Environment and Development Economics (2024), 1–21
doi:10.1017/S1355770X24000329

RESEARCH ARTICLE

Gender disparities in the prevalence of undernutrition in India: the unexplored effects of drinking contaminated water

Khushboo Aggarwal,[1] and Rashmi Barua[2]*

[1]Delhi School of Public Policy and Governance, Institution of Eminence, University of Delhi, New Delhi, India and [2]Centre for International Trade and Development (CITD), Jawaharlal Nehru University, New Delhi, India
*Corresponding author: Rashmi Barua; Email: barua.bhowmik@gmail.com

(Submitted 14 August 2022; revised 21 August 2023, 31 May 2024; accepted 19 August 2024)

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 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 underdeveloped bodies.

Keywords: arsenic; water; child health; birth order; gender

JEL classification: I3; I15; Q53; Q54

1. Introduction

Numerous studies have investigated the relation between gender and child growth indicators as determined by their respective share in households' available resources. However, in addition to adequate nutrition, safe drinking water acts as an indispensable input to child health. More than 2,000 children under the age of five die every day from gastrointestinal diseases, 90 per cent of which is 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 underdeveloped bodies. To the best of our knowledge, no study has addressed the role that gender plays in the relation between child stunting and access to safe drinking water.

The first aim of this study is to investigate the impact of exposure to arsenic contaminated groundwater on child growth outcomes in India. Overconsumption of arsenic can lead to fatal health outcomes such as bone diseases, kidney and heart failure, 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.[2] Second, we argue that, in the presence of gender bias, girls born in higher birth orders 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 pollutants on health outcomes. While arsenic is known to readily cross the placenta, exclusive breastfeeding protects infants against arsenic (Fängström *et al.*, 2008; Samiee *et al.*, 2019).[3] Thus, if girls, particularly those born in higher birth orders, 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. Consistently, Gardner *et al.* (2013) find an inverse association between arsenic exposure and growth outcomes of children in Bangladesh with significantly larger effects among girls.[4] They find that nutritional deficiencies act as a primary factor for adverse effects among girls from low socioeconomic status (SES) households.

Using geographical variation in arsenic concentration in water, we estimate the association between arsenic levels and child health outcomes (height-for-age (HAZ) and weight-for-age (WAZ) z-scores) in India using data from 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 a region.[5] To overcome this identification challenge, we use an instrumental variable (IV) framework. We use the variation in fraction of clayey soil textures across districts to instrument for arsenic levels in groundwater. 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 (Madajewicz *et al.*, 2007; Huang *et al.*, 2020).

While ordinary least squares (OLS) estimates are imprecise, IV estimates indicate that exposure to arsenic in groundwater has a negative and significant impact on HAZ and

---

[1] Arsenic is also known to impact educational outcomes and cognitive skills (Aggarwal *et al.*, 2024).

[2] While there is ample epidemiological evidence that arsenic affects child growth outcomes (Watanabe *et al.*, 2007; Rahman *et al.*, 2017), 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; Bae *et al.*, 2021). There is also strong evidence that arsenic crosses the placenta and adversely impacts health in utero and later in life (Kile *et al.*, 2016).

[3] Consistent with this, following an arsenic awareness campaign in Bangladesh, Keskin *et al.* (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.

[4] Exposure was based on urinary concentrations of arsenic in urine samples collected from pregnant women and their children.

[5] 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).

WAZ among children less than five years of age, regardless of gender. This is an important finding as child stunting and wasting, 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, underperform in school and have reduced future earnings (Glewwe and Miguel, 2007; Case and Paxson, 2008).

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 unit increase in arsenic levels in groundwater leads to a modest reduction in HAZ and WAZ by 0.65 and 0.63 standard deviations for a later born girl child, respectively, relative to a male child born at first birth order.[6]

Existing studies find that children born at higher birth orders have a higher probability of being from a large size family (Behrman and Taubman, 1986; Spears *et al*., 2022). 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. While we acknowledge that it is not possible to check the exclusion restriction directly, we conduct several falsification and robustness checks to confirm that we are measuring the causal effect of drinking contaminated water on child health. The findings are also robust to the inclusion of district level controls for health infrastructure, weather, a host of other water contaminants, sex ratio, literacy, and income.

Our study contributes to the under-studied link between gender, environmental pollutants, and child growth measures. To the best of our knowledge, this is the first paper to explore the role of gender in the relation between environmental pollutants and child health outcomes. We also make significant contributions to the literature on social gradients in health and environmental inequality. Our findings are policy relevant as they suggest that economically disadvantaged groups are at greater risk of environmental hazards owing to their weaker immunity, lower nutrition and poor resource access.

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, followed by robustness checks and falsification tests in section 6. Lastly, in section 7 we provide 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 (Lundberg, 2005; Chung and Das Gupta, 2007; Fledderjohann and Channon, 2022). Studies find that the height disadvantage among girls increases with steeper birth order gradient, which can be explained by biased preferences of parents towards their eldest sons (Garg and Morduch, 1998; Jayachandra and

---

[6] 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, the birth of late parity daughters acts as a negative income shock and, as such, limited resources will be spent on them.

Kuziemko, 2011; Jayachandran and Pande, 2017).[7] Others attribute the differential pattern of investment in child rearing and health inputs due to differences in resources available with parents and their preferences (Becker and Tomes, 1976; Behrman *et al.*, 1986; Vogl, 2016). Some studies show that women's nutritional status may be worse off due to lack of access to formal healthcare and differential childcare practices (DeRose *et al.*, 2000).[8]

We also contribute 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 (Gómez-Roig *et al.*, 2021). In economics, most studies have focused on the negative health outcomes of air pollution (Foster *et al.*, 2009; Arceo *et al.*, 2016; Goyal and Canning, 2018). 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, 2010).

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. Brainerd and Menon (2014) find that exposure to fertilizers via contaminated groundwater during pregnancy has a negative impact on child health outcomes.[9]

Finally, our study adds to the well-established literature on social gradients in health. Socially and economically disadvantaged groups may experience increased susceptibility to all forms of environmental hazards owing to weaker immunity, lower nutrition, and poor accessibility of resources (Lynch *et al.*, 2006). More vulnerable communities are disproportionately more likely to suffer from environmental hazards (Fecht *et al.*, 2015; Deguen *et al.*, 2022).[10]

---

[7] Jayachandran and Pande (2017) find that parents allocate more prenatal inputs during a pregnancy when they do not have any sons. 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 their newborn 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.

[8] 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 (Bahrami-Rad, 2021). Bardhan (1974) finds that the neglect of girl children in northern regions of India could be attributed to the lower participation of females in agricultural activities that leads to their lower economic value.

[9] Some studies evaluate environmental water policies, for instance, Greenstone and Hanna (2014) 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.

[10] Consistent with this, Neidell (2004) estimates the impact of air pollution on asthma related child hospitalizations in the United States and finds large negative effects for low SES children who cannot afford to live in cleaner areas. Brainerd and Menon (2014) find that the child health impact of fertilizer agrichemicals in water is largest among children of uneducated poor women. Tanaka (2015) finds that air pollution regulations led to reductions in infant mortality among children born to mothers with low educational levels in China.

3. Data and data source

The data for our analysis comes from the Demographic and Health Survey (National Family Health Survey, NFHS-4, 2015–16), administered by the Ministry of Health and Family Welfare. NFHS is a nationally representative dataset that comprises 111,667 children between 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 in the age group of 0–5 years using z-scores calculated in accordance with World Health Organization (WHO) guidelines. To assess the impact of water pollution on child health, we use two measures of child health. First, we study 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 Drèze, 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 later life outcomes. 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 underweight measured by WAZ z-scores. Underweight is a symptom of acute malnutrition and is a dire consequence of inadequate intake of food or high incidence of infectious diseases. Stunting and underweight are aspects of malnutrition that are closely linked to each other. The presence of both stunting and underweight in a child intensifies the risk of mortality (Thurstans *et al*., 2022).

Figures A1 and A2 in the online appendix plot the HAZ and WAZ scores, respectively, by birth order among boys and girls. The percentage of girls who are moderately or severely stunted increases with birth order.[11] For instance, at first birth order approximately 10 per cent of girls suffer from severe stunting which increases to 12 per cent and 15 per cent for 2^{nd} and $3^{\text{rd}+}$ birth order, respectively. A similar pattern is visible for boys. The summary statistics of the variables that are included in our analysis are shown in table 1.[12]

Data for rainfall is provided by the Indian Meteorological Department (IMD) at the 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 per cent, respectively. To control for district prosperity, we use data on monthly per capita expenditure (MPCE) from the 68th round of the National Sample Survey Office. Production of rice and wheat (in million tons) for 2011 is obtained from the Ministry of Agriculture and Farmers Welfare. 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 has set a standard of $50\,\mu\text{g}\text{L}^{-1}$ (microgram per liter) for arsenic in drinking water. The level of arsenic in

---

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

[12] The data is gender balanced with girls comprising 48 per cent of the sample with an average age of 27 months. While 34 per cent of the sample consists of children at first birth order, 29 per cent are at second birth order and 37 per cent of children are at higher than second birth order. Uneducated mothers comprise 37 per cent of the sample and only 9 per cent have post-secondary education. The average age of mothers in the sample is 27 years and approximately 15 per cent of mothers suffer from severe to moderate anemia. More than three-fourths of our sample comprises rural households.

Table 1. For caption see next page

Variable	Mean	Std. Dev.
Height-for-age (*z* scores)	−1.64	1.65
Weight-for-age (*z* scores)	−1.66	1.18
Arsenic (μg/l)	94	444
Clayey soil (percentage)	27.86	7.82
Groundwater usage for drinking	0.78	0.41
Child characteristics		
Birth order (first)	0.34	0.47
Birth order (second)	0.29	0.45
Birth order (third)	0.37	0.48
Age	2.23	1.49
% Girls	0.48	0.50
Maternal characteristics		
Mother's education		
Illiterate	0.37	
Primary	0.14	
Secondary	0.40	
Higher and above	0.09	
Mother's age Maternal Anemia	27	4.93
Severe	0.008	
Moderate	0.148	
Mild	0.424	
Not anemic	0.418	
Household characteristics		
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 variables		
Sex ratio (Female/male)	925	44
Rainfall (millimeters)	76.7	42
Iron (mg/l)	1.58	2.43
Health centers	315	153

Continued.

Table 1. Descriptive statistics and district level control variables

Variable	Mean	Std. Dev.
Rice (million tons)	51,433	70,530
Wheat (million tons)	71,285	73,794
Monthly per capita expenditure (rupees)	168,990	66,660
% Literacy	68.4	8.24

Note: Sample size is $N = 85{,}520$.

groundwater, a continuous variable, is aggregated at the district level from block level data.[13]

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 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.[14] 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 Harmonized World Soil Database (HWSD) established in 2008 by the Food and Agricultural Organization and International Institute for Applied System Analysis. HWSD is a 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 per cent.

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:

$$Y_{ids} = \alpha_1 Ars_{ds} + \alpha_2 Ars_{ds} * Girl_{\mathrm{ids}} + X_{ids} + D_{ds} + S + e_{ids}. \qquad (1)$$

We are interested in measuring the effect of arsenic on two outcome variables: HAZ and 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. α_2 captures the interaction effects of arsenic with gender (girl = 1). X_{ids} represents a vector of controls for individual level characteristics (gender, age and age square), mother characteristics (age, education, maternal anemia, standardized height for age and weight for height), family background and socio-economic characteristics (religion, caste, family size, wealth index[15] and place of residence). Children born at higher birth orders have a higher probability of being from a large family. Moreover, family size and resources allocated to each child are highly correlated, which might in turn affect the health outcomes of children (Booth and Kee, 2009; Kugler and Kumar, 2017). Thus, we control family size in all regressions. We also control for district level controls (D_{ds}) for rainfall, presence of other contaminants (iron), per capita

---

[13] Figures A5 and A6 in the online appendix provide maps of arsenic affected regions of India.

[14] These nine states are Punjab, Uttar Pradesh, Chhattisgarh, and Haryana in the North; Assam, West Bengal, Jharkhand and Bihar in the East and North-East; and Karnataka in the South.

[15] The NFHS reports a wealth index that ranges from poorest (coded 1) to richest (coded 5).

consumption expenditure, sex ratio, number of public health facilities, rice and wheat production and literacy. Finally, we include state fixed effects in our regression analysis. Heteroskedasticity robust standard errors are clustered at the primary sampling unit (PSU) level.[16]

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. Hence, to overcome the potential endogeneity of arsenic levels, we use an IV 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 which facilitates arsenic concentration in groundwater (McArthur *et al.*, 2001; Madajewicz *et al.*, 2007).[17] 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. Since arsenic concentration is higher in clayey relative to coarse soil, we exploit the variation in percentage of clayey soil across districts within a state to instrument for groundwater arsenic contamination. The first stage equation is given by:

$$Ars_{ds} = \beta_1 Soil_{ds} + \beta_2 Soil_{ds} * Girl_{\text{ids}} + X_{ids} + D_{ds} + S + \varepsilon_{ids}. \tag{2}$$

In this just-identified specification, the two endogenous variables Ars_{ds} and Ars_{ds} * *Girl* are instrumented by $Soil_{ds}$ and $Soil_{ds}$ * $Girl_{\text{ids}}$, respectively, where $Soil_{ds}$ is the percentage of clayey soil in district *d* and state *s*. The remaining specification is the same as in equation (1). The main identifying assumption is that soil texture fractions affect health outcomes only through the impact on the level of arsenic in groundwater.[18] To check if health effects of arsenic exposure vary by gender and birth order, we also estimate the following OLS (equation (3)) and first stage equations (equation (4)):

$$\begin{aligned} Y_{ids} = {} & a_1 Ars_{ds} + a_2 girl_{\text{ids}} + a_3 2ndchild_{\text{ids}} + a_4 3rd^{+} child_{\text{ids}} \\ & + a_5 (Ars_{ds} * girl_{ids} * 2ndchild_{ids}) + a_6 (Ars_{ds} * girl_{ids} * 3rd^{+} child_{ids}) \end{aligned}$$

---

[16] PSUs are unique and the smallest working unit in NFHS-4 survey. A PSU 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.

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

[18] 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 to specific crops such as fruit trees, potatoes, vegetables and berries. Use of such pesticides might alter some properties of superficial soil (uppermost layer of soil), but not the subterranean soil used in our analysis.

$$+\,a_7(Ars_{ds} * 2ndchild_{ids}) + a_8(Ars_{ds} * 3rd^{+}child_{ids}) + a_9(Ars_{ds} * girl_{ids})$$

$$+\,a_{10}(2ndchild_{\mathrm{ids}} * girl_{ids}) + a_{11}(3rd^{+}child_{ids} * girl_{ids}) + a_{12}X_{ids} + D_{ds} + S + e_{ids}, \quad (3)$$

where $2ndchild$ 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 order. Children born at the first birth order are taken as the base category in our analysis:

$$\begin{aligned} Ars_{ds} = {} & \pi_1 Soil_{ds} + \pi_2 girl_{ids} + \pi_3 2ndchild_{ids} + \pi_4 3rd^{+}child_{ids} \\ & + \pi_5 Soil_{ds} * girl_{ids} * 2ndchild_{ids}) + \pi_6(Soil_{ds} * girl_{ids} * 3rd^{+}child_{ids}) \\ & + \pi_7(soil_{ds} * 2ndchild_{ids}) + \pi_8(soil_{ds} * 3rd^{+}child_{ids}) + \pi_9(Soil_{ds} * girl_{ids}) \\ & + \pi_{10}(2ndchild * girl_{ids}) + \pi_{11}(3rd^{+}child_{ids} * girl_{ids}) + \pi_{12}X_{ids} + D_{ds} + S + \varepsilon_{ids}. \quad (4) \end{aligned}$$

here, the main coefficient of interest to be estimated is $a5$ and $a6$ which are associated with the three-way interaction (Ars_{ds} * $girl_{ids}$ * $2ndchild_{ids}$) and (Ars_{ds} * $girl_{ids}$ * $3rd^{+}$ $child_{ids}$) respectively. X_{ids} account 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.

4.2 Instrument validity

The validity of the instrument hinges on clayey soil not varying with other weather, geographic or demographic factors which may in turn affect economic outcomes. While there is geographical variation in weather and soil chemical composition, they do not vary by proportion of clayey soil across districts within the same state.

In table A1 in the online appendix, we show the correlation between proportion of clayey soil and several district level indicators of weather (rainfall and temperature), contaminants found in fertilizers and groundwater (arsenic, iron, nitrate, nitrogen, phosphorous, potassium, lead and fluoride), economic and demographic factors (MPCE, rice and wheat production, literacy, sex ratio, male and female employment in agriculture). The table reports the coefficient on clayey soil, from the regression of reported district level variables on the percentage of clayey soils in a district conditional on state fixed effects. Within a state, variation in percentage of clayey soil is uncorrelated with all other contaminants except arsenic and iron. Further, most coefficients have zero magnitude. We also find no evidence that variation in clayey soil across districts within a state is correlated with weather patterns, district income, soil productivity or its suitability for a certain crop.[19]

There is a significant positive correlation between soil permeability and iron levels. However, this would be against finding a negative impact of arsenic on health outcomes

---

[19] A plausible threat to the identification assumption is that income might be affected by the 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. However, we control rice and wheat production in all regressions noting that the results do not change when these variables are excluded.

and, if anything, underestimating our findings as groundwater with a high iron concentration is associated with a decreased risk of childhood anemia. There is also a positive correlation between rainfall and clayey soil. 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.[20] Further, we control for iron levels and rainfall in all regressions noting that the results do not change when we exclude these two variables.[21]

5. Results

5.1 Arsenic, child health and gender

We first show results for OLS estimates using equation (1). Columns 1 and 2 (table 2) show OLS estimates of the effect of arsenic on HAZ and WAZ, respectively. For HAZ scores, OLS estimates are insignificant. OLS estimates for WAZ show that one unit increase in arsenic is associated with a 0.03 standard deviation (SD) increase in WAZ. In the remaining columns we add the interaction effect of gender (girl = 1) and arsenic. Coefficients of the interaction term, though negative, are statistically significant for HAZ scores but not WAZ scores.

To overcome the issue of endogeneity, we use an IV 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 show a positive and statistically significant relationship between arsenic and soil texture (clayey soil).[22]

The IV results for HAZ, shown in table 3, indicate that the OLS is severely downward biased. A one unit increase in arsenic leads to a decrease in HAZ and WAZ, both decreasing by 0.83 SD units. We further analyze whether the effect of arsenic on child growth outcomes varies by gender. As is evident from the remaining columns of table 3, there is no difference by gender in the effect of arsenic contamination on HAZ scores though all the interaction terms are negative. While the IV results show that arsenic has an adverse effect on stunting and underweight as measured by lower HAZ and WAZ scores, the interaction effects 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. In the last two columns, we show that these estimates are not sensitive

---

[20] 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.

[21] A wide body of literature has studied the adverse impact of consumption of lead via groundwater on a range of health and behavioral outcomes in the United States. See for example Billings and Schnepel (2018) and Trejo *et al.* (2021). A report from the Central Ground Water Board (2022) shows high levels of lead in India in the following states: Telangana, Jammu & Kashmir, Jharkhand, Delhi, Haryana, Kerala, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal. Among these states, four states also have a prevalence of arsenic, namely, Punjab, Haryana, West Bengal and Uttar Pradesh. Even after dropping these four states from the sample, the main IV results stay robust. Results are available upon request.

[22] The first stage result is available in table A2 in the online appendix. The Montiel-Pflueger robust weak instrument test, which is valid under heteroskedasticity, yields a high F-statistic (74) suggesting that clayey soil texture is a strong instrument for arsenic levels. Table A2 shows the first stage results for the simplest specification with one endogenous variable (arsenic) and one instrument (soil). For the models with interaction terms (and thus, multiple endogenous variables), the associated first-stage test statistics, namely, the Cragg-Donald Wald F statistic and the Kleibergen- Paap Wald F statistic are reported in table 3.

Table 2. Arsenic and child anthropometric measures: by gender (OLS estimates)

Anthropometric measures (*z* scores)	HAZ (1)	WAZ (2)	HAZ (3)	WAZ (4)
Arsenic	−0.029 (0.019)	0.029 (0.014)	−0.002 (0.023)	0.037 (0.019)
Arsenic*Girl			−0.054 (0.029)	−0.016 (0.024)
Girl	0.083 (0.011)	0.025 (0.008)	0.087 (0.011)	0.027 (0.008)
Individual controls	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes
Family background	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes
State F.E.	Yes	Yes	Yes	Yes
Observations	88,106	88,106	88,106	88,106

Notes: Standard errors clustered at the PSU level in parentheses. 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, rice and wheat production and MPCE, individual level controls (age, age square and gender), maternal controls (mother's age, mother's education, maternal anemia (severe, moderate, mild, non-anemic), mother's standardized height for age (HAZ) and weight for height (WHZ)) and family background controls (religion, caste, family size, wealth index and place of residence).

to including any other soil and water contaminant. While the magnitude of the effect of arsenic exposure drops marginally, the results stay negative and significant for both health outcomes.

5.2 Heterogeneous effect of household wealth

As discussed earlier, 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 the poorest and low SES households should exhibit larger detrimental effects of arsenic exposure since they are more likely to suffer from nutritional deficiencies.

In table 4, we study the relation between arsenic and health outcomes separately for the poorest and relatively richer households.[23] The results clearly show that among children belonging to lowest income households, a one SD increase in arsenic is associated with a 1.19 SD decrease in HAZ and a 1.15 SD decrease in WAZ scores. The coefficients are lower for the high wealth group and imprecisely estimated (insignificant for HAZ and significant at 10 per cent for WAZ scores). Thus, the negative effect of arsenic exposure on height for age is largest among children from poorest households, consistent with the literature (Lynch *et al.*, 2006).

These results capture the effect of arsenic on health among children with poor nutritional intake, however, they do not necessarily imply that the effect is driven by gender

---

[23] 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 category (poorest). The richer and richest categories are defined as *high wealth*.

Table 3. Arsenic and child anthropometric measures (IV estimates)

Anthropometric measures (*z* scores)	HAZ (1)	WAZ (2)	HAZ (3)	WAZ (4)	HAZ (5)	WAZ (6)
Arsenic	−0.829 (0.248)	−0.822 (0.189)	−0.797 (0.294)	−0.803 (0.221)	−0.520 (0.302)	−0.589 (0.225)
Arsenic*Girl			−0.066 (0.335)	−0.040 (0.261)	−0.118 (0.320)	−0.039 (0.247)
Girl	0.082 (0.011)	0.025 (0.008)	0.088 (0.029)	0.029 (0.023)	0.096 (0.028)	0.028 (0.022)
Kleinbergen-Paap test-stat	74	74	37.88	37.88	29	29
Cragg-Donald Wald *F*-stat	842.8	842.8	421.21	421.21	317.5	317.5
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
Other contaminants	No	No	No	No	Yes	Yes
State F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	88,106	88,106	88,106	88,106	79,305	79,305

Notes: Standard errors clustered at the PSU level. Arsenic is measured in milligrams per liter. 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, rice and wheat production and MPCE. Individual level controls (age, age square and gender), maternal controls (mother's age, mother's education, maternal anemia (severe, moderate, mild, non- anemic), mother's standardized height for age (HAZ) and weight for height (WHZ)) and family background controls (religion, caste, family size, wealth index and place of residence). The last two columns include other soil and water contaminants, namely, nitrate, nitrogen, phosphorous, potassium, lead and fluoride

specific breastfeeding patterns 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 this hypothesis.

5.3 *Arsenic and child health across gender and birth order*

We study the interaction between arsenic exposure, gender, and birth order in table 5. We show OLS results for HAZ and WAZ, respectively, 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 5 for the triple interaction terms (arsenic*girl*birth order) suggest that girls in arsenic affected regions have a 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 unit change in arsenic leads to a decrease in HAZ (stunting) for third (or later) born girls by 0.65 SDs. The significance of our estimate for third (or later) born girls indicates that arsenic induced stunting in girls increases with steeper birth gradient. We find similar IV results for WAZ as shown in column 6. IV estimates on WAZ indicate that a one unit increase in arsenic leads to a decrease in WAZ (underweight) for second and third (or later) born girls by 0.51 and 0.63 SDs, respectively. When the concentration of arsenic in groundwater increases, later born girls (born at higher birth order) experience

Table 4. Heterogeneous effects by household wealth (IV estimates)

Anthropometric measures (*z*-score) Wealth index	HAZ (1) Low	HAZ (2) High	WAZ (3) Low	WAZ (4) High
Arsenic	−1.192 (0.423)	−0.783 (0.648)	−1.148 (0.319)	−0.894 (0.502)
Observations	27,715	24,586	27,715	24,586
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

Notes: Standard errors clustered at the PSU level. 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, rice and wheat production and MPCE. Individual level controls (age, age square), maternal controls (mother's age, mother's education, maternal anemia (severe, moderate, mild, non-anemic), mother's standardized height for age (HAZ) and weight for height (WHZ)) and family background controls (religion, caste, family size and place of residence).

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; Chaudhry *et al*., 2021).[24]

6. Robustness and falsification tests

6.1 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 neonatal 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. Further, to overcome the issue of endogeneity of sibling size, we use the gender of the first child as an instrument for sibling size. Evidence suggests that having a girl as the first child is positively associated with fertility, particularly in the presence of son preference, as parents will continue to have more children until the desired number of boys are born in a family (Aksan, 2021; Pörtner, 2022). Further, gender of the first child is exogenously determined and should affect child health outcomes only through fertility (Kugler and Kumar, 2017).

---

[24] It is worth noting that resource allocation may vary by birth order due to many reasons. For instance, as the family size increases, maternal time investment may decline, or families may face limited resources. However, in the absence of gender bias, boys and girls should be equally likely to face limited resources in larger families or face parental time constraints. Thus, both boys and girls born in higher birth orders should be equally likely to face a disadvantage and we should observe no gender specific birth order effects of drinking contaminated water on health outcomes.

Table 5. 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	0.032 (0.056)	−0.507 (0.304)	−0.458 (0.298)	0.076 (0.056)	−0.559 (0.218)	−0.507 (0.216)
Arsenic*girl*BO3	−0.041 (0.073)	−0.761 (0.312)	−0.649 (0.304)	−0.036 (0.052)	−0.742 (0.232)	−0.627 (0.228)
Arsenic	0.046 (0.034)	−0.492 (0.384)	−0.364 (0.358)	0.062 (0.028)	−0.461 (0.282)	−0.492 (0.269)
Birth order 2	−0.032 (0.013)	−0.028 (0.034)	−0.024 (0.033)	−0.026 (0.009)	−0.020 (0.025)	−0.016 (0.024)
Birth order 3	−0.074 (0.014)	0.081 (0.046)	0.062 (0.043)	−0.058 (0.011)	0.065 (0.034)	0.060 (0.032)
Girls	0.081 (0.011)	0.067 (0.032)	0.088 (0.030)	0.021 (0.008)	0.031 (0.025)	0.028 (0.023)
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	Yes	Yes	Yes	Yes	Yes
Observations	85,520	100,958	85,520	85,520	100,041	85,520

Notes: Standard errors clustered at the PSU level. 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, rice and wheat production and MPCE. Individual level controls (age, age square), maternal controls (mother's age, mother's education, maternal anemia (severe, moderate, mild, non-anemic), mother's standardized height for age (HAZ) and weight for height (WHZ)) and family background controls (religion, caste, family size and place of residence).

In this specification, we estimate the regressions separately by gender.[25] We estimate the following OLS and first stage regressions for boys and girls, separately:

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

$$\begin{aligned} Ars_{ids} &= \lambda_1 Soil_{ds} + \lambda_2 2ndchild_{ids} + \lambda_3 3rd^{+}child_{ids} + \lambda_4(Soil_{ds} * 2ndchild_{ids}) \\ &\quad + \lambda_5(Soil_{ds} * 3rd^{+}child_{ids}) + \lambda_6 gender\_first + \lambda_7 X'_{ids} + D_{ds} + S + e_{ids}, \end{aligned} \tag{6}$$

where sib_size_{ids} is the number of children under the age of five in a household. This variable is instrumented by *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 the same as in the previous regressions.[26]

---

[25] The instrument for sibling size (gender of first child) will otherwise be perfectly collinear with our main explanatory variables.

[26] We are cautious in interpreting the coefficient on sibling size in table 6. Angrist and Pischke (2009) explain in their textbook *Mostly Harmless Econometrics* that it is difficult to interpret models with multiple endogenous variables. This is commonly done, for instance, in models where education externalities (like peer effects) are measured. Here the researcher introduces an aggregate measure of schooling and a private

Table 6. Arsenic, sibling size and birth order (IV estimates)

Anthropometric measures	HAZ Girls (1)	HAZ Boys (2)	WAZ Girls (3)	WAZ Boys (4)
Arsenic*birth order2	−0.253 (0.288)	0.564 (0.296)	−0.842 (0.223)	0.330 (0.218)
Arsenic*birth order3^{rd+}	−1.046 (0.292)	−0.117 (0.289)	−1.387 (0.225)	−0.214 (0.214)
Sibling size (below age 5)	−0.110 (0.066)	−0.227 (0.065)	−0.109 (0.051)	−0.109 (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	40,778	44,227	40,778	44,227

Notes: Standard errors clustered at the PSU level. 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, rice and wheat production and MPCE. Individual level controls (age, age square), maternal controls (mother's age, mother's education, maternal anemia (severe, moderate, mild, non-anemic), mother's standardized height for age (HAZ) and weight for height (WHZ)) and family background controls (religion, caste, family size, wealth index and place of residence). Also includes controls for arsenic and birth order.

IV results from this specification are shown in table 6. Columns 1 and 3 show IV estimates for health outcomes (HAZ and WAZ, respectively) for girls. HAZ and WAZ for boys are reported in columns 2 and 4, respectively. After controlling for sibling size, a one SD increase in arsenic exposure leads to a significant decrease in HAZ and WAZ for girls born at third (or higher) birth order by 1.05 SD and 1.39 SD, respectively. For boys, we find insignificant effects on HAZ and WAZ for third birth order. Interestingly, column 2 shows a positive effect for HAZ scores among boys born in the second birth order. Looking at the coefficient on sibling size, a greater number of children in the household has a negative and significant effect on growth outcomes after accounting for birth order effects.

6.2 Non-arsenic states

We have shown that exposure to arsenic contaminated water impacts child growth and that this process disproportionately affects girls born at a later birth order. Our main regression exploits variation in arsenic levels across Indian states, thus the sample includes the nine states where arsenic is present in groundwater. However, comparing results in arsenic-affected states to those not in arsenic-affected states should also 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 above. Instead, we exploit the variation in source of drinking water and gender and compare results by birth order for arsenic and non-arsenic states. Thus, we run the following

---

measure of schooling, both of which are instrumented (see, for instance, Acemoglu and Angrist, 2000). Thus, we have instrumented both variables, but we note in passing that the estimates do not change if we treat sibling size as exogenous and only as an instrument for arsenic in equation (6).

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

$$Yids = a1swaterids + a2girlids + a3(swaterids * girlids) + a4Xids + D + eids, \qquad (7)$$

where *swater* is the source of drinking water[27] where we categorize it as a binary variable which takes the value of 1 if the primary source of drinking water is unsafe and 0 for safer sources of drinking water.[28] The main variable of interest is $a3$ which captures the health effects of drinking groundwater among females. This specification also allows us to control for district fixed effects, so the results are not confounded by geographical variation in agricultural patterns, irrigation, soil type or irrigation. Further, we control a host of household characteristics including income. 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.

The results in table 7 show that for non-arsenic states there is no effect on health outcomes; all coefficients are statistically insignificant at conventional levels. 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. Though the results may not necessarily capture the causal effect of drinking contaminated groundwater (as access to safe water is correlated with household education, income and information), the lack of a significant effect among non-arsenic states gives us further confidence regarding the birth order effects.

6.3 Soil quality and female labor supply

In a recent study, Carranza (2014) argues that loamy soils allow for deep tillage and thereby reduce 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 subsection 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

---

[27] We have excluded 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 since 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 per cent of the sample procured drinking water through this source while 78 per cent of households rely on groundwater sources for drinking.

[28] Unsafe sources of drinking water refers to groundwater sources including tube-wells, wells, protected and unprotected springs. Safer sources of drinking water refer to piped into dwelling, piped to yard, public tap/standpipe, rainwater, tanker truck, cart with small tank, bottled water and community plants that supply water purified through the reverse osmosis (RO) process.

Table 7. Interaction effect of groundwater and gender on health outcomes by birth order (arsenic and non-arsenic states)

	Non-Arsenic States				Arsenic States			
	HAZ BO1	HAZ BO3/4	WAZ BO1	WAZ BO3/4	HAZ BO1	HAZ BO3/4	WAZ BO1	WAZ BO3/4
Groundwater	0.045 (0.019)	0.066 (0.023)	0.026 (0.014)	0.034 (0.016)	0.020 (0.050)	0.208 (0.054)	0.063 (0.0372)	0.063 (0.024)
Girl	0.087 (0.017)	0.013 (0.022)	0.050 (0.013)	−0.030 (0.015)	0.128 (0.048)	0.165 (0.060)	0.083 (0.037)	0.051 (0.023)
girl*groundwater	−0.019 (0.023)	−0.032 (0.027)	−0.021 (0.017)	−0.001 (0.018)	−0.036 (0.057)	−0.176 (0.066)	−0.058 (0.042)	−0.052 (0.026)
Observations	65,855	58,743	65,855	58,743	13,971	16,203	13,971	16,203
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level in parentheses. All regressions include district fixed effects and individual level controls (age and age square), maternal controls (mother's age, mother's education), maternal anemia (severe, moderate, mild, non-anemic), mother's standardized height for age (HAZ) and weight for height (WHZ) 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, bottled water, rainwater harvesting).

labor force participation in agriculture across Indian districts.[29] Labor 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. As shown in table A3 (online appendix), the main results are robust to the inclusion of both the male and female employment in agriculture. Our results are also robust to dropping one state at a time from the analysis, and one region at a time after dividing the nine arsenic states into three regions: North, South and East (there are no western states with arsenic).[30] Finally, it is worth mentioning that the exogeneity of the instruments is hard to show in practice. At the same time, soil texture is an important determinant of arsenic levels. We also check the robustness of the IV results by estimating the equations simultaneously in an iterated seemingly unrelated regression model. While we do not show the results here, our results are robust to this alternative strategy.

7. Conclusion and policy implications

Gender inequality is a fundamental challenge to sustainable development. While considerable efforts have been made to explore the impacts of gender inequality on women, less is 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 (Guilmoto *et al.*, 2018). 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, fewer post-natal health inputs such as vaccinations 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

---

[29] 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).

[30] These results are available upon request.

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. While we acknowledge that finding a good IV is very hard in practice, we address the endogeneity of arsenic levels using an IV strategy which stands robust to several checks for internal validity and to alternate specifications. IV 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 are largest in poorer households, suggesting that nutritional deficiencies in childhood exacerbate the adverse effects of arsenic exposure.

Our results show heterogeneous effects of arsenic exposure by birth order, highlighting the role played by son biased preferences in magnifying the negative impact of unsafe water on health. 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 WHO, lack of accessibility of safe water is the 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 (Asadullah and Chaudhury, 2011). But any government policy that solely aims to provide safe drinking water will not deliver desired goals unless these policies are accompanied by equitable distribution of food and 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.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355770X24000329.

Acknowledgements. We are grateful to seminar participants at the EFD Annual Meetings, Nairobi, 2024; the NEUDC 2021 and the 3rd annual workshop of women in the economy, ISI Delhi. We thank the editor, coeditor, two anonymous referees and Professor Nidhiya Menon for their feedback and suggestions. All errors are our own.

Statements and declaration. The authors declare that no funds, grants or other support were received for this work and the authors have no financial relationships with any organizations that might have an interest in this work. There are no potential conflicts of interest to declare. This research is original and has not been simultaneously submitted for publication in any other journal. All authors contributed to the study conception and design.

Competing interest. The authors declare none.

References

Acemoglu D and Angrist J (2000) How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual* **15**, 9–59.

Aggarwal K, Barua R and Vidal-Fernandez M (2024) Still waters run deep: groundwater contamination and education outcomes in India. *Economics of Education Review* **100**, 102525.

Aksan AM (2021) Son preference and the fertility squeeze in India. *Journal of Demographic Economics* **87**, 67–106.

Angrist JD and Pischke JS (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Arceo E, Hanna R and 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**, 257–280.

Asadullah MN and 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**, 873–888.

Bae S, Kamynina E, Guetterman HM, Farinola AF, Caudill MA, Berry RJ, Cassano PA and Stover PJ (2021) Provision of folic acid for reducing arsenic toxicity in arsenic-exposed children and adults. *Cochrane Database of Systematic Reviews* **10**, CD012649.

Bahrami-Rad D (2021) Keeping it in the family: female inheritance, inmarriage, and the status of women. *Journal of Development Economics* **153**, 102714.

Bardhan PK (1974) On life and death questions. *Economic and Political Weekly* **9**, 1293–1304.

Becker GS and Tomes N (1976) Child endowments and the quantity and quality of children. *Journal of Political Economy* **84**, 143–162.

Behrman JR and Taubman P (1986) Birth order, schooling, and earnings. *Journal of Labor Economics* **4**, S121–S145.

Behrman JR, Pollak RA and Taubman P (1986) Do parents favor boys? *International Economic Review* **27**, 33–54.

Billings SB and Schnepel KT (2018) Life after lead: effects of early interventions for children exposed to lead. *American Economic Journal: Applied Economics* **10**, 315–344.

Booth AL and Kee HJ (2009) Birth order matters: the effect of family size and birth order on educational attainment. *Journal of Population Economics* **22**, 367–397.

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

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

Case A and Paxson C (2008) Stature and status: height, ability, and labor market outcomes. *Journal of Political Economy* **116**, 499–532.

Central Ground Water Board (2022) *Dynamic Ground Water Resources of India, River Development & Ganga Rejuvenation*. New Delhi: Ministry of Jal Shakti, Government of India.

Chaudhry TT, Khan M and Mir AS (2021) Son-biased fertility stopping, birth spacing, and child nutritional status in Pakistan. *Review of Development Economics* **25**, 712–736.

Chung W and Das Gupta M (2007) The decline of son preference in South Korea: the roles of development and public policy. *Population and Development Review* **33**, 757–783.

Clougherty JE (2010) A growing role for gender analysis in air pollution epidemiology. *Environmental Health Perspective* **118**, 167–176.

Coffey D and Spears D (2021) Neonatal death in India: birth order in a context of maternal undernutrition. *The Economic Journal* **131**, 2478–2507.

Deaton A and Drèze J (2009) Food and nutrition in India: facts and interpretations. *Economic and Political Weekly* **44**, 42–65.

Deguen S, Amuzu M, Simoncic V and Kihal-Talantikite W (2022) Exposome and social vulnerability: an overview of the literature review. *International Journal of Environmental Research and Public Health* **19**, 3534.

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

Do QT, Joshi S and Stolper S (2018) Can environmental policy reduce infant mortality? Evidence from the Ganga pollution cases. *Journal of Development Economics* **133**, 306–325.

Dyson T and Moore M (1983) On kinship structure, female autonomy, and demographic behavior in India. *Population and Development Review* **9**, 35–60.

Fängström B, Moore S, Nermell B, Kuenstl L, Goessler W, Grandér M, Kabir I, Palm B, Arifeen SE and Vahter M (2008) Breast-feeding protects against arsenic exposure in Bangladeshi infants. *Environmental Health Perspectives* **116**, 963–969.

Fecht D, Fischer P, Fortunato L, Hoek G, De Hoogh K, Marra M, Kruize H, Vienneau, D, Beelen R and Hansell A (2015) Associations between air pollution and socioeconomic characteristics, ethnicity and age profile of neighbourhoods in England and the Netherlands. *Environmental Pollution* **198**, 201–210.

Fledderjohann J and Channon M (2022) Gender, nutritional disparities, and child survival in Nepal. *BMC Nutrition* **8**, 50.

Fledderjohann J, Agrawal S, Vellakkal S, Basu S, Campbell O, Doyle P, Ebrahim S and 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**, e107172.

Foster A, Gutierrez E and Kumar N (2009) Voluntary compliance, pollution levels, and infant mortality in Mexico. *American Economic Review* **99**, 191–197.

Gardner RM, Kippler M, Tofail F, Bottai M, Hamadani J, Grandér M, Nermell B, Palm B, Rasmussen KM and Vahter M (2013) Environmental exposure to metals and children's growth to age 5 years: a prospective cohort study. *American Journal of Epidemiology* **177**, 1356–1367.

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

Glewwe P and Miguel EA (2007) The impact of child health and nutrition on education in less developed countries. *Handbook of Development Economics* **4**, 3561–3606.

Gómez-Roig MD, Pascal R, Cahuana MJ, García-Algar O, Sebastiani G, Andreu-Fernández V, Martínez L, Rodríguez G, Iglesia I, Ortiz-Arrabal O, Mesa MD, Cabero MJ, Guerra L, Llurba E, Domínguez C, Zanini MJ, Foraster M, Larqué E, Cabañas F, Lopez-Azorín M, Pérez A, Loureiro B, Pallás-Alonso CR, Escuder-Vieco D and Vento M (2021) Environmental exposure during pregnancy: influence on prenatal development and early life: a comprehensive review. *Fetal Diagnosis and Therapy* **48**, 245–257.

Goyal N and 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**, 22.

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

Guilmoto CZ, Saikia N, Tamrakar V and Bora JK (2018) Excess under-5 female mortality across India: a spatial analysis using 2011 census data. *The Lancet Global Health* **6**, e650–e658.

Herath I, Vithanage M, Bundschuh J, Maity JP and Bhattacharya P (2016) Natural arsenic in global groundwaters: distribution and geochemical triggers for mobilization, *Current Pollution Reports* **2**, 68–89.

Huang B, Yuan Z, Li D, Zheng M, Nie X and Liao Y (2020) Effects of soil particle size on the adsorption, distribution, and migration behaviors of heavy metal(loid)s in soil: a review. *Environmental Science: Processes & Impacts* **22**, 1596–1615.

Jayachandran S and Kuziemko I (2011) Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India. *The Quarterly Journal of Economics* **126**, 1485–1538.

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

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

Kile ML, Cardenas A, Rodrigues E, Mazumdar M, Dobson C, Golam M, Quamruzzaman Q, Rahman M and Christiani DC (2016) Estimating effects of arsenic exposure during pregnancy on perinatal outcomes in a Bangladeshi cohort, *Epidemiology (Cambridge, Mass.)* **27**, 173–181.

Kugler AD and Kumar S (2017) Preference for boys, family size, and educational attainment in India. *Demography* **54**, 835–859.

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

Lynch J, Smith GD, Harper S and Bainbridge K (2006) Explaining the social gradient in coronary heart disease: comparing relative and absolute risk approaches. *Journal of Epidemiology & Community Health* **60**, 436–441.

Madajewicz M, Pfaff A, van Geen A, Graziano J, Hussein I, Momotaj H, Sylvi R and Ahsan H (2007) Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh, *Journal of Development Economics* **84**, 731–754.

McArthur JM, Ravenscroft P, Safiulla S and Thirlwall MF (2001) Arsenic in groundwater: testing pollution mechanisms for sedimentary aquifers in Bangladesh. *Water Resources Research* **37**, 109–117.

Neidell MJ (2004) Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics* **23**, 1209–1236.

Pörtner CC (2022) Birth spacing and fertility in the presence of son preference and sex-selective abortions: India's experience over four decades. *Demography* **59**, 61–88.

Rahman A, Granberg C and Persson LÅ (2017) Early life arsenic exposure, infant and child growth, and morbidity: a systematic review. *Archives of Toxicology* **91**, 3459–3467.

Samiee F, Leili M, Faradmal J, Torkshavand Z and Asadi G (2019) Exposure to arsenic through breast milk from mothers exposed to high levels of arsenic in drinking water: infant risk assessment. *Food Control* **106**, 106669.

Spears D, Coffey D and Behrman JR (2022) Endogenous inclusion in the demographic and health survey anthropometric sample: implications for studying height within households. *Journal of Development Economics* **155**, 102783.

Tanaka S (2015) Environmental regulations on air pollution in China and their impact on infant mortality. *Journal of Health Economics* **42**, 90–103.

Thurstans S, Sessions N, Dolan C, Sadler K, Cichon B, Isanaka S, Roberfroid D, Stobaugh H, Webb P and Khara T (2022) The relationship between wasting and stunting in young children: a systematic review. *Maternal & Child Nutrition* **18**, e13246.

Trejo S, Yeomans-Maldonado G and Jacob B (2021) The psychosocial effects of the Flint water crisis on school-age children. Working Paper 29341, National Bureau of Economic Research, Cambridge, MA.

UNICEF (2013) *Improving Child Nutrition: The Achievable Imperative for Global Progress*. New York: United Nations Children's Fund (UNICEF).

Vogl TS (2016) Differential fertility, human capital, and development. *The Review of Economic Studies* **83**, 365–401.

Watanabe C, Matsui T, Inaoka T, Kadono T, Miyazaki K, Bae MJ, Ono T, Ohtsuka R and Bokul AT (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**, 1835–1841.

Cite this article: Aggarwal K, Barua R (2024). Gender disparities in the prevalence of undernutrition in India: the unexplored effects of drinking contaminated water. *Environment and Development Economics* 1–21. https://doi.org/10.1017/S1355770X24000329