The Australian Research Council Centre of Excellence for Children and Families over the Life Course
Phone +61 7 3346 7477 Email lcc@uq.edu.au lifecoursecentre.org.au

WORKING PAPER SERIES

STILL WATERS RUN DEEP: GROUNDWATER ARSENIC CONTAMINATION AND EDUCATION OUTCOMES IN INDIA

Khushboo Aggarwal

Rishihood University

Rashmi Barua

Centre for International Trade and Development – Jawaharlal Nehru University

Marian Vidal-Fernandez

The University of Sydney

A more recent version of this paper was published as Aggarwal K, Barua R, Vidal-Fernandez M. (2024) Still Waters Run Deep: Groundwater Contamination and Education Outcomes in India. Economics of Education Review, 100, 102525. DOI: 10.1016/j.econedurev.2024.102525

No. 2022-08

May 2022

NON-TECHNICAL SUMMARY

The water crisis is a pressing issue in developing countries because climate change reduces the rate of rainwater seeping underground, increasing the concentration of toxins in groundwater. India and Bangladesh have the largest population in the world exposed to arsenic in groundwater. Approximately 70 million people in India are exposed to groundwater arsenic that is above permissible levels recommended by the WHO. Children comprise nearly half of the affected population and consuming arsenic contaminated water is likely a contributor to India's high child mortality rate. Moreover, very little is known about the effects of arsenic consumption on children's education outcomes.

We find that prolonged exposure to arsenic-contaminated water consumption in India significantly increases school absenteeism, grade retention, and decreases test scores of primary and secondary students. Moreover, our newly collected dataset shows that detrimental effects are larger among girls, exacerbating an already underlying health gender disadvantage in childhood.

ABOUT THE AUTHORS

Khushboo Aggarwal graduated from her PhD at Jawaharlal Nehru University in 2021. She is currently an assistant professor at Rishihood University in New Dehli. She was awarded with the best research paper at 2ndInternational Conference at FLAME University in association with Indian Econometrics Society and the Wage Indicator Foundation. Her paper also received the best paper award in the International Conference on "Sustainable Development in International Higher Education: Strategies to achieve Gender Equality and Empowerment" organised by Guru Gobind Singh Indraprastha University & Jawaharlal Nehru University. Email: Khushboojnu14@gmail.com

Rashmi Barua* teaches at the Centre for International Trade and Development, Jawaharlal Nehru University (New Delhi). Prior to that she has worked at the School of Economics, Singapore Management University and the Economics and Planning Unit of the Indian Statistical Institute, New Delhi. Her research spans across the economics of education and labor markets with specific interest in early childhood human capital investments (health and education), female labor supply and economics of crime. More recently she has worked on financial education among international migrants. Most of her research has been a mix of methodological contributions, including randomized evaluations, and empirical applications that are guided by sound economic theory. Email: rashmibarua@mail.jnu.ac.in

Marian Vidal-Fernandez is an Associate Professor at the University of Sydney and an LCC Associate Investigator. She is an applied human capital economist interested in learning what prevents disadvantaged populations to achieve their full potential. For instance, she has studied how setting minimum academic requirements to enrol into athletic activities or obtaining a driving license can improve high school graduation and crime rates, the increasing relevance of grandmothers as childcare givers, the birth order effect, how personality traits affect productivity, and the dynamic and heterogeneous impacts of experiencing the death of a sibling. She is an award-winning teacher and experienced engaging with the media and general public. Email: m.vidal-fdez@sydney.edu.au

Acknowledgements

We are grateful to Anant Khanikar, Nipendra Sharma, Ajoy Lahon, Runti Chowdhury, and Dr. Ratul Mahanta for support in the initial stages of this project. We thank Professor Abhiroop Mukhopadhyay for feedback. Thanks to Sultana Khan who was an excellent project coordinator and the entire team of research assistants in Jorhat. We thank participants in seminars at Jawaharlal Nehru University, Indian Statistical Institute, the 2018 ISI Conference on Economic Growth and Development, the CECFEE 2019 conference in Tezpur, Assam, and the AASLE 2021 participants. Marian Vidal Fernandez gratefully acknowledges funding from the University of Sydney's India Development Fund. IERB JNU Ref. No. 2018/174

ABSTRACT

We study the effect of exposure to arsenic in groundwater, the main source of arsenic contamination, on a range of education outcomes among children in India. First, 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 school attendance. Second, to support our initial findings, we surveyed students in public schools in one of the most arsenic contaminated regions of India and exploited variation in the geographical coverage and timing of construction of government water supply schemes. We find that prolonged exposure to contaminated water sources increases school absenteeism, grade retention, and decreases test scores. Moreover, detrimental effects are larger among girls, exacerbating an already underlying health gender disadvantage in childhood.

Keywords: Arsenic, children, water, test scores, India

Suggested citation: Aggarwal, K., Barua, R. & Vidal-Fernandez, M. (2022). 'Still Waters Run Deep: Groundwater Arsenic Contamination and Education Outcomes in India', Life Course Centre Working Paper Series, 2022-08. Institute for Social Science Research, The University of Queensland.

1. Introduction

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 and inadequate sanitation (UNICEF, 2006). The water crisis is a particularly pressing issue in developing countries because climate change reduces the rate at which rainwater seeps underground, gradually increasing the concentration of toxins in water (Mc Arthur et al., 2001). In the presence of toxic metals in groundwater, common safety practices such as boiling do not turn water potable (Central Ground Water Board, 2006). One such contaminant – arsenic– poses a threat to the availability of clean drinking water in various emerging countries worldwide.

India and Bangladesh have the largest population in the world exposed to arsenic consumption through drinking water.[1] More than 70 million people across 40 districts of India are exposed to arsenic in groundwater that is above the permissible levels recommended by the World Health Organisation (Khurana and Sen, 2008). Children comprise nearly half of the affected population and consuming arsenic contaminated water is likely to contribute to India's high child mortality rate (Asadullah and Chaudhury, 2011). Infants 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).

It is well established that birth outcomes and health in childhood play a crucial role in determining education and adult economic outcomes.[2] While the effect of drinking contaminated water on a range of health outcomes, including growth outcomes among children, is well documented,[3] less is known about its *causal* effects on education outcomes of children.

A causal relation between drinking arsenic-contaminated water and educational outcomes may arise for at least four reasons. First, prolonged exposure to arsenic has been linked to cognitive delays. Such

---

[1] See Appendix Figure A.2.1

[2] See for instance Case, Lubotsky, and Paxson (2002), Case, Fertig and Paxson (2005) and Behrman and Rosenzweig (2014) for studies in developed countries. Miguel and Kremer (2004) show that the provision of deworming drugs in Kenya led to an increase in educational attainment. Bobonis, Miguel and Puri-Sharma (2006) find a 20 percent decrease in school absenteeism rates following an iron supplementation and deworming intervention among 2–6 year-old preschoolers in the slums of Delhi, India.

[3] See for instance; Del Razo et al. (2011), Minamoto et. al. (2005), Tseng (2007) and Watanabe et. al. (2007)

impairment is likely to be more severe among children who have been also exposed to contaminants in utero and/or suffer an additional nutritional disadvantage in childhood (Asadullah and Chaudhury, 2011). Second, arsenic-induced illnesses, such as gastroenteritis, lead to school absenteeism. Third, visible symptoms of arsenic poisoning on the skin such as hyperkeratosis, and melanosis can be misinterpreted as contagious (Hassan et al., 2005) and lead to social isolation. Additionally, children are likely to miss school if they need to support or care for other sick relatives (Carson et. al., 2010).

Thus we expect arsenic contamination to affect educational outcomes not only due to adverse consequences on health but also due to its effect on social ostracization by peers. At the same time, most of the aforementioned papers estimate the effects of arsenic on education using cross sectional data and therefore are subject to selection bias and identification challenges. This paper addresses this gap in the literature by estimating the effects of exposure to arsenic in groundwater, the main source of arsenic contamination, on children's educational outcomes.

Estimating the effects of groundwater contaminants on educational outcomes, using either the regional variation in contaminant levels or the variation in a households access to safe water, is problematic because 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 because naturally occurring arsenic dissolves out of rock formations when groundwater levels drop significantly (Madajewicz et al., 2007). Similarly, relying on variation across households in access to safe water is misleading because socio-economic status is positively correlated both with educational levels and the availability of safe water at home.

In this paper, we employ two empirical strategies and databases to overcome both identification challenges. First, we use the India Human Development Survey II (IHDS-II) and the variation in the fractions of clayey soil texture across districts within the same state as an instrument for arsenic levels in groundwater to measure its impact on school absenteeism. Finer soils such as clay have relatively higher particle density and are less porous than coarse sandy soil which increases the concentration of contaminated groundwater (Brammer & Ravenscroft, 2009; Madajewicz et al., 2007). Instrumental Variable (IV) estimates indicate that, on an average, being exposed to an additional microgram per liter of arsenic in groundwater increases school absenteeism by 0.13 percentage points. This is a conservative estimate because absenteeism information is only available for children enrolled in school. This finding is robust to the inclusion of state fixed effects and district level controls for weather, type of crops, water quality measures, education quality and socio-economic variables.

An important question for policy is whether provision of safe government piped water can help counter the adverse effects of drinking contaminated water in arsenic prone areas. To address this question and to support the previous results, we collected school-level data in one of the most arsenic affected district of India (Jorhat) in the state of Assam. The identification strategy relies on two government programs, the National Rural Drinking Water Program (NRDWP) and the Swachh Bharat Mission (SBM) that provide time and cohort variation in access to safe, piped water.[4]

Differences-in-differences estimates indicate that more number of years of exposure to contaminated drinking water (before government water supply was made available) is associated with lower Cumulative Grade Point Average (CGPA) in the prior grade attended and higher rates of grade repetition. In addition, exposure to contaminated groundwater leads to higher rates of absenteeism and lower numeracy scores. An additional year of exposure to unsafe water increases school absenteeism by 6 percentage points. Retention rates increase by 3.6% with an additional one year exposure to contaminated water for the entire sample with higher effects among females (4.6%) compared to males (3%).

We also find that negative effects on test scores and grade retention in the survey are larger among girls. We argue that a plausible explanation is that girls are more likely to face a nutritional disadvantage in childhood, including shorter periods of breastfeeding, that is likely to exacerbate the detrimental effects of arsenic consumption. This is consistent with the statistics that India is the only country in the world where the under-five mortality rates are worse among girls than boys.[5] Further, a number of studies have shown linkages between the incidence and severity of arsenic induced health effects and nutritional status that suggest that people with poor nutrition are more susceptible to ill-effects of arsenic (Vahter, 2007). Similarly, Aggarwal and Barua (2021) find that arsenic exposure beyond the safe threshold level is negatively associated with Height-for-age and Weight-for-age and the negative effects are larger for girls who are born at higher birth orders relative to the elder ones.

---

[4] These programs were implemented by Public Health Engineering Department (PHED) of Assam in 2008-2009. The PHED is the main government agency responsible for water supply.

[5] UN Inter-agency Group for Child Mortality Estimation (2013) Child mortality estimates. Available: http://www.childmortality.org/.

We make relevant contributions to the existing literature. First, to the best of our knowledge, no previous study has been able to identify the causal effect of arsenic exposure on education outcomes. Our two complementary identification strategies and newly collected dataset are unique and able to identify the causal effect of arsenic exposure at the regional level as well as at individual level. Second, we are the first study to analyse the role of gender in the relation between child health and access to safe drinking water. Our results suggest that lack of adequate nutrition and health care during early childhood can make girls more vulnerable to an external environmental hazard. Third, global warming is expected to reduce the rate at which rainwater seeps underground, increasing the level of arsenic concentration. In the backdrop of climate change and over exploitation of groundwater, this paper makes a significant contribution to the literature on environmental health and education outcomes.

The remainder of the paper is structured as follows. Section 2 reviews the related literature. In Section 3 we introduce our two empirical strategies for each dataset. Section 4 describes our two data sources: the IHDS-II data and our newly conducted survey in the Indian state of Assam. In Section 5 we report the main results of the study. Concluding remarks and policy implications of our analysis are presented in Section 6.

2. Literature Review

Ingesting arsenic can lead to serious health problems. Arsenic poisoning, or so-called arsenicosis, can become a chronic illness if arsenic is consumed regularly through drinking and consuming food cooked with contaminated water. Arsenicosis has been linked to kidney and heart failure, mental illnesses, cancer, skin-related diseases, and adverse pregnancy outcomes (Tseng 2007). Children are more susceptible to arsenic because of their lower immunity levels and relatively higher proportion of body water compared to adults (Saing and Cannonier 2017). Moreover, epidemiological evidence suggests that arsenic crosses the placenta and adversely impacts pregnancy outcomes such as birth weight (Kile et al. 2016). As discussed in the previous section, it is well established that birth outcomes and health in childhood play a crucial role in determining education and adult economic outcomes.

A myriad of observational studies find a negative link between arsenic contamination in drinking water and education outcomes (see, for instance, Hassan et al. 2005, Murray and Sharmin 2015 and Wasserman et al. 2004; 2007). However, the source of drinking water, i.e., whether households use surface water or groundwater depends on unobserved household characteristics such as health safety knowledge and the availability of resources to access safe water that are also related to education. Thus, the correlational negative effects of arsenic exposure and educational outcomes are likely overestimated.

To the best of our knowledge, only two studies have attempted to tackle this identification challenge. Asadullah and Chaudhury (2011) find that living closer to a contaminated well is associated with lower mathematical scores among school-age children in Bangladesh. However, Pitt et al. (2015) argue that proximity to arsenic contaminated sources is not exogenous to educational outcomes because households are likely to switch to surface water sources in areas where widespread information campaigns are conducted. Therefore, families located near contaminated wells are more informed about the consequences of contaminated drinking water and are more likely to switch to alternative water sources.

A second identification strategy used by Saing and Cannonier (2017) exploits regional variation in the measured levels of arsenic in groundwater as a proxy for arsenic contamination exposure. The authors measure the impact of arsenic contamination on school enrolment of children in Cambodia and find that a one standard deviation increase in average arsenic levels in groundwater reduces the probability of having ever been enrolled in school, particularly among girls.

While arsenic occurs naturally in soil and water, over-exploitation of groundwater increases concentration levels (Mc Arthur et al. 2001). Therefore, arsenic levels in soil are also likely correlated with regional differences in economic outcomes related to education such as agricultural practices, population density, and industrialization. This geographical relationship also challenges the identification of the effects of arsenic on educational outcomes.

Our paper fills a gap in the literature by employing a new dataset and empirical strategy to analyse the causal effects of groundwater arsenic exposure on educational outcomes of children. We therefore contribute to the relatively thin literature on the effect of environmental contaminants on education outcomes of children in developing countries. [6]

---

[6] While a myriad studies have looked at the relationship between pollution and education in developed countries (see, for instance, Currie et al.(2009) for the US and Lavy, Ebenstein and Roth (2014) for Israel), in the developing country context the evidence is scarce. Bharadwaj et al. (2017) compare sibling outcomes to find that exposure to air pollution in the womb has a negative impact on mathematics and language skills among fourth grade students in Santiago, Chile.

3. Datasets and Empirical Strategy

3.1. Effects of Arsenic in Groundwater on School Absenteeism Using a National Survey

3.1.1 IHDS

We first use district level data from the India Human Development Survey –II (2011-12). The IHDS covers around 42,152 households and 204,568 individuals across 1,503 villages and 971 urban neighbourhoods in India. The IHDS is the only nationally representative dataset in India providing information on education outcomes such as school absenteeism and numeracy scores together with a wide range of individual, household and family background characteristics.[7]

Our main outcome variable, school absenteeism*,* is defined as the number of days a child was absent from school in the previous month. There are approximately 12,941 children in our sample between the ages of 5-19 enrolled in primary and secondary schools. Table 1 summarizes the key variable used in the IHDS analysis including the district level controls. On average, children missed 4.4 days of school in the previous month or approximately one school day a week. The average child in the sample is 11.8 years old and the sample is gender balanced.

3.1.2 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[8] which facilitates arsenic concentration in groundwater (Mac Arthur et al. 2001; Madajewicz et al. 2007).

Herath et. al (2016) find that in the Ganges–Meghna–Brahmaputra basin of India and Bangladesh, aquifers covered by finer sediments (clay) contain greater concentrations of arsenic in groundwater, whereas arsenic concentrations are significantly lower in aquifers with permeable sandy materials at the surface.

---

[7] The ASER (Annual Status of Education Report) is an alternative annual survey that also focuses on learning outcomes and schooling status of children for rural districts in India, but it only provides information on rural areas. Because arsenic levels increase with anthropogenic activities, urban areas could potentially have high levels of arsenic in groundwater and hence the use of this dataset is not appropriate.

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

Because arsenic concentration is higher in clayey relative to coarse soil, we exploit the variation in soil texture across districts within a state to instrument for ground water arsenic contamination. Figure 1 provides visual evidence supporting a positive association between arsenic (measured in micrograms per litre) and percentage of clayey soil across districts of India.

Our estimating and first stage equations, respectively, are as follows:

$$\boldsymbol{Y_{ids} = \delta Ars_{ds} + \gamma X_{ids} + D_{ds} + S + e_{ids}} \quad \textbf{(1)}$$

$$\boldsymbol{Ars_{ds} = \pi Soil_{ds} + \mu X_{ids} + D_{ds} + S + \epsilon_{ids}} \quad \textbf{(2)}$$

We are interested in measuring the effect of arsenic on the education outcomes $\boldsymbol{Y_{ids}}$ of child *i* in district *d* of state *s* in Equation (1). The main explanatory variable Ars_{ds} indicates the concentration level of arsenic in groundwater. We instrument arsenic soil contamination using $Soil_{ds}$ i.e., the percentage of clayey soil in districts *d*. X_{ids} is a vector of controls including individual characteristics (age, age-squared and a gender dummy) and family background characteristics (parental education, religion and caste dummies). S denotes state fixed effects. D are the district-specific controls including sex ratio, ratio of rice to wheat production, other contaminants in groundwater (iron and fluoride), monthly per capita consumption expenditure in rupees, average rainfall in millimetres (5-year average), and urbanization rate. We also control for district-level education variables, namely, total of student enrolments and the ratio of teachers to schools in private and public schools.

The identifying assumption is that soil texture fractions should affect education outcomes only through the impact on the level of arsenic in groundwater. A threat to identification is that soil texture impacts crop suitability and therefore, income. In particular, clayey-rich soil regions are suitable for growing water intensive crops such as rice relative to wheat that requires relatively less irrigation. Thus, we control for the district-level ratio of rice to wheat production (measured in millions tonnes) in all the regressions. We also control for district level male to female ratio in all regressions to control for the possibility that soil texture can affect economic outcomes through relative female to male employment rates (Carranza, 2014). As we show in the subsequent sections, the IV estimates do not change with inclusion of these additional control variables.

Note that while 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 which might alter some properties of superficial soil but not the subterranean soil used in our analysis.

However, there might be further threats to our identification strategy if clayey soil is correlated with other geographic or demographic characteristics that impact economic outcomes. We do not find any direct correlation between the proportion of clayey soil in a district on weather, other contaminants (iron and fluoride), economic or demographic factors, conditional on state fixed effects.[9] While we do find a significant difference by soil permeability (and thus more arsenic) in the ratio of teachers to schools in government schools, this would be against finding a negative impact of arsenic on educational outcomes and if anything, underestimate our results.

We next discuss the school level survey and the accompanying alternative identification strategy.

3.2. Effects of Safe Water Access on Education Outcomes: School Survey in Assam

3.2.1. Background: Arsenic Contamination and Water Sanitation Programs in Assam

Among the arsenic affected states of India, Assam is one of the most severely impacted (Government of Assam, 2013). According to the 2011 census, access to safe drinking water is only available to 9.2 percent of the population compared to the national average of 32 percent because over 50 percent of households use groundwater from tube wells and hand pumps as their primary sources of drinking water.

Out of the 27 districts in Assam, Jorhat has the largest amount of contaminated habitations, namely 815 out of a total of 963 (Singh, 2004).[10] In Jorhat, concentration of arsenic varies between 194 to 491 microgram per litre in most of the habitations, which is far beyond the safety limit of 50 micrograms per litre as recommended by the Bureau of Indian Standards in accordance with the guidelines issued by WHO.

Arsenic can be removed using either arsenic-removal treatment plants or filters. Filters consist of an inlet connected to a household or communal hand pump or tube well. While they are relatively easy to

---

[9] Results are reported in Table A.3.1. in the Appendix.

[10] There are 35 states and union territories in India. Each state is further administratively divided into districts (also known as Zila). Further, these districts are categorized into sub-districts where the lowest administrative unit is a town (urban areas) or a village (rural areas). The primary sampling unit for our survey is a habitation, where a group of households collectively form habitations.

assemble, filtering systems are often faulty over time because they need regular maintenance. Alternatively, safe rain water can be harvested or irrigation systems can lower the concentration of arsenic in ground water. However, our survey evidenced that high maintenance costs made the practice of rainwater harvesting and filtration techniques unsuitable.

Another commonly known water contaminant in Assam is iron. A widely used and effective tool to remove iron from drinking water are sand filters, but unfortunately they fail to remove arsenic.[11]

To tackle the rise in arsenic levels in drinking water, in 2008, two central government sponsored flagship programs were introduced in Jorhat by the Assam Public Health Engineering Department (PHED), namely, the National Rural Drinking Water Program (NRDWP) and the Swachh Bharat Mission (SBM). Both schemes provide access to safe surface water drinking sources from two neighbouring rivers: Doyang and Dhansiri.

Clean water is made available via community taps (and private connections) to a cluster of households withing a villages, also called habitations. NRDWP was planned to cover 507 habitations with approximately 40,000 individuals distributed across 17 councils (Gram Panchayats) of the Titabor block in Jorhat.[12] The scheme provided two safe water access options. Households could pay a private connection or alternatively, access a more affordable shared community level connection.

To ensure smooth functioning of the scheme, water committees consisting of around 15 members representing each habitations were formed. Households had to sign an agreement with the president of the water committee (known as *Panee Samitee*) to be connected to piped water. In addition, beneficiary households have to pay set-up fee of INR 1,000 (approximately USD 14) plus INR 100 monthly for connection charges.[13] Informal conversations with PHED officials revealed that the majority of households

---

[11] According to the 2015-16 round of the National Family Health Survey (NFHS-4), 48% percent of households across 27 districts of Assam treat their drinking water to make it potable using ceramic, sand or other water filter.

[12] As explained in the following sections, our sample consists of 283 habitations in the Titabor block of Jorhat district in the state of Assam. Figure Map A.2.2 in the Appendix shows the geographical location of Titabor.

[13] The average monthly per capita income in Jorhat district is approximately INR 3,200 (Human Development Report, Assam 2014).

in Titabor opted for cheaper communal connections that shared costs among members.[14] Safe water is supplied twice a day at both homes and community taps for one to two hours.

We use PHED administrative data on government water supply schemes at habitation level between April 2009 and March 2018. This data includes population by habitation, year in which the safe piped water scheme was installed, number of piped versus unsafe groundwater schemes,[15] start and completion date for the project. To create the main explanatory variable we use data on year reported which is defined as the first year in which a particular pipe water scheme was reported to be installed in a habitation.

3.3.2. Newly Collected School Survey

Titabor, one of the most contaminated blocks of the Jorhat district, has 162 villages with a total population of 201,791 individuals (2011 Census) out of which approximately 90% live in rural areas. In 2018, we surveyed 117 primary and secondary schools in Titabor and merged this newly collected student and school data with school administrative data. Figure A.2.3. in the Appendix depicts the details of the sampling process.

In May 2018 we first conducted focus group discussions followed by a pilot survey. The main survey was conducted across all 3rd, 5th, and 8th graders present in school and distributed across 283 habitations. According to administrative data provided by the education department, 4,316 students were enrolled across the three grades in the selected schools. We surveyed 3,065 students on the first visit and an additional 446 students during a follow-up, significantly reducing attrition from 29% to 19%.[16]

We administered a principal and a student questionnaire. The latter captures household background characteristics, whether the child faced any unfriendly atmosphere at school and why, number of days absent in previous month (verified with administrative data), and reasons for absenteeism. The

---

[14] Unfortunately there is no information available about who or how many chose this option. Moreover, since private connections required annual maintenance costs borne by households, a majority of households switched to community taps.

[15] Groundwater scheme include shallow and deep tube wells installed. The tube wells that were constructed under these schemes were found to be contaminated due to the presence of high levels of metals particularly arsenic.

[16] School-based surveys face potential selection bias resulting from absenteeism on the day of the survey. This could underestimate the results from our study as the children who are absent on any particular school day are precisely the ones most likely to be affected by adverse health conditions related to arsenic contamination. To minimize the bias due to absenteeism, we revisited schools on the final exam day which were scheduled in the last week of July 2018. Since schools were closing for summer vacations on the last day of exam, giving us a short time frame to conduct revisits, we revisited only those schools where the number of absent children were high on the initial date of survey. On the second visit, we surveyed only those students who were absent on the initial day of survey.

questionnaire also provides specific information on child's awareness of water contaminants, and the primary source of drinking water at home. We categorised tap water, filter/sand filters, rain water harvesting and piped water supply as safe and hand pumps, tube well, and/or pond water as unsafe.[17] The school principal survey was designed to capture school quality characteristics such as playground, library and toilet availability, student-teacher ratio, teachers' experience, enrolment and class size.

Our three main dependent variables based on school administrative data are: days missed school in the past 30 days, CGPA in the previous academic year, and if the child has ever repeated a grade. In addition, we administered short grade-specific literacy and numeracy assessments mimicking the National Achievement Survey (NAS) that was administered nationally in 2017 by the National Council of Educational Research and Training (NCERT). NAS is a district-level survey that aims to assess the learning outcomes of approximately 220,000 students across 700 districts in India.[18]

Table 2 provides means and standard deviations of key outcomes variables in our analysis. Just about 68 percent of the sample has access to safe drinking water at home. Only 15 percent of children could answer correctly all three math questions while 17 percent could not answer any of the questions correctly. The scores are even lower in case of verbal test (3% and 37% respectively). The other indicator of education achievement, CGPA scored in previous grade is low, only 31 percent of students scored above 71 percent.

Table 3 provides the means of various control variables such as age and gender, caste, religion, mother's education, structure of house, land/home ownership, durable assets and ownership of heavy vehicles.

---

[17] We also collected information on the source of drinking water at school. However, due to constant switching of water source by school, it was difficult to trace out the actual source of drinking water while in school. For instance, there were some schools which considered water from rain water harvesting as their primary source of drinking water. At the same time, due to the failure in its maintenance, students ultimately consumed water from hand pumps and tube wells..

[18] To measure numeracy among 3rd, 5th and 8th graders, three questions about mathematical operations (addition, subtraction, multiplication, geometry and linear algebra) were assessed depending on the grade of the student. Similarly, verbal abilities were assessed for 5th and 8th graders based on three questions on reading comprehension (reading an advertisement or small passage in English). Verbal ability tests were not conducted among third graders based on the pilot revealing rather poor English language abilities among younger children. Our outcome variables are the fraction of questions answered correctly by the student. Appendix 3.4 lists the grade and subject specific questions that were administered.

3.2.3. Differences-in-Differences Approach

The identification strategy using the school-based survey exploits exogenous variation in timing, coverage of the government water supply scheme, and child year of birth to measure the effects of safe water access and exposure. In our estimating equation

$$Y_{igsh} = \beta_1 years_{igsh} + \beta_2 X_{igsh} + \beta_3 treat_h + G + S + H + \varepsilon_{igsh,} \quad (3)$$

Y_{igsh} refers to the education outcomes for child *i* in grade *g*, school *s* and habitation *h*. $Years_{igsh}$ is the number of years the child has been exposed to unsafe water. This variable is determined by the interaction of age of the child, the timing the safe water availability, and habitation of residence.

For instance, in a habitation that obtained access to a safe piped water in 2013, the number of years habitation had access to safe water in 2018 is 5 years. If a child living in this habitation is 12, she will have been exposed 7 years to unsafe water. X_{igsh} is a vector of individual and family and household background characteristics. G is a grad fixed effect, S is school fixed effect and H is habitation fixed effect. Standard errors are clustered at the habitation level.

The Differences-in-Differences (DD) identification strategy exploits two sources of variation. First, we compare children in the same grade and school living in different habitations that got access to water at different points in time. Second, we compare children in different grades living in the same habitation. This specification, that uses variation in years of exposure, also solves the issue of interpreting the Average Treatment Effect (ATE) in a standard two-way fixed effects DD when treatment timing varies (Goodman-Bacon, 2018). Further, in Section 5.4.1, we conduct an event study type analysis to further give credibility to our estimates and show that the treatment effect varies by years exposed to contaminated water.

One concern with the identification strategy is that parents who are aware of the water contamination problem may respond by campaigning for the water supply scheme to reach their village. In that case, we would observe that piped water is more likely to be supplied to habitations with more educated and thus more aware parents. However, survey data reveals very low levels of information on arsenic contamination among the population. Mahanta, Chowdhury and Nath (2016) find that 86% of households in arsenic affected habitations of Titabor did not know about the prevalence of arsenic in groundwater.

Similarly, our survey finds that only 7.5% of children studying in public schools in Titabor had heard about arsenic. [19]

5. Results

5.1 IV Results from the IHDS

Table 4, shows the first stage regression (Equation 2) and a positive and statistically significant relationship between arsenic and the fraction of clayey soil. The F-statistic is very high supporting soil texture as a strong instrument for arsenic levels. Table 5 reports OLS (column 1) and IV estimates for absenteeism measured in number of school days missed in the last month. While OLS estimates in column 1 are insignificant, the IV estimates suggest that arsenic has a positive and statistically significant (at the 5% level) effect on school absenteeism. The IV estimates in columns 2 and 3 show that results are not sensitive to the exclusion of geographic and economic control variables giving further credibility to the exogeneity of the instrument.

In particular, the estimate in Column 3 suggests that a 10 microgram per litre increase in arsenic in a district leads to a 0.06 days increase in monthly absenteeism. Given that the average number of school days missed in the previous months is 4.4, this estimate implies a 1.4 percent increase in number of school days missed in the previous month.

5.2 Heterogeneous Effects by Gender

The social medical literature suggests that the adverse effect of arsenicosis on school participation is larger for girls relative to boys because of social ostracization associated with visible skin lesions due to arsenic poisoning (Saing and Cannonier, 2017; Hassan et al., 2005). Contrary to this, in Table 5, columns 4 and 5 we find larger impact of arsenic on school absenteeism for boys than girls. The IV estimates indicate that, a ten microgram per litre increase in arsenic in groundwater in a district leads to a 0.1 days increase in absenteeism among boys (or 2.3 percent) and a statistically insignificant effect for girls.

---

[19] As an additional robustness, we also estimated the regressions controlling for the child's awareness related to presence of arsenic, fluoride or nitrates in groundwater. We implicitly assume that, if parents are informed about water contaminants, this should reflect in their children's awareness. Our results do not change with this additional control variable.

A concern with these estimates is that school absenteeism is conditional on enrolment. Children who are often sick, due to long term exposure to arsenic, will be less likely to be enrolled in school. Alternatively, older children or boys may need to work to substantiate the income of adult family members, particularly if family members are suffering from arsenic related health problems. Further, when faced with illnesses due to arsenic exposure, girls might be less likely to get treatment and hence more likely to drop out of school. Therefore, the results could be driven by selection due to differential school enrolment rates by gender.

To tackle selection into enrolment, we restrict the sample to a younger cohort (6-11 years) of children whose school enrolment is mandatory by law (97% of IHDS sample is enrolled in school in this age group). Results in Table 6 indicates that the gender gap in absenteeism is not due to differential enrolment rates as the estimates do not change when we restrict the sample to a younger age group with higher enrolment rates.

To sum up, we do not find evidence of higher social ostracization induced absenteeism among girls. At the same time, higher absenteeism among boys might be because they have to work to substantiate income of adult family members who are affected by arsenic contamination. More importantly, the IHDS results suggest that the impact of drinking contaminated water has gendered effects that needs more detailed examination. We explore this further in our school level survey.

5.4 Assam School Survey Results

5.4.1 Cohort Analysis and Identification Assumption

We provide first some visual evidence in support of the identification assumptions. If drinking arsenic contaminated water for longer periods of time has adverse effects on child health and therefore on education, then, the longer a child has been exposed to arsenic contaminated water, the worse should be her educational outcomes. Based on this intuition, consider the following variant of Equation 3:

$$\boldsymbol{Y_{ighs} = \beta_j \sum_{j=0}^{18} D_{ij} + \beta_2 X_{igsh} + \beta_3 treat_h + G + S + H + \varepsilon_{igsh}} \quad (4)$$

Where, $\boldsymbol{D_{ij}}$ is a dummy variable for whether student *i* has been exposed *j* number of years to unsafe water based on her birth year and the year when safe water became accessible in her habitation. Thus, the higher *j*, the longer a child has been exposed to unsafe water in her habitation. Students who have been exposed to unsafe water for up to 5 years are the omitted category. All other control variables are as in Equation (3) and described in Table 2. $Treat_h$ is a dummy variable equal to 1 if a habitation has access to safe drinking water. Figure 2 plots the coefficients $\boldsymbol{\beta_j}$ i.e. of the number of years a child has been exposed to unsafe water on school absenteeism, CGPA scored in prior grade, grade retention and numeracy. The vertical lines depict 95-percent confidence intervals around the estimates.

Results are striking; all coefficients increase with years exposed with the strongest results for school absenteeism and grade repetition. The magnitude of the effect on school absenteeism is relatively negligible until 9 years of exposure, and starts to increase thereafter though the estimates are noisy for most cohorts. Compared to children who have been exposed to contaminated water for less than 5 years, children with 6 years of exposure to unsafe water are approximately 10% more likely to repeat a grade while those with 10 years of exposure are 20% more likely to repeat a grade. Similar increasing trends are visible for numeracy scores yet coefficients are statistically significant only for children who are exposed to contaminated water for at least 13 years.

While the age of students is completely exogenous to the construction of water supply schemes, a concern is that the timing of construction might be driven by unobserved habitation level characteristics. Though we do not have data on the detailed characteristics at the habitation level, we can test if timing of construction is correlated with aggregate school quality measures at the habitation level. These results, shown in the appendix suggest that school quality measures are not correlated with the timing of water supply. Similarly we also find that village level characteristics are uncorrelated with the timing of construction of water supply schemes.

Our estimates will also be overestimating a positive impact of safe water access if safe water access was coincidentally being provided jointly with further policies targeted at improving educational outcomes.[20] Between 2011 and 2012, the Assamese government, distributed free bicycles to school going girls from low income households up to grade 10 who were studying in government schools to increase girls' enrolment rates. For this policy to bias our results, the timing of provision of bicycles would have to

---

[20] Examples of government policies implemented in the 2010s targeting children enrolled in grades 9-12 include Aarohan, Saptadhara, Cash or laptop award schemes. However, our sample is restricted to 3rd, 5th and 8th graders and were therefore directly unaffected by these schemes.

coincide with the timing of access to water across habitations. Though this is unlikely, we control for asset ownership including bicycles in all our regressions. [21]

5.4.2 Differences-in-Differences Results

Next, we show results for the differences-in-differences estimates corresponding to Equation 3 for the complete sample and separately by gender in Tables 7 and 8. We see that in table 7 that a one year increase in exposure to unsafe water is associated with a 0.2 day increase in absenteeism per month regardless of gender. This translates to a 6% increase in absenteeism with respect to the mean (3.4 days). The probability of repeating a grade increases by 3.6% with an additional one year exposure to contaminated water for the entire sample with higher effects among girls (4.6%) compared to boys (3%). Consistently, CGPA scored in the previous grade are lower for females but not males.

Table 8 shows results for numeracy and verbal assessments. We find that while there is no effect of years of exposure on verbal scores, an additional year of exposure to arsenic contaminated water leads to a 0.9% decline in the numeracy score. From columns 2 and 3, it is clear that the results are driven by girls. As expected, revisited students (i.e. those who were absent in the first round of surveying) obtain lower scores. The treatment dummy is positive for both genders suggesting that habitations with water access might be positively selected in terms of school quality and/or family background characteristics.

5.4.3 Explaining the Gender Differences in Effects of Arsenic Exposure

Girls exposed to unsafe water are more likely to repeat a grade and score lower on assessment tests than boys despite not having higher rates of absenteeism. We explore three hypotheses that can explain the differential effects by gender. First, because girls are more likely to face non-physical bullying such as social exclusion (Pells et al., 2016) and the impact that bullying can have on education (Gorman et al, 2019), visible skin lesions due to arsenic contamination could impact their test scores.

---

[21] In September 2012, the government of Assam jointly with UNICEF implemented sanitation and hygiene policies (WASH) in schools. This program provided access to toilets particularly for girls in schools. Nonetheless, the timing of implementation of this scheme does not coincide with timing of access to water across habitations. We also include school fixed effects in all regressions to control for differential access to facilities under WASH across schools. Results are robust to an alternative specification in dropping school fixed effects and controlling for school quality measures including provision of toilet and availability of safe drinking water, school infrastructure (library, daily hours of electricity, playground), teacher experience, class size and student teacher ratio. Results available upon request.

Second, these results could also be driven by gender differences in time use if, when a relative falls ill, girls are more likely to care for their family member than boys instead of studying (Motiram & Osberg, 2010).

Third, girls can be more severely affected by arsenic contamination because they tend to receive a poorer nutrition quality intake relative to boys (Jayachandran & Kuziemko, 2011) that translates into a poorer health status from a very young age. As shown also by Fledderjohann et al. (2014) , the gender differential treatment by parents starts in infancy as Indian girls are breastfed for shorter periods than boys and consume less milk.

We first test if girls are more likely to be exposed to an unfriendly atmosphere in school due to arsenicosis, at the bottom of Table 7. We do not find that being exposed to arsenic contaminated water for longer, and consequentially having more visible symptoms of arsenicosis, is related to facing an unfriendly atmosphere in school overall or by gender.

Because our school survey does not have nutrition or time use information, we rely on alternative datasets to test the rest of our hypotheses.[22] The IDHS records how many hours children spend doing homework per week during the last month. IV regressions using soil texture as instrument yield insignificant and close to zero estimates, indicating that the level of arsenic in groundwater in a district has no effect on time spent studying regardless of gender. We show these results in the appendix.

Finally, to explore gender differences in health outcomes, we take advantage of the 2011 Indian decennial census data which reports district level infant and child mortality rates by gender across India. We regress the ratio of girls (G) to boys (B) mortality rates $\mathbf{CMR_G/CMR_B}$ and $\mathbf{IMR_G/IMR_B}$ on a dummy equal to 1 if the district's groundwater is contaminated with arsenic and 0 otherwise. IMR is defined as deaths of children under one year of age per 1,000 live births while CMR is defined as deaths of children below five year of age per 1,000 live births. The regressions include state fixed effects and district level controls for rainfall, literacy, urbanization, pattern of cultivation, iron in groundwater and gross domestic product.

The results shown in Table 9 are striking. We find that in arsenic-affected districts, the relative infant and child mortality rates for girls are larger even after controlling for a host of district-specific factors. These results, even though not causal, do suggest that the effect of arsenic is magnified among girls due to

---

[22] Time use questions were not asked during the survey due to budgetary concerns

poorer health status starting early in life. These results are also in line with recent work by Barua and Aggarwal (2021) who find that higher levels of arsenic in groundwater is associated with worse health outcomes among girls born in higher birth orders.

A remaining concern with the interpretation of results is whether the effect of safe water supply schemes could have also reduced exposure to other water pollutants such as iron and fluoride found in groundwater or biological contaminants found in surface water. While high levels of natural fluoride have been found in other districts of Assam, the natural levels of fluoride are low in Titabor block. As aforementioned, there is a high awareness regarding the presence of iron in drinking water in Assam and households have traditionally used sand filters to treat it. This method effectively removes iron but not arsenic in water. In our sample, while only 7.5% of students were aware of arsenic, 51% were aware of the presence of iron. Due to the high awareness of iron and historical use of sand filters, we do not expect the water supply schemes to have a significant effect on iron exposure levels. Finally, while a reduced exposure to biological contaminants may explain the effects on absenteeism, it cannot explain the gender differences or the effects on cognitive test scores that we find in our study.

6. Discussion and Policy Implications

The leading cause of morbidity in India is the lack of access to safe drinking water affecting more than 75.8 million people (WHO 2009). The estimated yearly costs to the economy are 90 million days in production loss each year and 6 billion Rupees in production. Moreover, over exploitation of groundwater and climate change is steadily increasing the concentration of toxic metals in groundwater and thus increasingly adding to the costs. These figures do not include the costs associated with the loss in human capital and its intergenerational effects. Our paper contributes to the literature measuring the effects of environmental contamination on human capital.

Combining results from a large nationally representative household survey with a primary survey conducted across schools in one of the most arsenic affected regions of India, we study the effect of arsenic contamination in groundwater on education outcomes. We first show that levels of arsenic above permissible limits has a direct causal association with absenteeism among children. Then we study the education outcomes more closely in one of the most arsenic contaminated districts of India where a safe alternative to groundwater was made available via a government water supply project. We find that children exposed to clean water for a shorter period of time had higher school absenteeism, grade retention, and lower test scores and these adverse effects are larger among girls.

India is the only country in the world where the under-five mortality rates are worse among girls than boys. Thus, we hypothesize that our results could be driven by the higher childhood nutritional disadvantage faced by girls relative to boys. At the same time, in areas with contaminated drinking water and gender bias, large scale government water supply programs are an effective policy tool to address the health and education disparities between boys and girls.

Mahanta et. al. (2016) estimate that the annual health costs in Assam of a one microgram increase in arsenic concentration per litre to be equivalent to USD 0.01 million. The welfare gains from reducing the level of arsenic concentration to the safe limits is estimated to be USD 2.49 million. In addition to health costs, our study finds that the negative impact of Arsenic on education outcomes would imply substantial economic costs associated with decreased productivity and wages.

References

Aggarwal, K. & Barua, R. (2021). “Gender Disparities in the Prevalence of Undernutrition in India: The Unexplored Effects of Drinking Contaminated Water.” Working Paper.

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.

Behrman, J. R., & Rosenzweig, M. R. (2004). “Returns to Birthweight,” *The Review of Economics and Statistics*, 86, 586-601.

Bharadwaj, P., Matthew Gibson, Joshua Graff Zivin & Christopher Neilson (2017). “Gray Matters: Fetal Pollution Exposure and Human Capital Formation,” *Journal of the Association of Environmental and Resource Economists*, 4(2): 505-542.

Bobonis, G. J., Miguel, E., & Puri-Sharma, C. (2006). “Anaemia and school participation,” *Journal of Human Resources*, 41(4): 692-721.

Brammer, H. and Ravenscroft, P. (2009). “Arsenic in groundwater: A threat to sustainable agriculture in South and South-east Asia,” *Environment International*, 35: 647-654.

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.

Carson, R. T., Koundouri, P., & Nauges, C. (2010). “Arsenic mitigation in Bangladesh: a household labor market approach,” *American Journal of Agricultural Economics*, *93*(2): 407-414.

Case**,** A., Lubotsky, D., & Paxson, C. (2002). “Economic status and health in childhood: The origins of the gradient,” *American Economic Review*, *92*(5): 1308-1334.

Case, A., Fertig, A., & Paxson, C. (2005). “The lasting impact of childhood health and circumstance,” *Journal of Health Economics*, *24*(2): 365-389.

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

Currie, J., Hanushek, E.A., Kahn, E.M., Neidell, M. & Rivkin, S.G. (2009). “Does pollution increases school absences?” *Review of Economics and Statistics*, 91(4): 682-94.

Del Razo, LM. et al. (2011). “Exposure to arsenic in drinking water is associated with increased prevalence of diabetes: a cross-sectional study in the Zimapán and Lagunera regions in Mexico,” *Environment Health,* 10(1): 73.

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.

Goodman-Bacon, Andrew. (2018). “Difference-in-Differences with Variation in Treatment Timing.” NBER Working Paper 25018.

Gorman, E., Harmon C.P., Mendolia, S., Staneva, A., and Walker, I. (2019). “The Causal Effects of Adolescent School Bullying Victimisation on Later Life Outcomes,” IZA Working paper 12241.

Government of Assam (2013). Annual report, Public Health Engineering Department, Assam.

Hassan, M. M. et al. (2005). “Social implications of arsenic poisoning in Bangladesh,” *Social Science & Medicine*, 61(10): 2201-2211.

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

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

Khurana, I., Sen, R. (2008). “Drinking water quality in rural India: Issues and approaches,” *Water Aid India*, 2008, (288701), 31.

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*, *27*(2): 173–181.

Lavy, Victor, Avraham Ebenstein & Sefi Roth (2014). “The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance and Human Capital Formation” NBER Working Papers 20648.

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 behaviour? Response to arsenic contamination of groundwater in Bangladesh,” *Journal of Development Economics*, 84(2): 731-754.

Mahanta R, Chowdhury J, Nath HK (2016) “Health costs of arsenic contamination of drinking water in Assam, India,” *Economic Analysis and Policy* 49:30–42.

Miguel, E., & Kremer, M. (2004). “Worms: identifying impacts on education and health in the presence of treatment externalities,” *Econometrica*, *72*(1): 159-217.

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.

Motiram, S & Osberg, L. (2010). “Gender Inequalities in Tasks and Instruction Opportunities within Indian Families,” *Feminist Economics*, 16(3): 141-167.

Murray, M. P., Sharmin, R. (2015). “Groundwater arsenic and education attainment in Bangladesh,” *Journal of Health, Population and Nutrition*, 33(1).

Pitt, Mark M. and Rosenzweig, Mark Richard and Hassan, Nazmul (2015). “Identifying the Cost of a Public Health Success: Arsenic Well Water Contamination and Productivity in Bangladesh.” NBER Working Paper 21741.

Pells, K., Ogando Portela, M.J, Espinoza Revollo, P. (2016). “Experiences of Peer Bullying among Adolescents and Associated Effects on Young Adult Outcomes: Longitudinal Evidence from Ethiopia, India, Peru, and Viet Nam,” Office of Research-Innocenti. UNICEF. Discussion Paper 2016-03.

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.

Saing, C. H., & Cannonier, C. (2017). “Arsenic Exposure and School Participation in Cambodia,” SSRN Workinp Paper 2907461.

Tseng CH. (2007). “Metabolism of inorganic arsenic and non-cancerous health hazards associated with chronic exposure in humans,” *Journal of Environmental Biology,* 28(2): 349–357.

UNICEF. (2006). Progress for children: A report card on water and sanitation Number 5, September 2006.

Vahter, Marie E. (2007) “Interactions between Arsenic-Induced Toxicity and Nutrition in Early Life,” *The Journal of Nutrition*, Volume 137 (12), December 2007, Pages 2798–2804

Wasserman, G. A. et al. (2004). “Water arsenic exposure and children’s intellectual function in Araihazar, Bangladesh,” *Environmental Health Perspectives*, *112*(13): 1329.

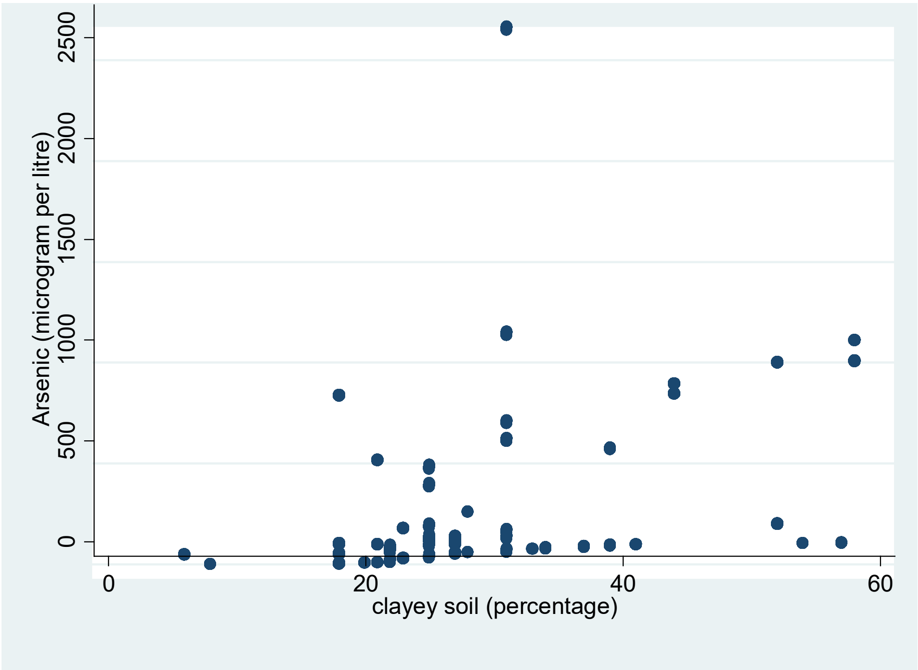
Wasserman, G. A. et al. (2007). “Water arsenic exposure and intellectual function in 6-year-old children in Araihazar, Bangladesh,” *Environmental Health Perspectives*, *115*(2).

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.

World Health Organization. (2009). Boron in drinking-water: Background document for development of WHO Guidelines for Drinking-water Quality (No. WHO/HSE/WSH/09.01/2). Geneva: World Health Organization.

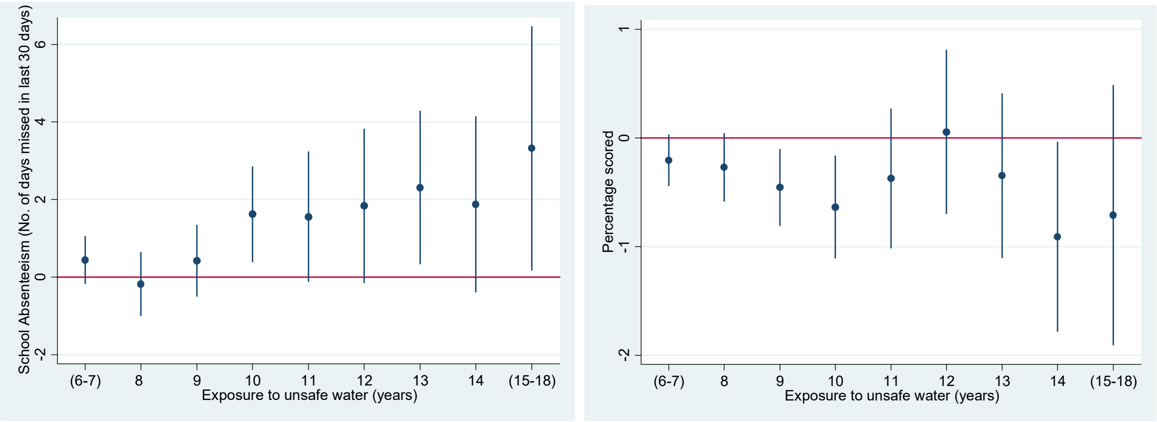
WHO, U. (2015). WHO/UNICEF joint monitoring programme for water supply and sanitation. Estimates on the use of water sources and sanitation facilities.

Figure 1: Relationship between Arsenic (microgram per litre) and percentage of clayey soil



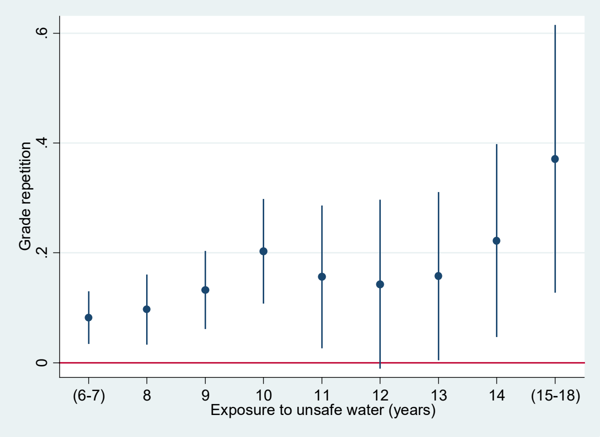
Source: Data on arsenic is taken from report of Central Ground Water Board under Ministry of Water Resources, River development and Ganga Rejuvenation, Gov. of India. Data on soil comes from the Harmonized world soil database.

Figure 2: Years of Exposure to Unsafe Water and Educational Outcomes[23]



---

[23] Due to small sample sizes, years 6-7 and 15-18 have been grouped as shown on the X-axis.



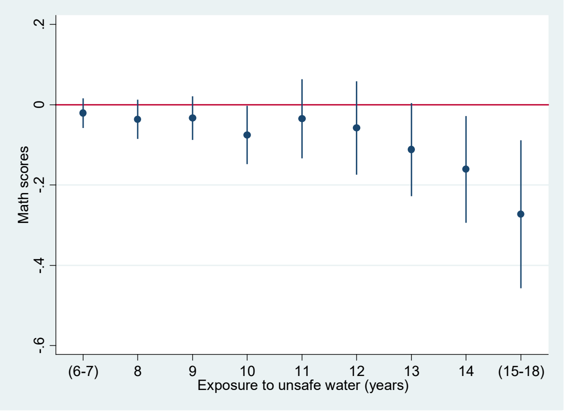


Table 1: IHDS Descriptive Statistics and District Level Control Variables

Variable	Mean	Std. Dev.
No. of days absent in last month	4.4	5.71
Arsenic (micrograms/L)	101.21	244.79
Percentage? Clayey soil	28.58	9.23
Family background characteristics		
Parental education in years	5.55	4.8
Hindu	0.75	0.43
Scheduled Caste/Scheduled Tribe	0.42	0.49
Other Backward Caste	0.33	0.47
Individual Characteristics		
Age	11.82	3.52
% Male	0.53	0.5
District level control variables		
Female per 1000 males	929.96	44.82
Iron contamination	1.61	2.23
Rainfall	81.94	54.21
Ratio rice to wheat (Million tonnes)	1,299.98	6,000.55
Urbanization Rate	28	8.2
Monthly per capita expenditure in Rupees	192,426.2	88,873.42
District level Education controls		
Gross enrolment in government schools	321,410.7	264,568
Gross enrolment in private schools	132,694.9	130,063.3
Ratio of teachers to schools (government)	4.39	1.24
Ratio of teachers to schools (private)	7.63	3.42

Sample size is N=12,941. Iron is measured in milligrams per litre. Expenditure refers to monthly per capita expenditure in a district

Table 2: Descriptive Statistics in the School Survey (Assam)

	Mean	Std. Dev.	N
Outcome Variables			
No. of days absent in last 30 days	3.4	4.69	3,500
Unfriendly atmosphere at school (=1 if yes)	0.08	0.27	3,500
Ever repeated a grade (=1 if yes)	0.15	0.36	3,500
% Scored in previous grade			
20 to 40 %	0.1	0.3	3,387
41to70 %	0.58	0.49	3,387
>71 %	0.31	0.46	3,387
Literacy			
None of questions answered correctly	0.37	0.01	948
One questions answered correctly	0.39	0.01	989
Two questions answered correctly	0.21	0.01	529
All questions answered correctly	0.03	0.01	85
Numeracy			
None of the questions answered correctly	0.17	0.01	594
One questions answered correctly	0.33	0.01	1,161
Two questions answered correctly	0.35	0.01	1,214
All questions answered correctly	0.15	0.01	529
Main Explanatory Variable			
Safe drinking water at home	0.68	0.47	3,500
Years exposed to unsafe water	6.68	2.97	3,500

Table 3: Descriptive Statistics for School Survey: Control Variables

Control variables	Mean	Std. Dev.	N
Individual characteristics			
Age	11.37	2.2	3,500
Male	0.49	0.5	3,500
Parental characteristics			
Religion (Hindu)	0.89	0.31	3,499
Solid house structure[24]	0.27	0.44	3,500
Caste			
General/Brahmins	0.18	0.01	594
Other Backward caste	0.67	0.01	2,228
Scheduled caste/Scheduled tribe	0.16	0.01	518
Assets			
Land/house	0.98	0.13	3,500
Durable Assets	0.87	0.33	3,500
Heavy vehicles	0.92	0.27	3,500
Mothers' education			
Illiterate	0.46	0.01	1,611
Primary	0.14	0.01	484
Secondary	0.33	0.01	1,158
University	0.07	0.01	247

Table 4: First Stage Regression in IDHS

	Arsenic (microgram/litre)
% Clayey soil	6.159***
	(0.213)
State F.E.	Yes
F-statistic	844.15
Observations	12,941

Robust SE clustered by PSU (village/neighborhood/town level). Significant at ***1%,**5%,*10%. Independent variable is percentage of clayey soil in a district. Model includes individual characteristics, all district level control variables shown in Table 1.

---

[24] Brick, concrete, stone, timber and cement.

Table 5: OLS and IV Estimates of Arsenic on Monthly School Absenteeism by Gender

	OLS Full Sample (1)	IV Full Sample (2)	IV Full Sample (3)	IV Females (4)	IV Males (5)
Arsenic	0.001	0.004**	0.006**	0.002	0.011**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
Individual characteristics	Yes	Yes	Yes	Yes	Yes
Parental characteristics	Yes	Yes	Yes	Yes	Yes
Education factors	Yes	Yes	Yes	Yes	Yes
Geographical factors	Yes	No	Yes	Yes	Yes
Economic factors	Yes	No	Yes	Yes	Yes
Observations	12,941	14,035	12,941	6,090	6,851

Notes: SE clustered by PSU (village/neighborhood/town level). *** Significant at 1%, ** significant at 5%, * significant at 10%. Absenteeism is measured in number of days. Regression includes state fixed effects and district level controls for rainfall, sex ratio, pattern of cultivation, iron, urbanisation, gross enrolment, per capital consumption expenditure. Other individual and family related controls are age, gender, caste, parental education.

Table 6: IV Estimates of Arsenic on Absenteeism in 6-11 year olds

	Females Only (6-11 yrs)	Males Only (6-11 yrs)
Arsenic	0.003	0.014**
	(0.003)	(0.007)
Observations	2,844	3,061

Notes: SE clustered by PSU level (village/neighborhood/town level). *** Significant at 1%, ** significant at 5%, * significant at 10%. Absenteeism is measured in number of days. Columns 2 to 5 report IV estimates. Regression includes state fixed effects and district level controls for rainfall, sex ratio, pattern of cultivation, iron, urbanisation, gross enrolment, per capital consumption expenditure. Other individual and family related controls are age, gender, caste, parental education.

Table 7: Exposure to Contaminated Drinking Water and Educational Outcomes

Educational outcomes	(1) Full sample	(2) Females	(3) Males
School absenteeism	0.212*	0.171	0.248
	(0.111)	(0.179)	(0.191)
Observations	3,341	1,678	1,663
Grade Repetition	0.036***	0.046***	0.030**
	(0.009)	(0.016)	(0.012)
Observations	3,341	1,678	1,663
CGPA scored in previous grade	-0.104***	-0.115**	0.058
	(0.031)	(0.046)	(0.051)
Observations	3,232	1,620	1,612
Unfriendly atmosphere	0.003	0.005	-0.000
	(0.005)	(0.006)	(0.006)
Observations	3,341	1,663	1,678
Grade F.E.	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes
Habitation F.E.	Yes	Yes	Yes

Robust standard errors clustered at the habitation level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All control variables, as shown in Table 3, are included in the regressions.

Table 8: Effect of Years of Contaminated Water Exposure on Test Scores

	Full sample	Female	Male
	Math scores (Panel A)		
Exposure to unsafe water (years)	-0.009*	-0.020**	-0.001
	(0.005)	(0.009)	(0.006)
	(0.036)	(0.049)	(0.048)
Observations	3,339	1,677	1,662
	Verbal scores (Panel B)		
Exposure to unsafe water (years)	0.000	0.010	-0.002
	(0.006)	(0.012)	(0.009)
Observations	2,447	1,239	1,208
Grade F.E.	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes
Habitation F.E.	Yes	Yes	Yes

Standard errors clustered at the habitation level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All control variables, as shown in Table 3, are included in the regressions.

Table 9: Arsenic & child health outcomes (2011 Census): Infant mortality and child mortality (OLS Estimates)

	(1)	(2)	(3)	(4)
Child health outcomes	**IMR_G /IMR_B**	**IMR_G /IMR_B**	**CMR_G /CMR_B**	**CMR_G /CMR_B**
Arsenic	0.0333*	0.0360*	0.0381**	0.0416**
	(0.0172)	(0.0190)	(0.01)	(0.0189)
Additional controls	No	Yes	No	Yes
Observations	147	132	148	134

Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for rainfall, literacy, urbanization, pattern of cultivation, iron in groundwater and gross domestic product. IMR (Infant Mortality Rate) is defined as deaths of children under one year of age per 1000 live births. CMR (Child Mortality Rate) is defined as deaths of children below five year of age per 1000 live births. The subindex B and G denotes boys and girls, respectively. Arsenic is a binary variable where 1 indicates regions where groundwater is contaminated by arsenic (irrespective of the concentration level) and 0 for non-arsenic regions. The sample includes only arsenic affected states and comprises of arsenic and non-arsenic affected districts within those states.

Appendix

A.1 Additional Data Sources

Control variables

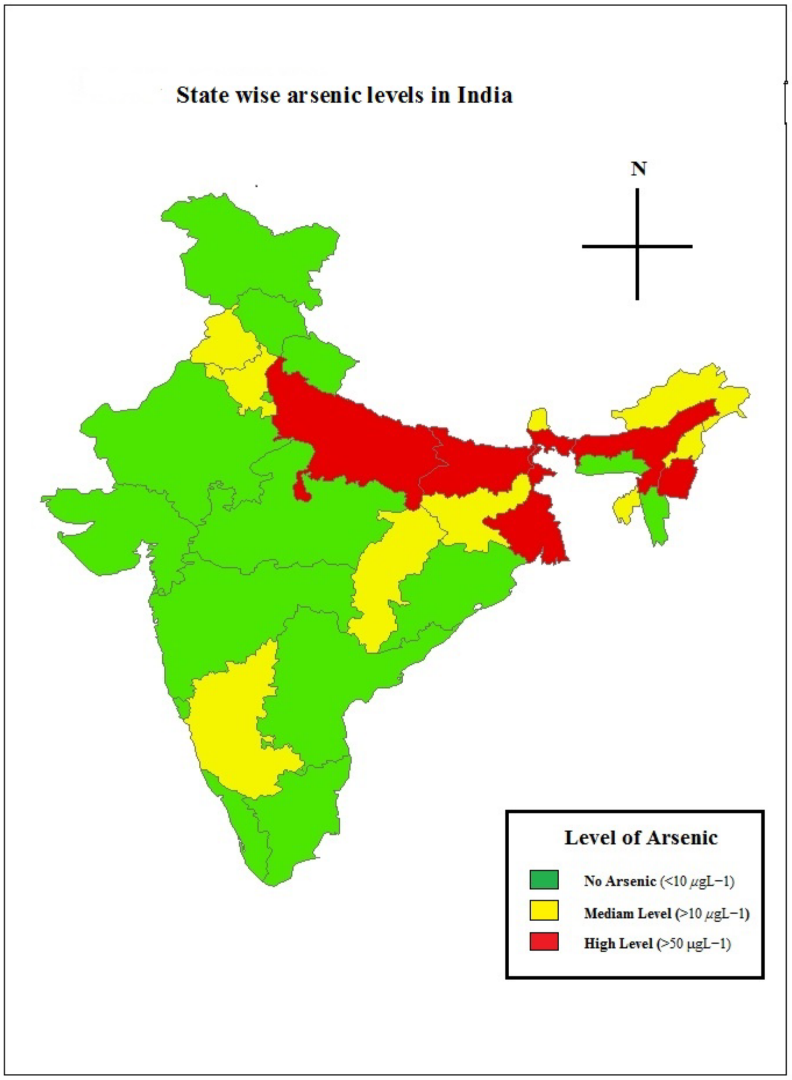
Data on total area under rice and wheat production (in millions tonnes) is obtained from the Ministry of Agriculture and Farmer's Welfare. Data for rainfall is provided by the Indian Meteorological Department (IMD) at district level in India. We take the 5 year average rainfall in millimetres. District level sex ratio and urbanization data is acquired from the 2011 Census of India. We also control for district level education factors such as gross enrolment and ratio of number of teachers to number of schools in government as well as private schools. Data for these variables is provided by District Information System for Education (DISE) under National University of Educational Planning and Administration, Government of India. To control for district level gross domestic product we also use data on Monthly Per Capita Expenditure (MPCE) from 68th round of NSSO as a proxy for district level GDP.

The data for level of arsenic, fluoride and iron in groundwater is provided by Central Ground Water Board. Following the WHO guidelines, the Bureau of Indian Standards (BIS) has notified a standard of 50 μgL-1 (microgram per litre) for arsenic in drinking water. The level of arsenic in groundwater is aggregated at the district level from block level data. 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 14,073 school going children across 13 arsenic affected states and 160 districts, where 41 districts are arsenic affected and 119 are non-arsenic affected districts. As shown in Table 1, the average level of arsenic is 100ug/l across districts in India, remarkably higher than the threshold limit.

The data on soil texture is generated 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. The average percentage of clayey soil across districts is approximately 28 percent. Among the district level characteristics, sex ratio is unbalanced at 929 females to every 1,000 males. These figures are reported in Table 1.

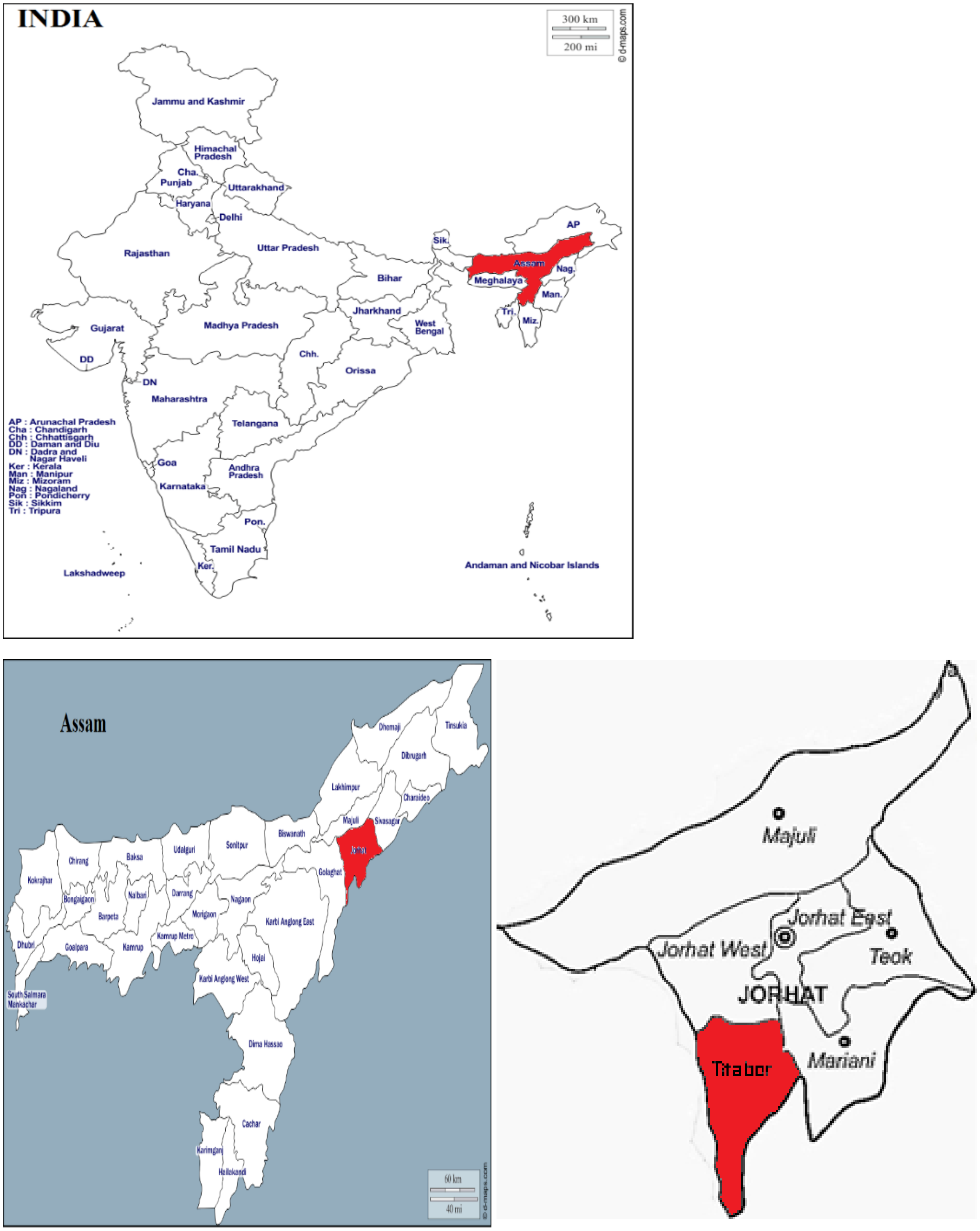
A.2 Figures and Maps

Figure A.2.1: Geographical Distribution of Arsenic Levels Across States of India

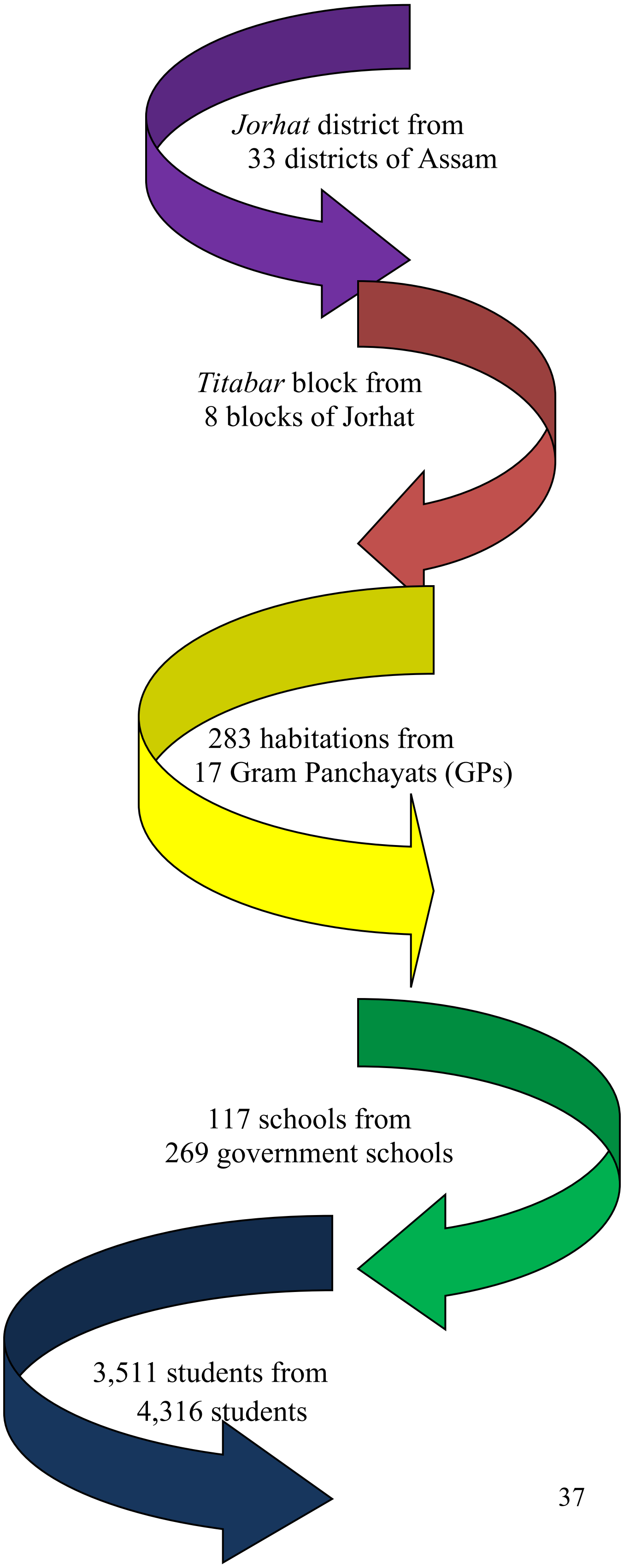


Source: Authors calculation using Central Ground Water Board report data (2016)

Figure A.2.2: Geographical location of Titabor Block in Jorhat District of Assam



A.2.3 Sampling Cascade: Geographical Distribution of Surveyed Sample



- Among the 13 arsenic affected states of India, Assam is one of the most severely impacted.

- Out of the 27 districts in Assam, Jorhat has the largest amount of contaminated habitations, namely 815 out of total 963.

- Amongst the 8 arsenic affected blocks of Jorhat, government schools in Titabor were surveyed. This selection of block was based on implementation of water supply project (SCADA) initiated by PHED, Gov. of Assam.

- Out of 269 government lower primary, primary, upper primary, higher secondary, and senior higher secondary, 117 schools were surveyed comprising of 3rd, 5th and 8th grade students. School selection criteria were based on the total number of student's enrolled in each grade being at least 10.

- As per the administrative data, 4,316 students were enrolled across three grades in surveyed schools. We were able to survey 3,511 students.

A.3 Additional Tables

Table A.3.1: Clayey soil and District Level Characteristics

	Clayey soil	N
Other Contaminants:		
Iron (mg/litre)	0.001	152
	(0.024)	
Fluoride (mg/litre)	0.370	149
	(0.544)	
Nitrate (mg/litre)	-0.000	149
	(0.008)	
Weather:		
Rainfall (millimetres)	-0.846	142
	(0.549)	
Maximum temperature (Celsius)	-0.107	42
	(0.201)	
Minimum temperature (Celsius)	0.180	36
	(0.306)	
Education:		
Public Student Enrolment	978.94	152
	(1810.39)	
Private Student Enrolment	1336.66	152
	(1124.50)	
Teacher/School (govt)	0.021**	152
	(0.009)	
Teacher/School (pvt)	0.028	143
	(0.021)	
Demographic& Economic Factors:		
Ratio Rice to Wheat (million tonnes)	-47.22	
	(82.96)	
Nitrogen (Kilogram/hectare)	0.000	139
	(0.000)	
Phosphorus (Kilogram/hectare)	0.000*	139
	(0.000)	
Potassium (Kilogram/hectare)	-0.000	139
	(0.000)	
Sex Ratio (per 1000 females)	-0.250	
	(0.365)	
Per Capital Expend.	1133.72	
	(765.59)	

*** Significant at 1%, ** 5%, *10%. Table reports the coefficient on clayey soil, from a regression of district level variables on the % of clayey soils in a district with state fixed effects.

In Table A.3.2, the response variable is an index that is calculated using the year reported for supply of safe water in a habitation. For instance, if a habitation got access to safe water in year 2009 then it is given value 1. Similarly, if the year reported is 2018 the variable will take value nine. Results show that all but one of the school quality measures are not correlated with the timing of water supply. Note that, since the right hand side school quality measures are aggregated at the habitation level and there are several habitations in the sample that do not have any schools, the sample size shrinks drastically in these regressions.

Table A.3.2: Correlation Between Year of Water Supply and Aggregate School Quality Measures

Aggregate School Quality Measures	*Years of safe water supply in habitation*
Teaching Experience (years)	0.000
	(0.014)
No. of teachers	0.025
	(0.031)
Class size	-0.005
	(0.010)
Playground availability	-0.007**
	(0.003)
Electricity Access (hours)	0.063
	(0.091)
Toilet	-0.011
	(0.263)
Library	0.002
	(0.004)
Proportion of Scheduled Caste	0.351
	(0.329)
Population	0.001*
	(0.000)
Observations	79

Note: Standard errors in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). The dependent variable is an index for the year in which a habitation got access to safe water ranging from 1 to 9 years.

There has been no large scale public awareness campaign in India about the adverse effects of drinking arsenic contaminated groundwater. Thus, it is not surprising that according to the 2011 decennial census, over 50 percent of households in Assam use groundwater sources for drinking purposes. Due to the lack of awareness and resources, we do not expect any endogenous migration into habitations where safe water was becoming available. Yet, to ensure that village level characteristics are uncorrelated with the timing of construction of water supply schemes, in Table A.3.3 we regress the index of year of construction on 2011 census village characteristics such as per capita population, literacy rate, percentage of scheduled caste and scheduled tribes and the population of marginal workers i.e. those workers who have not worked for the past 6 months. Reassuringly, all coefficients are statistically insignificant and close to zero.

Table A.3.3: Correlation Between Years Since Water Supply and Village Level Characteristics

Census Village Characteristics	*Years*
Population/Households	0.002
	(0.258)
Children Population (0-6 years)	0.001
	(0.002)
Scheduled Caste	0.000
	(0.001)
Scheduled Tribe	0.000
	(0.000)
Literacy	-0.003
	(0.007)
Total Workers	-0.001
	(0.001)
Marginal Worker	0.000
	(0.001)
Observations	158

Note: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The dependent variable is an index for the year in which a habitation got access to safe water.

Table A.3.4: IV Estimates of the Effect of Arsenic on Hours Spent Per Week Doing Homework (IHDS)

	(1) Girls	(2) Boys
Arsenic	0.001	0.003
	(0.004)	(0.005)
Observations	6,079	6,856

Note: Robust SE clustered by PSU level (village/neighbourhood/town). *** Significant at 1%, ** significant at 5%, * significant at 10%. Regression includes state fixed effects and district level controls for rainfall, sex ratio, pattern of cultivation, iron, urbanisation, gross enrolment, per capital consumption expenditure. Other individual and family related controls are age, gender, caste, parental education.

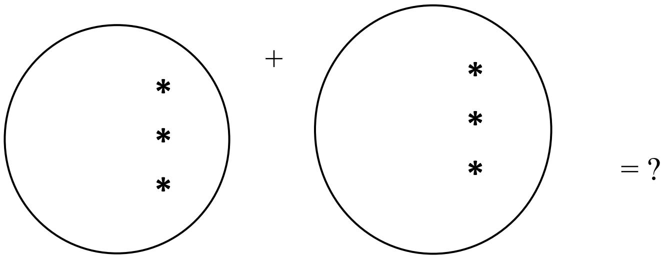
A.4 Titabor Survey Cognitive Skill Questions

Mathematical Ability of third grade students

Q1. Addition: 45+38=?
1) 73 2) 38 3) 90 4) 83

Q2. Multiplication: 30*2=?
1) 60 2)50 3)55 4)70

Q3. Counting



1) 5 2) 8 3) 6 4) 7

Verbal ability of fifth grade students

Read the passage below and answer the questions that follow:

She had some sweets that she wouldn't share,
She had a book that she wouldn't lend,
She wouldn't let anyone play with her doll,
She is nobody's friend!

He had some toffee and ate every bit
He had a tricycle he wouldn't lend,
He never let anyone play with his train,
He is nobody's friend!

But I will share all of my sweets with you,
My ball and my books and my games I will lend,
Here's half my apple and half my cake
\- *I am your friend!*
ENID BLYTON

Q1. What are the things that the girl does not want to share?

01) Book, doll and sweets 02) Balloons
03) Bicycle 04) None of the above

Q2. What are the things that the boy does not want to share?
01) Books 02) Tricycle, toffee and train
03) Balloons 04) None of the above

Q3. What does the child in the last paragraph want to share?
01) Sweets, apple and cake 02) Ball and Books
02) 03) Games 04) All of the above

Mathematical Ability of fifth grade students

Q1. Addition: 7010 + 2699=?
01) 9799 02)9699 03)9709 04)9609

Q2. What is the difference between 500.2 and 499.101?
01) 1.099 02)1.990 03)1.109 04)1.101

Q3. The digits 3 and 4 of the number 354 are inter-changed to form a new number. What is the difference between the new number and the original number?
01) 199 02)101 03)109 04)99

Verbal ability of 8th grade students

Analyse the picture given below and answer the following questions below:

Yellow Peas Dal
A nutritional way to a healthy diet at a reasonable cost

Only at Rs. 26/- per kg !

Benefits of Yellow Peas Dal:
- High Nutritional Value
- Facilitates lowering of cholesterol
- Rich in Iron
- Contains high fiber / roughage ... and many many more such benefits

Available in Select
Kendriya Bhandar , NAFED & NCCF Stores in Delhi
from 8th November, 2009.

For any help / clarification,
feel free to call
National Consumer Help Line 1800-11-4000
(Toll free : Monday-Saturday 9.30 am to 5.30 pm)
011-27662955-58 - Normal call charges apply

Issued in Public Interest by:

Ministry of Consumer Affairs, Food and Public Distribution
Department of Consumer Affairs
Government of India
Krishi Bhawan, New Delhi- 110 001 Website : www.fcamin.nic.in

Q1. The advertisement is about the importance of which of the following?

01) Benefits of nutritious food   02) Benefits of yellow peas Dal
03) Reasonably priced food   04) All of the above

Q2. Who has issued this advertisement?
01) Kendriya Bhandar
02) Ministry of Consumer Affairs, Food and Public Distribution
03) Central Food Technology Research Institute
04) Mother Dairy.

Q3. What is the meaning of the word "Nutrition"?
01) Process of purchase of costly food
02) Food rich in unhealthy fats
03) Process of providing or obtaining food for health and growth
04) Process of making handmade food

<u>Mathematical Ability of eighth grade students</u>

Q1. Three exterior angles of a quadrilateral are 70 degree, 80 degree and 100 degree. The fourth exterior angle is:
1) 70 degree   2) 80 degree   3) 100 degree   4) 110 degree

Q2. If $(x + 8) = 15$ then the value of x is:
1)10   2) 11   3) 7   4) 15

Q3. If $(2x + 1)/(x + 3)$ =1 then the value of x is:
1) 2   2) 3/2   3) 1   4) -1