EFFECT OF ARSENIC POISONING ON EDUCATION OUTCOMES: EVIDENCE FROM SELECT INDIAN DISTRICTS

Thesis submitted to Jawaharlal Nehru University for award of degree of

DOCTOR OF PHILOSOPHY

KHUSHBOO AGGARWAL

Centre for International Trade and Development

School of International Studies

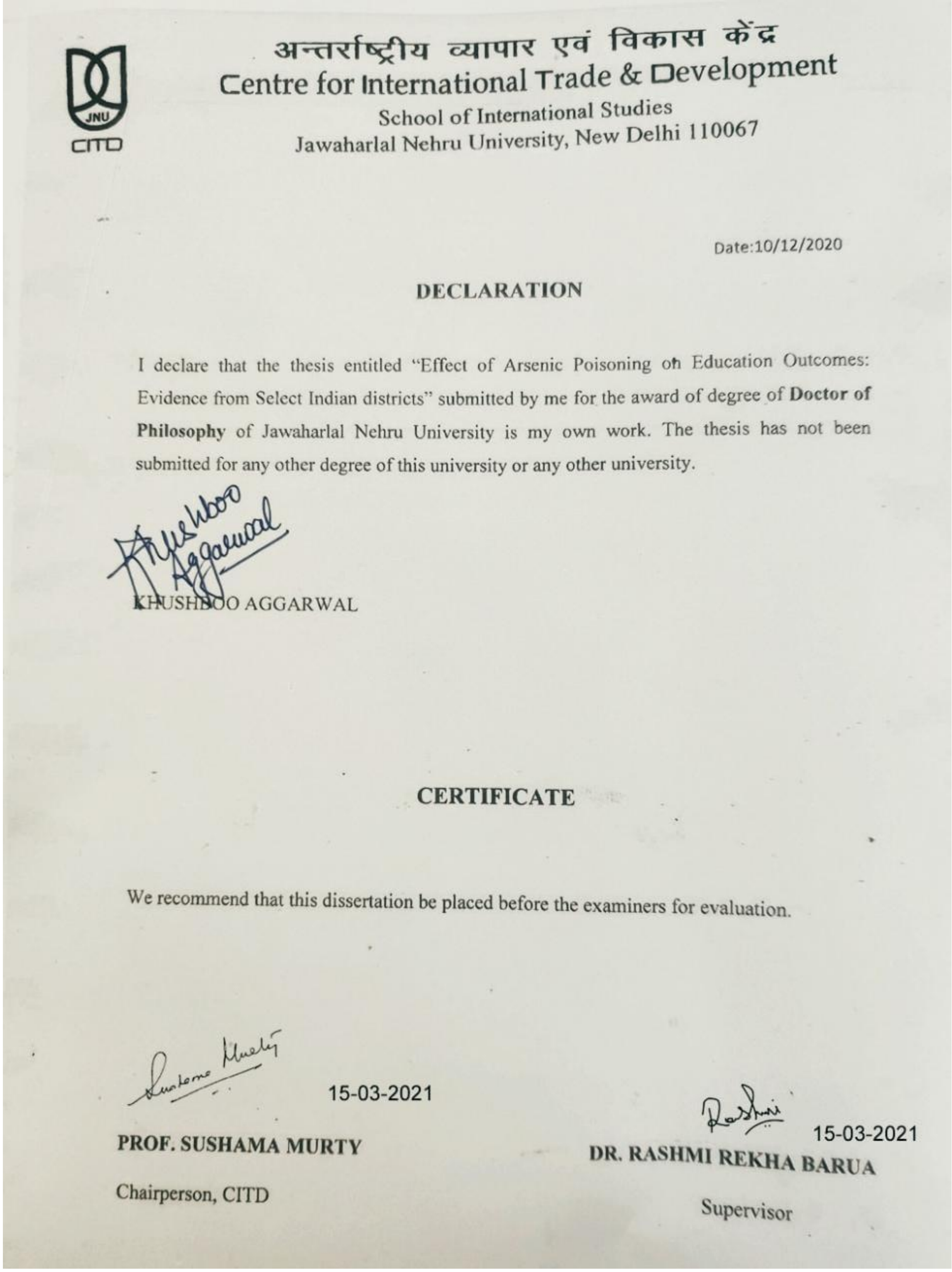
JAWAHARLAL NEHRU UNIVERSITY

New Delhi 110067

2021

To my loving parents, parent-in-laws, husband,
brothers, sister-in law and cute little Ahoo.

Thank you for your immense support
and encouragement throughout this
amazing and extraordinary journey.



अन्तर्राष्ट्रीय व्यापार एवं विकास केंद्र
Centre for International Trade & Development
School of International Studies
Jawaharlal Nehru University, New Delhi 110067

Date:10/12/2020

DECLARATION

I declare that the thesis entitled "Effect of Arsenic Poisoning on Education Outcomes: Evidence from Select Indian districts" submitted by me for the award of degree of **Doctor of Philosophy** of Jawaharlal Nehru University is my own work. The thesis has not been submitted for any other degree of this university or any other university.

Khushboo Aggarwal
KHUSHBOO AGGARWAL

CERTIFICATE

We recommend that this dissertation be placed before the examiners for evaluation.

Sushama Murty 15-03-2021
PROF. SUSHAMA MURTY
Chairperson, CITD

Rashmi 15-03-2021
DR. RASHMI REKHA BARUA
Supervisor

ACKNOWLEDGEMENT

I am really thankful and grateful to all those who have helped me and guide me throughout this journey. Firstly, I would like to express my gratitude to my supervisor, ***Dr. Rashmi Rekha Barua,*** *not just for her deep, qualitative and insightful suggestions but also for guiding me at all phases in my life. Without her constant support, encouragement and cooperation all this would not have been possible. She has helped me in developing research skills and boosted my interest in the field of research particularly in health, development and education economics.*

I would also like to thank Prof. Sushama Murty, Prof. Meeta Keswani Mehra, Prof. Aparna Sawhney, Prof. Sangeeta Bansal, Prof. Amit Shovon Ray and Associate Prof. Dr. Mandira Sarma, Dr. Brishtu Guha, Dr. Priya Bhagowalia at the Centre for International Trade and Development (CITD), Jawaharlal Nehru University for their immense support and encouragement.

I gratefully acknowledge the support and guidance of Dr. Marian Vidal Fernandez, University of Sydney. Her comments and suggestions were highly constructive in shaping the course of the study.

I would like to extend my earnest thanks to the organizers and participants at the conferences at Indian Statistical Institute (ISI), Delhi, Jawaharlal Nehru University, Calcutta University, Flame

University and CECFEE 2019 conference in Tezpur, for their valuable comments.

A word of thanks is also due to the non-teaching staff for their co-operative and helpful attitude and guiding me through the administrative formalities.

Finally, I would like to mention the name of some people without whom my dream of becoming a researcher would not have been possible. My mother, Mrs. Mithlesh Aggarwal and my husband, Nitin Kumar Mittal for their constant encouragement, positive attitude and support. Their presence has helped me to overcome all difficulties and problems in my life. I think, more than me it's their dream which has turned into reality.

I would like to extend my gratitude to my friends and other family members, Apoorva Gupta, Sultana Khan, Priyanka, Rajashree, Mukhlisa, Nazia, Imrana, Prathna Aggarwal, Yashaswini, Yashobanta Parida, Malvika. Finally, this chapter has come to an end but a new chapter of new beginning will start soon and I am really excited to apply all my knowledge and skills that I have learned in this long and enjoyable doctorate journey at the Centre for International Trade and Development (CITD), Jawaharlal Nehru University.

Khushboo Aggarwal
PhD Student
CITD/JNU

Table of Contents

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

WHO	World Health Organization
CGWB	Central Ground Water Board
UNICEF	United Nation Children’s Fund
IHDS	India Human Development Survey
BIS	Bureau of Indian Standards
CGPA	Cumulative Grade Point Average
NAS	National Achievement Survey
NCERT	National Council of Educational Research and Training
NRDWP	National Rural Drinking Water Programme
SBM	Swachh Bharat Mission
PHED	Public Health Engineering Department
NFHS	National Family Health Survey
HAZ	Height-for-Age (z scores)
WAZ	Weight-for-Age (z scores)
GDP	Gross Domestic Product
IV	Instrumental Variable
OLS	Ordinary Least Square
DD	Differences-in-Differences
ATE	Average Treatment Effect
WASH	Water, Sanitation and Hygiene
SC	Scheduled Caste
ST	Scheduled Tribes
IMR	Infant Mortality Rate
CMR	Child Mortality Rate
PSU	Primary Sampling Units
IMD	Indian Meteorological Department
DISE	District Information System for Education
MPCE	Monthly Per Capita Expenditure

HWSD	Harmonised World Soil Database
FAO	Food and Agricultural Organisation
GIS	Geographic Information System
IIASA	International Institute for Applied System Analysis
DHS	Demographic and Health Survey
MoHFW	Ministry of Health and Family Welfare
GoI	Government of India
NSSO	National Sample Survey Office

Chapter-1

Introduction

1.1 Water contamination, Health & Education: Overview

Across the globe, over 2,000 children below the age of five years die every day mainly due to water related diseases. According to UNICEF (2013), 90 percent of these deaths are primarily due to lack of accessibility to safe drinking water. India alone accounts for 24 percent of world's total child mortality (WHO, 2015). Inadequate availability of safe drinking water is a critical issue, particularly in underdeveloped and developing countries. Changing climatic condition might decrease percolation rate of rainwater in underground sources which could result in increasing concentration of hazardous toxins in groundwater (Mc Arthur et al., 2001). According to estimates by World Health Organization (WHO 2009), lack of accessibility to safe drinking water is the major cause of morbidity in India.

Early childhood health outcomes are an important indicator of a country's economic situation, including educational attainment and earnings. At the same time, exposure to air and water pollution also affect children's well-being. Several authors have assessed the negative impact of contaminated water on infant and child health (Currie et al., 2013). In India, Do et al. (2018) show that the curtailment of industrial pollution in the River Ganges led to lower incidence of infant mortality. Another study by Brainerd and Menon (2014) studied the impact of harmful chemicals released in water via fertilisers on children's health. They find that exposure to fertilisers during pregnancy has a negative impact on child health outcomes.

Environmental degradation might have severe and adverse health implications, particularly amongst children due to their lower immunity and developing bodies. This in turn might have unfavorable impact on school participation, learning abilities and other educational outcomes of children, especially the younger ones. There exists extensive literature that deals with analysing the association between health and education (Grossman, 2008; Cutler and Lleras-Muney, 2006). Fowler et al. (1992) used a sample of children aged 6-17 years in the U.S. and they found that children with asthmatic conditions have higher rates of school absenteeism relative to their peers without such conditions. In the context of developing countries, evidence shows that drastic reduction in school absenteeism is realised for children getting healthier, like through deworming in rural Kenya (Miguel and Kremer, 2004) or through provision of mid-day meal (Dreze and Kingdon, 2001) or iron tablets in India (Bobonis, Miguel & Sharma, 2006).

1.2 Arsenic and Health

Several regions in India have witnessed massive exploitation of groundwater for agriculture and drinking purposes that has led to increasing concerns of groundwater contamination. Due to the presence of hazardous toxins in groundwater such as arsenic, iron, fluoride or nitrate, common preventive measures like merely boiling of water does not make it potable (Central Ground Water Board, 2006). This study mainly focused on arsenic and its adverse health and economic outcomes (particularly education) among children.

Arsenic poisoning poses a serious threat to more than 140 million people worldwide (UNICEF, 2013). Presently India and Bangladesh together comprise of largest arsenic affected population in the world. In India, over 70 million people spanning over 40 districts are facing threat of presence of arsenic beyond threshold limit which are issued as per the guidelines prescribe by WHO (Khurana and Sen, 2008). Children constitutes the highest share of arsenic affected victims and exposure to arsenic is likely to be significant contributor to increasing child mortality in India (39 deaths per 1000 live births, (Assadullah and Chaudhary, 2011). Prolonged exposure to arsenic could result in adverse health outcomes.

Arsenicosis or arsenic poisoning relates to failure of heart and kidney, cancer, mental impairment, physical illness, skin lesions and still birth (Tseng, 2007). If arsenic is consumed regularly through drinking or consuming food cooked with contaminated water, it can cause chronic illness. Unlike adults, children are more susceptible to arsenic-related illness due to their less developed immunity system and comparatively greater percentage of body water (Saing and Cannonier, 2017). Heath implications of arsenicosis are well documented in literature but only a handful of studies have tried to analyse the impact of arsenic exposure on academic performance and other education outcomes among children (Hassan et al., 2005; Murray and Sharmin, 2015; Carson et al., 2010).

A myriad of observational studies has found 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 to be overestimated.

While arsenic occurs naturally in soil and water, over-exploitation of groundwater increases its 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 industrialisation. This geographical relationship also challenges the identification of arsenic's impact on educational development.

This study fills a gap in the literature by employing a new dataset and empirical strategy to analyse the effects of groundwater arsenic exposure on health outcomes and educational status of children. It therefore contributes to relatively thin literature on effects of water environmental contaminants on education and health of children in developing countries.[1] Additionally, it also explores the channels through which drinking arsenic-contaminated water may affect education outcomes.

---

[1] While several 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 context of developing countries, the evidence is scarce. Bharadwaj et al. (2017) compare sibling outcomes to find that exposure to elevated levels of air pollution in the womb has negative impact on mathematics and language skills among fourth grade students in Santiago, Chile.

1.3 Objectives of the Study

This thesis aims to explore the various emerging concerns related to children's health and education and its relation to accessibility to safe drinking water and adequate provision of nutrition. The three broad objectives of this study are listed below.

1.3.1 Impact of arsenic on education

The **first** objective is to analyse the impact of arsenic-contaminated water on school absenteeism of children using a large, nationally representative household survey data set (India Human Development Survey, 2011-12).

Estimating the effects of arsenic on educational outcomes using regional variation in arsenic levels is challenging. Relying on regional variation in groundwater arsenic levels 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 as naturally occurring arsenic dissolves out of rock formations when groundwater levels drop significantly (Madajewicz et al., 2007).

This study uses household survey data from 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 checking arsenic levels in groundwater and for measuring 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 (Madajewicz et al., 2007). Instrumental Variable (IV) estimates indicate that, on an average, being exposed to an additional microgram per litre of arsenic in groundwater increases school absenteeism by 0.13 percentage points amongst school going children. This is a conservative estimate because absenteeism information is only available for children enrolled in school. Finding of this study 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.

1.3.2 School level survey in Jorhat, Assam

Second objective is to substantiate the secondary data by comparing it with self-collected, school-level dataset in Jorhat, one of India's severely arsenic affected districts in Assam. In Jorhat, majority of habitation has face threat imposed by high level of arsenic concentration in groundwater that ranges between 194 and 491 microgram per litre, this holds to be far greater than the safe threshold limit suggested by BIS (Bureau of Indian Standards) in line with guidelines prescribed by WHO. Amongst all the arsenic affected blocks in Jorhat, Titabor block has the maximum population suffering from drinking contaminated water.

Sample survey and survey description

The survey was conducted among all children of 3rd, 5th and 8th grades studying in 117 primary and secondary schools, and distributed across 283 habitations in Titabor. This newly collected student and school data was merged with the school administrative data.

The survey comprised of principal/teacher-in-charge questionnaire and a student questionnaire. The latter captured 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 questionnaire also provides specific information on child's awareness of water contaminants, and the primary source of drinking water at home. 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 were: days missed school in the past 30 days, Cumulative Grade Point Average (CGPA) of 2017, and if the child 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.[2]

Administrative framework in Titabor

In 2008, two central government sponsored flagship programmes, namely the National Rural Drinking Water Programme (NRDWP) and the Swachh Bharat Mission (SBM) were initiated by the Assam Public Health Engineering Department (PHED) to tackle the high arsenic levels found in drinking water. Both of these government-initiated schemes aim to provide access to safe surface water sources from its neighboring rivers: Doyang and Dhansiri. We used administrative data on population by habitation, year in which safe piped water scheme was started, number of piped versus unsafe groundwater schemes, start and completion date of the project. To create the main explanatory variable, we used data on year reported which is defined as the first year in which a particular pipe water scheme was reported to be started in a habitation. Under this objective, the identification strategy relied on NRDWP and SBM that provide time and cohort variation in access to safe, piped water.

Differences-in-differences estimates indicate that more number of years of exposure to contaminated drinking water is associated with lower CGPA in the previous grade attended and higher rates of grade repetition. In addition, exposure to contaminated water leads to higher rates of absenteeism and lower numeracy scores. Consumption of unsafe water for an additional year might increase 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%).

---

[2] 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, as the pilot revealed rather poor English language abilities among younger children. Our outcome variables are the fraction of questions answered correctly by the student. Appendix 2A.3.5 lists the grade and subject-specific questions that were administered.

While both our survey and the IHDS-II results found no gender differences in school absenteeism, we found 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. We explore this further in the last chapter of the thesis.

1.3.3 National Family Health Survey-4: Arsenic, gender bias and undernutrition

The final objective of thesis is to assess the effect of exposure to arsenic-contaminated groundwater on child's growth (measured in terms of stunting and underweight) in India. We further explore the role played by nutritional deficiencies across gender in determining the health impact of exposure to arsenic-contaminated water.

Stunting amongst children has adverse implications on their health and overall development. Existing evidence on stunting clearly indicates that stunted children often fall sick, have lower learning abilities, under perform in school and thus, have reduced future earnings (Glewwe and Miguel, 2008; Barker et al., 1993; Case and Paxson, 2008). Another indicator of child's heath is low weight as indicated by Weight-for-Age (z scores). It deals with body mass of an individual relative to his/her chronological age. Poor nutritional intake during early childhood leads to stunting and low weight among children. Several studies in India shows that height and weight disadvantage among children increase with steeper birth order[3] gradient, particularly amongst girls. Formerly, the adverse implication of consuming unsafe water should not differ by gender. But due to biased preference of parents towards male child, girls are more susceptible to environmental hazards. Poor nutritional intake and inadequate health care received by girls make them more vulnerable to environmental pollutants, leading to lower immunity and under-developed bodies. Despite wide prevalence of gender discrimination practiced in developing countries like India, exiting studies make no attempt to analyse the role played by gender bias in the relation between child health and accessibility to safe water. This

---

[3] It refers to the order in which a child is born in the family, like first born, second born and so on.

study contributes to the existing research by exploring the link between gender, environmental pollutants and child growth measures. To the best of our knowledge, ours is the first study to analyse the role of gender in the relation between environmental pollutants and child health outcomes.

The empirical framework in this analysis exploits the geographical variation in the levels of arsenic in groundwater and we have tried to estimate the relation between arsenic and health outcomes (stunting and underweight) amongst children under the age of five years. But our analysis could lead to biased estimation by merely relying on geographical variation in arsenic since there exists a correlation between the presence of arsenic in groundwater in a region and economic activity in that region.[4] To overcome the issue of identification in our estimation, we relied on instrumental variable framework. To assess the impact of arsenicosis on stunting and underweight, we used variation in fraction of clayey soil textures across districts within a state to study the arsenic levels in groundwater. This analysis is based on secondary data source from National Family Health Survey-4 (NFHS-4, 2015-16).

Our findings suggest that arsenic exposure has adverse health implication for girls less than five years of age. A second aim of this study is to analyse if the adverse effects amongst girls are mainly due to an underlying nutritional disadvantage (Jayachandra and Kuziemko, 2011). Thus, we study the effect of birth order on the association between arsenic and health outcomes. Major findings of this study show that increase in arsenic by an additional unit (microgram per liter) leads to a significant reduction in height-for-age (stunting) and weight-for-age (underweight) amongst second born and third born (or higher) girls relative to male child born at first birth order. All our findings are robust to accounting for state fixed effects and vector of district level controls including weather, water quality measures, pattern of cultivation, education quality and economic growth (district GDP). Hence, our findings contribute to existing literature by showing that the adverse health implication of

---

[4] For instance, agriculturally dominant regions in India have higher levels of arsenic contamination in groundwater. This is primarily due to overexploitation of groundwater, as arsenic which occurs naturally dissolves out of rock formation when groundwater level drops considerably (Madajewicz et al., 2007).

arsenicosis amongst girls might mainly arise due to lower nutritional intake and poor health care, including shorter periods of breastfeeding. The effect holds for districts with poor sex ratio where the phenomenon of son preference is more predominant (Gupta, 1987).

1.4 Order of Chapters

The remaining chapters of the thesis are structured as follows:

Chapter 2 titled "Still Waters Run Deep: Arsenic Contamination of Groundwater and Education Outcomes in India" studies the effect of drinking arsenic-contaminated water on education outcomes of children in arsenic affected regions of India. This chapter is classified into *two parts* - the *first part* deals with examining the effect of arsenic contamination on school absenteeism of children enrolled in primary and secondary schools; and the *second part* aims to complement secondary analysis results with a newly collected survey of students in public schools located in the state of Assam – one of the worst arsenic affected regions of India. Exploiting variation in the geographical coverage and timing of government water supply schemes allows us to identify the effect of arsenic on student outcomes. We find that children with prolonged exposure to contaminated water sources experience higher rates of absenteeism, grade retention and lower test scores. We find that the detrimental effects are more among girls, likely due to an already underlying health disadvantage in childhood.

Chapter 3 on "Gender Disparities in the Prevalence of Under-nutrition in India: The Unexplored Effects of Drinking Contaminated Water" aims to evaluate water pollution as the cause of stunting and underweight in arsenic contaminated regions of India. Using NFHS data and by exploiting the variation in soil textures across districts as an instrument for arsenic, we show that exposure to arsenic beyond the safe limit is inversely associated with height-for-age (stunting) and weight-for-age (underweight). Negative effects are more for girls who are born at higher birth orders relative to the elder ones. Within India, the analysis shows that the number of underweight girls is higher in regions with high preference for sons as indicated by low sex ratios. The effects are also heterogenous by place of residence (urban/rural) and religion. We argue that this is because nutritional deficiencies, including shorter periods of breastfeeding, exacerbate the adverse impact of environmental exposure to arsenic on health outcome.

The last chapter 4 is "Summary of Findings and Policy Implications". It collates a brief summary of findings and conclusions of all the chapters of this research work. It also focuses on policy implications so as to seek governmental intervention to prevent arsenic-related water contamination.

Chapter-2

Still Waters Run Deep: Groundwater Arsenic Contamination and Education Outcomes in India

2.1. Introduction

Across the world, over 2,000 children below the age of five years die every day mainly due to water related diseases. Inadequate availability of safe drinking water is a critical issue, particularly in underdeveloped and developing countries. Changing climatic condition might decrease percolation rate of rainwater in underground sources which could result in increasing concentration of hazardous toxins in groundwater (Mc Arthur et al., 2001). Arsenic is one amongst such contaminants that poses a severe threat to quality of drinking water in various emerging economies worldwide.

Presently India and Bangladesh together comprise of largest arsenic affected population in the world. In India, over 70 million people spanning over 40 districts are facing threat of presence of arsenic beyond threshold limit which are issued as per the guidelines prescribe by WHO (Khurana and Sen, 2008). Children constitutes the highest share of arsenic affected victims and exposure to arsenic is likely to be significant contributor to increasing child mortality in India which stands to be at 39 deaths per 1000 live births (Assadullah and Chaudhary, 2011). Prolonged exposure to arsenic could result in adverse health outcomes.

While the effect of drinking contaminated water on a range of health outcomes is well documented,[5] less is known about its causal effects on education outcomes of children. This study 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.

Drinking arsenic-contaminated water may affect education outcomes through several channels. First, prolonged exposure to arsenic has been linked to cognitive delays. Such

---

[5] See for instance, Del Razo et al. (2011) and Tseng (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, absenteeism may increase further due to social ostracization since visible symptoms of arsenic poisoning (such as skin lesions, hyperkeratosis and melanosis) are often misinterpreted as contagious (Hassan et al., 2005). Additionally, children are likely to miss school if they need to support or care for household members also suffering from arsenic induced diseases (Carson et. al., 2010).

Estimating the effects of arsenic on educational outcomes, using either the regional variation in arsenic levels or the variation in a household's access to safe water, is challenging. Relying on regional variation in groundwater arsenic levels 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 since arsenic that occurs naturally dissolves out of rock formations during significant drop in level of groundwater (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 education and the availability of safe water at home. In this chapter, we employ two empirical strategies and databases to overcome both identification challenges.

First, we use 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 (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.

Second, we substantiate the study by collecting a new and rich school-level dataset in one of the most arsenic affected districts of India (Jorhat) in the state of Assam. The identification strategy relies on two government programs, NRDWP and SBM that provide time and cohort variation in access to safe, piped water.[6]

Differences-in-differences estimates indicate that more number of years of exposure to contaminated drinking water 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 water leads to higher rates of absenteeism and lower numeracy scores. Exposure to unsafe water for an additional year might lead to increase in 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%).

While both our survey and the IHDS-II results find no gender differences in school absenteeism, we 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.[7]

We make important 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 are unique and able to identify the causal effect of arsenic exposure at the regional level as well as at

---

[6] These programs were implemented by PHED of Assam in 2008-2009. The PHED is the main government agency responsible for water supply.

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

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 study makes a significant contribution to the literature on environmental health and education outcomes.

The remainder of this chapter is structured as follows. Section 2.2 reviews the related literature. In Section 2.3 we introduce our two empirical strategies for each dataset. Section 2.4 describes our two data sources: the IHDS-II data and our newly conducted survey in the Indian state of Assam. In Section 2.5 we report the main results of the study. Concluding remarks and policy implications of our analysis are presented in Section 2.6.

2.2. 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 or consuming food cooked with contaminated water. Arsenic poisoning can also lead to failure of heart and kidney, skin lesion, still births and cancer (Tseng 2007). Children are more susceptible to arsenic due to their growing immunity system and greater percentage of water in body as compared to adults (Saing and Cannonier 2017). Several epidemiological evidences finds that arsenic could adversely affect fetus as it could pass through placenta and might lead to low birth weight of newly born children (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.[8]

---

[8] 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.

A myriad of observational studies finds 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 study 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. [9]

2.3. Data and Methods

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

2.3.1.1 IHDS (India Human Development Survey –II, 2011-12).

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 neighborhoods 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.[10]

Our main outcome variable, school absenteeism, is defined as the number of days a child was absent in the previous month. There are approximately 12,941 children in our sample between the age of 5-19 enrolled in primary and secondary schools. Table 2.1 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.

---

[9] 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.

[10] Apart, from these datasets, ASER (Annual Status of Education Report) is an annual survey that focuses primarily on learning outcomes and schooling status of children for rural districts in India. Unfortunately, it only provides information on rural areas. Because arsenic levels increase with anthropogenic activities, urban areas would be more prone to arsenic contamination and hence the use of this dataset is not appropriate.

2.3.1.2 Instrumental Variable Approach

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. Finer soils have relatively more particle density and lower porosity levels and as a result, their permeability level is relatively lower than loamy soil[11] which facilitates arsenic concentration in groundwater (Mac Arthur et al. 2001; Madajewicz et al. 2007). Figure 2.1 provides visual evidence supporting a positive association between arsenic (measured in microgram per litre) and percentage of clayey soil in a district.

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 $\mathbf{Y_{ids}}$ 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 millimeters (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 affect education outcomes only through the impact on the level of arsenic in groundwater.[12] A threat to identification is

---

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

[12] 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. Use of such

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 million tonnes) in all the regressions. We also control for district level male to female ratio in all regressions since soil texture can affect economic outcomes through relative female to male employment rates (Carranza,2014).

There might be further threats to our identification strategy if clayey soil is correlated with other geographic or demographic characteristics and such features 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.[13] There is a significant difference by soil permeability in the ratio of teachers to schools in government schools. However, because districts with higher proportion of clayey soil (and thus more arsenic) have more teachers per government school, 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 identification strategy.

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

2.3.2.1. 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). Recent estimates by Census (2011) show that merely 9.2 percent of population in Assam has accessibility to clean drinking water which is minimally low as compared to nation's average (32 percent) as more than half of its population still use groundwater (tube-well and handpumps) to meet their drinking requirements

Out of the 27 districts in Assam, Jorhat has the largest number of contaminated habitations, namely 815 out of a total of 963 (Singh, 2004).[14] In Jorhat, majority of the habitation has

---

pesticides might alter some properties of superficial soil (upper most layer of soil), but not the subterranean soil used in our analysis.

[13] Results are reported in Table 2A.3.1. in the Appendix.

[14] 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

marked with wide range of arsenic in groundwater (194 to 491 micrograms per litre), which exceeds the limit suggested by BIS (50 microgram per litre) in line with norms given 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 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.[15]

To tackle the rise in arsenic levels in drinking water, in 2008, two central government sponsored flagship programs were initiated by PHED (Gov. of Assam) in Jorhat namely, NRDWP and SBM. These schemes aim to provide accessibility to clean drinking water from its two neighboring rivers: Doyang and Dhansiri. The availability of clean water is tracked by clusters of households within a village, 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.[16] The scheme provided two safe water access options. Households could pay a private connection or alternatively, access a more affordable shared community level connection.

Water committees also known as *Panee Samitee* were formed which mainly comprises of 15 members each representing their corresponding habitation so as to facilitate hurdle free functioning of such scheme. In order to get access to piped water connection, each

---

(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.

[15] 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.

[16] 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 2A.2.2 in the Appendix shows the geographical location of Titabor.

household was required to sign an agreement with head/president of the committee. In addition, beneficiary households have to pay set-up fee of INR 1,000 (approximately USD 14) plus INR 100 monthly for connection charges.[17] Informal conversations with PHED officials revealed that the majority of households in Titabor chose communal connection that was relatively cheaper than private connection as the cost is collectively borne by all community members[18] in case of community connection. Pipe water is provided on hourly basis (1-2 hours) both in case of private as well as community connection.

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[19], 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.

2.3.3.2. 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 2A.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 children present and in 3rd, 5th, and 8th grades 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

---

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

[18] Unfortunately, there is no information available about who or how many chose this option.

[19] 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.

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%.[20]

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 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.[21] 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 NCERT. NAS is a district-level survey that aims to assess the learning outcomes of approximately 220,000 students across 700 districts in India.[22]

---

[20] 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.

[21] 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. Similarly, due to lack of appropriate number of filters/sand filters at school, students drink groundwater. Hence, to study the effect of consuming unsafe water on children's academic achievements, we focus on child's source of drinking water at home, conditional on school fixed effects.

[22] 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

Table 2.2 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 2.3 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. Cross tabulation in the Appendix confirms that the source of drinking water is related to household characteristics (Table 2A.3.2). Children with access to safer source of drinking water are significantly more advantaged than their counterparts as evident from the t-test of differences in means.

2.3.2.3. Differences-in-Differences Approach

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

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

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

---

questions answered correctly by the student. Appendix 2A.3.5 lists the grade and subject specific questions that were administered.

individual and family and household background characteristics. G is a grade 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 2.5.4.1, we conduct an cohort 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. Note that the identification strategy exploits variation in the *timing of construction of piped water supply* and the *age of the child*, both of which are arguably exogenous. Yet, to address the concern that treated habitations may be endogenously chosen, we include a treatment dummy (T) in the regressions that is equal to 1 if the habitation has access to safe water and 0 if there is no clean water supplied to the habitation.

2.4. Results

2.4.1 IV Results from the IHDS

Table 2.4, 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 844 supporting soil texture as a strong instrument for arsenic levels.

Table 2.5 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 significant but with small coefficients, the IV estimates suggest that arsenic has a positive and statistically significant (at 5% level) effect on school absenteeism.[23] 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 one microgram per litre increase in arsenic in a district leads to 0.006 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 0.14 percent increase in number of school days missed in the previous month.

2.4.2 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 2.5, columns 4 and 5 we find larger impact of arsenic on school absenteeism for boys than girls. The IV estimates indicate that, a one microgram per litre increase in arsenic in groundwater in a district leads to 0.01 days increase in absenteeism among boys (or 0.23 percent) and a statistically insignificant effect for girls.

A concern with these estimates is that school absenteeism is conditional on enrolment. Therefore, the results could be biased due to differential enrolment by gender. Children who are often sick, due to long term exposure to arsenic, will be less likely to be enrolled in school. Similarly, older children 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

---

[23] IHDS also captures a proxy of numeracy among of 8-11 year old children enrolled in primary schools. Scores are classified into four categories (0=unable to recognize numbers, 1=basic arithmetic, 2= can do subtraction but not division, and 3=able to divide). The same questions are administered to only 8-11 year olds regardless of their age/grade. Arguably due to measurement error and relatively small sample size, do not find any effect of arsenic on this outcome. Results from ordered probit and IV probit models are shown in Appendix Table 2A.3.3.

likely to get treatment and hence more likely to drop out of school. Thus, differential enrolment of boys and girls may bias our results.

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). Table 2.6 confirms that the gender gap in absenteeism is not due to differential enrolment rates as the results do not change when we restrict the sample to a younger age group with high enrolment rates.

2.4.3 Assam School Survey Results

2.4.3.1 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:

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

Where, $\boldsymbol{D}_{\mathbf{ij}}$ 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 2.2. $Treat_h$ is a dummy variable equal to 1 if a habitation has access to safe drinking water. Figure 2.2 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. In Table 2.7, 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.

Unlike in Bangladesh, there has been no large-scale public awareness campaign in India about the adverse effects of drinking arsenic contaminated groundwater. Therefore, 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 2.8 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.[24] Reassuringly, all coefficients are statistically insignificant and close to zero.

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.**[25]** 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 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.

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[26].

2.4.3.2 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 2.9 and 2.10. We see that in Table 2.9 that a one-year increase in exposure to unsafe water is associated with

---

[24] Marginal workers are those workers who had not worked for major part of the reference period i.e., for less than 6 months (Census 2011).

[25] 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.**

[26] Results available upon request.

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 2.10 shows results for numeracy and verbal scores. 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. Looking at 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.

These results reflect that more informed or educated water committee members and/or parents may lobby for early access to clean water. Though all models control for asset ownership and parental education, as an additional robustness, we further include a control variable 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. As shown in Table 2.11, our results do not change with this additional control variable.[27]

2.4.3.3 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

---

[27] We also explore if the effect of exposure is larger among disadvantaged groups by including in the regressions interaction terms of exposure with family background characteristics. We find that the negative effect of additional year of exposure to contaminated water on school attendance is larger among children whose mothers have lower education and those belonging to Scheduled *Caste* (SCs) and Scheduled Tribes (STs). There is no heterogenous effect of family background characteristics on any other measures of education. Results available upon request.

impact that bullying can have on education (Gorman et al, 2019), visible skin lesions due to arsenic contamination could impact their test scores.

Second, these results could also be driven by gender differences in time use if, when a relative fall 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 girls in India are breastfed for relatively shorter duration as compared to boys and thus consume lesser quantity of 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 2.9. 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.[28] The IDHS records how many hours children spend doing homework per week during the last month. Table 2.12 shows IV regressions using soil texture as instrument that 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.

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. As per Census (2011), IMR is defined as deaths

---

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

of children under one year of age per 1000 live births while CMR is defined as deaths of children below five years of age per 1000 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 2.13 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 give further credibility to the hypothesis that the effect of arsenic is magnified among girls due to poorer health status starting early in life.[29]

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. 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.

2.5. 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 is steadily increasing the concentration of

---

[29] We also explore gender differences in resources being provided to children. In Table 2A.3.4 in the Appendix, we regress a dummy variable that equals one if a child goes to school in a bicycle, bus or scooter, and 0 if she walks to school on a gender dummy and all the individual and family background characteristics included in previous regressions as well as school and habitation fixed effects. We find that males are 5% more likely to use a bicycle or other vehicles to school while girls are more likely to walk to school. Though this does not show that there is gender discrimination in nutritional intake directly, it does point towards more resources being spent towards male children. We need, however, to be cautious in interpreting these estimates as causal since it could simply reflect preferences i.e., if girls prefer to walk to school while boys prefer to commute by bicycles or motor vehicles.

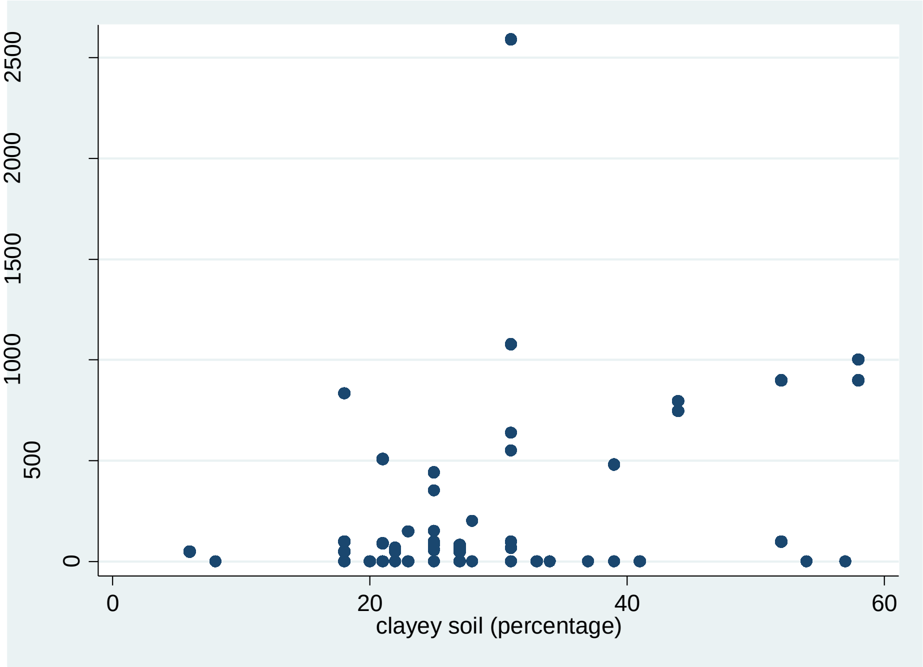
toxic metals in groundwater and thus increasingly adding to the costs. These figures do not include the costs associated with the loss in human capital and its intergenerational effects. Our analysis contributes to the literature measuring the effects of environmental health on human capital.

Combining data from a large nationally representative household survey with a primary survey conducted across schools in one of severely arsenic affected regions of India, we study the effect of arsenic contamination in groundwater on education outcomes. We find that exposure to arsenic contaminated water leads to large increase in school absenteeism and grade retention, and a decline in test scores and CGPA scored in the prior grade attended.

Our results show heterogenous effects of drinking arsenic contaminated water by gender and socio-economic status highlighting the importance of targeting water policy towards the most vulnerable groups. We find that the adverse effect of drinking contaminated water on numeracy and literacy scores is 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 among girls compared to boys.

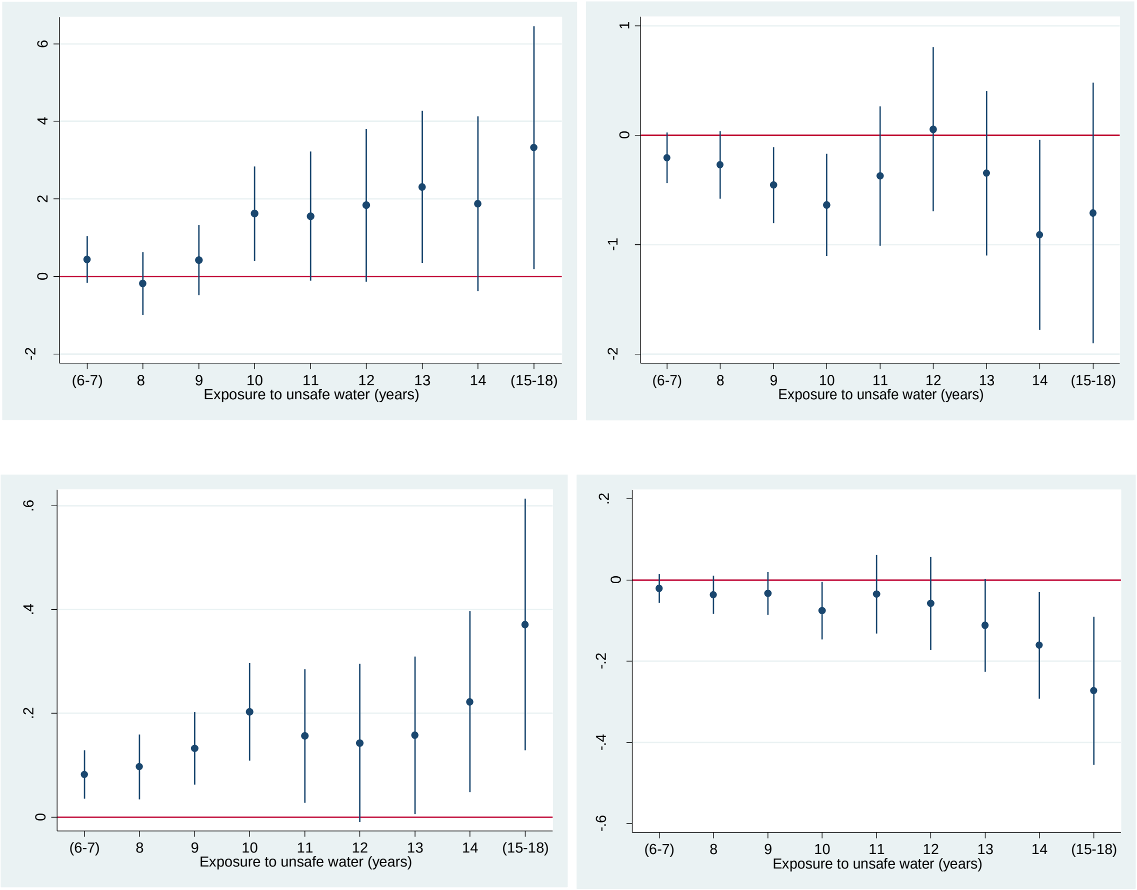
Mahanta et. al. (2016) estimates that an additional increase in arsenic concentration by one microgram per litre might impose an annual health cost of USD 0.01 million in Assam. By reducing the concentration of arsenic to safe limit might generate welfare gain of approximately USD 2.49 million. In addition to these 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.

Figure 2. 1:Relation between Arsenic (microgram per litre) and percentage of clayey soil type.



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

Figure 2. 2: Years of Exposure to Unsafe Water and Educational Outcomes[30]



---

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

Table2. 2: IHDS Descriptive Statistics and District Level Control Variables

Variable	Mean	Std. Dev.
No. of days absent in last month	4.4	5.71
Arsenic	101.21	244.79
Clayey soil	28.58	9.23
Family background characteristics		
Parental education (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/Male	929.96	44.82
Iron contamination	1.61	2.23
Rainfall	81.94	54.21
Ratio of rice to wheat (production)	1299.98	6000.55
Urbanization	28	8.2
Monthly per capita expenditure (rupees)	192426.2	88873.42
District level Education controls		
Gross enrolment in government schools	321410.7	264568
Gross enrolment in private schools	132694.9	130063.3
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 2.2: Descriptive Statistics for School Survey

	Mean	Std. Dev.	N
Outcome Variables			
No. of days absent in last 30 days	3.4	4.69	3500
Unfriendly atmosphere at school	0.08	0.27	3500
Ever repeated a grade	0.15	0.36	3500
Percentage scored in previous grade			
Percentage scored (20-40)	0.1	0.3	3387
Percentage scored (41-70)	0.58	0.49	3387
Percentage scored (>71)	0.31	0.46	3387
Literacy			
None of questions answered correctly	0.37	0.01	948
One question 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 question answered correctly	0.33	0.01	1161
Two questions answered correctly	0.35	0.01	1214
All questions answered correctly	0.15	0.01	529
Main Explanatory Variable			
Safe drinking water at home	0.68	0.47	3500
Years exposed to unsafe water	6.68	2.97	3500

Table 2.3: Descriptive Statistics for School Survey: Control Variables

Control variables	**Mean**	**Std. Dev.**	**N**
Individual characteristics			
Age	11.37	2.2	3500
Gender (Male)	0.49	0.5	3500
Parental characteristics			
Religion (Hindu)	0.89	0.31	3499
Structure of house (Pucca house[31])	0.27	0.44	3500
Caste			
General/Brahmins	0.18	0.01	594
Other Backward caste	0.67	0.01	2228
Scheduled caste/Scheduled tribe	0.16	0.01	518
Assets			
Land/house	0.98	0.13	3500
Durable Assets	0.87	0.33	3500
Heavy vehicles	0.92	0.27	3500
Mothers education			
Illiterate	0.46	0.01	1611
Primary	0.14	0.01	484
Secondary	0.33	0.01	1158
University	0.07	0.01	247

Table 2.4: 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. Regression includes individual characteristics, all district level control variables shown in Table 2.1 and state fixed effects.

---

[31] Pucca house refers to those dwellings that are made up of substantial materials (brick, concrete, stone, timber and cement).

Table 2.5: 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 2.6: IV Estimates of Arsenic on Absenteeism in 6-11 years old

	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 2.7: 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.

Table 2.8: 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 2.9: 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 2.10: 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 2.3, are included in the regressions.

Table 2.11: Exposure to Contaminated Drinking Water and Education Outcomes: Robustness Checks

Cognitive scores	(1) **Math**	(2) **Verbal**	(3) **Absenteeism**	(4) **Repetition**	(5) **CGPA**	(6) **Unfriendly atmosphere**
Exposure (Years)	-0.010**	0.001	0.211*	0.037***	-0.106***	0.003
	(0.005)	(0.006)	(0.112)	(0.009)	(0.032)	(0.005)
Observations	3,338	2,446	3,340	3,340	3,231	3,340
Grade F.E.	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Habitation F.E.	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors clustered at the habitation level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. This table controls for a binary variable that is equal to 1 if the child was informed about the presence of either arsenic or nitrate or fluoride in drinking water. All control variables, as shown in Table 2.3, are included in the regressions.

Table 2.12: 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/neighborhood/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.

Table 2.13: 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 years 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

2A.1 Additional Data Sources

Control variables

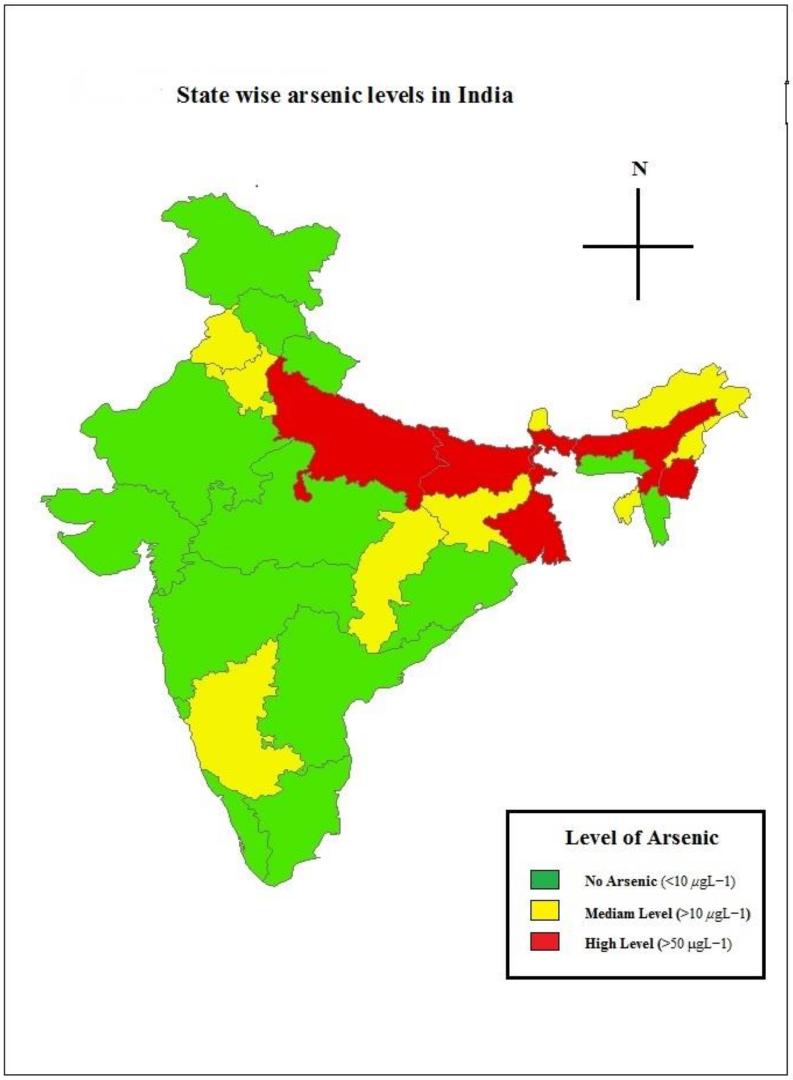
Data on total area under rice and wheat production (in million tonnes) is obtained from the Ministry of Agriculture and Farmer's Welfare. Data for rainfall is provided by the Indian Meteorological Department (IMD) at district level in India. We take the 5-year average rainfall in millimeters. District 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 1000 males. These figures are reported in Table 2.1.

2A.2 Figures and Maps

Figure 2A.2. 1:Geographical distribution of Arsenic across various States of India



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

Figure 2A.2. 2:Geographical location of Titabor Block in Jorhat District of Assam

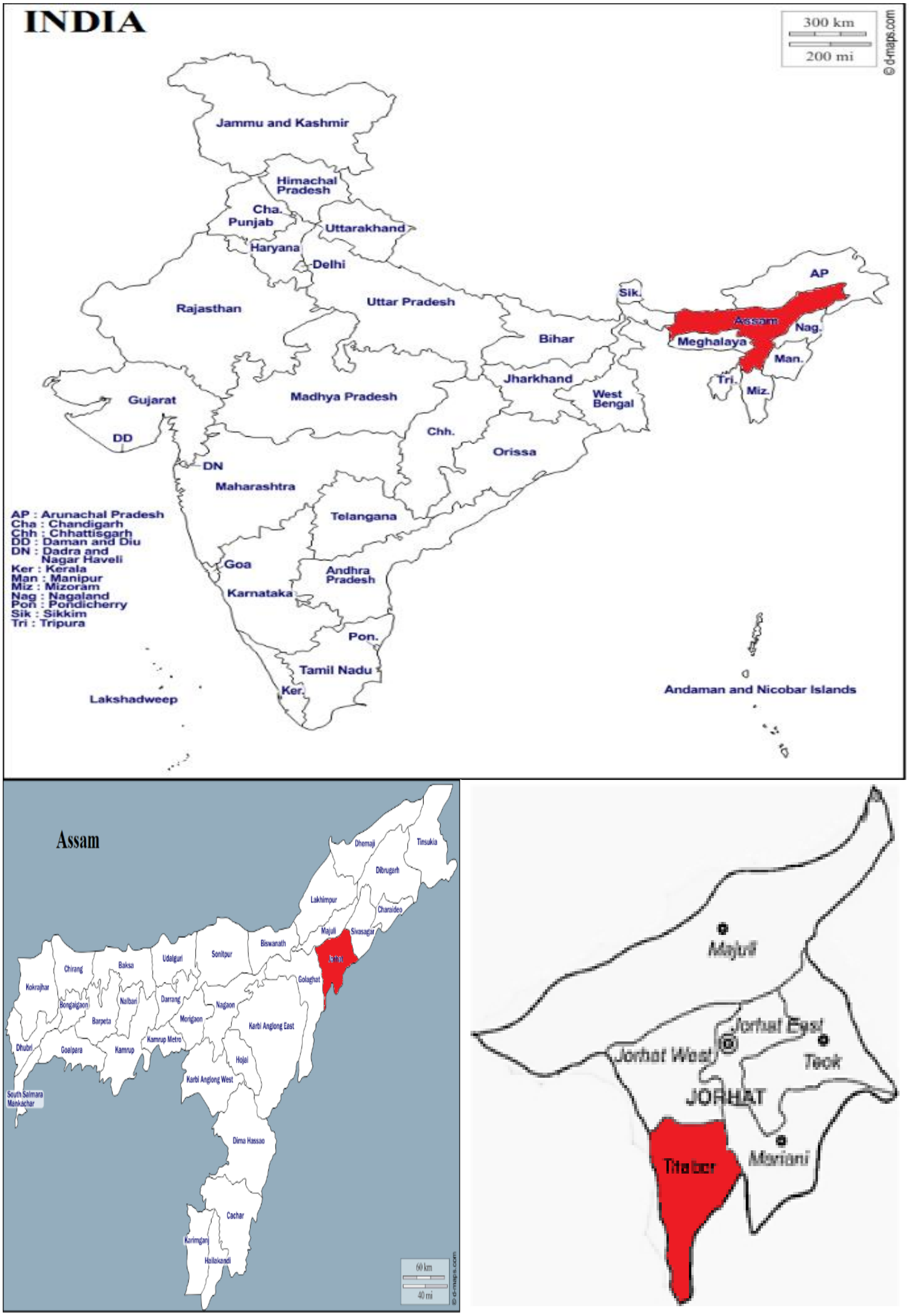
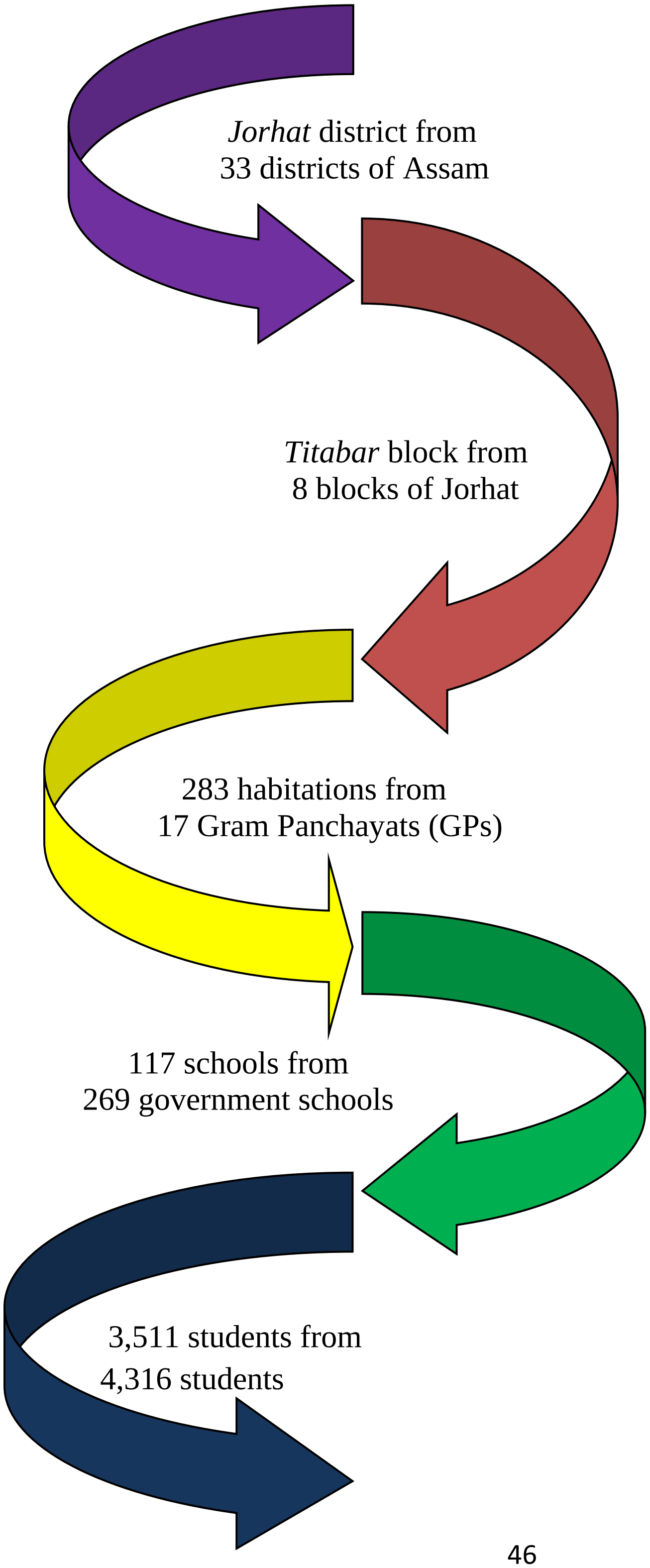


Figure 2A.2.3:Sampling Cascade: Geographical Distribution of Surveyed Sample



- 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.

2A.3 Additional Tables

Table 2A.3.1: Clayey soil and District Level Characteristics

	Clayey soil
Other Contaminants:	
Iron	0.001
	(0.024)
Fluoride	0.011
	(0.015)
Weather:	
Rainfall	-0.846
	(0.549)
Education:	
Public Student Enrolment	978.94
	(1810.39)
Private Student Enrolment	1336.66
	(1124.50)
Teacher/School (govt)	0.021**
	(0.009)
Teacher/School (pvt)	0.028
	(0.021)
Demographic & Economic Factors:	
Ratio of Rice to Wheat	-47.22
	(82.96)
Sex Ratio	-0.250
	(0.365)
Monthly Per Capita Expenditure (rupees)	1133.72
	(765.59)
State fixed effects	Yes

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

Table 2A.3.2: Household Characteristics By Source of Water: School Survey

Variable	Mean	N	Mean	N	t-stats
	Safe Sources		Unsafe Sources		
Pucca house	0.29	2384	0.22	1116	-4.38***
	-0.46		-0.42		
Owns a house or land	0.98	2384	0.99	1116	1.98**
	-0.14		-0.1		
Owns durable good/s	0.89	2384	0.84	1116	-4.88***
	-0.31		-0.37		
Owns heavy vehicle/s	0.94	2384	0.89	1116	-4.6***
	-0.24		-0.31		
Mother illiterate	0.42	2384	0.54	1116	6.32***
	-0.49		-0.5		
Mother primary ed.	0.12	2384	0.18	1116	4.5***
	-0.33		-0.38		
Mother secondary ed.	0.37	2384	0.25	1116	-6.92***
	-0.48		-0.43		
Mother higher ed.	0.09	2384	0.03	1116	-5.65***
	-0.28		-0.18		

Notes: A Pucca house is defined as the dwelling that is made up of substantial material such as stone, brick, cement, concrete, or timber.

Table 2A.3.3: Ordered Probit Estimates of Arsenic on Numeracy (Marginal effects)

Panel A: Ordered Probit	Numeracy=0	Numeracy=1	Numeracy=2	Numeracy=3
	(1)	(2)	(3)	(4)
Arsenic	0.0001**	0.0000**	-0.0000**	-0.0001**
	(0.000)	(0.000)	(0.000)	(0.000)
State F.E.	Yes	Yes	Yes	Yes
Observations	2,635	2,635	2,635	2,635
Panel A: IV Ordered Probit	**Numeracy=0**	**Numeracy=1**	**Numeracy=2**	**Numeracy=3**
	(1)	(2)	(3)	(4)
Arsenic	0.0000	0.0000	-0.0000	-0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
State F.E.	Yes	Yes	Yes	Yes
Observations	2,635	2,635	2,635	2,635

Notes: Robust standard errors clustered by state in parentheses. *** Significant at 1%, ** significant at 5%, * significant at 10%. Numeracy is a count variable where 0=unable to recognise numbers, 1=basic arithmetic, 2= can do subtraction but not division, and 3=able to divide. 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 2A.3.4: Gender Differences in Resources by Gender: School Survey

	(1) Mode of transport to school other than walking	(2) Arsenic awareness	(3) Iron awareness
Male	0.050***	-0.015	0.003
	(0.016)	(0.011)	(0.024)
Observations	3,341	3,341	3,341
Controls	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes
Habitation F.E.	Yes	Yes	Yes

Standard errors clustered at habitation level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The outcome variable in Column 1 is equal to one if the student goes to school on a bicycle, bus, or scooter.

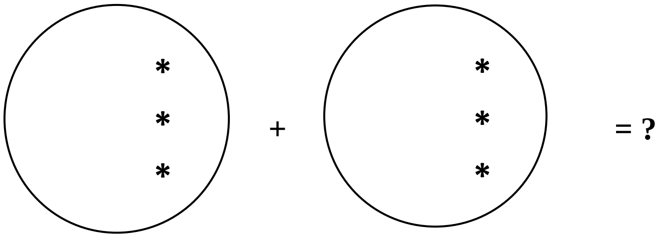
2A.3.5 Titabor Survey Cognitive Skill Questions[32]

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

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

---

[32] Note that surveys were administered in both Assamese and English, except the sections that deals with comprehensive part of cognitive test[sh]. Students were given the option of choosing their preferred language.

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
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

JAGO GRAHAK JAGO

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

Mathematical Ability of eighth grade students

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

Chapter-3

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

3.1. Introduction

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

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

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

In this analysis, we investigate the impact of exposure to arsenic contaminated groundwater on child health outcomes in India. Overconsumption of arsenic can lead to fatal health outcomes such as kidney and heart failure, mental illnesses, cancer, skin-related diseases, and adverse pregnancy outcomes.[33] Unlike adults, children are more exposed to threat imposed by arsenic owing to underdeveloped immunity and comparatively significant percentage of water in their bodies. Further, epidemiological evidence indicates that arsenic could adversely affect the health of fetus by passing through placenta and can lead to fata health outcomes in its later stages (Rahman et al. 2009; Kile et al. 2016).

While there is epidemiological evidence that arsenic affects child growth outcomes (Watanabe et al. 2007; Minamoto et al. 2005), we argue that in the presence of gender bias, girls may be more likely than boys to be adversely impacted by drinking arsenic contaminated water. This is because nutritional deficiencies, including shorter periods of breastfeeding, might exacerbate the adverse impact of environmental exposure to arsenic on health outcomes. While arsenic is known to readily cross the placenta, exclusive breastfeeding protects infants against arsenic (Fängstrom et al. 2008).[34]Thus, if girls are

---

[33] Arsenic poisoning, or Arsenicosis, is defined as state of chronic illness that results from consuming arsenic contaminated water for longer time interval.

[34] Consistent with this, following an arsenic awareness campaign in Bangladesh, Keskin, Shastry and Willis (2017) show that infants were breastfed for longer duration by their mothers and consequently they have lower mortality rates and lower incidence of diarrhea during their early childhood.

less likely to be breastfed or given adequate nutrition in childhood, the adverse health effects of arsenic exposure can be more severe among girls.

Using geographical variation in arsenic concentration in water, we estimate the association between arsenic levels and child health outcomes (stunting and underweight) in India. But relying on regional variation in groundwater arsenic levels is problematic due to the correlation between concentration levels of arsenic in groundwater and economic activity of region. For instance, agriculturally dominant regions in India have higher levels of arsenic contamination in groundwater. This is primarily due to overexploitation of groundwater, since naturally occurring arsenic dissolves out of rock formation when groundwater level drops significantly (Madajewicz et al. 2007). To overcome this identification challenge, we use an instrumental variable framework in our analysis. Our data is sourced from NFHS-4 (2015-16).

We use the variation in fraction of clayey soil textures across districts within a state to instrument for arsenic levels in groundwater to measure its impact on child health. Finer soils such as clay have relatively higher particle density and are less porous than coarse sandy soil which increases the concentration of contaminated water (Mc Arthur et al, 2001).

Instrumental variable estimates indicate that exposure to arsenic in groundwater has negative and significant impact on HAZ among children less than five years of age, regardless of gender. To test if the effects are larger among girls due to a nutritional disadvantage, following Jayachandra and Kuziemko (2011), we study the effect of birth order on the association between arsenic and health outcomes. We find that a one standard deviation increase in arsenic levels in groundwater leads to a reduction in HAZ by 0.035 (2.11 percent) and 0.060 (3.61 percent) standard deviations for the second born and third born girl child, respectively, relative to a male child born at first birth order.[35]These

---

[35] 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

findings are robust to the inclusion of district level controls for weather, water quality measures, pattern of cultivation, education quality and income. Further, we find similar negative effects of arsenic on weight-for-age (also called underweight) among later born girls as compared to the eldest son. These results hold consistent even after accounting for all district level controls.

The gender bias in nutrition and resource allocation has been extensively researched in India (Gupta 1987; Behrman 1988; Jayachandran and Kuziemko 2011; Jose 2011; Fledderjohann et al. 2014; Jayachandran and Pande, 2017; Pande 2019). Son preference in India can be explained by a combination of economic, religious and sociocultural factors such as patrilineality and patrilocality associated with the Hindu Kinship system (Dyson and Moore 1983). Moreover, inheritance rights are in favor of sons and religious rites in Hinduism, including death rituals, are conducted only by the male heir (Arnold, Choe and Roy 1998).

Consistent with this hypothesis, we find heterogenous effects of arsenic exposure by sex ratio, religion, caste and urban/rural status. In particular, the adverse effects among later born girls are relatively larger in districts with negatively skewed sex ratio where patrilineal Hindu kinship system is most likely followed (Gupta 1987). Adverse effects are also larger for Hindu households, compared to Muslim households, and households located in rural locations.

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

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

---

family with no boys. Parents with only daughters would be keen on having a son, irrespective of their desired family size. Hence, birth of late parity daughters' acts as a negative income shock and thus limited income will be spent on them.

the role of gender in the relation between environmental pollutants and child health outcomes.

The remainder of this chapter is structured as follows: Section 3.2 reviews the existing literature. In section 3.3 we provide detailed description of dataset followed by the empirical framework presented in section 3.4. In section 3.5 we report the primary findings of our study including heterogenous effects and robustness checks. Lastly, in section 3.6 we give concluding remarks and policy implications of our analysis.

3.2. Literature review

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

Jayachandran and Pande (2017) examine variation in provision of pre-natal and post-natal health inputs across birth order gradient. Their findings suggest that during pregnancy more prenatal inputs are allocated by parents in case of having no sons in family. Surprisingly, the authors find a reverse pattern for post-natal inputs such as vaccination and duration of breastfeeding, when the elder child is a girl. Jayachandra and Kuziemko (2011) show that mothers, with no sons or fewer sons, who want to conceive again would limit their breastfeeding duration for new born daughter. The authors argue that lower rate of breastfeeding for girls increases their vulnerability to water related contaminants and thus, in turn increases their mortality rate.

Our study is also related to the literature on the effect of environmental pollutants on health outcomes of children. Epidemiological studies have established that early-life environmental exposure plays a role in growth outcomes (Walker et. al. 2007). In economics, most studies have focused on the negative health outcomes of air pollution (Arceo-Gomez et al. 2012). Foster et al. (2009) evaluates the impact of clean industry certification program on pollution and consequentially on respiratory diseases among infants in Mexico. Goyal and Canning (2017) find negative impact of air pollution on in-utero health and other child growth indicators in Bangladesh.

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

3.3. Data and Data Source

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

To assess the impact of water pollution on child health, we use two measures of child health. First, we study Height-for-Age (HAZ) for children in the age group of 0 to 5 years. HAZ is a commonly used yardstick to measure stunting or nutritional status of children

(Deaton and Dreze 2009). It is a cumulative measure of nutritional dearth from birth or conception onwards and is the best aggregate measure of malnutrition among children that is correlated with outcomes at later stages of life. Stunting is linked to underdeveloped brains, lower retention and reduced learning ability that adversely affects productivity and earning capacity of an individual.

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

Figure 3.1 plots the HAZ scores by birth order among boys and girls. It is clear from the figure that HAZ among girls decreases with increasing birth order. In particular, the percentage of girls who are moderately or severely stunted increases with birth order[36]. For instance, at first birth order approximately 10 percent of girls suffers from severe stunting which at later birth order increases to 12 percent and 15 percent for 2nd and 3rd+ birth order, respectively. Similar pattern is visible for boys. Similarly, figure 3.2 shows that the percentage of girls with moderate to severe underweight increases with increasing birth order. Birth order effects on stunting and underweight reflects the poor nutritional status among girls and boys particularly at higher birth order.

The average HAZ and WAZ for our sample is -1.66 and -1.64, respectively. The NFHS data also includes a host of individual, household and family background characteristics. The summary statistics of the variables that are included in our analysis are shown in Table 3.1. The data is gender balanced with girls comprising 48 percent of the sample with an average age of 27 months. While 34 percent of the sample consists of children at first birth order, 29 percent are at second birth order and 37 percent of children are at higher than second birth order. 38 percent of mother’s are uneducated while 68 percent are educated

---

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

(14 percent-primary, 39 percent-secondary and 9 percent-higher and above). The average age of mothers in the sample is 27 years. More than three fourth of our sample comprises of rural households with 31 % scheduled caste (SC) and scheduled tribes (ST) and 50 % belong to other backward classes.

3.3.1 District level Control variables

Data on total production under rice and wheat production is obtained from Ministry of Agriculture and Farmer's Welfare. Data for rainfall is provided by the Indian Meteorological Department (IMD) at district level in India, with a mean value of 79.24 mms. District level sex ratio and literacy data is acquired from the 2011 Census of India. The average sex ratio and literacy rate in our estimation sample is 928 and 68 percent respectively. To control for district level gross domestic product, we also use data on Monthly Per Capita Expenditure (MPCE) from 68th round of NSSO (National Sample Survey Office) as a proxy for district level GDP.

Data for the level of arsenic and iron in groundwater is provided by the Central Ground Water Board. Following the WHO guidelines, the Bureau of Indian Standards (BIS) has notified a standard of 50 $\mu gL-1$ (microgram per liter) for arsenic in drinking water. The level of arsenic in groundwater is aggregated at the district level from block level data. The appendix[37] provides a map of arsenic affected regions of India. We restrict the analysis to only those states where the presence of arsenic is measured beyond the threshold limit in at least one district. The final dataset comprises of 73,160 children under the age of five, across 9 arsenic affected states and 261 districts, where 105 districts are arsenic affected and 156 are non-arsenic affected districts. As shown in Table 3.1, the average level of arsenic is 107.32 microgram per liter across districts in India, remarkably higher than the threshold limit. Apart from arsenic, mean level of iron is 1.6 mg/l as indicated in Table 1.

The data on soil texture is obtained from Harmonised World Soil Database (HWSD) which was established in July 2008 by the Food and Agricultural Organisation (FAO) and International Institute for Applied System Analysis (IIASA). HWSD is global soil database

---

[37] See appendix 3A.1.3

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 3.1, the average clayey soil across districts is approximately 28 percent.

3.4. Empirical Model

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

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

We are interested in measuring the effect of arsenic on two outcome variables: height-for-age and weight-for-height of child *i* in district *d* of state *s* as given in equation (1). The main explanatory variable is Ars_{ds} which indicates the concentration level of arsenic in groundwater in district *d* and state s. $\boldsymbol{X_{ids}}$represents vector of controls for individual level characteristics (gender, age and age square), mother characteristics (mother's education and age), family background characteristics and socio-economic characteristics (religion, caste, family size and place of residence). We also control for district level controls ($\boldsymbol{D_{ds}}$) for rainfall, pattern of cultivation, presence of other contaminants (iron),[38] per capita consumption expenditure, sex ratio and literacy. Finally, we include state fixed effect in our regression analysis. Heteroskedasticity robust standard errors are clustered at the PSU (Primary Sampling Unit) level[39].

Estimating the effects of arsenic on nutritional outcomes in equation (1), using regional variation in arsenic levels, is problematic since the intensity of economic activities in a region may be correlated with arsenic concentration levels. In areas with high economic activity, overexploitation of groundwater is a major cause of arsenic contamination since

---

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

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

naturally occurring arsenic dissolves out of rock formations when groundwater levels drop significantly (Madajewicz et al., 2007). Hence, to overcome the issue of endogeneity, we use an instrumental variable approach.

3.4.1 Instrumental Variable Approach

Arsenic concentration is higher in clayey relative to coarse soil; thus, we exploit the variation in soil texture across districts within a state to instrument for arsenic groundwater contamination. Finer soils have relatively more particle density and lower porosity levels and as a result, their permeability level is relatively lower than loamy soil[40] which facilitates arsenic concentration in groundwater (Mc Arthur et al. 2001; Madajewicz et al. 2007). Figure 3.3 provides the visual evidence for the positive correlation between arsenic and soil texture.

The first stage equation is given by:

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

We instrument arsenic soil contamination using $Soil_{ds}$ i.e. the percentage of clayey soil in district *d.* Rest of the specification is same as in equation (1) above. The main identifying assumption is that soil texture fractions affect health outcomes only through the impact on the level of arsenic in groundwater.[41]

A threat to identification is that income might be affected by pattern of cultivation which is determined by soil texture. For instance, in India, water intensive crops (rice) are cultivated in areas with clayey soil due to its water retention capacity unlike sandy soil. As we show in the results, our regressions are robust to the inclusion of district-level ratio of rice to wheat production. Further, we also include sex ratio (measured at district level) as

---

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

[41] Note that while groundwater arsenic levels could also rise through increased use of fertilizers, the literature suggests that use of fertilizers does not alter the physical properties of soil (Carranza 2014). Unlike commercial crops like rice and wheat, arsenic-based pesticides are applied in specific crops such as fruit trees, potatoes, vegetables and berries. Use of such pesticides might alter some properties of superficial soil (upper most layer of soil), but not the subterranean soil used in our analysis.

soil texture can affect economic outcomes through relative female to male employment rates (Carranza 2014).

An additional threat to our identification strategy exists if clayey soil varies with other weather, geographic or demographic factors and these might affect economic outcomes. We provide evidence of no correlation between proportion of clayey soil and several district level indicators of weather (rainfall and temperature), other contaminants (nitrate and fluoride), economic or demographic factors (monthly per capita expenditure, rice to wheat production, literacy, sex ratio, usage of fertilizers), conditional on state fixed effects.[42]There is significant difference by soil permeability in iron, as districts with higher iron also have higher proportion of clayey soil (and thus more arsenic). However, this would be against finding a negative impact of arsenic on health outcomes and if anything, underestimate our findings. There is also a positive correlation between rainfall and clayey soil. Though there is no direct effect of rainfall on soil permeability levels as both are exogenous in nature, but both can combinedly determine the level of groundwater and presence of contaminated metals in groundwater.[43]

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

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

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

Where, $\boldsymbol{2ndchild}$ is an indicator for a child i whose birth order is 2. Similarly, $3rd^{+}$child indicates whether the child born is at 3rd or higher birth orders. Children born at first birth

---

[42] Results are reported in Table 3A.2.1 in the Appendix.

[43] If the amount of rainfall is less than the soil can absorb, it will all infiltrate; there will be no run-off or no discharge of water in the ground. But if rainfall is more than the absorption capacity of soil (defined by soil permeability level), there will be more discharge of water.

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

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

As a robustness check, we also control for family size measured by the number of children under the age of five in the household in the main regressions. [44] To overcome the issue of endogeneity of family size, we use the gender of first child as an instrument for family size. Having a girl as first child is positively associated with family size, particularly in the presence of son preference, as parents will continue to have more children until desired number of boys are born in a family (Pande and Astone, 2007). Further, gender of first child is exogenously determined and should affect child health outcomes only through family size (Kugler and Kumar 2017).

To determine if the health outcomes of arsenic exposure vary by family size and birth order, we control for family size and run the regressions separately by gender[45]. We use the gender of first child as an instrument for family size. We estimate the following OLS and first stage regressions separately by gender:

---

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

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

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

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

Where, Fam_size is the size of the family measured by the number of children under the age of five in a household**.** This variable is instrumented by $Gender\_First$**,** a binary variable for the gender of the first-born child in a household which takes the value of 1 for girls and 0 for boys. All other variables are same as in the previous regressions.

3.4. Results

3.4.1 Arsenic and child health by gender

We first show results for OLS estimates using equation (1). Column 1 and Column 2 (Table 3.2) shows OLS estimates of the effect of arsenic on HAZ and WAZ, respectively. Children who belong to arsenic affected districts have lower height-for-age. OLS estimates are significant but small and a one standard deviation increase in arsenic, reduces HAZ by 0.002 standard deviation units (Column 1, Table 3.2). OLS Estimates for WAZ shows that arsenic exposure is associated positively with weight-for-age (0.003 standard deviation).

In Table 3.3 we study whether the impact of arsenic on stunting and underweight varies by gender. Higher exposure to arsenic contaminated water is positively associated with stunting among girls (0.003 standard deviation) but not boys. However, in case of underweight no such evidence of gender difference is found in Column 3 and column 4 (Table 3.3). A one standard deviation increases in arsenic leads to increase in Weight-for-age by 0.003 standard deviation units regardless of gender.

To overcome the issue of endogeneity, we use an instrumental variable approach, where variation in soil texture across districts within a state is used as an instrument for arsenic levels in groundwater. The first stage regression results in Table 3.4 shows a positive and statistically significant relationship between arsenic and soil texture (clayey soil). The F-statistic (685.48) suggests that soil texture is a strong instrument for arsenic levels.

The IV results for HAZ, shown in table 3.5, indicate that the OLS is severely downward biased. A one standard deviation increases in arsenic leads to decrease in height-for-age by 0.028 standard deviation units which translates to a 1.69 percent decline relative to the mean. Column 2 of Table 3.5 indicates that higher level of arsenic exposure is positively linked to underweight as well (0.018 standard deviation – 1.09 percent).

We further analyze whether the effect of arsenic on stunting and underweight varies by gender. As is evident from columns 1 and 2 of Table 3.6, there is no difference by gender in the effect of arsenic contamination on HAZ. On the other hand, column 3 and 4 suggests that arsenic has a negative and significant impact on WAZ (1.09 percent- significant at 5 percent level of significance) unlike the OLS results. But there are no gender differences in the effect of arsenic on underweight.

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

3.4.2 Arsenic and child health across gender and birth order

Unlike OLS estimates, IV results in column 3 of Table 3.7 suggest that girls in arsenic affected regions have higher height disadvantage than boys, and the effects are magnified for later born girls relative to the eldest. A one standard deviation unit change in arsenic leads to decrease in height-for-age (stunting) for second and third (or later) born girls by 0.015 and 0.046 standard deviation units which is equivalent to 0.90 percent and 2.77 percent respectively. Significance of our estimate for third (or later) born girls indicates that stunting in girls increases with steeper birth gradient. These estimates are robust to the inclusion of various district level controls[46]. The IV coefficients are not sensitive to the

---

[46] All our findings in this study are robust to inclusion of family size, though we address this in detail in section 3.4.6.

inclusion of district level controls, as can be seen from comparing columns 2 and 3. This gives further credibility to the exogeneity of the instrument.

We find similar IV results for Weight-for-age as shown in column 4 of Table 3.7. IV estimates on WAZ indicates that a standard deviation unit increase in arsenic is accompanied with decrease in weight-for-age (underweight) for second and third (or later) born girls by 0.024 and 0.041 standard deviation units which is equivalent to 1.46 percent and 2.5 percent respectively.

When the concentration of arsenic in groundwater increases, later born girls (born at higher birth order) experience more stunting and underweight than their older sibling (lower birth order), particularly if the elder sibling is male. This could be explained by the *sibling rivalry effect* i.e. having an older brother limits the availability of essential nutrients along with other health inputs to later born daughters in the family (Fledderjohann et al., 2014; Victoria et al. 1987). To support this hypothesis, we next study the heterogenous effects of arsenic exposure on health outcomes.

3.4.3 Heterogeneous impact by Sex Ratio

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

We divide the sample into two groups based on the median district sex ratio and estimate the main regressions separately for households located in districts below median sex ratio and those located above median sex ratio[47]. IV estimates in Column 3 of Table 3.8 shows

---

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

that exposure to higher levels of arsenic leads to significantly lower weight-for-age for second (0.016 standard deviation- 0.98 percent) and third (or later) born girls (0.143 standard deviation which translates to 8.72 percentage-significant at 5 percent level of significance). The coefficient for third born girls is more than three times larger compared to the regressions estimates in Table 3.7. On the other hand, the effects are zero and insignificant in regions with above median sex ratio. At the same time, we find negative but insignificant association between arsenic exposure and height-for-age for later born girls in regions with sex below median sex ratio.

3.4.4 Heterogeneous impact by religious & cultural beliefs

We further explore the role of social and cultural identity manifested via different religious beliefs practiced in India. Table 3.9 examines difference in stunting across two different religion groups (Hindu and Muslims). Relative to Islam, Hinduism places greater emphasis on having a male heir so as to fulfil social rituals. Son preference is less prevalent amongst Muslims as evident by less skewed sex ratio and lower gender gap in child mortality amongst Muslims relative to Hindus (Borooah and Iyer 2005; Bhaltora, Valente and Soest 2010).

Our findings in Column 1 of Table 3.9 suggests that for Hindu community, height disadvantage amongst girls materializes at higher birth order (4.04 percent- significant at 5 percent level of significance) compared to Muslims who have muted birth order effect on stunting as shown in column 2. Further, column 3 also indicates that the adverse effect of arsenic on Weight-for-age aggravates with increasing birth order for Hindu girls (3.05 percent- significant at 5 percent level of significance). Thus, the detrimental effect of arsenic contamination on stunting and underweight in India could partially be explained by heterogeneity in cultural and religious beliefs practiced across various regions of India.

3.4.5 Heterogeneous impact by household location (rural or urban)

Next, we study whether the prevalence of stunting is higher in rural areas relative to urban regions. Our results in Column 1 and Column 3 (Table 3.10) suggests that nutritional status (as measured by Height-for-age and Weight-for-age) amongst girls is worse in rural regions. The adverse impact of arsenicosis on stunting (3.43 percent) and underweight (2.8 percent) magnifies for elder daughters in rural households relative to urban households. The primary reason for rural-urban difference in impact of arsenicosis on nutritional status of children particularly girls might be economic inequality. Inadequate provision of basic amenities such as adequate food, drinking water and health care might worsen their long-term economic outcomes at later stages of life.

3.4.6 Arsenic, birth order and family size

A key concern with the analysis thus far is that children born at higher birth orders have a higher probability of being from a large size family. Moreover, family size and resource allocated to each child are highly correlated, which might in turn could affect the health outcomes of children. To address this potential problem, we conduct a robustness check where we control for family size and instrument it with gender of first child. As discussed earlier, we estimate the regressions separately by gender.

The first stage coefficient on the IV for family size is high at 13.28 (statistically significant at the 1% level). The F-statistics is 685 suggesting that gender of the first child is a valid instrument for family size.[48] IV results from this specification is shown in Table 3.11, where column 1 and column 3 provides the IV estimates for health outcomes (HAZ and WAZ) for girls. Height-for-age and weight-for-age for boys are reported in column 2 and column 4. Our estimates in column 1 shows that one unit increase in arsenic exposure leads to significant decrease in height-for-age for girls born at second and third (or higher) by 0.021 (translates to 1.27 percent) and 0.122 standard deviation (translates to 7.35 percent- significant at 1 percent level of significance) respectively. Similar estimates are shown for arsenic exposure on underweight amongst girls born at second (2.62 percent- significant at 1 percent level of significance) and third or higher birth order (6.83 percent- significant at 1 percent level of significance). We find negative impact of arsenicosis on height-for-age (0.004 standard deviation- second born and 0.057- third or higher birth order) and weight-for-age amongst boys, but the impact of arsenic exposure on health outcomes (height-for-

---

[48] Table not shown but available upon request

age and weight-for-age) amongst girls is three times greater than for boys even after accounting for family size and other related factors. Looking at the coefficient on family size, a greater number of children in the household has a negative and significant effect on HAZ but not WAZ.

To sum up, exposure to higher levels of arsenic has adverse impact on health outcomes for children, but the effect is considerably higher for girls as compared to boys, particularly the younger ones. This can be attributed to the lack of adequate nutrition and health care provided to girls during their early childhood that would result in poor immunity and growth, which in turn increases their vulnerability to diseased environment. This is consistent with our findings that the effect of arsenic exposure is magnified in areas with worse sex ratio and in rural locations, where gender norms are likely to be distorted. The negative effect is also larger among Hindu households, that follow Hindu kinship systems.

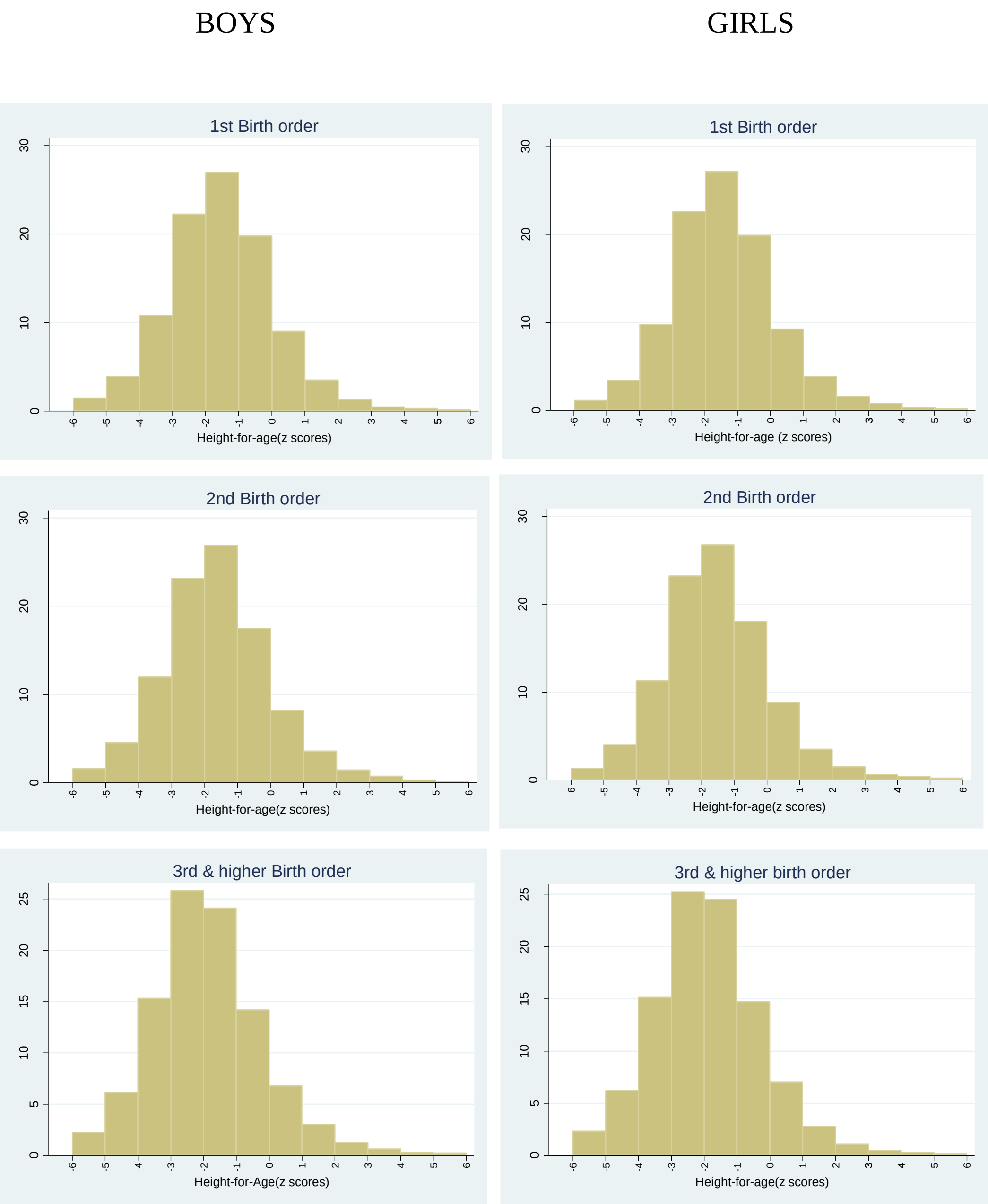
3.5. Conclusion and Policy Implications

Gender inequality is one of the most fundamental challenges to sustainable development. While considerable efforts have been made to explore the impacts of gender inequality on women, lesser is still known regarding its impact on child health. India is the only developing country where the under-five child mortality rates are worse among girls than boys (Census, Government of India, 2011). This might be due to discrimination in resource allocation by parents at early stages of their lives, in the form of shorter duration of breastfeeding, lesser post-natal health inputs such as vaccination and supplementary food items.

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

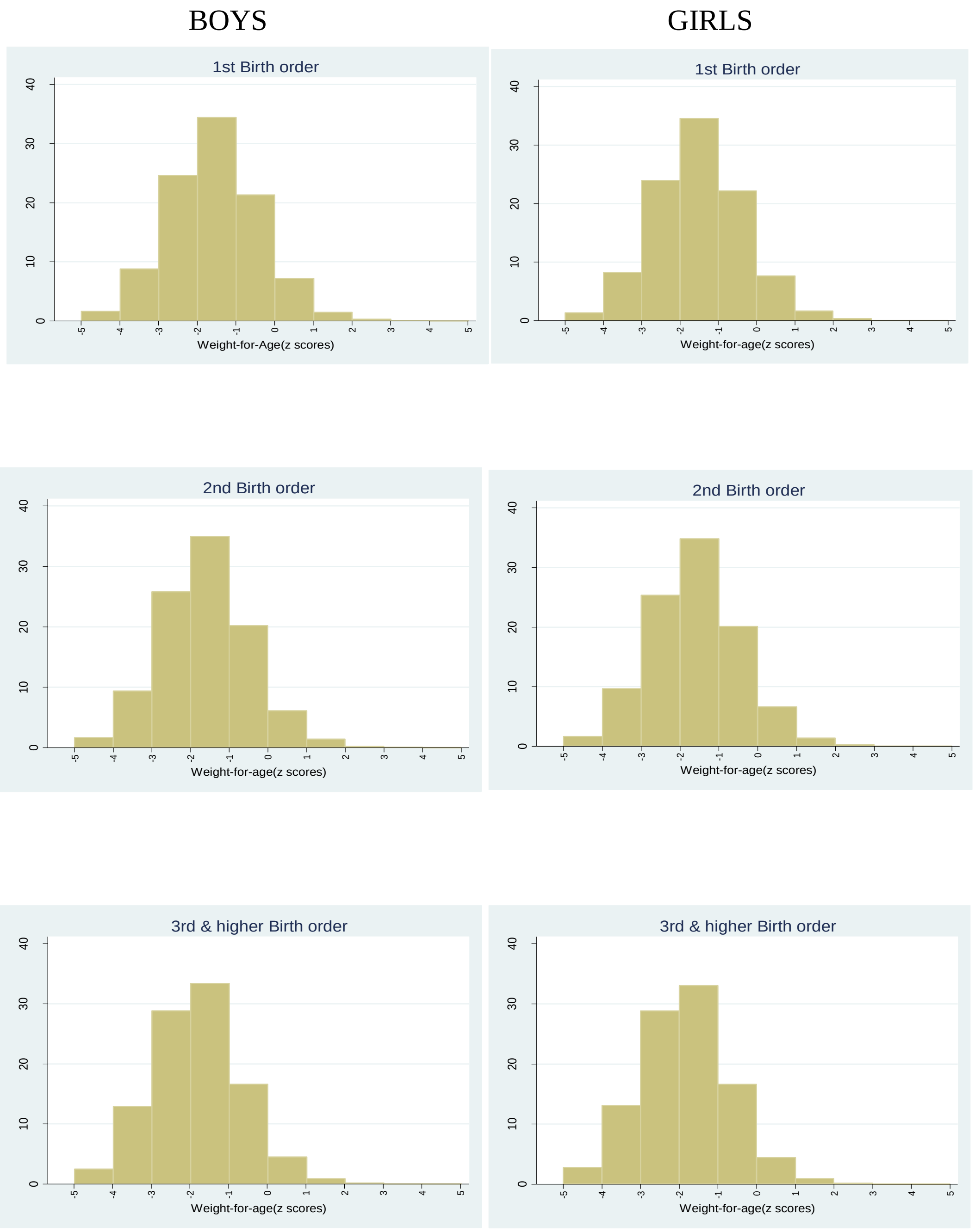
Our results show heterogeneous effect of arsenic exposure across cultural norms and socio-economic status, highlighting the role played by son biased preferences in magnifying the negative impact of unsafe water on health for girls. Despite safe water being an indispensable input to human health, to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

Figure 3. 1:Histogram for Heights-for-age(z scores) for boys & girls, by birth order



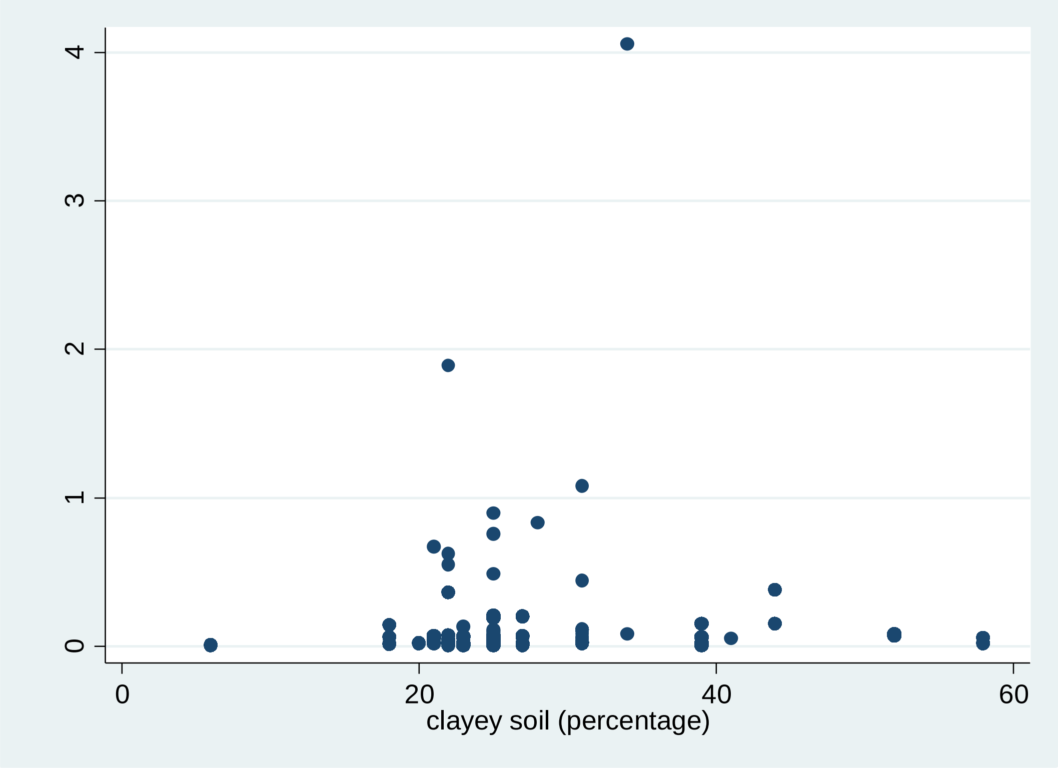
Note: The above figure shows percentage of boys and girls whose z scores for Height-for-age (z scores) decrease with increasing birth order, as represented on the horizontal axis. As per guidelines issued by World Health Organization children whose z scores are below -2 and above -3 indicates moderate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

Figure 3. 2:Histogram for Weight-for-age (z scores) for boys & girls, by birth order



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

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



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

Table 3.1: Descriptive Statistics and District level control variables

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

Sample size is N=73,160

Table 3.2: Arsenic and Child's anthropometric measures (OLS Estimates)

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

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size and place of residence).

Table 3.3: Arsenic and Child Anthropometric Measures: By Gender (OLS Estimates)

	HAZ	**HAZ**	**WAZ**	**WAZ**
Anthropometric measures (z scores)	**Girls (1)**	**Boys (2)**	**Girls (3)**	**Boys (4)**
Arsenic	-0.003*	-0.001	0.003**	0.003**
	(0.002)	(0.001)	(0.001)	(0.001)
Individual controls	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes
State F.E	Yes	Yes	Yes	Yes
Observations	36,198	39,173	36,198	39,173

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

Table 3.4: First Stage Regression

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

Note: Robust SE (*** Significant at 1%, ** significant at 5%, * significant at 10 %.). Independent variable is defined as percentage of clayey soil present in district. Regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size and place of residence) and included.

Table 3.5: Arsenic and Child Anthropometric Measures (IV Estimates)

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

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size and place of residence) and included.

Table 3.6: Arsenic and Child Anthropometric Measures (IV Estimates)

	HAZ	**HAZ**	**WAZ**	**WAZ**
Anthropometric measures	**Girls**	**Boys**	**Girls**	**Boys**
(z scores)	**(1)**	**(2)**	**(3)**	**(4)**
Arsenic	-0.029**	-0.029**	-0.018**	-0.018**
	(0.011)	(0.012)	(0.008)	(0.008)
Individual controls	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes
State F.E	Yes	Yes	Yes	Yes
Observations	36,198	39,173	36,198	39,173

*Robust Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). All regressions include state fixed effects and district level controls for sex ratio, rainfall, literacy, pattern of cultivation, iron and gross domestic product. Individual level controls (age, age square and gender), maternal controls (mother's age and mother's education) and family background controls (religion, caste, family size and place of residence).

Table 3.7: Arsenic, gender and birth order gradient in Height-for-age and Weight-for-age (IV Estimates)

	HAZ (OLS) (1)	HAZ (IV) (2)	HAZ (IV) (3)	WAZ (OLS) (4)	WAZ (IV) (5)	WAZ (IV) (6)
Arsenic*girl*BO2	0.003	-0.035	-0.015	0.005	-0.043**	-0.024
	(0.004)	(0.026)	(0.023)	(0.004)	(0.018)	(0.017)
Arsenic*girl*BO3	-0.003	-0.060**	-0.046*	-0.004	-0.062**	-0.041**
	(0.005)	(0.027)	(0.025)	(0.004)	(0.019)	(0.018)
Arsenic*BO2	-0.006**	-0.009	-0.022	-0.003	-0.002	-0.013
	(0.003)	(0.018)	(0.016)	(0.003)	(0.013)	(0.011)
Arsenic*BO3	-0.009*	-0.094***	-0.090***	-0.006**	-0.072***	-0.065***
	(0.003)	(0.020)	(0.018)	(0.003)	(0.014)	(0.013)
Arsenic*girls	0.000	0.104***	0.068***	-0.001	0.049***	0.029***
	(0.002)	(0.016)	(0.015)	(0.003)	(0.012)	(0.010)
Arsenic	0.002	-0.049**	-0.023	0.006**	-0.023**	0.004
	(0.002)	(0.015)	(0.012)	(0.002)	(0.011)	(0.009)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes	Yes	Yes	Yes
District level controls	Yes	No	Yes	Yes	No	Yes
Observations	73,160	102731	73,160	73,160	102731	73,160

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

Table 3.8: Sex ratio, Child’s Height-for-age and Weight-for-age (IV Estimates)

	HAZ	HAZ	WAZ	WAZ
Anthropometric measures (z scores)	**(1) Sex ratio (below 922[49])**	**(2) Sex ratio (above 922)**	**(3) Sex ratio (below 922)**	**(4) Sex ratio (above 922)**
Arsenic*girls*birth order2	0.061	-0.035	-0.016	-0.031
	(0.081)	(0.028)	(0.061)	(0.020)
Arsenic*girls*birth order3	-0.075	-0.031	-0.143**	0.006
	(0.076)	(0.033)	(0.059)	(0.024)
Arsenic*birth order2	-0.073	-0.015	-0.060	-0.003
	(0.055)	(0.019)	(0.042)	(0.014)
Arsenic*birth order3	-0.198**	-0.083**	-0.100**	-0.085***
	(0.057)	(0.026)	(0.044)	(0.019)
Arsenic*girls	0.115**	0.068***	0.040	0.032**
	(0.056)	(0.018)	(0.043)	(0.013)
Arsenic	0.032	-0.037**	0.158***	-0.034**
	(0.052)	(0.019)	(0.042)	(0.013)
Individual controls	Yes	Yes	Yes	Yes
Maternal controls	Yes	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes	Yes
District level controls	Yes	Yes	Yes	Yes
State F.E	Yes	Yes	Yes	Yes
Observations	38,579	33,012	38,579	33,012

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

---

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

Table 3.9: Heterogeneity across religious groups (Within India Evidence: IV Estimates)

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

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

Table 3.10: Heterogeneity across rural or urban regions (Within India Evidence: IV Estimates)

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

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

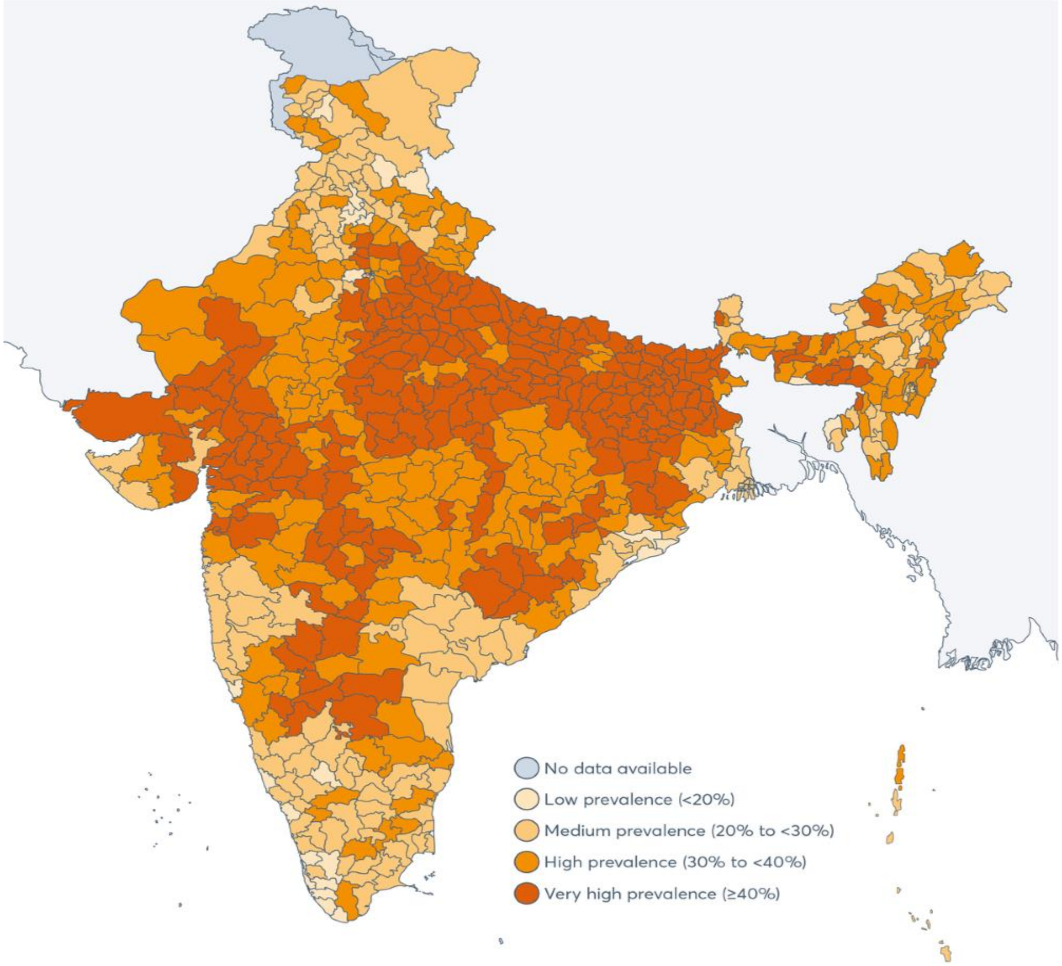
Table 3.11: Arsenic, family size and birth order (IV Estimates)

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

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

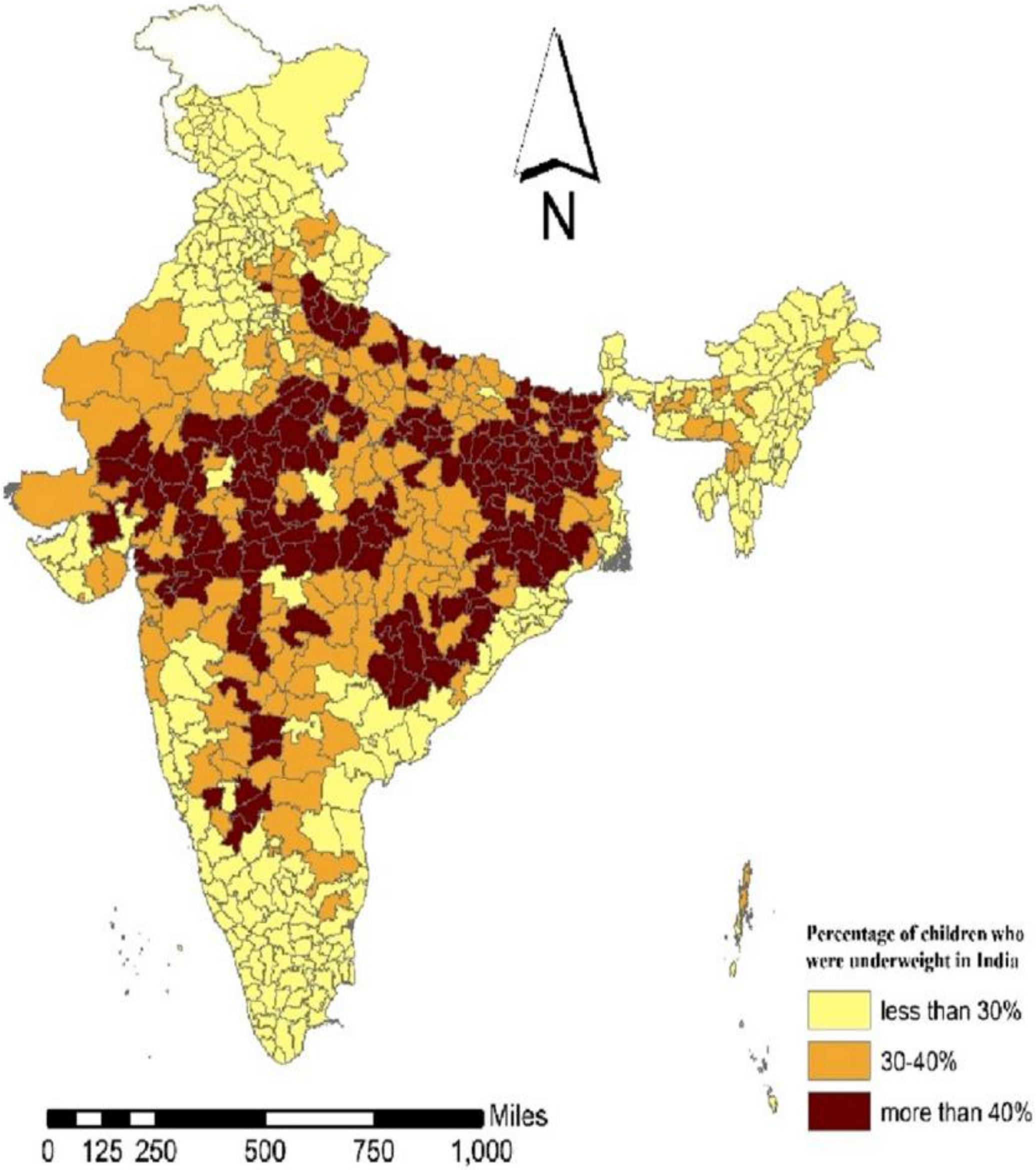
Appendix 3A.1: Figures and Maps

Figure 3A.1. 1:Prevalence of stunting across districts of India



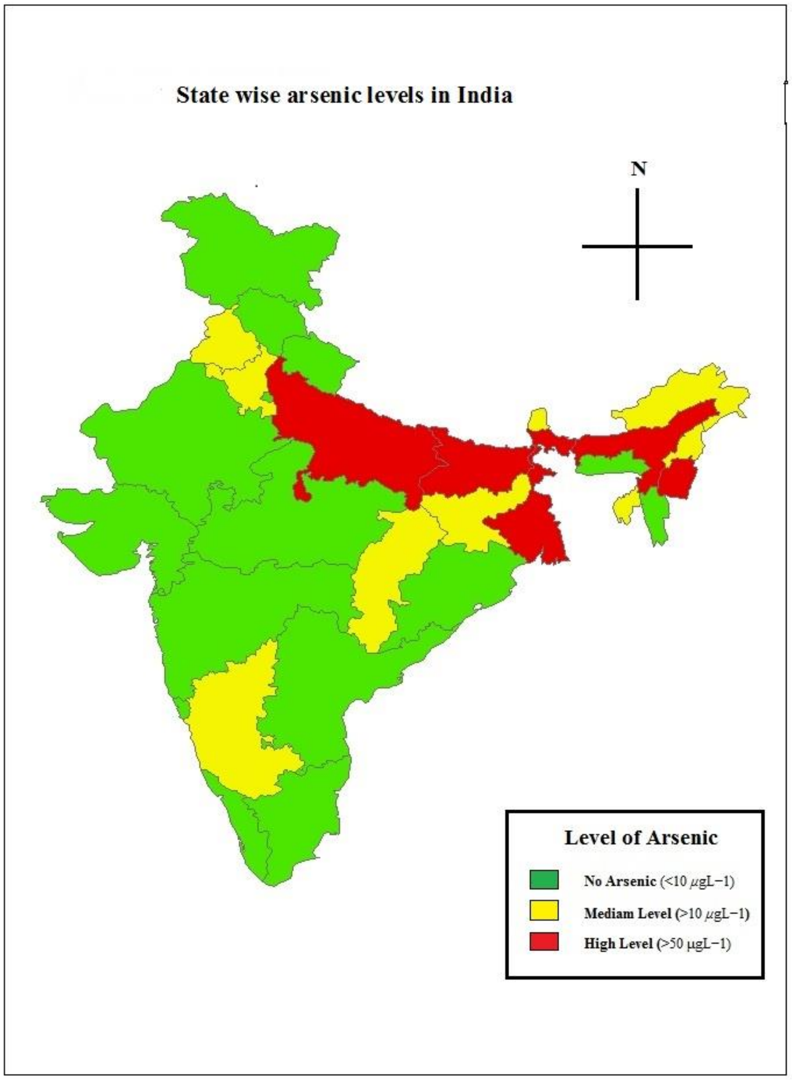
Source: Menon et al. (2018)

Figure 3A.1.2: Prevalence of underweight across districts of India



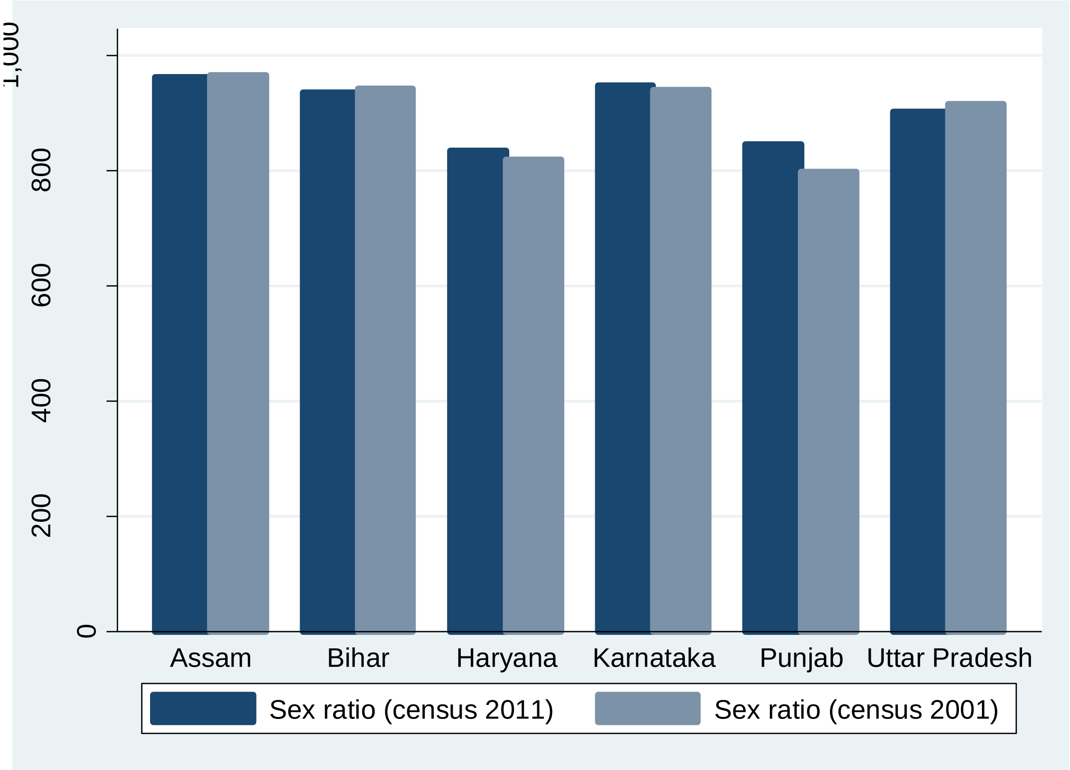
Source: Sharma et al. (2020)

Figure 3A.1. 3:Geographical distribution of Arsenic levels across States of India



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

Figure 3A.1. 4:Bar Graph for child sex ratio: Central Northern, North Eastern and Southern states of India



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

Appendix 3A.2 Additional Tables

Table 3A.2.1 Cross Tabulation of district level characteristics by arsenic contamination

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

Table 3A.2.2: Clayey soil and District Level Characteristics

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

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

Chapter-4

Conclusion

Despite India's sustained economic growth, it continues to perform poorly in terms of child nutrition as reflected by wide prevalence of stunting. Estimates by UNICEF (2013) suggest that more than 30 percent of world's stunted children live in India. Shorter height is an outcome of poor health and lower nutritional intake in early childhood, which later inhibits the growth and development of their mental as well as physical abilities.

Although several efforts have been taken by the government and international organisations in the form of anti-poverty and nutritional programs,[50] but these schemes have failed to deliver their targets efficiently. This is primarily due to the difficulties faced in accurate recognition of targeted beneficiaries, as these beneficiaries are identified on the basis of income criterion which is quite burdensome to be estimated correctly. Presence of wide intra-household disparities makes it more complex (Deaton, 2016). For instance, children born at first birth order are allocated relatively significant share of household's budget as compared to later born children, especially if the first born is a male child (Jayachandran and Pande, 2017). Unequal intra-household allocation of health services, made available on the basis of gender and birth order, is thus a major cause of poor nutritional status among children, especially girls. Poor nutritional intake amongst girls in India is clearly reflected in the highest under-five child mortality rate as compared to any other developing nation (Census 2011, Government of India). child sex ratio is another crucial indicator which declined drastically from 927 in 2001 to 919 in 2011 (Census 2011, Government of India).

Lack of adequate nutrition and health care during early childhood can make the girls more vulnerable to external environmental hazards due to their lower immunity and under-developed bodies. Epidemiological studies have established that early-life environmental

---

[50] Mid-day Meal scheme, Integrated Child Development Services scheme, Applied Nutrition Program, National Nutritional Anemia Prophylaxis Program, Wheat Based Nutrition program, National Goiter Control Program.

exposures play a role in growth outcomes (Walker et al., 2007). Therefore, adequate food and healthy environment are key to a child's growth and overall development.

This thesis makes important contributions to the existing literature in this regard. 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 are unique and able to identify the causal effect of arsenic exposure at the regional as well as individual levels. Second, ours is 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 external environmental hazards. 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 study makes a significant contribution to the literature on health and education outcomes.

The first objective of our study is evaluated with the help of secondary data sourced from India Human Development Survey II (IHDS-II, 2011-12). To assess the impact of arsenicosis on school absenteeism, instrumental regression approach has been used, where we exploit the geographical variation in permeability levels of soil across districts as an instrument to check the presence of arsenic level in groundwater. For instance, clayey soil has high water retention capacity and is relatively less porous unlike sandy soil, that leads to decreasing water table and increasing arsenic concentration in groundwater (Madajewicz et al., 2007; Mc Arthur et al., 2001). Our findings on IV suggest that on an average, an additional increase in arsenic by one microgram per litre is associated with increase in school absenteeism by 0.13 percentage points. But the estimate is conservative since the information available in the dataset corresponds to enrolled children in school.

The second objective of this study to substantiate the secondary data analysis (using IHDS-II) with self-collected school level survey in one of the most severely affected districts of Assam (Jorhat). We assess the impact of consuming unsafe water on several educational outcomes (CGPA, grade repetition, school absenteeism and numeracy scores) amongst children enrolled in schools. Our identification strategy relies on exploiting exogeneous variation in timing, coverage of implementation of government water supply schemes and child's year of birth to measure the effects of safe water access and exposure.

Our difference-in-difference results suggest that exposure to unsafe water for additional years is associated with lower Cumulative Grade Point Average (CGPA) in the previous grade attended and higher rates of grade repetition. Further, consumption of contaminated water leads to lower numeracy score and higher rate of school absenteeism. An additional increase in exposure to unsafe water is accompanied with increase in retention rate by 3.6 percent for children with greater effect among girls (4.6%) compared to boys (3%).

Using both the primary (school level survey) as well as secondary datasets (IHDS-II), our findings point towards the absence of gender difference in school absenteeism, but the survey data suggests that the adverse effects of consuming contaminated water on test scores and grade retention are greater among girls. This might be due to parental gender-based discrimination in terms of unequal provision of nutrition and health care provided to female child as witnessed mostly in rural parts of India. This holds to be consistent with the recent estimates which show that India is the only country in the world where the under-five mortality rate is worse among girls than boys.[51]

The final objective of this study is to analyse the primary cause of poor performance of girls in school owing to arsenic exposure. To do so, we use secondary data from National Family Health Survey-4 (NFHS-4, 2015-16). Our empirical analysis relies on exploiting the geographical variation in arsenic across various districts, and estimates its association with health outcomes (stunting and underweight) amongst children under the age of five years. But our Ordinary Least Square estimate could be biased as the presence of arsenic in groundwater is correlated with economic activity of the region. Hence, to overcome this issue of endogeneity, we use instrumental variable approach in our analysis.

---

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

Our primary findings suggest that increase in arsenic by one microgram per litre is associated with significant reduction in height-for-age (HAZ) and weight-for-age (WAZ) amongst children less than five years of age, regardless of gender. Further, we test whether lower nutritional intake amongst girls leads to significantly higher impact of arsenicosis. To do so, we analyse the effect of birth order on the association between arsenic and health outcomes. Our instrumental variable estimates indicate that an increase in arsenic level by one microgram per litre in groundwater leads to significant reduction in HAZ by 0.015 standard deviations (equivalent to 0.9 percentage) and 0.046 standard deviations (equivalent to 2.8 percentage) for the second born and third born girl child, respectively, relative to a male child born at first birth order. Similar effect is observed for low weight (measured by WAZ) amongst girls of higher birth order as compared to lower birth order. IV estimates show that additional exposure to arsenic (one microgram per litre) is linked to decrease in weight-for-age for second born (1.46 percent) and third born (or higher birth order - 2.5%) girls. All our findings are based on district level controls for weather, water quality measures, pattern of cultivation, education quality and income.

To conclude, our overall estimates on poor health outcomes amongst girls suggest that lack of adequate nutritional care exacerbates the detrimental effect of arsenic consumption. Girls are breastfed for shorter duration relative to boys, and are, thus, more exposed to arsenic, as arsenic does not pass-through breast milk (Shastry et al., 2017). This gender bias in nutrition and resource allocation has been extensively researched in India (Gupta, 1987; Fledderjohann et al., 2014). Cultural and religious norms such as patrilineality and patrilocality associated with Hindu kinship system promote son preference in India (Dyson and Moore, 1983; Arnold, Choe and Roy, 1998). Consistent with this, we find that states with negatively skewed sex ratios have higher incidences of low weight amongst girls, particularly later born girls.

The thesis concludes that one of the leading causes of morbidity in India is lack of access to safe drinking water affecting more than 75.8 million people (WHO, 2009). Moreover, over exploitation of groundwater 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 study contributes to the literature by measuring the effects of environmental health on human capital.

Merely focusing on water related policies would reduce the burden of diseases to some extent, but lower immunity of girls would remain a challenge. Thus, the government should target both its nutrition and water related strategies simultaneously to approach the above-mentioned issue of child health in a holistic manner. Given the emerging crisis of stunting and contaminated water in India, there is dire need of government interventions especially in states like Haryana, Uttar Pradesh, Bihar and other arsenic affected regions with regressive socio and cultural norms, as the adverse consequences of arsenicosis become more severe in these states due to negligence towards the health of young and newborn girls.

Chapter-5

References

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

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

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

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

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

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

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.

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

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

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

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

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

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

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

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

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

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.

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

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

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

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.

Cutler, D. M., & Lleras-Muney, A. (2006). *Education and health: evaluating theories and evidence* (No. w12352). National bureau of economic research.

Deaton, A. (2016). Measuring and understanding behavior, welfare, and poverty. *American Economic Review*, *106*(6), 1221-43.

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

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.

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

Drèze, J., & Kingdon, G. G. (2001). School participation in rural India, *Review of Development Economics*, *5*(1), 1-24.

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

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

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

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

Fowler, M. G., Davenport, M. G., & Garg, R. (1992). School functioning of US children with asthma. *Pediatrics*, *90*(6), 939-944.

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

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

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 Adolscent School Bullying Victimisation on Later Life Outcomes," IZA Working paper 12241.

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

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

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

Grossman, M. (2008). The relationship between health and schooling, *Eastern Economic Journal*, 34(3), 281-292.

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

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

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

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

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

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

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.

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

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, National Bureau of Economic Research, Inc.

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

Madajewicz, M. et al. (2007). "Can information alone change 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

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.

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

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

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

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

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): 1.

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

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.

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 No. w21741.

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

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

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

Saing, C. H., & Cannonier, C. (2017). "Arsenic Exposure and School Participation in Cambodia", *Available at SSRN 2907461*.

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

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. (2013) Children dying daily because of unsafe water supplies and poor sanitation and hygiene. Retrieved from UNICEF website:https://www.unicef.org/media/media_68359.html

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

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

Victoria, C. G., J. P. Vaughan, C. Lombardi, S. M. C. Fuchs, L. P. Gigante, P. G. Smith, L. C. Nobre, A. M. Texiera, L. B. Moreira, and F. C. Barros (1987). "Evidence for a Strong

Protective Effect of Breastfeeding against Infant Death due to Infectious Diseases in Brazil. *Lancet, 330(8554), 319–322.*

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

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.

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

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

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

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.