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

Khushboo Aggarwal[a] and Rashmi Barua[b]

January 2021

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 stunting and underweight in arsenic contaminated regions of India. Using a large nationally representative household survey 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 elder ones. Within India analysis shows that the effect of arsenic contamination on health outcomes among girls is comparatively higher in regions and communities with high preference for son. 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.

---

[a] Centre for International Trade and Development (CITD), Jawaharlal Nehru University (JNU), 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 thus, 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 might be due to 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.

In this paper, 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.[1] Children are more susceptible to arsenic because of their lower immunity levels and relatively higher proportion of body water compared to adults. Moreover, epidemiological evidence suggests that arsenic crosses the placenta and adversely impacts health in utero and later in life (Rahman et al. 2009;Kile et al. 2016).

While there is epidemiological evidence that arsenic affects child growth outcomes (Watanabe et al. 2007; Minamoto et al. 2005), 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, including shorter periods 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).[2]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 (stunting and underweight) in India. 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. Our data is sourced from the 2015-16 round of the National Family Health Survey (NFHS-4).

---

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

[2] 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.

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 (Mc Arthur et al, 2001).

Instrumental variable estimates indicate that exposure to arsenic in groundwater has negative and significant impact on HAZ 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 by 0.035 (2.11 percent) and 0.061 (3.67 percent) standard deviations for the second born and third born girl child, respectively, relative to a male child born at first birth order.[3]These findings are robust to the inclusion of district level controls for weather, water quality measures, pattern of cultivation, education quality and income. Further, we find similar negative effects of arsenic on weight-for-age (also called underweight) among later born girls as compared to the eldest son. These results hold consistent even after accounting for all district level controls.

The gender bias in nutrition and resource allocation has been extensively researched in India (Gupta 187; Behrman 1988; Jayachandran and Kuziemko 2011; Jose 2011; Fledderjohann et al. 2014; Jayachandran and Pande, 2017; Pande 2019). 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).

Consistent with this hypothesis, we find heterogenous effects of arsenic exposure by sex ratio, religion, caste and urban/rural status. In particular, the adverse effects among later born girls are relatively larger in districts with negatively skewed sex ratio where patrilineal Hindu kinship

---

[3] 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.

system is most likely followed (Gupta 1987). Adverse effects are also larger for Hindu households, compared to Muslim households, and households located in rural locations.

Finally, existing studies finds that children born at higher birth orders have a higher probability of being from a large size family. We conduct robustness checks where we separately account for family size and birth order effects. The results are robust to the use of gender of the first born as an instrument for family size.

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. Literature review

This paper is related to the literature that studies the impact of gender discrimination, measured by unequal parental investment, on child health. 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.

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). 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 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[4]. For instance, at first birth order approximately 10 percent of girls suffers from severe stunting which at later birth order increases to 12 percent and 15 percent for 2nd and 3rd+ 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. 38 percent of mother's are uneducated while 68 percent are educated (14 percent-primary, 39 percent-secondary and 9 percent-higher and above). 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.

---

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

3.2 District level Control variables

Data on total production under rice and wheat production is obtained from Ministry of Agriculture and Farmer's Welfare. Data for rainfall is provided by the Indian Meteorological Department (IMD) at district level in India, with a mean value of 79.24 mms. District level sex ratio and literacy data is acquired from the 2011 Census of India. The average sex ratio and literacy rate in our estimation sample is 928 and 68 percent respectively. To control for district level gross domestic product, we also use data on Monthly Per Capita Expenditure (MPCE) from 68th 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 μgL-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. The final dataset comprises of 73,160 children under the age of five, across 9 arsenic affected states and 261 districts, where 105 districts are arsenic affected and 156 are non-arsenic affected districts. As shown in Table 1, the average level of arsenic is 107.32 microgram per liter across districts in India, remarkably higher than the threshold limit. Apart from arsenic, mean level of iron is 1.6 mg/l as indicated in Table 1.

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 districts is approximately 28 percent.

4. Empirical Model

We investigative whether exposure to arsenic has an impact on growth of children as measured by Height-for-Age (z scores) and Weight-for-age (z scores). Thus, we estimate the following OLS regression separately for boys and girls:

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

We are interested in measuring the effect of arsenic on two outcome variables: height-for-age and weight-for-height 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 and place of residence). We also control for district level controls ($\boldsymbol{D_{ds}}$) for rainfall, pattern of cultivation, presence of other contaminants (iron),[5] per capita consumption expenditure, sex ratio 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[6].

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 issue of endogeneity, we use an instrumental variable approach.

4.1 Instrumental Variable Approach

Arsenic concentration is higher in clayey relative to coarse soil, thus we exploit the variation in soil texture across districts within a state to instrument for arsenic groundwater contamination. Finer soils have relatively more particle density and lower porosity levels and as a result, their permeability level is relatively lower than loamy soil[7] which facilitates arsenic concentration in groundwater (Mc Arthur et al. 2001; Madajewicz et al. 2007). Figure 3 provides the visual evidence for the positive correlation between arsenic and soil texture.

The first stage equation is given by:

---

[5] Additional robustness checks in the appendix controls for fluorides and nitrates

[6] 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.

[7] Loamy soil mainly consists of higher proportion of sandy and silty soil relative to clayey soil.

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

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 that soil texture fractions affect health outcomes only through the impact on the level of arsenic in groundwater.[8]

A threat to identification 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 district-level ratio of rice to wheat production. Further, we also include sex ratio (measured at district level) as soil texture can affect economic outcomes through relative female to male employment rates (Carranza 2014).

An additional threat to our identification strategy exists if clayey soil varies with other weather, geographic or demographic factors and these might affect economic outcomes. We provide evidence of no correlation between proportion of clayey soil and several district level indicators of weather (rainfall and temperature), other contaminants (nitrate and fluoride), economic or demographic factors (monthly per capita expenditure, rice to wheat production, literacy, sex ratio, usage of fertilizers), conditional on state fixed effects.[9]There is significant difference by soil permeability in iron, as districts with higher iron also have higher proportion of clayey soil (and thus more arsenic). However, this would be against finding a negative impact of arsenic on health outcomes and if anything, underestimate our findings. 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.[10]

---

[8] 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 soiled used in our analysis.

[9] Results are reported in Table A.2.1 in the Appendix.

[10] If the amount of rainfall is less than the soil can absorb, it will all 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 of water.

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:

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

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

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}) and state fixed effect (S). Heteroskedasticity robust standard errors are clustered at the PSU level.

There might be a potential source of bias in the above estimates due to family size. Existing studies finds that children born at higher birth orders have a higher probability of being from a large size family. Further, family size and resource allocated to each child are highly correlated and which might in turn could affect the health outcomes of children (Kugler and Kumar 2017; Booth and Kee 2005). For instance, siblings in a poor family are less likely to receive equal share of available resources allocated by parents towards their children health and education.

As a robustness check, we also control for family size measured by the number of children under the age of five in the household in the main regressions.[11] To overcome the issue of endogeneity of family size, we use the gender of first child as an instrument for family size. Having a girl as first child is positively associated with family size, particularly in the presence of son preference,

---

[11] The NFHS does not give information on the total number of children of all ages in the household or the total family size inclusive of adults. Anthropometric data is only measured for households with children less than five years of age.

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 family size (Kugler and Kumar 2017).

To determine if the health outcomes of arsenic exposure vary by family size and birth order, we control for family size and run the regressions separately by gender[12]. We use the gender of first child as an instrument for family size. We estimate the following OLS and first stage regressions separately by gender:

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

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

Where, $\boldsymbol{Fam\_size}$ is the size of the family measured by 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.

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. Children who belong to arsenic affected districts have lower height-for-age. OLS estimates are significant but small and a one standard deviation increase in arsenic, reduces HAZ by 0.002 standard deviation units (Column 1, Table 2). OLS Estimates for WAZ shows that arsenic exposure is associated positively with weight-for-age (0.003 standard deviation).

In Table 3 we study whether the impact of arsenic on stunting and underweight varies by gender. Higher exposure to arsenic contaminated water is positively associated with stunting among girls

---

[12] We run the regressions separately by gender since our instrument for family size (gender of first child) will otherwise be perfectly collinear with our main explanatory variables $\boldsymbol{Ars_{ds} * girl_{ids} * nth\ child_{ids}}$.

(0.003 standard deviation) but not boys. However, in case of underweight no such evidence of gender difference is found in Column 3 and column 4 (Table 3). A one standard deviation increases in arsenic leads to increase in Weight-for-age by 0.003 standard deviation units regardless of gender.

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 4 shows a positive and statistically significant relationship between arsenic and soil texture (clayey soil). The F-statistic (685.48) suggests that soil texture is a strong instrument for arsenic levels.

The IV results for HAZ, shown in table 5, indicate that the OLS is severely downward biased. A one standard deviation increases in arsenic leads to decrease in height-for-age by 0.028 standard deviation units which translates to a 1.69 percent decline relative to the mean. Column 2 of Table 5 indicates that higher level of arsenic exposure is positively linked to underweight as well (0.018 standard deviation – 1.09 percent).

We further analyze whether the effect of arsenic on stunting and underweight varies by gender. As is evident from columns 1 and 2 of Table 6, there is no difference by gender in the effect of arsenic contamination on HAZ. On the other hand, column 3 and 4 suggests that arsenic has a negative and significant impact on WAZ (1.09 percent- significant at 5 percent level of significance) unlike the OLS results. But there are no gender differences in the effect of arsenic on underweight.

While the results show that arsenic has an adverse effect on stunting and underweight, the simple gender segregated regressions do not indicate that girls are worse off than boys. To examine this further, we study if the effect of arsenic on height-for-age and weight-for-height varies across birth order and gender.

4.2 Arsenic and child health across gender and birth order

Unlike OLS estimates, IV results in column 3 of Table 7 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. A one standard deviation unit change in arsenic leads to decrease in height-for-age (stunting) for second and third (or later) born girls by 0.015 and 0.046 standard deviation units which is equivalent to 0.90 percent and 2.77 percent respectively. Significance of our estimate for third (or later) born girls indicates that stunting in girls increases with steeper birth gradient. These estimates are robust to the inclusion of various district level controls[13]. The IV coefficients are not sensitive to the inclusion of district level controls, as can be seen from comparing columns 2 and 3. This gives further credibility to the exogeneity of the instrument.

We find similar IV results for Weight-for-age as shown in column 4 of Table 7. IV estimates on WAZ indicates that a standard deviation unit increase in arsenic is accompanied with decrease in weight-for-age (underweight) for second and third (or later) born girls by 0.024 and 0.041 standard deviation units which is equivalent to 1.46 percent and 2.5 percent respectively.

When the concentration of arsenic in groundwater increases, later born girls (born at higher birth order) experience more stunting and underweight than their older sibling (lower birth order), particularly if the elder sibling is male. This could be explained by 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). To support this hypothesis, we next study the heterogenous effects of arsenic exposure on health outcomes.

4.3 Heterogeneous impact by Sex Ratio

Within India, we next examine the heterogenous impact of arsenic by sex ratio, where sex ratio is defined as the number of females per thousand males in the population. According to the 2011 Census of India, there were about 7.1 million fewer girls than boys under the age of six in India. Child sex ratios are skewed prenatally due to sex determination and sex selective abortions and postnatally through neglect of the girl child which leads to higher female mortality. Thus, a low sex ratio is one indicator of gender selection or gender bias in favor of the male child.

We divide the sample into two groups based on the median district sex ratio and estimate the main regressions separately for households located in districts below median sex ratio and those located above median sex ratio[14]. IV estimates in Column 3 of Table 8 shows that exposure to higher levels of arsenic leads to significantly lower weight-for-age for second (0.016 standard deviation- 0.97

---

[13] All our findings in this study are robust to inclusion of family size, though we address this in detail in section 4.6.

[14] The two groups are formed on the basis of median sex ratio, where first group represents those children who belong to regions with poor sex ratio (below 922) and the second group represents those children who belong to favorable sex ratio regions (above 922).

percent) and third (or later) born girls (0.143 standard deviation which translates to 8.72 percentage-significant at 1 percent level of significance). The coefficient for third born girls is more than three times larger compared to the regressions estimates in Table 7. On the other hand, the effects are zero and insignificant in regions with above median sex ratio. At the same time, we find negative but insignificant association between arsenic exposure and height-for-age for later born girls in regions with sex below median sex ratio.

4.4 Heterogeneous impact by religious & cultural beliefs

We further explore the role of social and cultural identity manifested via different religious beliefs practiced in India. Table 9 examines difference in stunting across two different religion groups (Hindu and Muslims). Relative to Islam, Hinduism places greater emphasis on having a male heir so as to fulfil social rituals. Son preference is less prevalent amongst Muslims as evident by less skewed sex ratio and lower gender gap in child mortality amongst Muslims relative to Hindus (Borooah and Iyer 2005; Bhaltora, Valente and Soest 2010).

Our findings in Column 1 of Table 9 suggests that for Hindu community, height disadvantage amongst girls materializes at higher birth order (4.04 percent- significant at 5 percent level of significance) compared to Muslims who have muted birth order effect on stunting as shown in column 2. Further, column 3 also indicates that the adverse effect of arsenic on Weight-for-age aggravates with increasing birth order for Hindu girls (3.05 percent- significant at 5 percent level of significance). Thus, the detrimental effect of arsenic contamination on stunting and underweight in India could partially be explained by heterogeneity in cultural and religious beliefs practiced across various regions of India.

4.5 Heterogeneous impact by household location (rural or urban)

Next, we study whether the prevalence of stunting is higher in rural areas relative to urban regions. Our results in Column 1 and Column 3 (Table 10) suggests that nutritional status (as measured by Height-for-age and Weight-for-age) amongst girls is worse in rural regions. The adverse impact of arsenicosis on stunting (3.43 percent) and underweight (2.8 percent) magnifies for elder daughters in rural households relative to urban households. The primary reason for rural-urban difference in impact of arsenicosis on nutritional status of children particularly girls might be economic inequality. Inadequate provision of basic amenities such as adequate food, drinking water and health care might worsen their long-term economic outcomes at later stages of life.

4.6 Arsenic, birth order and family size

A key concern with the analysis thus far is that 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. To address this potential problem, we conduct a robustness check where we control for family size and instrument it with gender of first child. As discussed earlier, we estimate the regressions separately by gender.

The first stage coefficient on the IV for family size is high at 13.28 (statistically significant at the 1% level). The F-statistics is 685 suggesting that gender of the first child is a valid instrument for family size.[15] IV results from this specification is shown in Table 11, where column 1 and column 3 provides the IV estimates for health outcomes (HAZ and WAZ) for girls. Height-for-age and weight-for-age for boys are reported in column 2 and column 4. Our estimates in column 1 shows that one unit increase in arsenic exposure leads to significant decrease in height-for-age for girls born at second and third (or higher) by 0.021 (translates to 1.27 percent) and 0.122 standard deviation (translates to 7.34 percent-significant at 1 percent level of significance) respectively. Similar estimates are shown for arsenic exposure on underweight amongst girls born at second (2.62 percent- significant at 1 percent level of significance) and third or higher birth order (6.83 percent- significant at 1 percent level of significance). We find negative impact of arsenicosis on height-for-age (0.004 standard deviation- second born and 0.057- third or higher birth order) and weight-for-age amongst boys, but the impact of arsenic exposure on health outcomes (height-for-age and weight-for-age) amongst girls is three times greater than for boys even after accounting for family size and other related factors. Looking at the coeffient on family size, more number of children in the household has a negative and significant effect on HAZ but not WAZ.

To sum up, exposure to higher levels of arsenic has adverse impact on health outcomes for children, but the effect is considerably higher for girls as compared to boys, particularly the younger ones. This can be attributed to the lack of adequate nutrition and health care provided to

---

[15] Table not shown but available upon request

girls during their early childhood that would result in poor immunity and growth, which in turn increases their vulnerability to diseased environment. This is consistent with our findings that the effect of arsenic exposure is magnified in areas with worse sex ratio and in rural locations, where gender norms are likely to be distorted. The negative effect is also larger among Hindu households, that follow Hindu kinship systems.

5. 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 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 indicate 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.

Our results show heterogeneous effect of arsenic exposure across cultural norms and socio-economic status, 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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.*

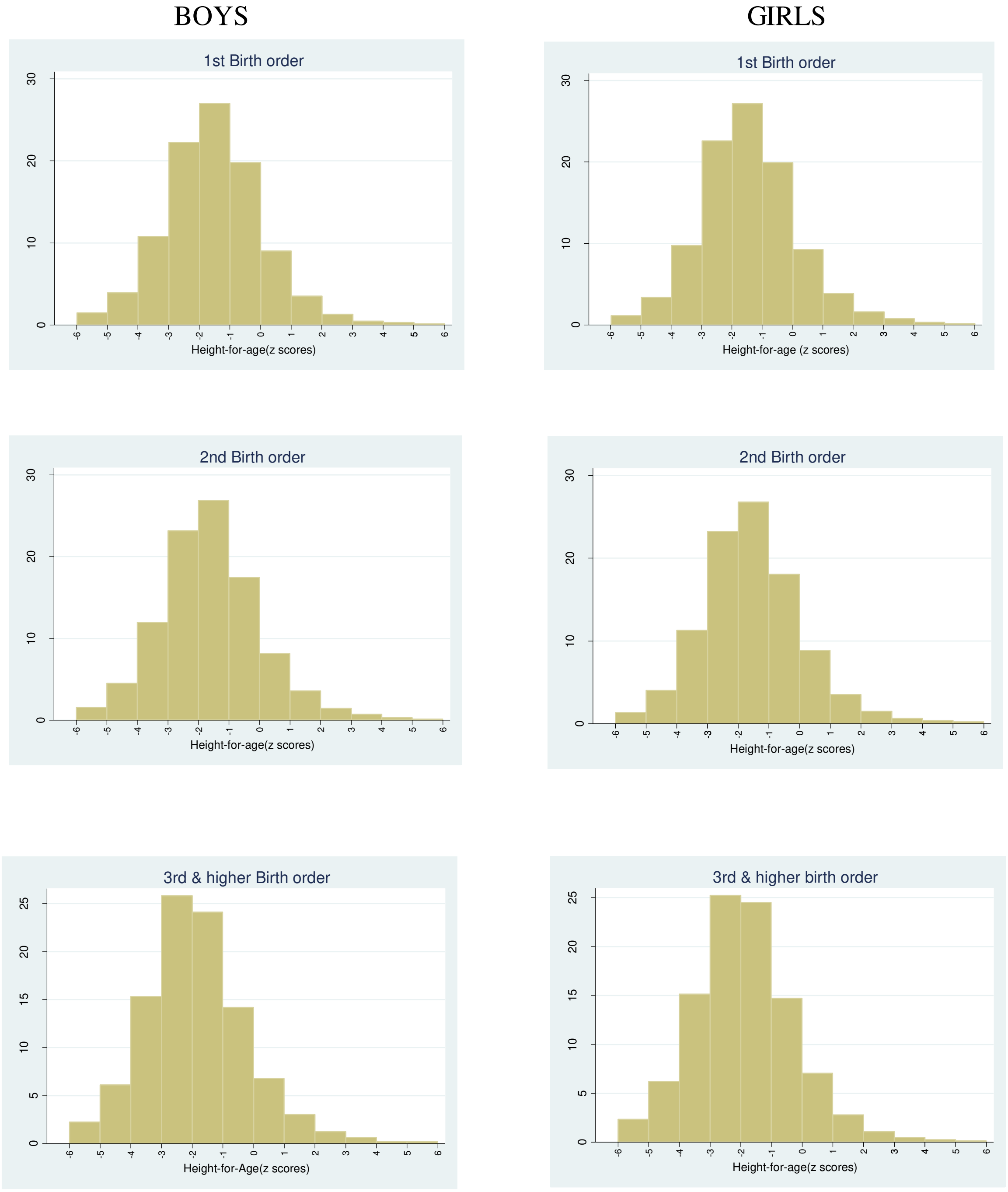
Walker SP, Wachs TD, Gardner JM, et al. (2007). Child development: risk factors for adverse outcomes in developing countries, *Lancet*, vol. 3699556, 145-157.

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.

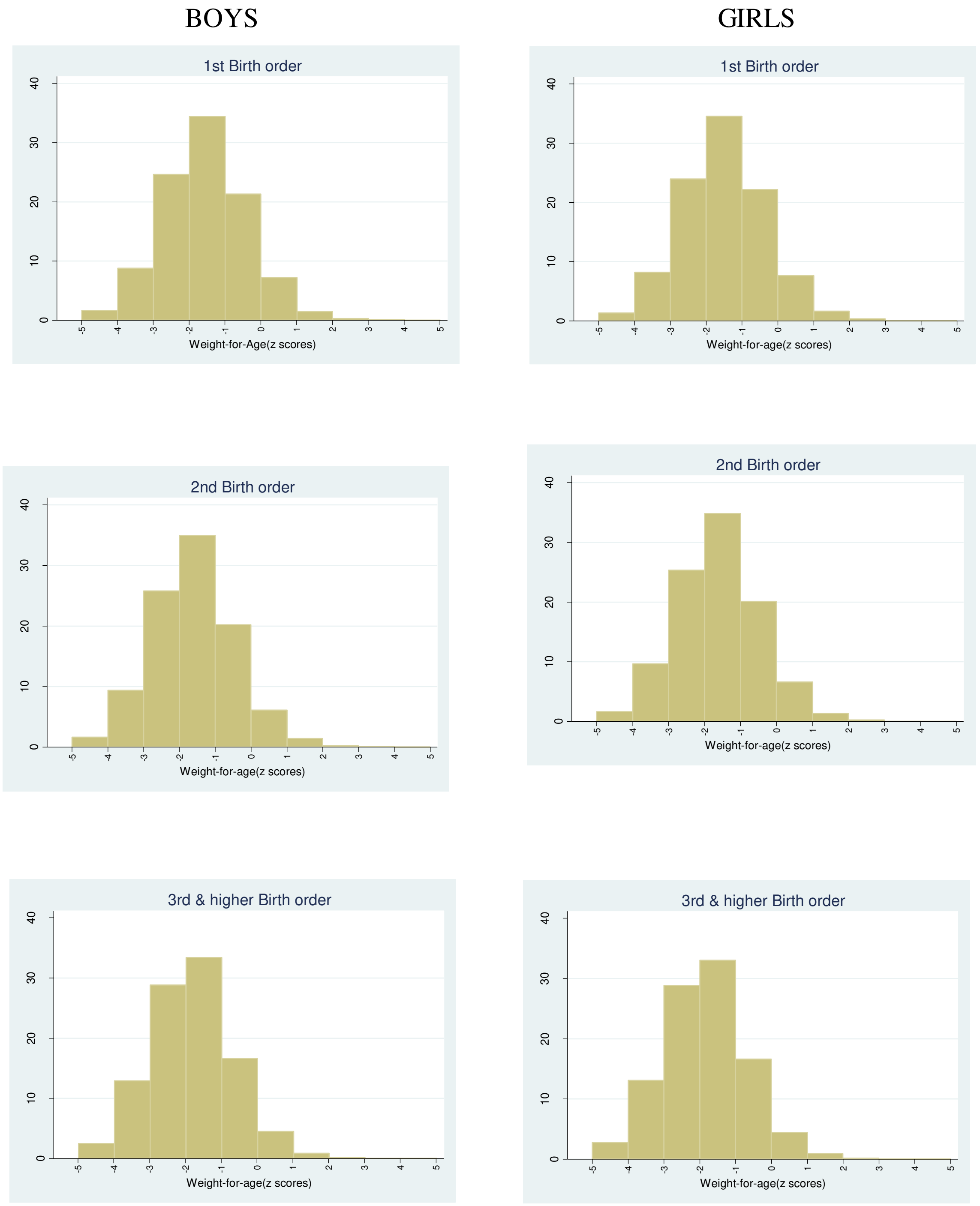
World Health Organization (2000): "Effect of Breastfeeding on Infant and Child Mortality Due to Infectious Diseases in Less Developed Countries: A Pooled Analysis," Lancet, 355(9202), 451–455.

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



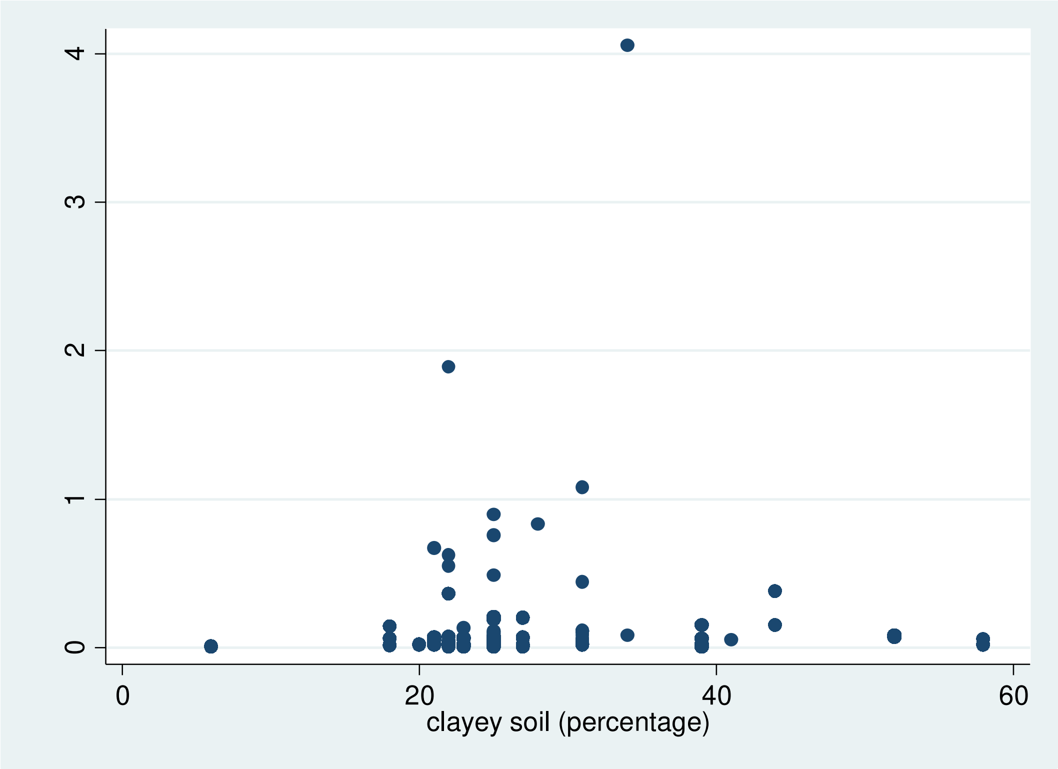
Note: The above figure shows percentage of boys and girls whose z scores for Height-for-age (z scores) decrease with increasing birth order, as represented on the horizontal axis. As per guidelines issued by World Health Organization children whose z scores are below -2 and above -3 indicates moderate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

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



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.

Figure 3: Relation between Arsenic (microgram per liter) and (percentage)of clayey soil type.



Source: Data on arsenic is taken from report of Central Ground Water Board under Ministry of Water Resources, River development and Ganga Rejuvenation, Gov. of India. Data on soil comes from the Harmonized world soil database.

Table 1: Descriptive Statistics and District level control variables

Variable	Mean	Std. Dev.
Height-for-age (z scores)	-1.66	1.63
Weight-for-age (z scores)	-1.64	1.18
Arsenic (ug/l)	107.32	481.13
Clayey soil (percentage)	28.23	7.58
Individual characteristics		
Birth order (first)	0.34	0.47
Birth order (second)	0.29	0.45
Birth order (third)	0.37	0.48
Age	2.30	1.49
% Girls	0.48	0.50
Maternal characteristics		
Mother's education		
Illiterate	0.38	0.49
Primary	0.14	0.35
Secondary	0.39	0.49
Higher & above	0.09	0.28
Mother's age	27.15	4.75
Family background characteristics		
Hindu	0.77	0.42
Muslim	0.18	0.39
Others	0.04	0.20
Scheduled Caste/Scheduled Tribe	0.31	0.46
Other Backward caste	0.50	0.50
Higher/Upper castes	0.19	0.38
Urban	0.21	0.40
District level control variables		
Sex ratio (Female/male)	927.59	44.20
Rainfall (millimeters)	79.24	43.42
Iron (mg/l)	1.60	2.53
Monthly per capita expenditure (rupees)	163816.6	59202.86
Ratio of rice to wheat (production)	741.27	10930.95
% Literacy	68.17	8.11

Sample size is N=73,160

Table 2: Arsenic and Child's anthropometric measures (OLS Estimates)

	Height-for-age (HAZ)	Weight-for-age (WAZ)
Anthropometric measures (z scores)	**Full sample (1)**	**Full sample (2)**
Arsenic	-0.002*	0.003**
	(0.001)	(0.001)
Individual controls	Yes	Yes
Maternal controls	Yes	Yes
Family background controls	Yes	Yes
District level controls	Yes	Yes
State F.E	Yes	Yes
Observations	75,371	75,371

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. 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 and place of residence).

Table 3: Arsenic and Child Anthropometric Measures: By Gender (OLS Estimates)

	HAZ	HAZ	WAZ	WAZ
Anthropometric measures (z scores)	**Girls (1)**	**Boys (2)**	**Girls (3)**	**Boys (4)**
Arsenic	-0.003*	-0.001	0.003**	0.003**
	(0.002)	(0.001)	(0.001)	(0.001)
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	36,198	39,173	36,198	39,173

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. 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 and place of residence) are included in all regressions.

Table 4: First Stage Regression

	Arsenic (microgram/liter)
Clayey soil (sub)	13.275***
	(0.294)
First stage F-statistics	685.48
Anderson Rubin Wald Statistics (p value)	0.000***
Observations	75,371

Note: Robust SE (*** 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, rainfall, literacy, pattern of cultivation, iron and gross domestic product. 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 and place of residence) and included.

Table 5: Arsenic and Child Anthropometric Measures (IV Estimates)

Anthropometric measures **(z scores)**	**Full sample** **(HAZ)**	**Full sample** **(WAZ)**
Arsenic	-0.028**	-0.018**
	(0.008)	(0.006)
Individual controls	Yes	Yes
Maternal controls	Yes	Yes
Family background controls	Yes	Yes
District level controls	Yes	Yes
State F.E	Yes	Yes
Observations	75,371	75,371

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. 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 and place of residence) and included.

Table 6: Arsenic and Child Anthropometric Measures (IV Estimates)

	HAZ	HAZ	WAZ	WAZ
Anthropometric measures (z scores)	**Girls (1)**	**Boys (2)**	**Girls (3)**	**Boys (4)**
Arsenic	-0.029**	-0.029**	-0.018**	-0.018**
	(0.011)	(0.012)	(0.008)	(0.008)
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	36,198	39,173	36,198	39,173

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. 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 and place of residence).

Table 7: 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	0.003	-0.035	-0.015	0.005	-0.043**	-0.024
	(0.004)	(0.026)	(0.023)	(0.004)	(0.018)	(0.017)
Arsenic*girl*BO3	-0.003	-0.060**	-0.046*	-0.004	-0.062**	-0.041**
	(0.005)	(0.027)	(0.025)	(0.004)	(0.019)	(0.018)
Arsenic*BO2	-0.006**	-0.009	-0.022	-0.003	-0.002	-0.013
	(0.003)	(0.018)	(0.016)	(0.003)	(0.013)	(0.011)
Arsenic*BO3	-0.009*	-0.094***	-0.090***	-0.006**	-0.072***	-0.065***
	(0.003)	(0.020)	(0.018)	(0.003)	(0.014)	(0.013)
Arsenic*girls	0.000	0.104***	0.068***	-0.001	0.049***	0.029***
	(0.002)	(0.016)	(0.015)	(0.003)	(0.012)	(0.010)
Arsenic	0.002	-0.049**	-0.023	0.006**	-0.023**	0.004
	(0.002)	(0.015)	(0.012)	(0.002)	(0.011)	(0.009)
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	73,160	102731	73,160	73,160	102731	73,160

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). BO stands Birth order. All regression includes state fixed effects. Column 1, column 3, column 4 and column 6 also includes district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. All regressions include individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste and place of residence).

Table 8: Sex ratio, Child's Height-for-age and Weight-for-age (IV Estimates)

	HAZ	HAZ	WAZ	WAZ
	(1)	(2)	(3)	(4)
Anthropometric measures (z scores)	**Sex ratio (below 922[16])**	**Sex ratio (above 922)**	**Sex ratio (below 922)**	**Sex ratio (above 922)**
Arsenic*girls*birth order2	0.061	-0.035	-0.016	-0.031
	(0.081)	(0.028)	(0.061)	(0.020)
Arsenic*girls*birth order3	-0.075	-0.031	-0.143**	0.006
	(0.076)	(0.033)	(0.059)	(0.024)
Arsenic*birth order2	-0.073	-0.015	-0.060	-0.003
	(0.055)	(0.019)	(0.042)	(0.014)
Arsenic*birth order3	-0.198**	-0.083**	-0.100**	-0.085***
	(0.057)	(0.026)	(0.044)	(0.019)
Arsenic*girls	0.115**	0.068***	0.040	0.032**
	(0.056)	(0.018)	(0.043)	(0.013)
Arsenic	0.032	-0.037**	0.158***	-0.034**
	(0.052)	(0.019)	(0.042)	(0.013)
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	38,579	33,012	38,579	33,012

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Column 1 and column 3 includes sample of those children who belong to districts with poor sex ratio (below 922). Column 2 and column 4 includes children who belong to districts with favorable sex ratio (greater than 922). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, family size and place of residence).

---

[16] Median sex ratio in our sample is 922.

Table 9: Heterogeneity across religious groups (Within India Evidence: IV Estimates)

	HAZ	HAZ	WAZ	WAZ
Anthropometric measures	**(1)**	**(2)**	**(3)**	**(4)**
(z scores)	**Hindus**	**Muslims**	**Hindus**	**Muslims**
Arsenic*girls*birth order2	-0.015	0.001	-0.024	-0.033
	(0.028)	(0.044)	(0.020)	(0.032)
Arsenic*girls*birth order3	-0.067**	0.040	-0.050**	-0.010
	(0.029)	(0.050)	(0.021)	(0.036)
Arsenic*birth order2	-0.013	-0.031	-0.011	-0.005
	(0.019)	(0.033)	(0.014)	(0.024)
Arsenic*birth order3	-0.077***	-0.099**	-0.055***	-0.064
	(0.021)	(0.037)	(0.015)	(0.027)
Arsenic*girls	0.070***	0.031	0.024*	0.042
	(0.018)	(0.026)	(0.013)	(0.020)
Arsenic	-0.019	-0.015	-0.001	0.000
	(0.014)	(0.027)	(0.010)	(0.019)
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	56,453	13,520	56,453	13,520

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Column 1 and Column 3 includes sample of those children who belong to Muslim communities. Column 2 and column 4 includes children who belong to Hindu society. All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother’s age and mother’s education) and family background controls (caste, family size and place of residence).

Table 10: Heterogeneity across rural or urban regions (Within India Evidence: IV Estimates)

	HAZ	HAZ	WAZ	WAZ
Anthropometric measures	**(1)**	**(2)**	**(3)**	**(4)**
(z scores)	**Rural**	**Urban**	**Rural**	**Urban**
Arsenic*girls*birth order2	-0.010	-0.037	-0.031	-0.011
	(0.027)	(0.046)	(0.019)	(0.035)
Arsenic*girls*birth order3	-0.057**	0.000	-0.046**	-0.027
	(0.028)	(0.060)	(0.020)	(0.043)
Arsenic*birth order2	-0.023	-0.011	-0.005	-0.030
	(0.019)	(0.030)	(0.014)	(0.022)
Arsenic*birth order3	-0.068**	-0.172***	-0.054***	-0.095
	(0.021)	(0.045)	(0.015)	(0.030)
Arsenic*girls	0.067***	0.077**	0.029	0.038
	(0.017)	(0.030)	(0.013)	(0.023)
Arsenic	-0.024	-0.025	-0.007	0.004
	(0.015)	(0.022)	(0.011)	(0.016)
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	58,082	15,078	58,082	15,078

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). *SES indicates Socio Economic Status. Column 1 includes sample of those children who belong to scheduled caste/schedule tribe or other backward communities. Column 2 includes children who belong to forward/upper class of society. All regressions include state fixed effects and district level controls for sexratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste and family size).

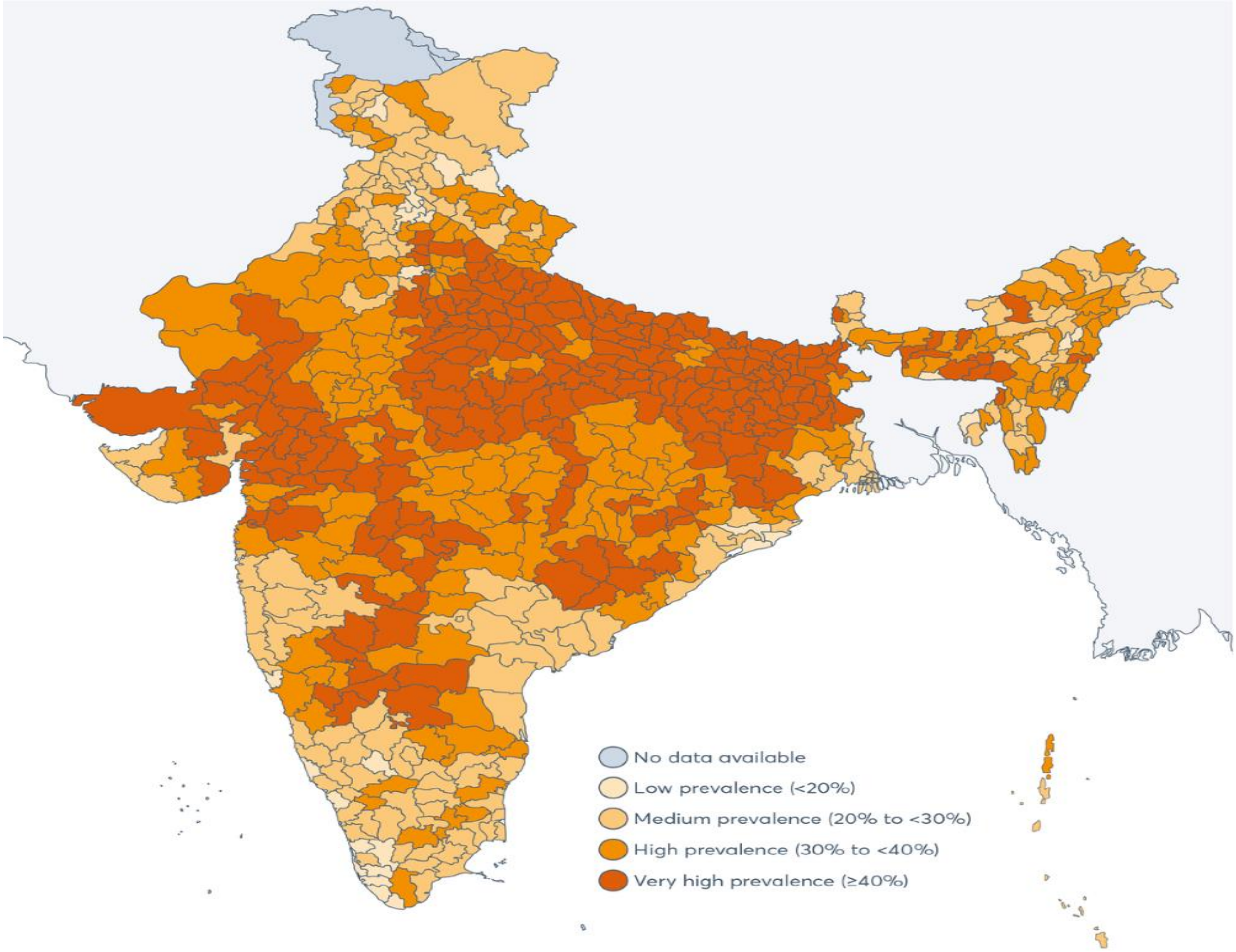
Table 11: Arsenic, family size and birth order (IV Estimates)

	HAZ	**HAZ**	**WAZ**	**WAZ**
Anthropometric measures	Girls	Boys	Girls	Boys
	(1)	(2)	(3)	(4)
Arsenic*birth order2	-0.021	0.004	-0.043***	-0.009
	(0.019)	(0.019)	(0.014)	(0.014)
Arsenic*birth order3^{rd+}	-0.122***	-0.057**	-0.112***	-0.062***
	(0.023)	(0.022)	(0.017)	(0.017)
Arsenic	0.008	-0.013	0.022*	-0.000
	(0.017)	(0.017)	(0.013)	(0.012)
Family size	-6.808*	-9.009**	2.583	-1.126
	(3.479)	(3.969)	(2.653)	(2.968)
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	34,916	37,772	34,916	37,772

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Column 1 and column 3 includes the sample of girls. Column 2 and column 4 incudes the sample of boys. All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (family size, religion, caste and place of residence) are included in all regressions.

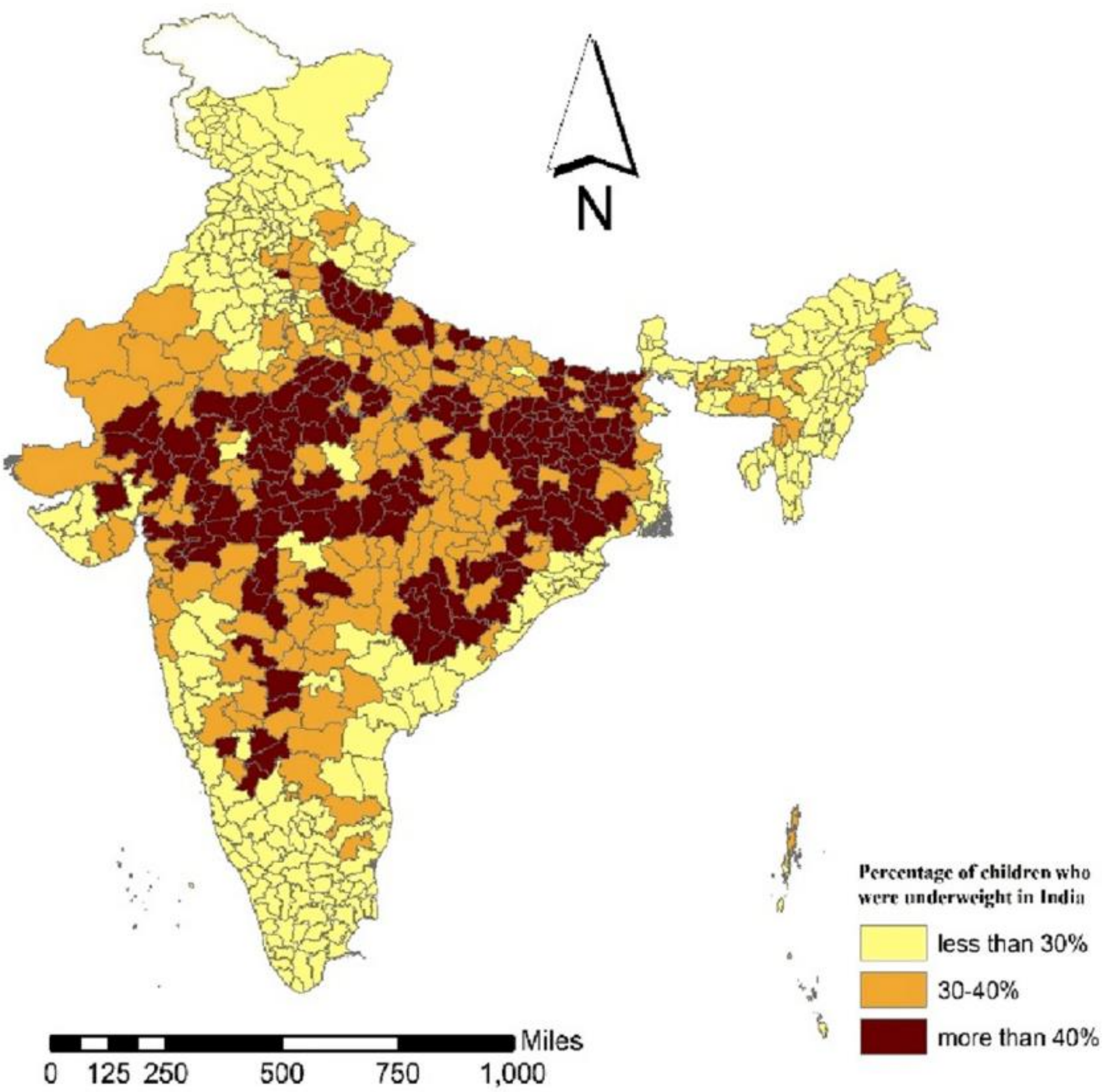
Appendix A.1: Figures and Maps

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



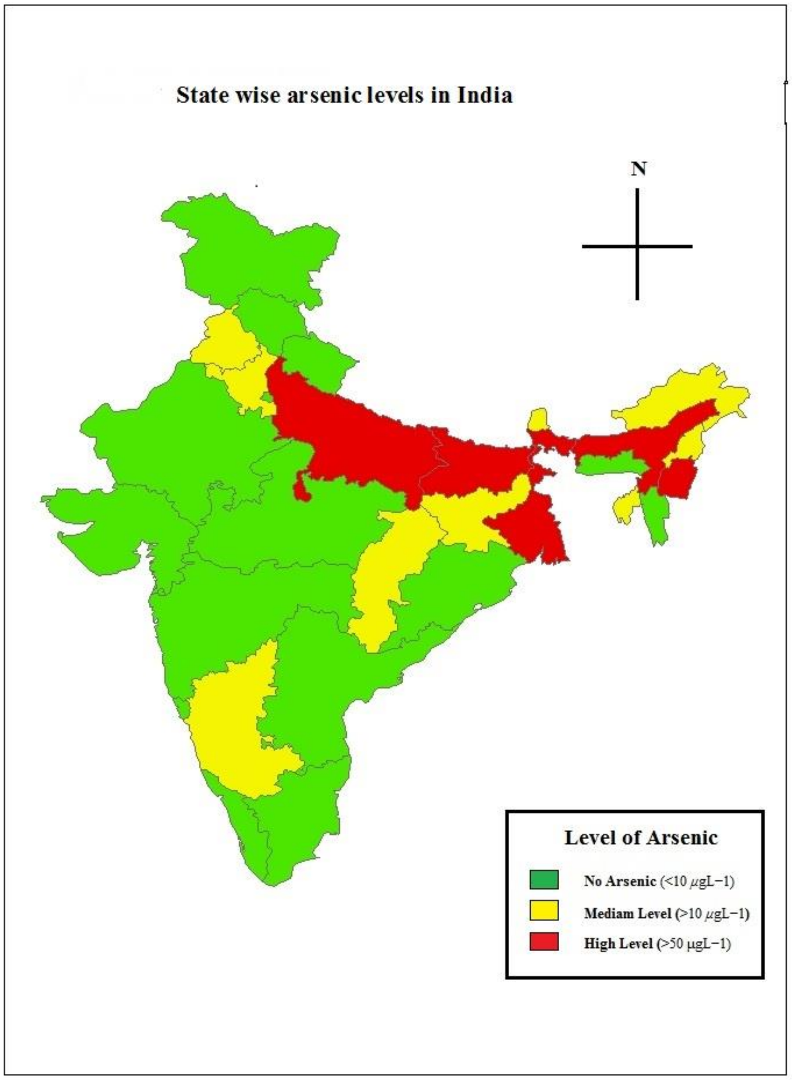
Source: Menon et al. (2018)

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



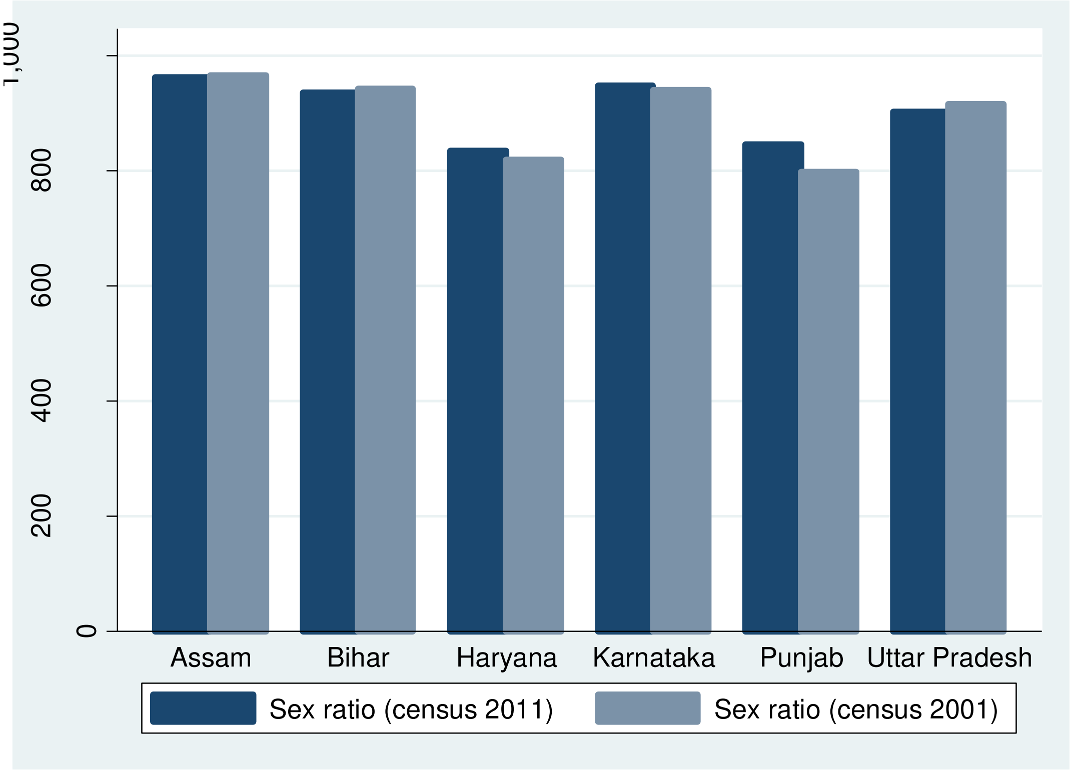
Source: Sharma et al. (2020)

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: Bar Graph for child sex ratio: Central Northern, North Eastern and Southern states of India



Source: Authors calculation using Census data, GOI (2001 & 2011)

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	N
Other Contaminants:		
Iron (mg/liter)	0.472**	261
	(0.207)	
Fluoride (mg/liter)	0.345	257
	(0.462)	
Nitrate (mg/liter)	-0.007	257
	(0.005)	
Weather:		
Rainfall (millimeters)	-0.031**	228
	(0.012)	
Maximum temperature (degree Celsius)	0.105	58
	(0.237)	
Minimum temperature (degree Celsius)	0.152	39
	(0.330)	
Education:		
Literacy	0.091	259
	(0.060)	
Demographic& Economic Factors:		
Ratio of Rice to Wheat (million tonnes)	-0.000	207
	(0.000)	
Nitrogen (Kilogram/hectare)	0.000**	236
	(0.000)	
Phosphorus (Kilogram/hectare)	0.000**	236
	(0.000)	
Potassium (Kilogram/hectare)	-0.000	236
	(0.000)	
Sex Ratio (per 1000 females)	0.002	259
	(0.015)	
Per Capital Expend.	0.000**	261
	(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.