YOUNG LIVES STUDENT PAPER

Does Access to Sanitation and Drinking Water Facilities Relate to Students' Learning Outcomes? Evidence from Andhra Pradesh, India

Divyanshi Wadhwa

April 2016

Thesis submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University in partial fulfilment of the requirements for the degree of Master of Public Policy in Public Policy.

The data used come from Young Lives, an international study of childhood poverty, following the lives of 12,000 children in four countries (Ethiopia, India, Peru and Vietnam) over 15 years. **www.younglives.org.uk**

Young Lives is core-funded by UK aid from the Department for International Development (DFID).

The views expressed here are those of the author. They are not necessarily those of the Young Lives project, the University of Oxford, DFID or other funders.

Does access to sanitation and drinking water facilities relate to students' learning outcomes? Evidence from Andhra Pradesh, India

A Thesis
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Public Policy
in Public Policy

By

Divyanshi Wadhwa, B.A.

Washington, DC
April 12, 2016

Does access to sanitation and drinking water facilities relate to students' learning outcomes? Evidence from Andhra Pradesh, India

Divyanshi Wadhwa, B.A.

Thesis Advisor: Adam T. Thomas, Ph.D.

Abstract

Indian students fall substantially behind students from other countries in academic performance. A majority of students in India also do not have access to usable toilets and drinking water facilities. Some literature suggests that access to sanitation and drinking water facilities is related to education outcomes, such as school enrollment rates, school attendance rates, and students' cognitive functions. I assess whether students who have greater access to usable toilets and drinking water facilities at school and home have higher test scores in Mathematics, English and Telugu. In order to test this hypothesis, I estimate an Ordinary Least Squares multivariate regression model using household-level and school-level survey data collected in Andhra Pradesh, India by the Young Lives Project. I find no relationship between access to household sanitation and drinking water facilities and test scores. However, I find that having at least one usable toilet and a drinking water source in schools is positively associated with English test scores.

Index words: Sanitation, Water, Test scores, India

Acknowledgments

I would like to express my sincerest gratitude to my thesis adviser, Prof. Adam Thomas for his endless support and guidance, and immense knowledge. I would also like to thank Prof. Jeffrey Mayer for his most helpful writing suggestions throughout the year.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Even though India is rapidly moving towards universal primary school enrollment, students' learning outcomes have been extremely poor (Gelda, Narayan, Mudiyam, Raturi, & Seshan, 2013). More than half of Grade V students in India cannot read a Grade II level text (ASER, 2015). In 2009, the two Indian states that participated in the Programme for International Student Assessment (PISA) test ranked second to last of the 73 participating countries (PISA, 2009). Another standardized test, the Trends in International Mathematics and Science Study (TIMSS), ranked Indian performance in mathematics 44^{th} out of the 51 participating countries (Das & Zajonc, 2008). With almost one-third of the Indian population in 2009 younger than 15 years (Central Bureau of Health Intelligence, 2011), one might argue that improving primary educational outcomes is imperative in order to improve the country's economic and political prospects.

Evidence suggests that educational outcomes are important indicators of economic well-being (Hanushek & Woßmann, 2010). Learning outcomes such as test scores have been found to be associated with economic growth and productivity (Delgado, Henderson, & Parmeter, 2014; Hanushek & Woßmann, 2010). On the other hand, indicators such as years of schooling have been found to be poor proxies for the quality of a country's human capital (Delgado, Henderson, & Parmeter, 2014). These findings suggest that Indian policymakers should consider focusing not only on boosting school enrollment, but also on improving conditions that enhance learning outcomes.

One such condition may be the quality of infrastructure for sanitation and drinking water at schools and in students' households (Adukia, 2014). Literature suggests that access to sanitation or drinking water may improve student participation or cognitive functions and therefore lead to better educational outcomes (Adukia, 2014; Greene, et al., 2012; Hunter, et al., 2014; Koolwal & van de Walle, 2010; Lorntz, et al., 2006; Newman, et al., 2002; Spears & Lamba, 2011). Hence, understanding the relationship between sanitation and water infrastructure and educational outcomes can have important policy implications for national and state governments seeking to redirect their resources in an appropriate and effective manner. If the provision of improved sanitation and drinking water facilities does improve learning outcomes for students, governments can focus on redirecting funds towards building better toilets and drinking water facilities in schools and households, and on expanding existing policies that provide them.[1]

[1]Improved sanitation facilities and drinking water source refer to facilities that are uncontaminated, specifically by fecal matter (UNICEF, 2015). I define these categories in my analysis by using indicators for "usable toilets" and "usable water sources."

Chapter 2

Background

In recent years, the Indian government has pursued several initiatives to improve sanitation and water supply. One recent and ongoing initiative is an extensive preventive healthcare program called the *'Swachh Bharat Abhiyan'* ('Clean India campaign') (Ministry of Human Resource Development, Government of India, 2014). As part of this campaign, a smaller program focuses on *'Swachh Vidyalaya'*, meaning 'clean schools,' with one of its aims being to provide toilet facilities in every Indian school (Ministry of Human Resource Development, Government of India, 2014). A previous initiative, the "Total Sanitation Campaign" launched in 1999, also focused on school sanitation and hygiene (Ministry of Human Resource Development, Government of India, 2014).

As a result of these initiatives, the Indian government has made progress towards achieving the Millennium Development Goals (MDGs) for clean drinking water. The MDGs are a set of time-bound goals adopted by the United Nations in 2000, and determined to be achieved by the year 2015. Over this period, all countries were to halve the proportion of people without access to improved sanitation and drinking water facilities. While India met this goal for improving access in drinking water facilities, it fell acutely short in terms of improved sanitation facilities (World Bank, 2015). For example, more than one-third of the schools in India do not have usable toilets and about one-fourth do not have access to drinking water (ASER, 2015). Moreover, less than one-third of Indian rural households have access to domestic

toilets and tap water (World Bank, 2015). Some researchers suggest that sanitation facilities are still lacking in urban areas because the population has been rapidly increasing and infrastructure development has not kept pace with the demographic change (Kumar, Kar, & Jain, 2011), and in rural areas because the local government institutions do not have access to sufficient financial resources (Kumar, Kar, & Jain, 2011) and therefore cannot invest in infrastructure development.

The MDG agenda has recently come to an end, but the newly established Sustainable Development Goals (SDGs) targeted for achievement by 2030 also focus on increasing access to sanitation and water – this time, for every person (United Nations Department of Economic and Social Affairs, 2015). The new goals reflect an increased consensus regarding the importance of access to improved sanitation and drinking water facilities. However, even though international organizations such as the UN and the World Bank recognize the benefit of such facilities to education (World Bank, 2015), there is a lack of evidence about their precise effects on students' learning outcomes.

I focus on this gap in the literature by studying the Indian states of Andhra Pradesh and Telangana, which formerly comprised the single state of Andhra Pradesh.[2] In 2011, Andhra Pradesh, a state in the southern peninsula of India, was the country's fourth largest in terms of size and the fifth largest in terms of population (Aiyar, 2011). The state was divided and reorganized by the Andhra Pradesh Reorganization Act, 2014 (AP Reorganisation Portal, 2014). One region, Telangana, was carved out of the former state of Andhra Pradesh in order to form two separate states, Andhra Pradesh and Telangana (AP Reorganisation Portal, 2014) (I shall use the term "former Andhra Pradesh" when referring to the state

[2]The Young Lives Project, which is my data source, chose Andhra Pradesh as its study site because its government instituted several initiatives focusing on eradicating poverty (Young Lives, 2014).

before the Reorganization Act, and the term "Andhra Pradesh" when referring to the reorganized version of the state).

My data were collected before the enactment of the reorganization. For the purpose of this paper therefore, I study former Andhra Pradesh. Although the former state's Gross Domestic Product (GDP) was higher than that of the average Indian state (Directorate of Economics and Statistics, 2014), its sanitation and drinking water infrastructure was still lacking. In 2014, over 22 percent of schools in former Andhra Pradesh did not have access to usable toilets or water supply (ASER, 2014). Among households, the situation was not very different. According to the 2011 Census, former Andhra Pradesh was one of the six worst performing states in India in terms of providing households access to safe drinking water (Kumar & Das, 2014). The 2011 Census data suggest that over 10 percent of former Andhra Pradesh households did not have safe drinking water available to them (Kumar & Das, 2014). The 2011 Census also estimated that more than half of households in former Andhra Pradesh did not have access to private latrine facilities (Kumar & Das, 2014). These households either had access only to public latrines, or they practiced open defecation (Kumar & Das, 2014). In view of these problems, this paper addresses the question of whether the installation of water and toilet facilities in India's schools and households is associated with students' learning outcomes.

Chapter 3

Literature Review

The global community has increased efforts to boost access of usable toilets and drinking water for children in developing countries in order to improve their school participation and ultimately their academic achievement levels (Greene et al., 2012). However, much of the literature regarding children's access to sanitation and drinking water facilities focuses on health outcomes; only a few studies examine education outcomes. Those studies that have looked at education outcomes commonly focus on student absenteeism rates, school participation, enrollment rates and completion rates. Very few operationalize education outcomes by using test scores.

Other studies have assessed the effects on education outcomes of broader public health interventions, including hygiene education, disease prevention through handwashing, and deworming. Still others focus on either the status of health and hygiene in schools *or* in households. However, there is a dearth of research on the education effects of overall access to sanitation and hygiene in *both* schools *and* households.

This review focuses on three aspects of the literature: the association between education outcomes and the quality of sanitation and water infrastructure in schools, the association between education outcomes and other public health interventions in schools, and the association between education outcomes and public health interventions in households.

3.1 Sanitation and Water Infrastructure and Schools

The central hypothesis motivating many public health and education studies has been that improvement in health and hygiene infrastructure leads to better health outcomes and, therefore, to reduced school absenteeism (UNICEF, 2012). Other studies have pointed out that, in India, as in many other developing countries, the burden of providing water for the household is borne mainly by women, and that a water supply is not readily available to households, women or girls are expected to travel considerable distances to reach one (National Commission for Women, India, 2005). As a result, women and girls are less able to attend school or participate in other productive work (Koolwal & van de Walle, 2010).

A study by the World Bank in nine developing countries evaluated the response of women to improved water infrastructure (Koolwal & van de Walle, 2010). The authors found that, in countries with a large gender gap in education, (such as Pakistan, Nepal, Yemen and Morocco), boys and girls in households that experienced improvement in water facilities also had higher school enrollment (Koolwal & van de Walle, 2010). Similarly, a study in Cambodia found that availability of free and safe drinking water in schools was strongly and negatively correlated with absenteeism during the dry season (Hunter, et al., 2014).

In regard to sanitation facilities, Adukia (2014) found that, in India, latrine construction in sample schools was associated with an eight-percent increase in enrollment rates in upper primary schools and a twelve-percent increase in primary schools, and with reduced dropout rates in all schools (Adukia, 2014). Adukia's study is one of the few that also examines the relationship between sanitation infrastructure and academic achievement. To assess this relationship, she performs additional analysis specific to Uttar Pradesh, the most populous state in India, but she uses an admit-

tedly imprecise measure for academic achievement: the probability that students will graduate on time and with honors. She finds no statistically significant relationship between sanitation infrastructure and student achievement (Adukia, 2014).

Recently, researchers have been paying increased attention to the effect of availability of sanitation facilities among girls in particular. Multiple studies have tested the "menstruation hypothesis," which suggests a correlation between school absenteeism and the onset of puberty for girls, since girls may leave school after the onset of their menses if the school does not have separate toilets for them (Herz & Sperling, 2004). In this regard, Adukia (2014) produces separate estimates by age and gender, and she finds that the presence of sex-specific latrines is strongly associated with improved education outcomes for older, pubescent-aged girls.

However, using data from a randomized controlled trial conducted in Nepal, Oster and Thornton (2009) cast a measure of doubt on the "menstruation hypothesis." The trial distributed menstrual cups to girls at random (Oster & Thornton, 2009), but found no statistically significant impact on school attendance (Oster & Thornton, 2009). Another study in Kenya also found that the difference in enrollment rates between girls and boys was not associated with the quality of sanitation and water facilities (Mensch & Lloyd, 1998). However, this study did suggest that higher performing schools had better access to water supply than lower performing schools. The Kenya study is also one of the few to evaluate academic achievement in relation to health infrastructure (Mensch & Lloyd, 1998).

3.2 Other Public Health Interventions and Schools

Schools in developing countries have been frequent recipients of public health interventions to improve education outcomes. Consequently, a number of randomized con-

trolled trials have been conducted to study the effect of handwashing on school participation, including one in Egypt (Talaat et al., 2011) and another in China (Bowen et al., 2007). Both studies showed reduced absenteeism in schools where children were either mandated to wash hands or monitored in the way that they washed hands every day.

Conversely, an evaluation of the School Water, Sanitation and Hygiene Plus program (also known as the SWASH+ program), conducted in the Nyanza province of Kenya, found no evidence of an association between implementation of the program and absentee rates (O'Reilly, et al., 2006). However, the Kenya study was only a difference-in-differences evaluation while the studies in Egypt and China were randomized controlled trials. Thus, the strongest evidence seems to suggest that handwashing programs are associated with lower absenteeism.

Other evidence on the relationship between public health interventions and education outcomes was provided by a randomized controlled trial of a deworming program in the Busia district of Kenya (Miguel & Kremer, 2004). This study found evidence of improvement in student participation (measured by attendance rates) for the treatment schools, but no significant impact on academic test scores (Miguel & Kremer, 2004). However, the Busia study focused only on the relationship between deworming and academic achievement, and did not address public health interventions relating to sanitation and drinking water infrastructure.

3.3 Public Health Interventions and Households

Like schools, households can play an important role in children's academic performance. Public health research has shown that early childhood health conditions are associated with the development of students' cognitive abilities (Spears & Lamba,

2011). Studies conducted in Jamaica (Nokes, Grantham-Mcgregor, Sawyer, Cooper, & Bundy, 1992) and Brazil (Lorntz, et al., 2006) found that whipworm infection and diarrhea - diseases that are most common among pre-school children with poor access to water and sanitation infrastructure - are associated with impaired cognitive function. This suggests that children's access to sanitation and water facilities is important not only in schools but also in households.

On the other hand, a study by the World Bank in rural communities in Bolivia showed no significant relationship between educational outcomes and a small-scale investment project in household health, water, and sanitation (Newman, et al., 2002). However, a more recent study in India that analyzed a toilet construction campaign showed a modest, positive and statistically significant relationship between household access to improved sanitation and the reading abilities of pre-school children (Spears & Lamba, 2011).[3]

One of the only studies to assess the relation of education outcomes to improved sanitation and drinking water facilities in schools *and* in households was an evaluation of the One Million Initiative, which sought to improve sanitation and drinking water infrastructure in households and schools in rural Mozambique (UNICEF, 2011). Researchers performed a mid-term evaluation of the project and found no association between school enrollment rates and improvements in facilities in schools and households (UNICEF, 2011). The final evaluation for this project, however, is not yet available.

3.4 The Present Study

Overall, the bulk of the existing research suggests that sanitation and drinking water facilities may have an effect on school participation, as well as on cognitive abilities.

[3]Reading abilities were measured by testing letter recognition.

Moreover, since school participation and cognitive skills are closely associated with educational outcomes, it is reasonable to assume that access to improved sanitation and water facilities at schools and in households might have an effect on educational outcomes and, in particular, on test scores. The present study tests this assumption.

My thesis differs from other research in this area in two ways. First, much of the relevant literature focuses on students' enrollment and attendance rates - that is, on the quantity of education. I focus instead on test scores, which is a key measure of academic performance. I conjecture that test scores capture not only school participation, but also the cognitive abilities of children. Both of these factors have been shown to be affected by sanitation and drinking water infrastructure. Second, while most research focuses only on the effect of access to sanitation and water infrastructure at schools, I study the effects of access in schools *and* in households. Thus, I examine the relationship between overall availability of improved sanitation and drinking water facilities and academic achievement.

Chapter 4

Conceptual Framework

The existing literature suggests that children with access to improved sanitation and drinking water infrastructure will be more likely to participate in school and/or will be less likely to experience cognitive impairment. Therefore, I hypothesize that children with improved access to sanitation and drinking water facilities will perform better academically.

The literature also suggests that the association between education outcomes and access to sanitation and water infrastructure is different for girls than for boys. And since girls tend to have worse education outcomes, I also hypothesize that there will be a stronger correlation between test scores and availability of improved sanitation and water facilities for girls than for boys.

My empirical model includes a rich set of controls that are plausibly associated with students' test scores and with the availability of sanitation and water infrastructure. These controls include child characteristics, family characteristics, school characteristics and community characteristics as shown in Figure 4.1 and discussed in the following subsections.

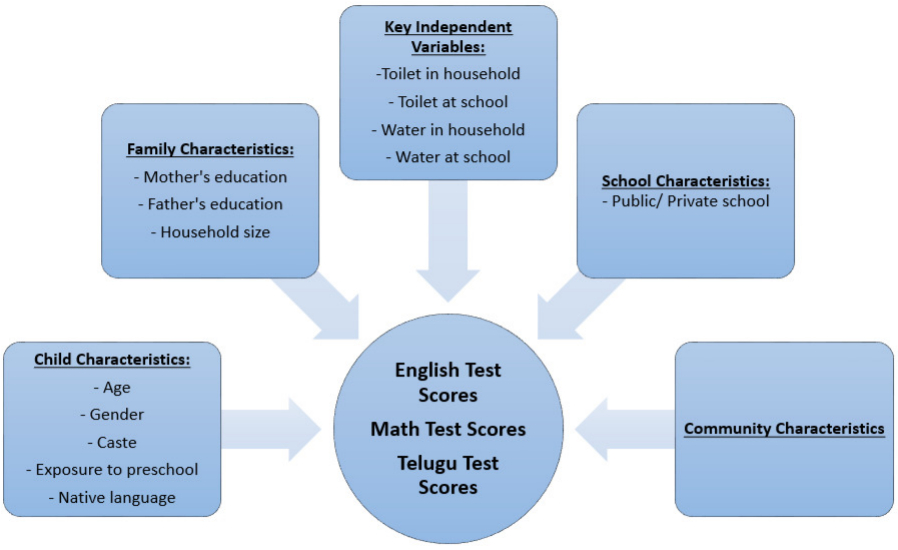

Figure 4.1: Conceptual Framework

4.1 Child Characteristics

India suffers from an educational gender gap: girls generally complete fewer years of school than boys (Kingdon, 2007). Evidence from a nationally representative survey (ASER, 2014) also suggests that girls tend to perform worse than boys in mathematics and language. In addition, as suggested by evidence from Adukia's research (2014), girls may be affected differently by the absence of improved sanitation and water access.

Similarly, evidence suggests that children belonging to castes that are subject to discrimination have poorer education outcomes. For example, Dreze and Kingdon (1999) find lower school participation rates in India among socially disadvantaged castes. In addition, studies of the relationship between cognitive ability and non-

school environmental factors often include controls for the age of the child, which may affect a child's cognitive development and, therefore, be related to the student's test scores (Kingdon, 2005; Adukia, 2014; Simons, 2011).

Much of the literature on early childhood education also suggests that preschool attendance can help to prepare students for school and is an important predictor of academic performance (Magnuson, et al., 2004). In addition, academic performance is determined to some degree by the medium of instruction. Research suggests that students who are taught in their first language tend to perform better than students who are taught in their second language (Saville-Troike, 2012).

In keeping with this literature, I control for age, gender, caste, pre-school exposure and whether the medium of instruction is same as the child's native language.

4.2 Household Characteristics

A study in northern India finds that parental education is positively correlated with children's schooling outcomes (Dreze & Kingdon, 1999). Additionally, large family size – which may require the dilution of family resources available to each child – may affect not only the quality of infrastructure available at home, but also children's academic performance (Downey, 1995). I therefore control for the level of education of the child's mother and the child's father, and also for household size.[4]

[4]Although a family's income is also likely to be associated with the quality of infrastructure at home and with the student's academic performance, I am unable to control for this factor because the survey does not administer questions about family income. Other studies have suggested that family income is likely to be strongly and positively correlated with parents' education and negatively correlated with household size (Kaur, 2000). Thus, my controls for parental education and household size may help to proxy for family income.

4.3 School Characteristics

Earlier studies in India suggest that students in private schools have better academic performance than students in public schools (Kingdon, 2005). Other studies suggest a gap between public and private schools in providing access to infrastructure. In rural India, private schools tend to have poorer infrastructure than public schools (Muralidharan & Kremer, 2007). Therefore, I control for whether the respondent's school is private or public.

4.4 Community Characteristics

Various community characteristics, such as the average income level of the community, the number of schools, infrastructural development, and whether the community is urban or rural, are associated with access to sanitation and water infrastructure, and with education outcomes. For example, research suggests that schools in rural areas tend to have poorer teacher quality, higher student-teacher ratios, and poorer school infrastructure (Cheney, Ruzzi, & Muralidharan, 2005). Therefore, in keeping with studies similar to mine (Adukia, 2014), I control for community fixed effects, such that comparisons in academic outcomes are only made between students who live in the same community. In the Young Lives project, a community is defined by a mandal which is an administrative unit of governance in India. Generally, each mandal is comprised of twenty to sixty villages.[5]

[5]India follows the Panchayati Raj system of governance in rural areas - a three-tier system with villages as the basic level, blocks as the intermediate level, and districts as the highest level of administration (Mondal, 2015). A block, which is the intermediate level, is commonly called a "mandal" in Andhra Pradesh and Telangana. I control for community characteristics at the mandal level.

Chapter 5

Data

My analysis uses data collected by the Young Lives Project, a longitudinal study of childhood poverty administered by the Department for International Development, University of Oxford (Young Lives, 2015).[6] The project has followed 12,000 children in four developing countries since 2002, and has collected household-level data four times (2002, 2006, 2009 and 2014) and school-level data once (2010-11). The survey is conducted at the student level for the school and household surveys. I use the school survey data from 2010-11 and the household survey data from 2009 for India (conducted in the former state of Andhra Pradesh).[7] I use the household survey data to supplement the school survey data. Since the two surveys were held only one year apart, I make a simplifying assumption that the household-level information collected in 2009, such as parents' education and household size, also applies for the time period from December 2010 to March 2011 (i.e. the time period when the school survey was conducted).

The Young Lives Project uses a sentinel site monitoring system to select participating areas (Young Lives, 2011).[8] Sentinel sites are locations that are representative

[6] Unless otherwise stated, all factual assertions in this paragraph are taken from Young Lives (2015).

[7] I excluded any observation from my sample for whom I did not have sufficient data to include them in my empirical analysis.

[8] Unless otherwise stated, all factual assertions in this paragraph are taken from Young Lives (2011).

of a specific type of population for a larger geographical area so that overall sampling is representative of the state's population. In this case, sentinel sites are at the mandal level. Mandals are selected non-randomly; poorer areas are over-sampled and rich areas are excluded. The methodology for choosing mandals ensures the selection of both poor and non-poor districts, and that these districts are representative of the population and geographical areas of the former Andhra Pradesh. Children are then randomly sampled from within each mandal.[9]

5.1 Dependent Variables

I use test scores to measure the academic performance of the students in my sample.[10] The 2010-11 Young Lives school survey administered a test to all survey respondents on three subjects, each measuring attainment of age-appropriate mathematics and language skills:

[9]In 2002, 2000 infants and 1000 children aged 7 to 8 years were randomly selected across all the sentinel sites. Of these, 954 children (who were infants in 2002) were surveyed in 2010-11 as part of the school survey component. My empirical model includes only the children surveyed for the school survey component. Of these 954 observations, three observations were dropped due to duplication of data for the same children (Unique IDs for the children were the same). Among the remaining 951 observations, five were deleted as there were missing data for test scores for one or more subjects. However, some students were not administered the English test as they were never taught English. These students have not been excluded from my analytical sample for reasons specified in footnote 13. Another 21 observations were dropped due to missing values for at least one of the key independent variables. Of the remaining 925 observations, 33 more were dropped because they had missing values for at least one of the control variables. Thus, the final analytic sample includes 892 observations.

[10]My analysis uses test scores to operationalize learning outcomes, as these scores are a commonly used indicator of student achievement (Scheerens, Luyten, & Ravens, 2010). While test scores are not a perfect measure of a student's learning outcome, they are able to capture a different dimension of a student's academic performance as compared to other education outcomes, such as school attendance, enrollment rates, and completion rates, as these latter metrics capture only the quantity of education.

- Mathematics: This test included questions about number recognition, sequences and series, and basic arithmetic operations and their application (Young Lives, 2010).

- English: This test included questions about word and object recognition, matching of sentences with their pictorial representations, and reading comprehension (Young Lives, 2010).

- Telugu: Telugu is the native language of Andhra Pradesh, and is a common medium of instruction in Andhra Pradesh schools. The content of the Telugu test was similar to that of the English test (Young lives, 2010).

5.2 Independent Variables

Among the many variables that measure access to and/or quality of sanitation and drinking water facilities in schools and households, I choose variables that reflect the availability of at least one usable toilet and one usable source of drinking water in the school and household. A majority of the students in my sample attend schools that have no usable toilets, and also live in households with no usable toilets, making access to even a single toilet a useful measure of access to sanitation facilities in my sample.

However, one usable toilet only captures a binary relationship for availability of toilets in schools, and doesn't capture change in test scores as number of toilets increase. Therefore, I also conduct a secondary analysis using a continuous measure of number of toilets in schools. Here, my analysis is divided among students who attend schools that have sex-specific toilets and students who attend schools that do not have sex-specific toilets. For students who attend schools that have sex-specific

toilets, I include a variable measuring the number of boys' toilets and a variable measuring the number of girls' toilets.

For the control variables, I measure child characteristics, school characteristics and community characteristics using the 2010-11 school survey data, and I measure family characteristics using the 2009 household survey data.[11] Table 5.1 describes each variable included in my model.

[11]For my mother's education and father's education variables, I treated values labeled as "Adult Literacy" (coded equal to 29) and "Religious Education" (coded equal to 28) as missing, and used single imputation to redefine values for these observations.

Table 5.1: Variable definitions

Variables	Definitions
Dependent Variable	
English Test Scores	A continuous variable measuring the proportion of answers that were correct for the English test.
Math Test Scores	A continuous variable measuring the proportions of answers that were correct for the Mathematics test.
Telugu Test Scores	A continuous variable measuring the proportions of answers that were correct for the Telugu test.
Key Independent Variables- Primary Analysis	
Toilet in household	A dichotomous variable that indicates whether the child has access to a usable toilet at home.
Toilet at school	A dichotomous variable that indicates whether the child has access to at least one usable toilet at school.
Water in household	A dichotomous variable that indicates whether the child has access to clean drinking water at home.
Water at school	A dichotomous variable that indicates whether the child has access to clean drinking water at school.
Key Independent Variables- Secondary Analysis	
Number of working toilets at school – for boys only	A continuous variable that measures the number of working toilets for boys at schools that have sex-specific toilets.
Number of working toilets at school – for girls only	A continuous variable that measures the number of working toilets for girls at schools that have sex-specific toilets.
Number of working toilets at school – if not segregated	A continuous variable that measures the number of working toilets at schools that do not have sex-specific toilets.
Control Variables	
Child Characteristics	
Age	A continuous variable that measures the age of the child.
Female	A dichotomous variable that indicates whether child is female.
Socially disadvantaged	A dichotomous variable that indicates whether the child belongs to a Scheduled Caste/Scheduled Tribe/ Other Backward Class.*
Native language same as medium of instruction	A dichotomous variable that indicates whether the native language of the child is same as the medium of instruction.
Exposure to preschool	A dichotomous variable that indicates whether the child attended pre-school.

Variables	**Definitions**
Control Variables	
Family Characteristics	
Mother's education	A continuous variable that measures the number of completed years of schooling by the mother.
Father's education	A continuous variable that measures the number of completed years of schooling by the father.
Household Size	A continuous variable that measures the number of members of the household of the child.
School Characteristics	
Private school	A dichotomous variable that indicates whether the school is privately owned or is a government school.
Community Characteristics	
Mandal	A series of dichotomous variables for each of the 20 sentinel sites or mandals.
*Scheduled Castes, Scheduled Tribes and Other Backward Classes are categories defined by the Indian Constitution, comprising people belonging to socially discriminated castes or backgrounds (Planning Commission).	

Chapter 6

Methods

I estimate an Ordinary Least Squares (OLS) model with mandal fixed effects. The fixed effects specification controls for all mandal-level characteristics, which include whether the site is rural or urban, the size of the mandal and access to schools in the block. I estimate separate empirical models to study the association between students' test performance for each subject and their access to improved sanitation and drinking water facilities at school and at home. Specifically, I estimate the following regression equations:

$$(EnglishTestScore)_{mi} = \beta_0 + \beta_1(ToiletinHousehold)_{mi} + \beta_2(ToiletatSchool)_{mi} + \beta_3(WaterinHousehold)_{mi} + \beta_4(WateratSchool)_{mi} + \beta_5(ChildCharacteristics)_{mi} + \beta_6(FamilyCharacteristics)_{mi} + \beta_7(SchoolCharacteristics)_{mi} + \alpha_m + \mu_{mi} \quad \textbf{(1)}$$

$$(MathTestScore)_{mi} = \beta_0 + \beta_1(ToiletinHousehold)_{mi} + \beta_2(ToiletatSchool)_{mi} + \beta_3(WaterinHousehold)_{mi} + \beta_4(WateratSchool)_{mi} + \beta_5(ChildCharacteristics)_{mi} + \beta_6(FamilyCharacteristics)_{mi} + \beta_7(SchoolCharacteristics)_{mi} + \alpha_m + \mu_{mi} \quad \textbf{(2)}$$

$$(TeluguTestScore)_{mi} = \beta_0 + \beta_1(ToiletinHousehold)_{mi} + \beta_2(ToiletatSchool)_{mi} + \beta_3(WaterinHousehold)_{mi} + \beta_4(WateratSchool)_{mi} + \beta_5(ChildCharacteristics)_{mi} + \beta_6(FamilyCharacteristics)_{mi} + \beta_7(SchoolCharacteristics)_{mi} + \alpha_m + \mu_{mi}, \quad \textbf{(3)}$$

where α represents the series of mandal fixed effects dummy variables, μ represents the error term, and the subscripts "i" and "m" are indices for each individual child and mandal respectively.

Chapter 7

Descriptive Results

Table 7.1 below provides descriptive statistics for the dependent variables, key independent variables, and control variables included in my model.[12] My analytical sample includes 892 observations.

7.1 Dependent Variables

The tests administered by the Young Lives Project assessed language and mathematical skills. Test scores are converted into percentage terms with a maximum score of 100. The data show disparities in learning outcomes for each test.

7.1.1 Language Skills

English and the native language of Andhra Pradesh, Telugu, were tested using 29 questions and 32 questions, respectively.[13] On average, students in the sample tended to perform better in Telugu than in English, with an average score of 64.68 percent in Telugu and an average score of 48.03 percent in English. The standard deviation for Telugu is higher than for English (22.41 for Telugu; 19.11 for English). However,

[12]The empirical model does not use any weighting scheme because the Young Lives survey does not include weights for its sample.

[13]In this survey, the English test was administered only to those students who were either being taught English in school or had exposure to English outside school (Young Lives, 2010). Of the 892 students in my analytical sample, 39 did not take the English test. Therefore, the sample for the regressions that analyze English test scores is restricted to 853 students.

because there is a substantial difference in the means of the test scores, I calculate the coefficient of variation (given by mean/standard deviation) to assess the relative magnitude of variation in the sample. As a consequence, the relative ordering changes, as the coefficient of variation for English is higher than for Telugu (0.40 for English; 0.35 for Telugu). In addition, some students were able to achieve the maximum score in Telugu, but no one achieved the maximum score in English; 86.21 points (out of 100) was the highest English test score.

7.1.2 Mathematics Skills

The mathematics test included 21 questions. The average Mathematics score was 54.91 percent (23.488) with a standard deviation higher than for the Telugu and English test scores. In order to compare variations in test scores in the sample, I also calculate the coefficient of variation for Mathematics. In this case, the relative ordering of the tests remains the same, as the coefficient of variation for Mathematics is the highest (0.428). In addition, no one achieved the maximum score on the Mathematics test. The highest score was 95.24 percent.

7.2 Key Independent Variables

Figure 7.1 and table 7.1 indicate large disparities in access to infrastructure. Among the households surveyed, only 25.8 percent had access to a good quality toilet, but 96.3 percent had access to clean drinking water. Moreover, among the households surveyed, only 25.8 percent had access to a good quality toilet, but 96.3 percent had access to clean drinking water.

These data indicate that there is much greater access to drinking water facilities than to toilets. The data also suggest that households perform better than schools in

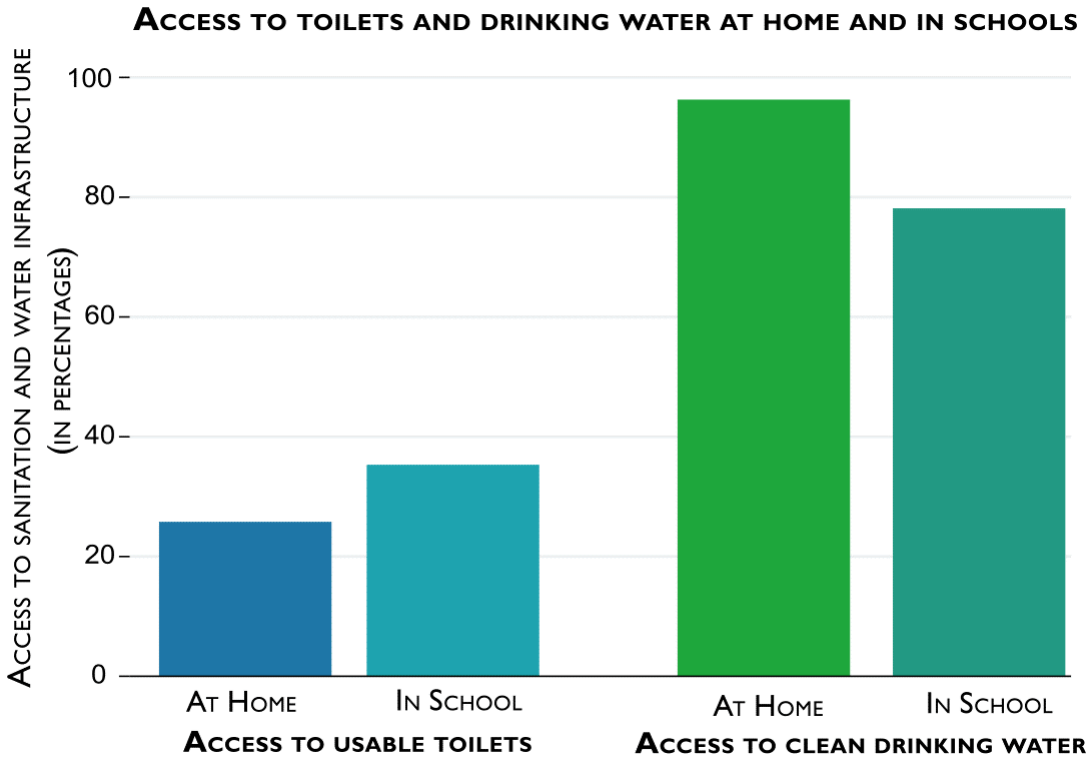

Figure 7.1: Water and Sanitation

terms of providing access to clean drinking water. On the other hand, according to my metric for toilet access (provision of at least one usable toilet), households perform worse than schools. Of course, one might argue that providing only one usable toilet in a school is not enough. My secondary analysis will therefore include a continuous variable that measures the number of usable toilets in a school.

Table 7.1: Descriptive statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Dependent Variables					
English Test Score	853	48.03	19.11	0	86.21
Math Test Score	892	54.91	23.49	0	95.24
Telugu Test Score	892	64.68	22.41	0	100
Key Independent Variables - Primary Analysis					
Toilet at Home (0,1)	892	0.26	0.44	0	1
Toilet in School (0,1)	892	0.35	0.48	0	1
Water at Home (0,1)	892	0.96	0.19	0	1
Water in School (0,1)	892	0.78	0.41	0	1
Key Independent Variables - Secondary Analysis					
Number of working toilets at school - for boys only	320	3.21	4.02	0	25
Number of working toilets at school - for girls only	342	3.69	6.91	0	65
Number of working toilets at school - if not segregated	215	0.98	1.25	0	8
Control Variables: Child Characteristics					
Age (years)	892	9.23	0.3	8.52	10.02
Female (0,1)	892	0.46	0.5	0	1
Disadvantaged Caste (0,1)	892	0.83	0.37	0	1
Native language same as medium of instruction (0,1)	892	0.67	0.47	0	1
Exposure to Pre-school (0,1)	892	0.87	0.34	0	1
Control Variables: Household Characteristics					
Mother's education (years)	892	3	3.99	0	14
Father's education (years)	892	4.96	4.68	0	14
Household Size	892	5.73	2.47	2	28
Control Variables: School Characteristics					
Private School (0,1)	892	0.37	0.48	0	1

Chapter 8

Regression Results

I employ a multiple linear regression model using my full suite of control variables and mandal fixed effects. My regression results are presented in Tables 8.1 through 8.4 below. I report regression coefficients and robust standard errors for each variable in the model, with the exception of mandal fixed effects indicator variables. In all tables, each column corresponds to one of my three dependent variables - English test score, Math test score, and Telugu test score.

The regressions whose results are presented in Table 8.1 include all four primary key independent variables - at least one usable toilet in school, a usable toilet at home, a usable drinking water source in school, and a usable drinking water source at home.[14] Overall, the results for these analyses indicate a significant correlation only between sanitation and drinking water facilities and English test scores. No correlation is observed for Math and Telugu test scores, which may be a function of the more limited scope for improvement in these subjects for the students in the sample; on average, students perform the worst in English. Therefore, there may be a greater margin for improvement in English than in Math or Telugu. In my conclusion,

[14] I also estimated regressions in which I included only the key independent variables related to toilet access (at home and school), and only the variables related to water (at home and school). The findings for these regressions were not qualitatively different from the results for the regressions that included all of my primary key independent variables. I also estimated separate regressions that had only key independent variables related to households and only key independent variables related to schools. The findings for these specifications were also similar to the results reported here. See tables 10.1 through 10.4 in the Appendix for results of these specifications.

I suggest other possible reasons for the difference in findings for English test scores and for Math and Telugu test scores.

I find English test scores to be significantly associated with the indicator for at least one toilet in school and approaching conventional significance levels of association with indicator for a usable drinking water source at school. The magnitude of the coefficients are of moderate size – having at least one usable toilet in school, as opposed to none, is associated with an improvement of 3.13 percentage points in English test scores for students on average; having a usable drinking water source in school, as opposed to none, is associated with an improvement of 2.54 percentage points in English test scores for students on average. As the average English test score is 48 percentage points, an addition of 2.54-3.13 points suggests a limited association.

The indicators for sanitation and drinking water facilities at home are insignificant, suggesting that school sanitation and water infrastructure may be of greater importance than household sanitation and water infrastructure for students' academic performance. Joint significance tests for sanitation and drinking water facilities at school, and for sanitation and drinking water facilities at home, validate this hypothesis - the variables measuring school infrastructure are jointly significant, while the variables measuring household infrastructure are jointly insignificant.

Joint significance tests for access to toilets at home and school, and for access to drinking water at home and school, also suggest that access to toilets may be more important than access to drinking water for students' academic performance. This finding may reflect the difference in extent of variation in access to toilets and extent of variation in access to drinking water. While most students have access to a usable drinking water source at home and school, a majority of the students have no usable toilets at home or attend schools with no usable toilets.

Table 8.1: Base specifications

VARIABLES	**English test scores**	**Math test scores**	**Telugu test scores**
Key Independent Variables			
Usable toilet at home	-0.74	1.93	-0.62
	(1.58)	(2.47)	(2.23)
At least one usable toilet in school	3.13**	1.49	-1.03
	(1.36)	(1.96)	(1.86)
Usable drinking water source at home	-1.50	-1.47	2.20
	(3.61)	(4.72)	(4.19)
Usable drinking water source in school	2.54^{+}	1.40	1.57
	(1.67)	(2.23)	(2.15)
Control Variables - Child Characteristics			
Age	3.97**	8.57***	9.84***
	(1.73)	(2.35)	(2.44)
Female	0.90	1.02	1.84
	(1.04)	(1.41)	(1.36)
Child belongs to a disadvantaged caste	1.05	-0.32	4.43**
	(1.37)	(2.14)	(1.94)
Medium of instruction is same as native language	-3.71***	1.44	3.16*
	(1.38)	(1.84)	(1.68)
Since 3yr of age has YL child attended pre-school	1.12	3.79	0.55
	(1.57)	(2.36)	(2.10)
Control Variables - Family Characteristics			
Mother's education	0.68***	1.07***	0.83***
	(0.18)	(0.25)	(0.23)
Father's education	0.40***	0.47**	0.48***
	(0.14)	(0.20)	(0.18)
Household size	0.12	0.40	0.56**
	(0.20)	(0.30)	(0.26)
Control Variables- School Characteristics			
Child goes to private school	13.03***	2.67	1.35
	(1.56)	(2.33)	(1.91)
Constant	0.25	-38.18*	-44.64*
	(16.68)	(22.40)	(23.15)

VARIABLES	English test scores	Math test scores	Telugu test scores
Joint Significance Tests			
Joint test, F-statistic (toilet at home, toilet in school)	2.79**	0.63	0.19
Prob > F	(0.06)	(0.53)	(0.83)
Joint test, F-statistic (water at home, water in school)	1.25	0.25	0.44
Prob > F	(0.286)	(0.78)	(0.65)
Joint test, F-statistic (water at home, toilet at home)	0.21	0.34	0.16
Prob > F (water at home, toilet at home)	(0.812)	(0.71)	(0.86)
Joint test, F-statistic (water in school, toilet in school)	3.48**	0.43	0.45
Prob > F	(0.03)	(0.65)	(0.64)
Observations	853	892	892
R-squared	0.42	0.26	0.23
Robust standard errors in parentheses for all coefficients, and p-values in parentheses for joint significance tests. *** p<0.01, ** p<0.05, * p<0.1, $^{+}$ p<0.15 Mandal Dummy variables have been included as fixed effects			

The regressions whose results are reported in Table 8.2 include interactions between gender and a usable toilet at home, and between gender and at least one usable toilet in school. These regressions test whether there is a different correlation between student test scores and having access to toilets for boys and girls. I estimate these models to test the "menstruation hypothesis", which suggests that girls are often absent from schools during menses due to lack of sanitation facilities in schools.

The findings from this analysis suggest that the association between access to toilets at school and English test scores is statistically significant for boys, as suggested by the fact that the coefficient on indicator for toilet in school is marginally significant; and for girls, as suggested by the fact that the coefficients on indicator for toilet in school and its interaction with indicator for girls are jointly significant. However, while these relationships may be statistically significant, the magnitude of the coefficients are of moderate substantive significance. Girls attending schools with at least one usable toilet score 3.9 percentage points higher in English on average than girls attending schools with no toilets, and boys attending schools with at least one usable toilet score 2.5 percentage points higher in English on average than boys attending schools with no toilets. However, the association of access to toilets with English test scores is not significantly different for boys and girls. Therefore, this finding contradicts the "menstruation hypothesis" to some extent, as girls do not appear to be harmed significantly more than boys by the absence of a toilet in school.

Table 8.1: Regressions with gender interactions

VARIABLES	**English test scores**	**Math test scores**	**Telugu test scores**
Key Independent Variables			
Usable toilet at home	-1.486	0.950	-3.752
	(1.761)	(3.031)	(2.801)
At least one usable toilet in school	2.515^{+}	-0.51	-2.975
	(1.648)	(2.347)	(2.262)
Usable drinking water source at home	-1.656	-1.909	1.669
	(3.614)	(4.743)	(4.169)
Usable drinking water source in school	2.557^{+}	1.488	1.558
	(1.673)	(2.227)	(2.149)
Interaction: female*toilet at home	1.471	1.796	6.227*
	(2.299)	(3.616)	(3.379)
Interaction: female*toilet in school	1.433	4.839^{+}	4.562^{+}
	(2.234)	(3.256)	(3.035)
Control Variables - Child Characteristics			
Age	3.963**	8.590***	9.824***
	(1.738)	(2.357)	(2.438)
Female	-0.0268	-1.152	-1.408
	(1.448)	(1.850)	(1.750)
Child belongs to a disadvantaged caste	0.933	-0.617	3.985**
	(1.370)	(2.127)	(1.926)
Medium of instruction is same as native language	-3.715***	1.455	3.161*
	(1.380)	(1.834)	(1.680)
Since 3yr of age has the child attended pre-school	1.115	3.735^{+}	0.506
	(1.574)	(2.357)	(2.081)
Control Variables - Family Characteristics			
Mother's education	0.675***	1.053***	0.803***
	(0.176)	(0.251)	(0.229)
Father's education	0.410***	0.488**	0.513***
	(0.136)	(0.196)	(0.175)
Household size	0.119	0.403	0.540**
	(0.196)	(0.299)	(0.261)
Control Variables - School Characteristics			
Child attends private school	12.92***	2.411	0.953
	(1.577)	(2.336)	(1.904)
Constant	1.038	-36.66	-41.89*
	(16.73)	(22.42)	(23.09)

VARIABLES	English test scores	Math test scores	Telugu test scores
Joint Significance Tests			
Joint test, F-statistic (toilet at home, interaction: female*toilet at home)	0.40	0.40	1.75
Prob > F	(0.67)	(0.67)	(0.18)
Joint test, F-statistic (toilet in school, interaction: female*toilet in school)	2.80*	1.36	1.28
Prob > F	(0.06)	(0.26)	(0.28)
Observations	853	892	892
R-squared	0.42	0.26	0.24
Robust standard errors in parentheses for all coefficients, and p-values in parentheses for joint significance tests. *** p<0.01, ** p<0.05, * p<0.1, $^{+}$ p<0.15 Mandal Dummy variables have been included as entity fixed effects			

The regressions presented in Table 8.3 and 8.4 employ a second, supplementary set of key independent variables. In Table 8.3, instead of an indicator for at least one toilet in school, I use a continuous variable reflecting the number of toilets in a school, segregated by gender, for students who attend schools with sex-specific toilets. Similarly, in Table 8.4, I use a continuous variable reflecting the number of toilets, not segregated by gender, for students who attend schools with common toilets for boys and girls.

The findings from the regressions shown in Table 8.3 suggest that the number of usable toilets is significantly associated with English, Math, and Telugu test scores for attending schools that have sex-specific toilets for boys. However, the magnitudes of the coefficients are quite small to be of any practical importance - the addition of one boys' toilet to a school is associated with increase in student test scores of 0.98

percentage points for Math and 0.79 percentage points for English and 1.1 percentage points for Telugu, all of which are only marginal increases. No such association is found for attending schools that have girls' toilets. This may be the case for various reasons - poor maintenance of girls' toilets (which require more resources than boys' toilets); absence of other sanitation facilities, such as running water and sanitary napkins; or concerns over a lack of privacy. Finally, in Table 8.4, no significant relation was observed between the number of toilets and students' test scores for those attending schools that do not have sex-specific toilets.[15]

[15]In an additional analysis, I also estimate the relation between test scores and access to at least one sex-specific toilet, and find no significance.

Table 8.1: Regressions with sex-specific toilets measures

VARIABLES	English test scores	Math test scores	Telugu test scores
Key Independent Variables			
Usable toilet at home	-2.189	2.001	-0.593
	(2.637)	(4.051)	(4.061)
Number of working latrines/ toilets - For boys only	0.979***	0.787*	1.085**
	(0.223)	(0.427)	(0.450)
Number of working latrines/ toilets - For girls only	-0.0619	-0.0348	-0.190
	(0.105)	(0.247)	(0.213)
Usable drinking water source at home	-2.039	5.303	9.364
	(4.024)	(16.11)	(12.69)
Usable drinking water source in school	-4.036	-6.495	-11.27**
	(4.063)	(6.509)	(5.524)
Control Variables - Child Characteristics			
Age	0.852	4.021	6.518^{+}
	(2.275)	(3.767)	(4.291)
Female	2.627*	5.588**	6.179**
	(1.579)	(2.397)	(2.398)
Child belongs to a disadvantaged caste	3.231*	2.188	10.32***
	(1.723)	(2.919)	(3.122)
Medium of instruction is same as native language	-2.811^{+}	0.978	4.133
	(1.937)	(2.915)	(2.915)
Since 3yr of age has the child attended pre-school	1.517	6.841*	2.961
	(2.141)	(3.788)	(3.411)
Control Variables - Family Characteristics			
Mother's education	0.680***	0.610^{+}	0.326
	(0.221)	(0.389)	(0.370)
Father's education	0.465**	0.808**	0.494
	(0.193)	(0.328)	(0.315)
Household size	0.199	0.558	0.355
	(0.246)	(0.484)	(0.320)
Control Variables - School Characteristics			
Child attends private school	8.582***	3.858	-2.333
	(2.976)	(5.381)	(4.598)
Constant	34.48	-5.230^{+}	-16.47
	(23.10)	(39.04)	(43.71)
Observations	318	320	320
R-squared	0.52	0.34	0.41

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, $^{+}$ p<0.15
Mandal Dummy variables have been included as fixed effects

Table 8.2: Regressions without sex-specific toilets measures

VARIABLES	English test scores	Math test scores	Telugu test scores
Key Independent Variables			
Usable toilet at home	6.448*	3.855	-0.171
	(3.528)	(5.553)	(4.326)
Number of working latrines/ toilets - For students and staff (if not segregated)	-0.421	-2.158	-1.441
	(1.130)	(1.779)	(1.794)
Usable drinking water source at home	-12.18**	-13.36*	-7.951
	(5.177)	(7.070)	(6.177)
Usable drinking water source in school	0.442	-4.037	2.745
	(3.981)	(5.698)	(4.460)
Control Variables - Child Characteristics			
Age	7.586*	12.00**	15.78***
	(4.032)	(5.201)	(5.024)
Female	-2.011	-1.507	0.648
	(2.339)	(3.147)	(2.980)
Child belongs to a disadvantaged caste	-4.398	-5.591	-1.874
	(4.076)	(5.702)	(4.680)
Medium of instruction is same as native language	-3.440	2.497	7.246*
	(3.343)	(4.714)	(3.822)
Since 3yr of age has the child attended pre-school	1.732	8.038^{+}	4.674
	(2.882)	(5.129)	(4.480)
Control Variables - Family Characteristics			
Mother's education	0.0393	0.670	1.216***
	(0.405)	(0.509)	(0.405)
Father's education	0.588**	0.715*	0.628*
	(0.278)	(0.405)	(0.329)
Household size	-0.312	0.245	0.0947
	(0.447)	(0.639)	(0.542)
Control Variables - School Characteristics			
Child attends private school	20.12***	7.700	4.391
	(4.873)	(6.876)	(5.998)
Constant	-15.55^{+}	-52.93	-90.74*
	(39.39)	(52.04)	(49.07)
Observations	209	215	215
R-squared	0.49	0.31	0.33

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, $^{+}$ p<0.15
Mandal Dummy variables have been included as fixed effects

In summary, my findings indicate that there is a limited relationship between access to toilets and drinking water at school and students' English test scores. This suggests that school infrastructure may play a small role, at best, in students' academic performance. Moreover, I find no association for most other regressions.

Chapter 9

Discussion

This study explores the relationship in India between access to improved sanitation and drinking water facilities at students' homes and schools and their academic performance in English, Mathematics, and Telugu. My findings might have helped to show whether allocation of resources for the current national program for improvement in access to sanitation and water infrastructure in India called *Swachh Bharat Abhiyan* is associated with improving education outcomes. I find only a modest, positive relationship between access to usable toilets in schools and students' English test scores. I find no relationship for other subjects, or between access to improved sanitation facilities at home and students' test scores, or between access to improved drinking water facilities and students' test scores.

My findings regarding access to improved sanitation infrastructure in schools are in line with my hypotheses. However, the lack of any relationship between access to improved sanitation facilities at home and students' test scores, and between access to improved drinking water facilities and students' test scores, contradict my hypotheses, as well as the majority of prior research in this area which suggests a modest, positive relationship between access to drinking water and sanitation infrastructure and education outcomes (Koolwal & van de Walle, 2010; Spears & Lamba, 2011). However, it is important to note that my study is different from the existing literature, since it examines the association between improved sanitation and drinking water infrastructure and students' test scores, rather than between sanitation and drinking

water infrastructure and students' enrollment rates, school participation rates, and other measures of academic performance. This may explain the contrast between my findings and those of prior research.

In addition, the most prominent prior study that established an association between household sanitation infrastructure and learning outcomes looked at children aged 0 to 5 years, thus assessing only early childhood behaviours (Spears & Lamba, 2011). It is possible that this association fades over time, such that I see no association when I look at school-going children between the ages of 8 and 10. My research also contrasts with another study that finds a positive association between access to drinking water at home and girls' education outcomes. The authors of that study interpret their results to suggest that in the absence of drinking water at home, girls have to travel long distances to collect water at the expense of attending school. Their analytical sample consists of girls aged 5 to 19 years, an age group that includes older-aged girls who can travel alone to collect water. I theorize that these older-aged girls may be driving this correlation. As my sample consists of children aged between 8 and 10 years old, it may not capture this mechanism, which could explain why my results are different.

My study has some important limitations. The association found between school sanitation infrastructure and English test scores is likely to be affected by omitted variable bias. I was unable to control for school quality, which may be positively associated with access to improved sanitation facilities in school. I theorize that students in high-quality schools are likely to perform better in English than students in low-quality schools. This is because the availability of good English teachers in India is limited compared with availability of good Math or Telugu teachers. Students who attend lower-quality schools are also likely to belong to families with comparatively limited education, especially in English. These students' lack of education in English

and the parents' inability to provide additional resources to support their education amplifies their weakness in English. As English test score is positively correlated with access to sanitation in school and likely to be positively correlated with school quality, there is likely to be an upward bias on the coefficients for English test scores. This means that my estimates of the correlation between English test scores and access to sanitation facilities in schools may be artificially higher than the true coefficient of correlation. On the other hand, school quality may not be as strongly associated with Telugu and Math scores because students may be more likely to have access to other resources for supporting their Math and Telugu education.

In addition, although the Young Lives provides information on a wide range of factors related to my research question, these data also have limitations. For example, I was unable to control for the family income of the students in my sample, which may also be a driver of omitted variable bias (Kaur, 2000). Family income may be positively associated with students' test scores, as students from higher income families may have additional resources for studying and a more supportive home environment. Family income may also be positively correlated with access to sanitation and water infrastructure, as students in higher income families are more likely to have usable toilets and clean drinking water at home, and are more likely to attend schools which have usable toilets and clean drinking water. Therefore, the estimated coefficients are likely to be biased upward. However, household size and parents' education, control variables included in my model, are likely to be strong predictors of family income, and I use these controls as proxies for family wealth. Nonetheless, these controls may not capture the full extent of student segregation by income in schools.

This review of the limitations of my study suggests that there is substantial scope for further research in assessing the relationship between access to improved drinking water and sanitation facilities and education outcomes. My study is restricted to a

small sample of sites in the former state of Andhra Pradesh. A comprehensive study of the importance for students of sanitation and water infrastructure in Indian schools and households would need to be countrywide.

My finding of a moderate association between improved sanitation infrastructure in schools and English test scores does provide some support for focusing more efforts to improve sanitation facilities in schools in India. However, policy makers should approach my conclusion cautiously due to the strong possibility of omitted variable bias. A nationwide study may be able to paint a more informative picture and a clearer guide to policy development.

Chapter 10

Appendix

Table 10.1: Regressions for sanitation and drinking water infrastructure at home

VARIABLES	**English test scores**	**Math test scores**	**Telugu test scores**
Key Independent Variables			
Usable toilet at home	-0.695	1.969	-0.516
	(1.607)	(2.475)	(2.242)
Usable drinking water source at home	-1.354	-1.456	2.161
	(3.550)	(4.717)	(4.164)
Control Variables - Child Characteristics			
Age	3.989**	8.581***	9.999***
	(1.737)	(2.353)	(2.435)
Female	0.768	0.941	1.753
	(1.047)	(1.417)	(1.368)
Child belongs to a disadvantaged caste	1.228	-0.226	4.450**
	(1.384)	(2.142)	(1.935)
Medium of instruction is same as native language	-3.843***	1.376	3.129*
	(1.384)	(1.831)	(1.669)
Since 3yr of age has the child attended pre-school	0.926	3.710	0.728
	(1.566)	(2.344)	(2.091)
Control Variables - Family Characteristics			
Mother's education	0.700***	1.078***	0.829***
	(0.177)	(0.251)	(0.227)
Father's education	0.416***	0.480**	0.485***
	(0.137)	(0.197)	(0.176)
Household size	0.143	0.411	0.540**
	(0.193)	(0.295)	(0.262)
Control Variables - School Characteristics			
Child attends private school	14.81***	3.556*	1.231
	(1.432)	(2.107)	(1.704)
Constant	2.310	-37.05*	-45.21*
	(16.62)	(22.38)	(23.14)
Observations	853	892	892
R-squared	0.42	0.26	0.23

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Mandal Dummy variables have been included as entity fixed effects

Table 10.2: Regressions for sanitation and drinking water infrastructure in school

VARIABLES	**English test scores**	**Math test scores**	**Telugu test scores**
Key Independent Variables			
At least one usable toilet in school	3.132**	1.466	-1.019
	(1.363)	(1.966)	(1.857)
Usable drinking water source in school	2.506	1.472	1.542
	(1.665)	(2.233)	(2.151)
Control Variables - Child Characteristics			
Age	4.002**	8.490***	9.882***
	(1.737)	(2.358)	(2.431)
Female	0.876	1.072	1.827
	(1.038)	(1.408)	(1.353)
Child belongs to a disadvantaged caste	1.080	-0.374	4.439**
	(1.370)	(2.132)	(1.936)
Medium of instruction is same as native language	-3.729***	1.410	3.204*
	(1.371)	(1.834)	(1.676)
Since 3yr of age has the child attended pre-school	1.056	3.713	0.669
	(1.599)	(2.350)	(2.092)
Control Variables - Family Characteristics			
Mother's education	0.669***	1.094***	0.827***
	(0.172)	(0.247)	(0.224)
Father's education	0.395***	0.491**	0.478***
	(0.135)	(0.195)	(0.174)
Household size	0.117	0.419	0.559**
	(0.195)	(0.297)	(0.264)
Control Variables - School Characteristics			
Child attends private school	12.95***	2.792	1.346
	(1.557)	(2.318)	(1.903)
Constant	-1.457	-38.60*	-43.16*
	(16.39)	(21.96)	(22.71)
Observations	853	892	892
R-squared	0.42	0.26	0.23

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Mandal Dummy variables have been included as entity fixed effects

Table 10.3: Regressions for sanitation infrastructure at home and school

VARIABLES	**English test scores**	**Math test scores**	**Telugu test scores**
Key Independent Variables			
Usable toilet at home	-0.678	1.966	-0.495
	(1.592)	(2.461)	(2.236)
At least one usable toilet in school	2.777**	1.290	-1.241
	(1.375)	(1.953)	(1.869)
Control Variables - Child Characteristics			
Age	4.133**	8.634***	9.956***
	(1.739)	(2.350)	(2.437)
Female	0.775	0.936	1.761
	(1.044)	(1.417)	(1.369)
Child belongs to a disadvantaged caste	1.161	-0.259	4.479**
	(1.379)	(2.140)	(1.938)
Medium of instruction is same as native language	-3.808***	1.368	3.149*
	(1.370)	(1.825)	(1.675)
Since 3yr of age has the child attended pre-school	1.147	3.755	0.727
	(1.598)	(2.354)	(2.096)
Control Variables - Family Characteristics			
Mother's education	0.687***	1.069***	0.839***
	(0.176)	(0.251)	(0.228)
Father's education	0.411***	0.477**	0.487***
	(0.137)	(0.197)	(0.176)
Household size	0.113	0.396	0.555**
	(0.195)	(0.297)	(0.264)
Control Variables - School Characteristics			
Child attends private school	13.65***	2.994	1.786
	(1.479)	(2.250)	(1.861)
Constant	-0.822	-39.07*	-42.68*
	(16.38)	(21.87)	(22.80)
Observations	853	892	892
R-squared	0.42	0.26	0.23

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Mandal Dummy variables have been included as entity fixed effects

Table 10.4: Regressions for drinking water infrastructure at home and school

VARIABLES	**English test scores**	**Math test scores**	**Telugu test scores**
Key Independent Variables			
Usable drinking water source at home	-1.511	-1.311	2.143
	(3.576)	(4.701)	(4.159)
Usable drinking water source in school	2.032	1.241	1.705
	(1.652)	(2.216)	(2.154)
Control Variables - Child Characteristics			
Age	3.874**	8.440***	9.902***
	(1.735)	(2.359)	(2.429)
Female	0.845	1.065	1.827
	(1.044)	(1.408)	(1.353)
Child belongs to a disadvantaged caste	1.175	-0.327	4.412**
	(1.376)	(2.134)	(1.932)
Medium of instruction is same as native language	-3.783***	1.405	3.186*
	(1.385)	(1.839)	(1.671)
Since 3yr of age has the child attended pre-school	0.830	3.655	0.641
	(1.566)	(2.342)	(2.085)
Control Variables - Family Characteristics			
Mother's education	0.685***	1.104***	0.817***
	(0.174)	(0.247)	(0.223)
Father's education	0.402***	0.494**	0.476***
	(0.135)	(0.195)	(0.175)
Household size	0.148	0.435	0.547**
	(0.194)	(0.295)	(0.262)
Control Variables - School Characteristics			
Child attends private school	14.35***	3.470	0.848
	(1.501)	(2.153)	(1.741)
Constant	2.065	-36.55	-45.55**
	(16.63)	(22.47)	(23.08)
Observations	853	892	892
R-squared	0.42	0.26	0.23

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Mandal Dummy variables have been included as entity fixed effects

Chapter 11

References

Adukia, A. (2014). Sanitation and Education. http://scholar.harvard.edu/files/adukia/files/adukia_sanitation_and_education.pdf

Aiyar, S. S. (2011). Andhra Pradesh Fastest Improver in Economic Freedom. In S. S. Aiyar, B. Debroy, & L. Bhandari, Economic Freedom of the States of India (pp. 41-65). Academic Foundation.

AP Reorganisation Portal. (2014). About Reorganization. Retrieved from Andhra Pradesh Reorganization: http://reorganisation.ap.gov.in/index.jsp

ASER. (2014). Andhra Pradesh + Telangana. Annual Status of Education Report. http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202014/ASER%20state%20pages%20Eng/andhratelangana.pdf

ASER. (2015). Annual Status of Education Report 2014: Main Findings. Pratham. http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202014/National%20PPTs/aser2014indiaenglish.pdf

Bowen, A., Long, T., MA, H., OU, J., BILLHIMER, W., Luby, S., INTZ, E. (2007, June). A Cluster-Randomized Controlled Trial Evaluating the Effect of a Handwashing-Promotion Program In Chinese Primary Schools. The American Journal of Tropical Medicine and Hygiene, 76(6), 1166-1173.

Central Bureau of Health Intelligence. (2011). National Health Profile. Retrieved from http://cbhidghs.nic.in/writereaddata/mainlinkFile/06%20Demographic%20Indicators%202011.pdf

Cheney, G. R., Ruzzi, B. B., & Muralidharan, K. (2005). India Education Profile. National Center on Education and the Economy.

Das, J., & Zajonc, T. (2008). India Shining and Bharat Drowning: Comparing Two Indian States to the Worldwide Distribution in Mathematics Achievement. The World Bank Development Research Group.

Delgado, M. S., Henderson, D. J., & Parmeter, C. F. (2014, June). Does Education Matter for Economic Growth? Oxford Bulletin of Economics and Statistics, Vol. 76, Issue 3, pp. 334-359, 2014, 76(3), 334-359, 2014.

Directorate of Economics and Statistics. (2014, May). Andhra Pradesh Economy.Retrieved from http://www.ap.gov.in/AP%20State%20Statistical%20Abstract%20May%202014/4_AP_Economy.pdf

Downey, D. B. (1995). When Bigger is not Better: Family Size, Parental Resources and Children's Educational performance. American Sociological Review, 746-761.

Dreze, J., & Kingdon, G. G. (1999). School Participation in Rural India. The Development Economics Discussion Paper Series.

Filmer, D. (2003). Determinants of Health and Education Outcomes. The World Bank.

Gelda, A., Narayan, V., Mudiyam, M., Raturi, K., & Seshan, N. (2013, January 02). 'Needs Improvement': Despite Progress, India's Primary Education System Has a Ways to Go. Retrieved from Knowledge@Whatron: http://knowledge.wharton.upenn.edu/article/needs-improvement-despite-progress-indias-primary-education-system-has-a-ways-to-go/

Government of India. (2014). The Andhra Pradesh Reorganization Act. Andhra Pradesh Legislature.

Greene, L. E., Freeman, M. C., Akoko, D., Saboori, S., Moe, C., & Rheingans, R. (2012). Greene, L. E., Freeman,Impact of a School-Based Hygiene Promotion and Sanitation Intervention on Pupil Hand Contamination in Western Kenya: A Cluster

Randomized Trial. The American journal of Tropical Medicine and Hygiene, 87(3), 385-393.

Hanushek, E. A., & Woßmann, L. (2010). Education and Economic Growth. International Encyclopedia of Education, 245-252.

Herz, B., & Sperling, G. (2004). What works in girls' education: evidence and policies from the developing world. New York: Council on Foreign Relations.

Hunter, P. R., Risebro, H., Yen, M., Lefebvre, H., Lo, C., Hartemann, P., Jaquenoud, F. (2014). Impact of the Provision of Safe Drinking Water on School Absence Rates in Cambodia: A Quasi-Experimental Study.

Kaur, H. (2000, April). Impact of income and education on fertility. Journal of Family Welfare, 46(1), 70-6.

Kingdon, G. G. (2005). Private and public schooling: The Indian experience. Retriebed from: http://www.ksg.harvard.edu/pepg/PDF/events/MPSPE/PEPG-05-15geeta.pdf

Kingdon, G. G. (2007). The progress of school education in India. Oxford Review of Economic Policy, 168-195.

Koolwal, G., & van de Walle, D. (2010). Access to Water, Women's Work and Child Outcomes. World Bank.

Koretz, D. (2008). In D. Koretz, Measuring Up: What Educational Testing Really Tells Us (pp. 7-14). Cambridge, MA: Harvard University Press.

Kumar, A., & Das, K. C. (2014, April). Drinking Water and Sanitation Facility in India and Its Linkages with Diarrhoea among Children under Five: Evidences from Recent Data. International Journal of Humanities and Social Science Invention, 3(4), 50-60.

Kumar, G. S., Kar, S. S., & Jain, A. (2011, Sep-Dec). Health and environmental sanitation in India: Issues for prioritizing control strategies. Indian Hournal of Occupational and Environmental Medicine, 15(3), 93-96.

Lorntz, B., Soares, A. M., Moore, S. R., Pinkerton, R., Gansneder, B., Bovbjerg, V. E., Guerrant, R. L. (2006). Early Childhood Diarrhea Predicts Impaired School Performance. The Pediatric Infectious Disease Journal, 513-520.

Magnuson, K. A., Meyers, M. K., Ruhm, C. J., & Waldfogel, J. (2004). Inequality in Preschool Education and School Readiness. American Educational Research Journal, 115-157.

Mensch, B. S., & Lloyd, C. B. (1998, June). Gender Differences in the Schooling Experiences of Adolescents in Low-Income Countries: The Case of Kenya. Studies in Family Planning, 29(2), pp. 167-184.

Miguel, E., & Kremer, M. (2004, January). Worms: Identifying Impacts On Education And Health in the Presence of Treatment Externalities. Econometrica, 72(1), 159-217.

Ministry of Human Resource Development, Government of India. (2014). Clean India: Clean Schools- A Handbook. Ministry of Human Resource Development, Government of India.

Mondal, P. (2015). The Three-Tier System of Panchayati Raj in India. Retrieved from Your Article Library: http://www.yourarticlelibrary.com/politics/the-three-tier-system-of-panchayati-raj-in-india/4827/

Muralidharan, K., & Kremer, M. (2007). Public and Private Schools in Rural India. http://econweb.ucsd.edu/~kamurali/papers/Published%20Edited%20Volume%20Chapters/Public%20and%20private%20schools%20in%20rural%20india%20(Final%20Pre-Publication).pdf

Nath, A. (2011, April- June). India's Progress Toward Achieving the Millennium Development Goals. Indian Journal of Community Medicine, 85-92. doi:10.4103/0970-0218.84118

National Commission for Women, India. (2005). Jalees, K. (2005). Water & women: a report by research foundation for science, technology, and ecology for national commission for women. Research Foundation for Science, Technology and Ecology.

Newman, J., Pradhan, M., Rawlings, L. B., Ridder, G., Coa, R., & Evia, J. L. (2002). An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund. The World Bank Economic Review, 241-274.

Nokes, C., Grantham-Mcgregor, S. M., Sawyer, A. W., Cooper, E. S., & Bundy, D. A. (1992). Parasitic Helminth Infection and Cognitive Function in School Children. Royal Society Publishing, 247(1319).

O'Reilly, C. E., Freeman, M. C., Ravani, M., Migele, J., Mwaki, A., Ayalo, M., . . . Quick, R. (2006). The impact of a school-based safe water and hygiene programme on knowledge and practices of students and their parents: Nyanza Province, western Kenya. Epidemiology and Infection, 80-91.

Oster, E., & Thornton, R. (2009). Menstruation and Educatio in Nepal. Cambridge: NBER.

PISA. (2009). PISA 2009 Key Findings. OECD.

Planning Commision. (n.d.). Socially Disadvantaged Groups. Retrieved from http://planningcommission.nic.in/plans/planrel/fiveyr/10th/volume2/v2_ch4_1.pdf

Popham, W. J. (1999, March). Why Standardized Tests Don't Measure Educational Quality. Educational Leadership: Using Standards and Assessments, 56(6), pp. 8-15.

Saville-Troike, M. (2012). What Really Matters in Second Language Learning for Academic Achievement? TESOL Quaterly, 199-219.

Scheerens, J., Luyten, H., & Ravens, J. v. (2010). Measuring Educational Quality by Means of Indicators. In Perspectives on Educational Quality (pp. 3-33). Springer Netherlands.

Simons, K. A. (2011, August). No Time to Thrive: Armed Conflict and Early Language and Cognitive Development in Ethiopia and Peru. Young Lives Student Paper. Young Lives.

Spears, D., & Lamba, S. (2011). Effects of Early-Life Exposure to Rural Sanitation on Childhood Cognitive Skills: Evidence from India's Total Sanitation Campaign.

Talaat, M., Afifi, S., Dueger, E., El-Ashry, N., Marfin, A., Kandeel, A., El-Sayed, N. (2011, April). Effects of Hand Hygiene Campaigns on Incidence of Laboratory-confirmed Influenza and Absenteeism in Schoolchildren, Cairo, Egypt. Emerging Infectious Diseases Journal, 17(4).

Tooley, J., Dixon, P., & Gomathi, S. V. (2007). Private schools and the millennium development goal of universal primary education: a census and comparative survey in Hyderabad, India. Oxford Review of Education, 539-560.

UNICEF. (2011). More Than Water. UNICEF - Government of The Netherlands Partnership for Water Supply, Sanitation and Hygiene. Retrieved from: https://www.oecd.org/countries/mozambique/49295401.pdf

UNICEF. (2012). Water, Sanitation and Hygiene (WASH)in Schools. New York: UNICEF. Retrieved from: http://www.unicef.org/publications/files/CFS_WASH_E_web.pdf

UNICEF. (2015). Access to water and sanitation: a few definitions. Retrieved from: http://www.unicef.org/wcaro/overview_2570.html

United Nations Department of Economic and Social Affairs. (2015). Tranforming Our World: The 2030 Agenda for Sustainable Development. United Nations Department of Economic and Social Affairs. Retrieved from Sustainable Development Knowledge Platform: https://sustainabledevelopment.un.org/topics

World Bank. (2015). Sanitation and Hygiene: Why They Matter. Retrieved from World Bank: http://water.worldbank.org/node/83311/

World Bank. (2015, March 24). World Bank Approves 248 Million dollars for Improving Rural Water and Sanitation Services in Punjab, India - Over 8 million people to benefit. Retrieved from World Bank: http://www.worldbank.org/en/news/press-release/2015/03/24/world-bank-improving-rural-water-sanitation-services-punjab-india

World Bank. (2015). World Development Indicators. Retrieved from: http://data.worldbank.org/sites/default/files/wdi-2015-frontmatter.pdf

Young Lives. (2010). Assessment of Children's Language Learning Experience. Retrieved from Young Lives: http://www.younglives.org.uk/files/questionnaires/school-survey/india-school-survey/questionnaires/IN_SS_Child_LANGUAGE_EXPERIENCES.pdf

Young Lives. (2010). Child English Test. Retrieved from Young Lives: http://www.younglives.org.uk/files/questionnaires/school-survey/india-school-survey/questionnaires/IN_SS__Child_ENGLISH_test.pdf

Young Lives. (2010). Child Maths Test. Retrieved from Young Lives: http://www.younglives.org.uk/files/questionnaires/school-survey/india-school-survey/questionnaires/IN_SS_Child_Maths_test.pdf

Young lives. (2010). Child Telugu Test (English Translation). Retrieved from Young Lives: http://www.younglives.org.uk/files/questionnaires/school-survey/india-school-survey/questionnaires/IN_SS_Child_telugu_TEST.pdf

Young Lives. (2011, December). Young Lives Methods Guide Sampling. Retrieved from Young Lives: http://www.younglives.org.uk/files/methods-guide/methods-guide-sampling

Young Lives. (2014, June 06). Young Lives study sites in India. Retrieved from Young Lives: http://www.younglives-india.org/our-work/young-lives-study-sites-in-india

Young Lives. (2015, August 14). What we do. Retrieved from Young Lives: http://www.younglives.org.uk/what-we-do