_2 2ndchild_{ids} + \lambda_3 3rd^{+}child_{ids} + \lambda_4(soil_{ds} * 2ndchild_{ids}) + \lambda_5(soil_{ds} * 3rd^{+}child_{ids}) + \lambda_6\, gender\_first + \lambda_7 X'_{ids} + D_{ds} + S + e_{ids}} \qquad \textbf{(6)}$$

[15] The instrument for sibling size (gender of first child) will otherwise be perfectly collinear with our main explanatory variables $\boldsymbol{Ars_{ds} * girl_{ids} * nth\ child_{ids}}$.

Where, $\boldsymbol{sib\_size}$ is the number of children under the age of five in a household. This variable is instrumented by $\boldsymbol{gender\_first}$, a binary variable for the gender of the first born child in a household which takes the value of 1 for girls and 0 for boys. All other variables are same as in the previous regressions.

IV results from this specification are shown in Table 7. Column 1 and column 3 show IV estimates for health outcomes (HAZ and WAZ, respectively) for girls. Height-for-age and weight-for-age for boys are reported in column 2 and column 4, respectively. After controlling for sibling size, a one SD increase in arsenic exposure leads to significant decrease in height-for-age and weight-for-age for girls born at third (or higher) birth order by 1.28 SD and 1.35 SD, respectively. Among boys, we find negative impact of arsenic exposure on weight-for-age (0.51 SD, birth order 3) and height-for-age (0.58 SD, birth order 2), but the effect on girls is more than twice the effect on boys. Looking at the coeffient on sibling size, more number of children in the household has a negative and significant effect on most growth outcome.

5.2 Non-Arsenic States

We have shown that arsenic contaminated water impacts child growth and that this process disproportionately affects girls born at later birth order. Our main regression exploits variation in arsenic levels across Indian states, thus the sample includes only nine states where arsenic is present in groundwater. However, comparing results in arsenic-affected states to those not in arsenic-affected states should provide a very useful test of our main hypotheses. Since there is no variation in arsenic levels across non-arsenic states, we cannot use the same identification strategy employed in equations 3 and 4. Instead, we exploit the variation in source of drinking water and gender and compare results by birth order for arsenic and non-arsenic states.

We run the following regression to estimate the impact of consuming groundwater on health outcomes:

$$\boldsymbol{Y_{ids} = a_1 swater_{ids} + a_2 girl_{ids} + a_3(swater_{ids} * girl_{ids}) + a_4 X_{ids} + D + e_{ids}}$$

Where, *swater* is the source of drinking water where we categorize it as a binary variable which takes the value of 1 if the primary source of drinking water is unsafe (groundwater sources including tube-well, wells, protected and unprotected springs) and 0 for safer sources of drinking water (piped into dwelling, piped to yard, public tap/standpipe, rainwater, tanker truck, cart with small tank, bottled water, community RO plant). We have excluded from the categorization drinking water that comes from surface sources such as rivers/dams/lakes/ponds/streams/canals. This is done to make a clear distinction between groundwater and safer water sources. Surface water is likely to be contaminated with biological contaminants, making the analysis complicated. However, dropping this sample is not a major cause of concern as only 0.7% of the sample procured drinking water through this source while 78% of households rely on groundwater sources for drinking.

The main variable of interest is $\boldsymbol{a_3}$ which captures the health effects of drinking groundwater among females. We estimate the equation separately for each birth order for both categories of states: arsenic contaminated and non-arsenic contaminated states. Note that groundwater contaminants can be via other forms such as agricultural chemicals and septic waste which may also have adverse implications on health outcomes of children. However, this should not lead to adverse birth order effects. On the other hand, as discussed in this paper, drinking arsenic contaminated groundwater could have adverse birth order effects.

Looking at the interaction effect of gender and groundwater, results in Table 8 show that for non-arsenic states there is no effect on health outcomes. This is true for both measures of health, HAZ and WAZ, and across birth order 1 and birth order 3 and 4. On the other hand, the right panel shows results for arsenic states. In these states being a later born girl in a household which consumes groundwater is associated with a large negative effect on health outcomes. At the same time, this effect is insignificant for those born in the first birth order.

5.3 Falsification Test

We next conduct a falsification test using the reduced form relation between the instrument (clayey soil) and health outcomes. If our results were driven by differences in agricultural patterns including crop type and irrigation availability or differences in female labor force participation rates, we should find a significant effect of soil type on health outcomes even in non-arsenic areas. On the other hand, if clayey soil is indeed exogenous and does not affect health outcomes through any other channels, then in areas where there is no evidence of arsenic in groundwater, clayey soil should have no effect on health outcomes. The exclusion restriction implies that this falsification test should have a zero effect on health outcomes in non-arsenic districts.

In Table 9 we present the OLS estimates for the effect of clayey soil on HAZ and WAZ scores for girls in districts where arsenic is absent (columns 1 and 2) and compare it to the results from districts where arsenic is present (columns 3 and 4). The results confirm that the identification assumptions are met as clayey soil has no effect on the girl child's HAZ and WAZ scores in areas where no arsenic is present.

5.4 Soil quality and female labor supply

In a recent study, Carranza (2014) argues that loamy soils allow for deep tillage and thereby reduces the need for female dominated agricultural tasks. As a result, in areas with a greater fraction of loamy relative to clayey soils women have a lower economic value. Consistent with this, she finds that the exogenous variation of soil quality (loamy soils) across districts in India can explain variation in the share of female agricultural labor participation and sex ratio. The falsification check in the previous sub-section addresses this concern by showing that clayey soil affects health outcomes only in arsenic prominent districts. Yet, we conduct a further check on the robustness of our results by controlling for male and female labor force participation in agriculture across Indian districts[16].

Labour force participation in agriculture for female (male) is calculated at the district level and is measured as the total female (male) employment in agriculture divided by the total cultivable land in the district. We control for both the male and female employment in agriculture and find that the results are robust to this inclusion (Table 10).[17]

6. Conclusion and Policy Implications

Gender inequality is one of the most fundamental challenges to sustainable development. While considerable efforts have been made to explore the impacts of gender inequality on women, lesser is still known regarding its impact on child health. India is the only developing country where the

[16] Data on total female employment in agriculture, total male employment in agriculture and total cultivable land is taken from the employment round of National Sample Survey data (2011-12).

[17] Our results are also robust to dropping one state at a time from the analysis, and, one region at a time after dividing the 9 arsenic states into three regions: North, South and East (There are no western states with arsenic). These results are available upon request.

under-five child mortality rates are worse among girls than boys (Census, Government of India, 2011). This might be due to discrimination in resource allocation by parents at early stages of their lives, in the form of shorter duration of breastfeeding, lesser post-natal health inputs such as vaccination and supplementary food items.

This paper adds to the literature on gender discrimination and child health by highlighting the importance of environmental factors in widening the gender gap in health outcomes. Using a large nationally representative sample of children in India (NFHS, 2015-16), we find that exposure to arsenic contaminated water leads to a height and weight disadvantage among girls that increases with birth order. These estimates suggest higher valuation of sons health than daughters health by their parents, since boys are perceived to yield better economic benefits than girls in later stages of their life. Due to paucity of resources, boys are given preference in terms of better health inputs than girls. We find that the detrimental effects of arsenic on HAZ exists only in poorer households suggesting that nutritional deficiencies in childhood exacerbate the adverse effects of arsenic exposure.

Our results show heterogeneous effect of arsenic exposure by birth order highlighting the role played by son biased preferences in magnifying the negative impact of unsafe water on health for girls. Despite safe water being an indispensable input to human health, to the best of our knowledge, there is no existing research that has studied the role of gender in the relation between access to safe water and child health. According to the World Health Organization, lack of accessibility of safe water is leading cause of morbidity in India. Consumption of arsenic contaminated water is likely to be a contributor to India's high child mortality rate of 39 deaths per thousand live births (Assadullah and Chaudhary 2011). But any government policy that solely aims to provide safe drinking water will not deliver desired goals unless and until these policies are accompanied by equitable distribution of food and other health care inputs to young children particularly girls. Water related policies would reduce the burden of diseases to some extent, but lower immunity of girls would remain a challenge.

References

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

Arnold, F., Choe, M. K., & Roy, T. K. (1998). Son preference, the family-building process and child mortality in India. *Population studies*, *52*(3), 301-315.

Asadullah, M. N., Chaudhury N.(2011), “Poisoning the mind: Arsenic contamination of drinking water wells and children's educational achievement in rural Bangladesh”, *Economics of Education Review* 30(5):873-888.

Bardhan, P. K. (1974). On life and death questions. *Economic and Political Weekly*, 1293-1304.

Barker, D. J., Godfrey, K. M., Gluckman, P. D., Harding, J. E., Owens, J. A., & Robinson, J. S. (1993). Fetal nutrition and cardiovascular disease in adult life. *The Lancet*, *341*(8850), 938-941.

Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of political Economy*, *84* (4), 143-162.

Behrman, Jere R and Paul Taubman (1986) “Birth order, schooling, and earnings,” *Journal of Labor Economics*, Vol. 4, pp. 121–145.

Behrman, J. R., Pollak, R. A., & Taubman, P. (1986). Do parents favor boys?. *International Economic Review*, 33-54.

Bhalotra, S., Valente, C., & Van Soest, A. (2010). The puzzle of Muslim advantage in child survival in India. *Journal of Health Economics*, *29*(2), 191-204.

Borooah, Vani K., and Sriya Iyer. (2005). "Religion, literacy, and the female-to-male ratio." *Economic and Political Weekly*, 419-427.

Booth, A. L., & Kee, H. J. (2009). Birth order matters: the effect of family size and birth order on educational attainment. *Journal of Population Economics*, *22*(2), 367-397.

Bozack, A. K., Saxena, R., & Gamble, M. V. (2018). Nutritional influences on one-carbon metabolism: effects on arsenic methylation and toxicity. *Annual review of nutrition*, *38*, 401-429.

Brainerd, E., & Menon, N. (2014). Seasonal effects of water quality: The hidden costs of the Green Revolution to infant and child health in India. *Journal of Development Economics*, *107*, 49-64.

Brammer, Hugh & Ravenscroft, Peter. (2009). Arsenic in groundwater: A threat to sustainable agriculture in South and South-east Asia. *Environment international*. 35. 647-54. 10.1016/j.envint.2008.10.004.

Briend, A., & Zimicki, S. (1986). Validation of arm circumference as an indicator of risk of death in one to four-year-old children. *Nutrition research*, *6*(3), 249-261.

Carranza, E. (2014). “Soil endowments, female labour force participation and the demographic deficit of women in India”, *American Economic Journal: Applied Economics*, *6*(4): 197-225.

Case, A., & Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of political Economy*, *116*(3), 499-532.

Central Ground Water Board (2006), *Dynamic Ground Water Resources of India (as on March, 2004)*, Ministry of Water Resources, Government of India, New Delhi.

Chakrabarti, Anindita, and Kausik Chaudhuri. (2007). “Antenatal and Maternal Health Care Utilization: Evidence from Northeastern States of India.” *Applied Economics* 39 (6): 683–95.

Clougherty JE. (2010) A growing role for gender analysis in air pollution epidemiology. *Environ Health Perspective* Feb;118(2):167-76.

Chung, W., & Gupta, M. D. (2007). The decline of son preference in South Korea: The roles of development and public policy. *Population and development review*, *33*(4), 757-783.

Deaton, A., & Drèze, J. (2009). Food and nutrition in India: facts and interpretations. *Economic and political weekly*, 42-65.

DeRose, LF, Das, M & Millman, SR. (2000). Does female disadvantage mean lower access to food? *Population and Development Review* 26(3), 517–547.

Do, Q. T., Joshi, S., & Stolper, S. (2018). Can environmental policy reduce infant mortality? Evidence from the Ganga Pollution Cases. *Journal of Development Economics*, *133*, 306-325.

Dyson, T., & Moore, M. (1983). On kinship structure, female autonomy, and demographic behavior in India. *Population and development review*, 35-60.

Fängström, B., Moore, S., Nermell, B., Kuenstl, L., Goessler, W., Grandér, M., ... & Vahter, M. (2008). Breast-feeding protects against arsenic exposure in Bangladeshi infants. *Environmental health perspectives*, *116*(7), 963-969.

Fledderjohann, J., Agrawal, S., Vellakkal, S., Basu, S., Campbell, O., Doyle, P.Stuckler, D. (2014). Do girls have a nutritional disadvantage compared with boys? Statistical models of breastfeeding and food consumption inequalities among Indian siblings, *PloS one*, *9*(9), e107172.

Foster, A., Gutierrez, E., & Kumar, N. (2009). Voluntary compliance, pollution levels, and infant mortality in Mexico. *American Economic Review*, *99*(2), 191-97.

Gardner RM, Kippler M, Tofail F, Bottai M, Hamadani J, Grandér M, Nermell B, Palm B, Rasmussen KM, Vahter M. (2013). Environmental exposure to metals and children's growth to age 5 years: a prospective cohort study. *American Journal of Epidemiology*. 2013 Jun 15;177(12):1356-67.

Garg, A., & Morduch, J. (1998). Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *Journal of Population Economics*, *11*(4), 471-493.

Glewwe, P., & Miguel, E. A. (2007). The impact of child health and nutrition on education in less developed countries. *Handbook of development economics*, *4*, 3561-3606.

Goyal, N., & Canning, D. (2018). Exposure to ambient fine particulate air pollution in utero as a risk factor for child stunting in Bangladesh. *International journal of environmental research and public health*, *15*(1), 22.

Greenstone, M., & Hanna, R. (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review*, *104*(10), 3038-72.

Gupta, Monica. Das. (1987). Selective discrimination against female children in rural Punjab, India. *Population and development review*, 77-100.

Heck JE, Chen Y, Grann VR, et al. (2008) Arsenic exposure and anemia in Bangladesh: a population-based study. *Journal of Occupational and Environmental Medicine,* 2008;50(1):80–87.

Herath, I., Vithanage, M., Bundschuh, J. *et al.* (2016). Natural Arsenic in Global Groundwaters: Distribution and Geochemical Triggers for Mobilization, *Current Pollution Reports* 2, 68–89.

Jayachandran, Seema & Ilyana Kuziemko. (2011). Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India, *The Quarterly Journal of Economics*, Volume 126, Issue 3, August 2011, Pages 1485–1538,

Jayachandran, S., & Pande, R. (2017). Why are Indian children so short? The role of birth order and son preference. *American Economic Review*, *107*(9), 2600-2629.

Karim, M. R., & Ahmad, S. A. (2014). Nutritional status among the children of age group 5-14 years in selected arsenic exposed and non-exposed areas of Bangladesh. *Journal of family & reproductive health*, *8*(4), 161.

Keskin, P., Shastry, G. K., & Willis, H. (2017). Water quality awareness and breastfeeding: evidence of health behavior change in Bangladesh, *Review of Economics and Statistics*, *99*(2), 265-280.

Kile, M. L., Cardenas, A., Rodrigues, E., Mazumdar, M., Dobson, C., Golam, M., … Christiani, D. C. (2016). Estimating Effects of Arsenic Exposure During Pregnancy on Perinatal Outcomes in a Bangladeshi Cohort, *Epidemiology (Cambridge, Mass.)*, *27*(2), 173–181.

Kugler, A. D., & Kumar, S. (2017). Preference for boys, family size, and educational attainment in India. *Demography*, *54*(3), 835-859.

Lundberg, S. (2005). Sons, daughters, and parental behaviour. *Oxford Review of Economic Policy*, *21*(3), 340-356.

Madajewicz, M. et al. (2007). Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh, *Journal of Development Economics*, 84(2): 731-754.

Marcoux, A, (2002). Sex differentials in undernutrition: A look at survey evidence. *Population and Development Review* 28(2), 275–284.

McArthur, J. M., Ravenscroft, P., Safiulla, S., & Thirlwall, M. F. (2001). Arsenic in groundwater: testing pollution mechanisms for sedimentary aquifers in Bangladesh, *Water Resources Research*, *37*(1): 109-117.

Milton, A. H., Attia, J., Alauddin, M., McEvoy, M., McElduff, P., Hussain, S., ... & Iyengar, V. (2018). Assessment of nutritional status of infants living in arsenic-contaminated areas in Bangladesh and its association with arsenic exposure. *International Journal of Environmental Research and Public Health*, *15*(1), 57.

Minamoto, K., Mascie-Taylor, C. G., Moji, K., Karim, E., & Rahman, M. (2005). Arsenic-contaminated water and extent of acute childhood malnutrition (wasting) in rural Bangladesh. *Environmental sciences: an international journal of environmental physiology and toxicology*, *12*(5), 283.

Mitra, S. R., Mazumder, D. G., Basu, A., Block, G., Haque, R., Samanta, S., ... & Smith, A. H. (2004). Nutritional factors and susceptibility to arsenic-caused skin lesions in West Bengal, India. *Environmental health perspectives*, *112*(10), 1104-1109.

Pande, R. P. (2003). Selective gender differences in childhood nutrition and immunization in rural India: the role of siblings. *Demography*, *40*(3), 395-418.

Price, J. (2008). Parent-child quality time does birth order matter? *Journal of human resources*, *43*(1), 240-265.

Rahman, A., Vahter, M., Smith, A. H., Nermell, B., Yunus, M., El Arifeen, S., ... & Ekström, E. C. (2009). Arsenic exposure during pregnancy and size at birth: a prospective cohort study in Bangladesh. *American journal of epidemiology*, *169*(3), 304-312.

Rosenzweig, M. R., & Schultz, T. P. (1982). Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India. *The American Economic Review*, *72*(4), 803-815.

Tarozzi, A. (2008). Growth reference charts and the nutritional status of Indian children. *Economics & Human Biology*, *6*(3), 455-468.

UNICEF. 2013. Improving Child Nutrition: The Achievable Imperative for Global Progress. New York: UNICEF.

UNICEF. 2017. Reducing stunting in children under five years of age: a comprehensive evaluation of UNICEF's strategies and program performance-India: UNICEF.

Victoria, C. G., J. P. Vaughan, C. Lombardi, S. M. C. Fuchs, L. P. Gigante, P. G. Smith, L. C. Nobre, A. M. Texiera, L. B. Moreira, and F. C. Barros (1987). "Evidence for a Strong Protective Effect of Breastfeeding against Infant Death due to Infectious Diseases in Brazil" *Lancet, 330(8554), 319–322.*

Walker SP, Wachs TD, Gardner JM, et al. (2007). Child development: risk factors for adverse outcomes in developing countries, *Lancet*, vol. 3699556, 145-157.

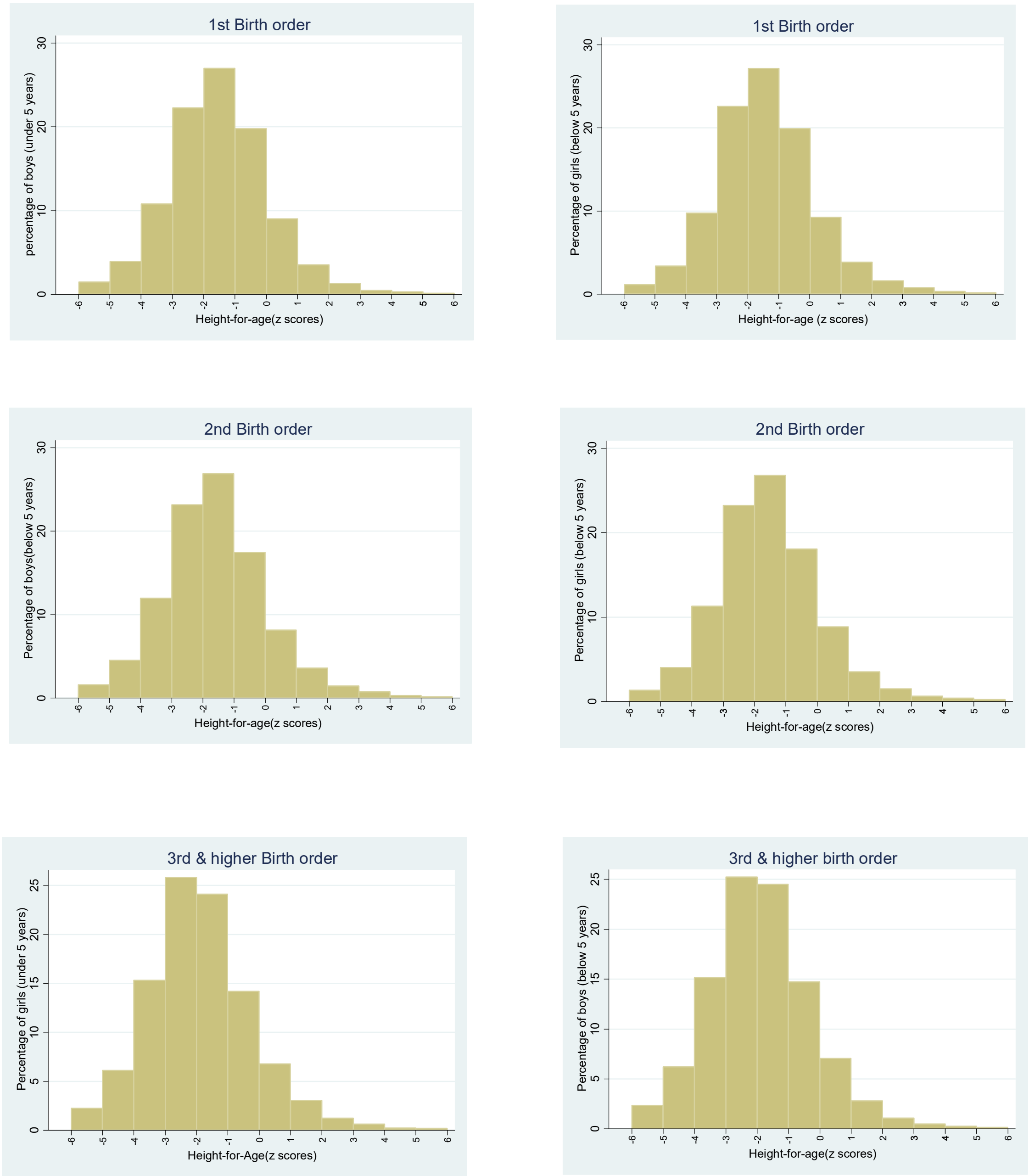
Watanabe, C., Matsui, T., Inaoka, T., Kadono, T., Miyazaki, K., Bae, M. J., ... & Mozammel Haque Bokul, A. T. M. (2007). "Dermatological and nutritional/growth effects among children living in arsenic-contaminated communities in rural Bangladesh" *Journal of environmental science and health, part a*, *42*(12), 1835-1841.

Waterlow, J. C. (1972). Classification and definition of protein-calorie malnutrition. *British medical journal*, *3*(5826), 566.

World Health Organization (2000): "Effect of Breastfeeding on Infant and Child Mortality Due to Infectious Diseases in Less Developed Countries: A Pooled Analysis," Lancet, 355(9202), 451–455.

Figure 1: Histogram for Height-for-age (z scores) for boys & girls, by birth order

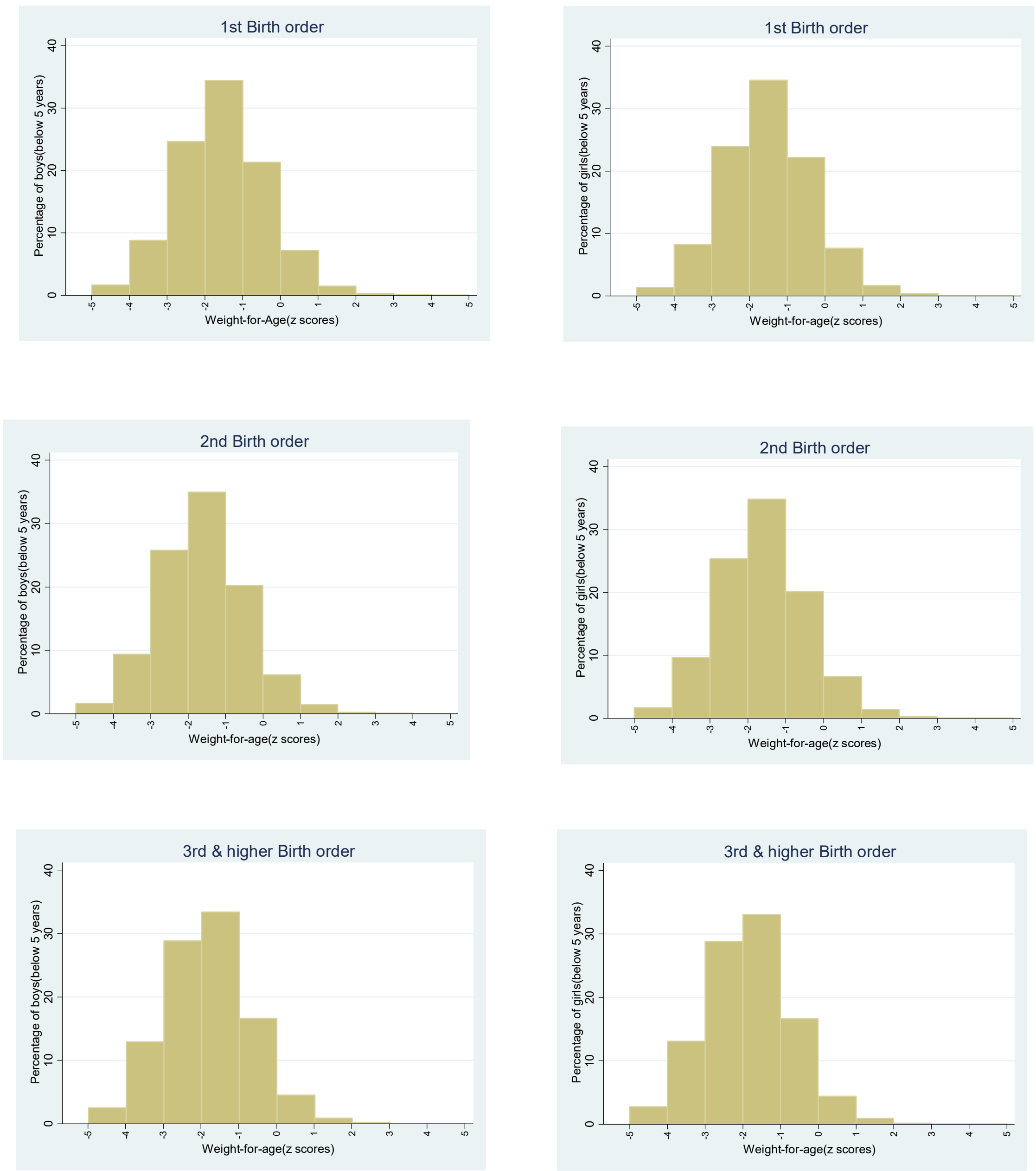
BOYS GIRLS



Note: The above figure shows percentage of boys and girls whose z scores for Height-for-age (z scores) decrease with increasing birth order, as represented on the horizontal axis. As per guidelines issued by World Health Organization children whose z scores are below -2 and above -3 indicates moderate stunting and z scores below -3 indicates severe stunting.

Figure 2: Histogram for Weight-for-age (z scores) for boys & girls, by birth order

BOYS GIRLS



Note: The above figure shows percentage of boys and girls whose z scores for Weight-for-age (z scores) decrease with increasing birth order, as represented on the horizontal axis. As per guidelines issued by World Health Organization children whose z scores are below -2 and above -3 indicates moderate underweight and z scores below -3 indicates severe underweight.

Table 1: Descriptive Statistics and District level control variables

Variable	Mean	Std. Dev.
Height-for-age (z scores)	-1.64	1.65
Weight-for-age (z scores)	-1.66	1.18
Arsenic (ug/l)	94	444
Clayey soil (percentage)	27.89	7.84
Groundwater usage for drinking	0.78	0.41
Individual characteristics		
Birth order (first)	0.34	0.47
Birth order (second)	0.29	0.45
Birth order (third)	0.36	0.48
Age	2.23	1.49
% Girls	0.48	0.50
Maternal characteristics		
Mother's education		
Illiterate	0.37	0.48
Primary	0.14	0.35
Secondary	0.40	0.49
Higher & above	0.09	0.28
Mother's age	27	4.93
Family background characteristics		
Hindu	0.78	0.42
Muslim	0.18	0.39
Others	0.04	0.20
Scheduled Caste/Scheduled Tribe	0.32	0.46
Other Backward caste	0.50	0.50
Higher/Upper castes	0.19	0.38
Urban	0.21	0.41
Household size	6.90	3.11
Wealth Index	2.55	1.40
District level control variables		
Sex ratio (Female/male)	925	44
Rainfall (millimeters)	76.7	42
Iron (mg/l)	1.59	2.53
Health centers	316	154
Monthly per capita expenditure (rupees)	169310	66931
% Literacy	68.4	8.24

Sample size is N=88,887

Table 2: Arsenic and Child Anthropometric Measures: By Gender (OLS Estimates)

	HAZ	**WAZ**	**HAZ**	**HAZ**	**WAZ**	**WAZ**
Anthropometric measures (z scores)	**Full**	**Full**	**Girls**	**Boys**	**Girls**	**Boys**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Arsenic	-0.020	0.035***	-0.027	-0.016	0.038**	0.031**
	(0.013)	(0.012)	(0.017)	(0.016)	(0.014)	(0.016)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes	Yes	Yes
Family background	Yes	Yes	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
State F.E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	88,887	88,887	42,625	46,262	42,625	46,262

*Standard errors clustered at the PSU level in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Arsenic is measured in milligrams per liter. All regressions include state fixed effects and district level controls for sex ratio, health facilities, rainfall, literacy, iron and MPCE, individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size, wealth index and place of residence).

Table 3: First Stage Regression

	Arsenic (microgram/liter)
Clayey soil (sub)	0.007***
	(0.000)
First stage F-statistics	63.6
Observations	88,887

Note: SE clustered at the PSU level (*** Significant at 1%, ** significant at 5%, * significant at 10 %.). Independent variable is defined as percentage of clayey soil present in district. Regressions include state fixed effects and district level controls for sex ratio, health facilities, rainfall, literacy, iron and MPCE. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size, wealth index and place of residence).

Table 4: Arsenic and Child Anthropometric Measures (IV Estimates)

	HAZ	WAZ	HAZ	HAZ	WAZ	WAZ
Anthropometric measures (z scores)	**Full (1)**	**Full (2)**	**Girls (1)**	**Boys (2)**	**Girls (3)**	**Boys (4)**
Arsenic	-0.585***	-0.566***	-0.632**	-0.553**	-0.677***	-0.475***
	(0.176)	(0.136)	(0.226)	(0.203)	(0.181)	(0.150)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes	Yes	Yes
Family background	Yes	Yes	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
State F.E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	88,887	88,887	42,625	46,262	42,625	46,262

Note: SE clustered at the PSU level (*** Significant at 1%, ** 5%, * 10 %.). Instrument for arsenic is defined as the % of clayey soil present in a district. Regressions include state fixed effects and district level controls for sex ratio, health facilities, rainfall, literacy, iron and MPCE. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size, wealth index and place of residence).

Table 5: Heterogeneous effects by household wealth (IV Estimates)

	HAZ	HAZ	WAZ	WAZ
Anthropometric measures (z-score)	**(1)**	**(2)**	**(3)**	**(4)**
Wealth Index	**Low**	**High**	**Low**	**High**
Arsenic	-0.639***	-0.332	-0.510***	-0.463***
	(0.219)	(0.223)	(0.162)	(0.170)
Observations	48,574	40,313	48,574	40,313
Individual controls	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes
State F.E	Yes	Yes	Yes	Yes

Note: SE clustered at the PSU level (*** Significant at 1%, ** significant at 5%, * significant at 10 %.). Instrument for arsenic is defined as the percentage of clayey soil present in a district. Regressions include state fixed effects and district level controls for sex ratio, health facilities, rainfall, literacy, iron and MPCE. Individual level controls (age,

age square), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size and place of residence).

Table 6: Arsenic, gender and birth order gradient in Height-for-age and Weight-for-age (IV Estimates)

	HAZ (OLS) (1)	HAZ (IV) (2)	HAZ (IV) (3)	WAZ (OLS) (4)	WAZ (IV) (5)	WAZ (IV) (6)
Arsenic*girl*BO2	0.028	-0.384	-0.334	0.047	-0.444**	-0.390**
	(0.041)	(0.257)	(0.250)	(0.041)	(0.185)	(0.182)
Arsenic*girl*BO3	-0.041	-0.678**	-0.582**	-0.045	-0.647***	-0.524***
	(0.056)	(0.288)	(0.279)	(0.044)	(0.216)	(0.212)
Arsenic	0.021	-0.348	-0.293	0.056**	-0.272	-0.312**
	(0.021)	(0.213)	(0.207)	(0.020)	(0.157)	(0.156)
Birth Order 2	-0.033***	-0.024	-0.025	-0.025***	-0.017	-0.012
	(0.013)	(0.024)	(0.025)	(0.009)	(0.018)	(0.018)
Birth Order 3	-0.077***	0.052	-0.038	-0.058***	0.043	0.049*
	(0.014)	(0.034)	(0.034)	(0.010)	(0.026)	(0.025)
Girls	0.080***	0.070***	0.088***	0.020***	0.031*	0.029
	(0.011)	(0.024)	(0.024)	(0.008)	(0.018)	(0.018)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes	Yes	Yes	Yes
District level controls	Yes	No	Yes	Yes	No	Yes
Observations	86,274	102,731	86,274	86,274	101,814	86,274

Note: SE clustered at the PSU level (*** Significant at 1%, ** significant at 5%, * significant at 10 %.). Instrument for arsenic is defined as the percentage of clayey soil present in a district. Regressions include state fixed effects and district level controls for sex ratio, health facilities, rainfall, literacy, iron and MPCE. Individual level controls (age, age square), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size, wealth index and place of residence). Also includes all double interaction terms.

Table 7: Arsenic, sibling size and birth order (IV Estimates)

	HAZ	**HAZ**	**WAZ**	**WAZ**
Anthropometric measures	Girls	Boys	Girls	Boys
	(1)	(2)	(3)	(4)
Arsenic*birth order2	-0.248	0.499**	-0.596***	0.184
	(0.249)	(0.247)	(0.190)	(0.180)
Arsenic*birth order3^{rd+}	-0.996***	-0.094	-1.118***	-0.296
	(0.263)	(0.250)	(0.203)	(0.185)
Sibling size (below age 5)	-0.089	-0.262***	0.066	-0.089*
	(0.065)	(0.067)	(0.050)	(0.049)
Individual controls	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes
State F.E	Yes	Yes	Yes	Yes
Observations	41,131	44,622	41,131	44,622

Note: SE clustered at the PSU level (*** Significant at 1%, ** significant at 5%, * significant at 10 %.). Instrument for arsenic is defined as the percentage of clayey soil present in a district. Regressions include state fixed effects and district level controls for sex ratio, health facilities, rainfall, literacy, iron and MPCE. Individual level controls (age, age square), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size, wealth index and place of residence). Also includes controls for arsenic and birth order.

Table 8: Interaction effect of groundwater and gender on health outcomes by birth order (Arsenic and Non-Arsenic States)

	Non-Arsenic States				*Arsenic States*			
	HAZ	HAZ	WAZ	WAZ	HAZ	HAZ	WAZ	WAZ
	BO1	BO3/4	BO1	BO3/4	BO1	BO3/4	BO1	BO3/4
groundwater	0.045**	0.066***	0.026*	0.034**	0.015	0.203***	0.062*	0.062***
	(0.019)	(0.023)	(0.014)	(0.016)	(0.049)	(0.054)	(0.037)	(0.024)
Girl	0.087***	0.013	0.050***	-0.030**	0.122***	0.165***	0.081**	0.050**
	(0.017)	(0.022)	(0.013)	(0.015)	(0.047)	(0.059)	(0.036)	(0.023)
girl*groundwater	-0.019	-0.032	-0.021	-0.001	-0.030	-0.179***	-0.057	-0.051**
	(0.023)	(0.027)	(0.017)	(0.018)	(0.056)	(0.066)	(0.042)	(0.026)
Observations	65,855	58,743	65,855	58,743	14,176	16,342	14,176	16,342
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered at the PSU level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include district fixed effects and Individual level controls (age and age square), maternal controls (mother's age and mother's education) and family background controls (caste, religion, family size, household wealth index and place of residence). Groundwater is a binary variable which takes value 1 if the primary source of drinking water is groundwater (tube-well, well, unprotected springs) and 0 for safer sources of drinking water (piped water, community RO plant, bottle water, rainwater harvesting).

Table 9: Reduced form relationship between clayey soil and health outcome among girls in Arsenic States

	(1)	(2)	(3)	(4)
	Non-Arsenic Districts		Arsenic Districts	
	HAZ	WAZ	HAZ	WAZ
Clayey Soil	-0.0015	-0.0029	-0.0061**	-0.0039**
	(0.0018)	(0.002)	(0.002)	(0.002)
State Fixed Effects	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
District Level Controls	Yes	Yes	Yes	Yes
Observations	19,854	19,854	17,450	17,450

Standard errors clustered at the PSU level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include state fixed effects and district level controls for sex ratio, health facilities, rainfall, literacy, iron and MPCE. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (caste, religion, family size, household wealth index and place of residence).

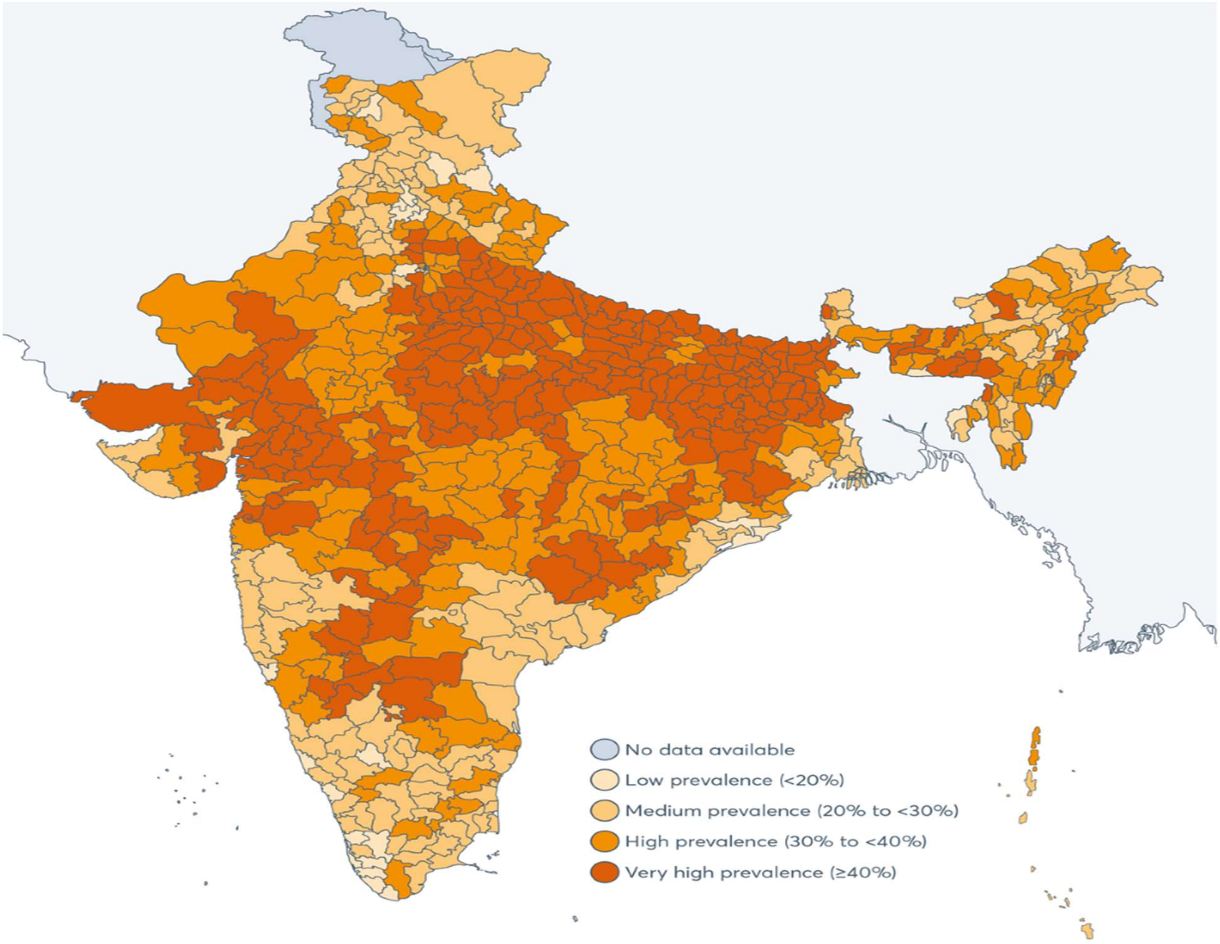
Table 10: Robustness Check: Controlling for male/female labor force participation in agriculture (IV Estimates)

	HAZ (IV) (3)	WAZ (IV) (6)
Arsenic*girl*BO2	-0.334	-0.390**
	(0.250)	(0.182)
Arsenic*girl*BO3	-0.582**	-0.524***
	(0.279)	(0.212)
Arsenic	-0.293	-0.312**
	(0.207)	(0.156)
Birth Order 2	-0.025	-0.012
	(0.025)	(0.018)
Birth Order 3	-0.038	0.049*
	(0.034)	(0.025)
girls	0.088***	0.029
	(0.024)	(0.018)
Individual controls	Yes	Yes
Maternal controls	Yes	Yes
Family background controls	Yes	Yes
District level controls	Yes	Yes
Observations	86,274	86,274

Note: SE clustered at the PSU level (*** Significant at 1%, ** significant at 5%, * significant at 10 %.). Instrument for arsenic is defined as the percentage of clayey soil present in a district. Regressions include state fixed effects and district level controls for male and female labor force participation rate in agriculture, sex ratio, health facilities, rainfall, literacy, iron and MPCE. Individual level controls (age, age square), maternal controls (mother’s age and mother’s education) and family background controls (religion, caste, family size, wealth index and place of residence). Also includes all double interaction terms.

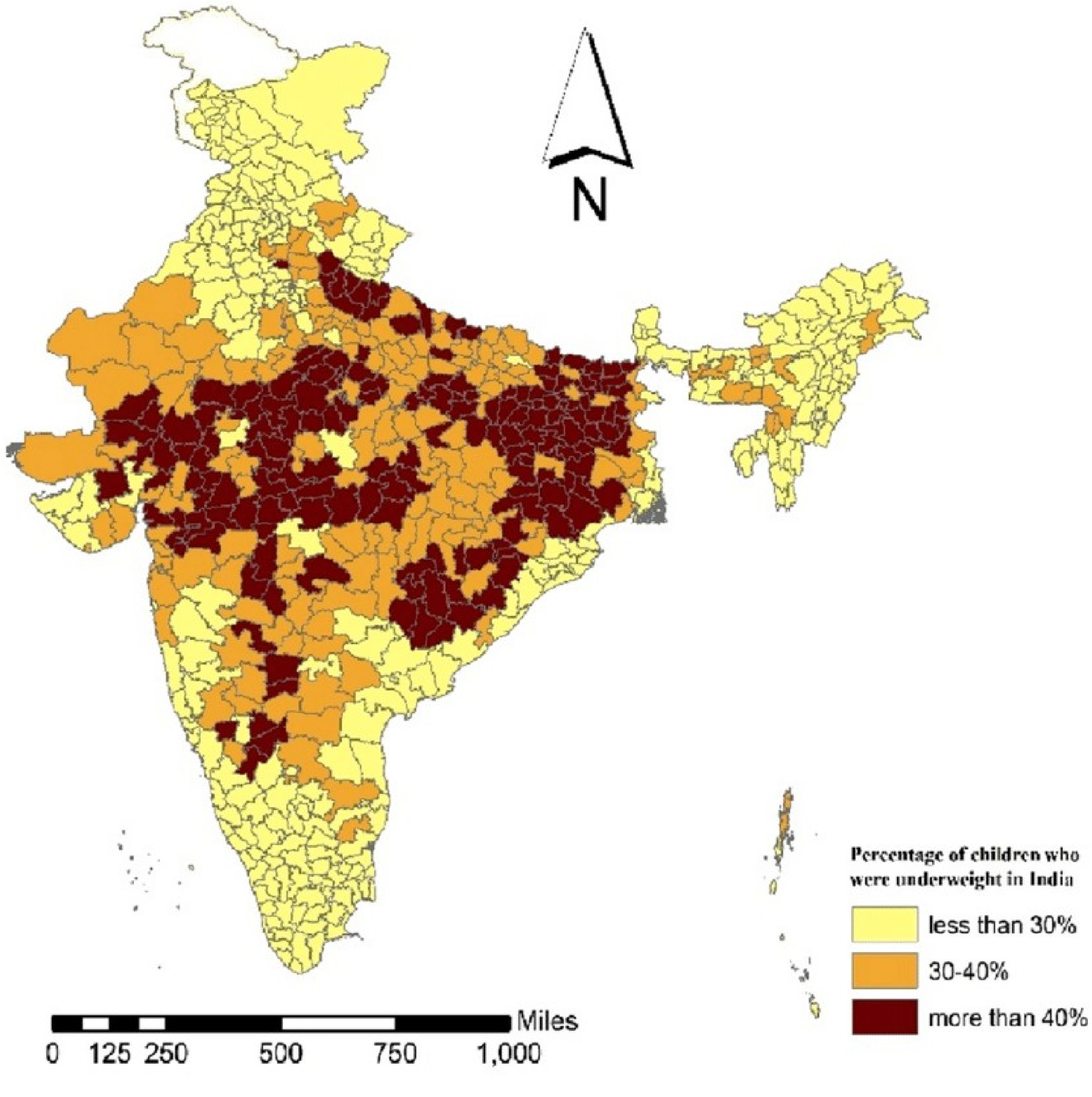
Appendix A.1: Figures and Maps

Figure A.1.1: Prevalence of stunting across districts of India



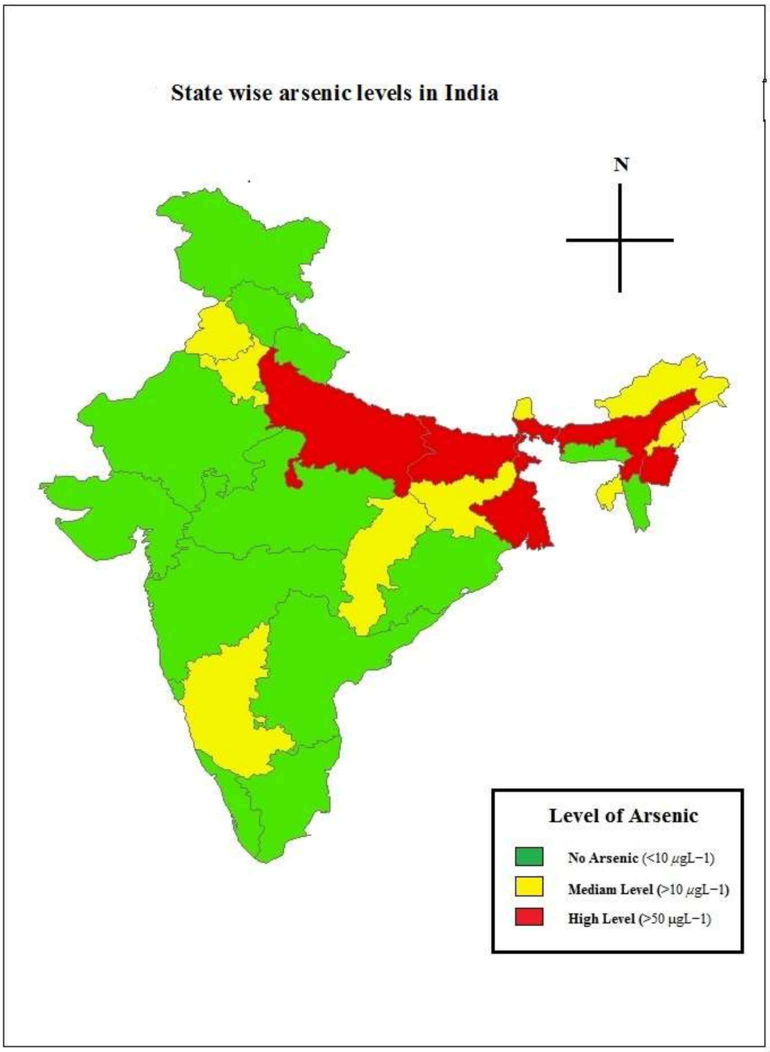
Source: Menon et al. (2018)

Figure A.1.2: Prevalence of underweight across districts of India



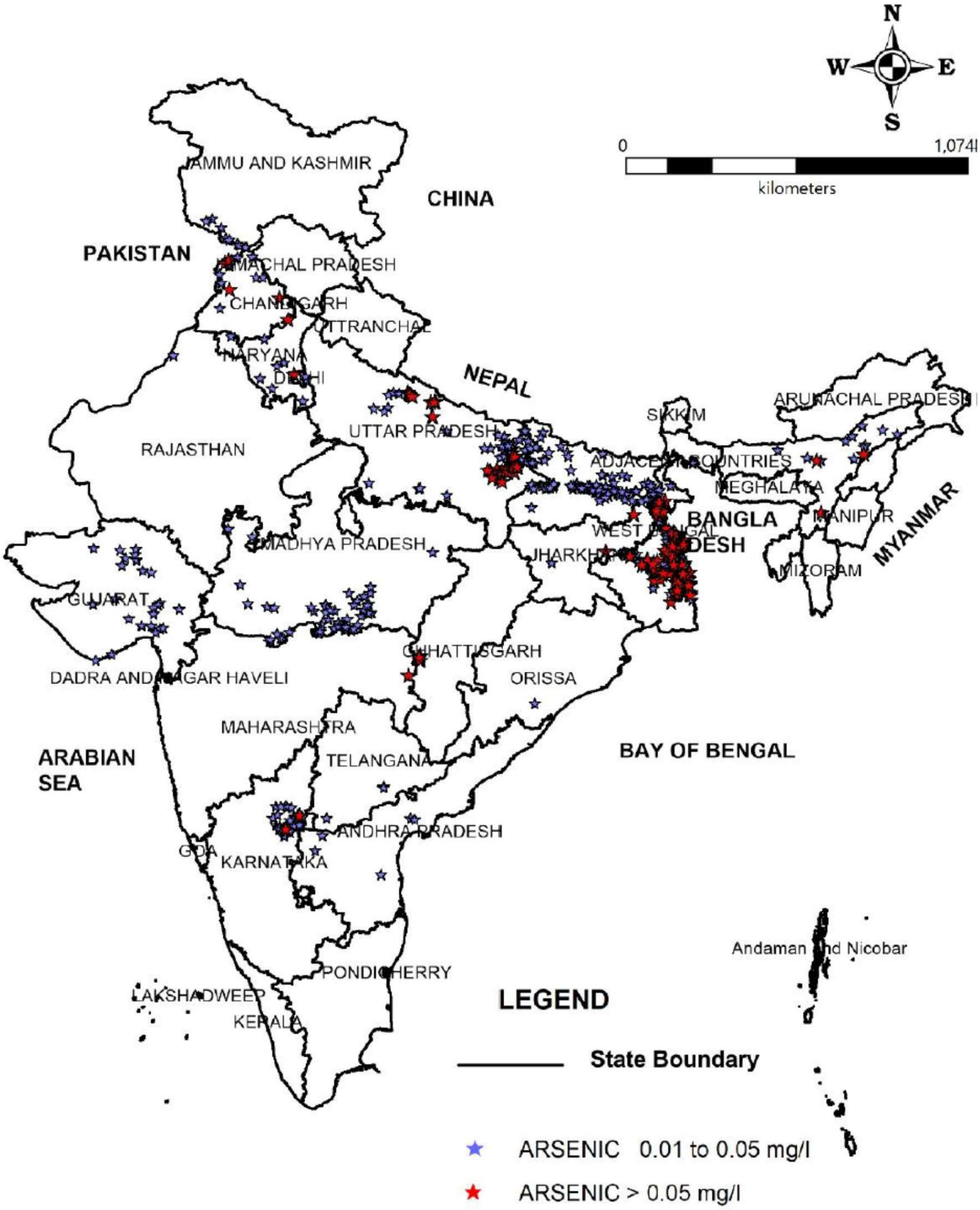
Source: Sharma et al. (2020)

Figure A.1.3: Geographical Distribution of Arsenic Levels across States of India



Source: Authors calculation using Central Ground Water Board report data (2016)

Figure A.1.4: Geographical Distribution of Arsenic Levels across Districts of India



Source: Central Groundwater Board

Appendix A.2 Additional Tables

Table A.2.1 Cross Tabulation of district level characteristics by arsenic contamination

Variable	Mean	N	Mean	N	T stats
	Non-arsenic districts		Arsenic districts		
Iron	1.47	156	1.74	105	-0.94
	2.37		2.22		
Fluoride	0.69	154	0.56	108	0.87
	0.10		0.11		
Nitrate	67.11	154	58.85	108	0.70
	7.74		8.83		
Rainfall	81.28	100	69.32	74	1.61
	56.8		33.94		
Maximum temperature	38.18	37	39.74	22	-1.10
	0.86		1.13		
Minimum temperature	12.09	25	10.36	14	0.91
	1.07		1.65		
Rice/Wheat(production)	1140.697	117	2580.698	90	-0.65
	6481.86		22881.09		
Nitrogen	25967.5	148	30756.68	93	-1.40
	2001.63		2901.66		
Phosphorus	11584.35	148	14138.46	93	-1.38
	1141.49		1466.88		
Potassium	3549.5	148	4762.68	93	-1.54
	426.73		724.76		
Literacy	69.162	154	69.163	105	0
	8.59		8.95		
Sex ratio	937.92	154	925.7	105	2.15*
	48.26		39.45		
Monthly per capita expenditure	175857.3	156	173518.5	105	0.26
	73933.62		63950.34		

Table A.2.2: Clayey soil and District Level Characteristics

	Clayey soil
Other Contaminants:	
Iron (mg/liter)	0.043**
	(0.019)
Fluoride (mg/liter)	0.345
	(0.462)
Nitrate (mg/liter)	-0.007
	(0.005)
Nitrogen (Kilogram/hectare)	0.000
	(0.000)
Phosphorus (Kilogram/hectare)	0.000
	(0.000)
Potassium (Kilogram/hectare)	-0.000
	(0.000)
Weather:	
Rainfall (millimeters)	-1.01**
	(0.449)
Maximum temperature (degree Celsius)	0.105
	(0.237)
Minimum temperature (degree Celsius)	0.152
	(0.330)
Demographic& Economic Factors:	
Ratio of Rice to Wheat (million tonnes)	-0.000
	(0.000)
Literacy	0.091
	(0.060)
Sex Ratio (per 1000 females)	0.002
	(0.015)
Per Capital Expend.	0.000
	(0.000)
Male labor participation (agriculture)	-0.000
	(0.000)
Female labor participation (agriculture)	-0.000
	(0.000)
State fixed effects	Yes

*** Significant at 1%, ** 5%, *10%.

Table reports the coefficient on clayey soil, from the regression of reported district level variables on the % of clayey soils in a district and state fixed effects. Standard errors are clustered at the district level. N=257 districts