Policy Research Working Paper 9059

Impact of Early Life Exposure to Environments with Unimproved Sanitation on Education Outcomes

Evidence from Bangladesh

George Joseph
Yi Rong Hoo
Nazia Sultana Moqueet
Gnanaraj Chellaraj

WORLD BANK GROUP
Water Global Practice
November 2019

Policy Research Working Paper 9059

Abstract

Despite Bangladesh's notable progress toward the eradication of open defecation, the country still faces severe deficits in the availability of improved sanitation. This paper analyzes the impact of exposure to unimproved sanitation early in childhood on primary school enrollment status, using pseudo-panel data for children ages six to nine years in Bangladesh. The results indicate that unimproved sanitation has a negative and significant impact on primary school enrollment. A child's early exposure to unimproved sanitation decreases the likelihood of being enrolled in primary school by eight to ten percentage points on average compared with a child with access to improved sanitation. The effect is particularly strong—a difference of 8 to 10 percentage points—for children ages six to seven. It is also strong in rural areas. The results are statistically robust to errors due to potential omitted variable bias.

This paper is a product of the Water Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at gjoseph@worldbank.org.

Impact of Early Life Exposure to Environments with Unimproved Sanitation on Education Outcomes: Evidence from Bangladesh

George Joseph[1], Yi Rong Hoo[1], Nazia Sultana Moqueet[2], Gnanaraj Chellaraj[1]

JEL Classification: I10; I20; Q25

Key words: Bangladesh, Fecal, Sanitation, Education, Enrollments

[1] World Bank, Washington DC, USA
[2] BRAC USA, New York USA

I. Introduction

In this paper, the impact of exposure to unimproved sanitation in Bangladesh on early childhood primary school enrollment is analyzed, using a pseudo-panel data for children ages 6 to 9 for the years 2007, 2011, and 2014. There is growing evidence that ill-health and poor nutrition in the early years of childhood have a lasting impact on later in life outcomes. This is especially true for the first 1,000 days of life. Failure to maximize a child's development during this phase could have negative impact later in life in areas such as education, earnings, as well as mental health (Walker et al., 2007). Evidence indicates that poor health conditions early in life are associated with higher mortality (Bengtsson & Lindstörm, 2000) and poor adult cognitive development (Rong et al., 2017; Vogl, 2014). Children who are healthy early in life have better educational outcomes (Case & Paxson, 2010) and subsequently better economic outcomes (Lawson & Spears, 2016; Vogl, 2014; Cutler et al., 2010).

Unsanitary conditions during childhood are likely to result in poor health and nutrition both in the short and long-run. Exposure to fecal matter and bacteria because of unsafe disposal of feces, poor sanitation, and subsequently contaminated water and soil cause environmental enteropathy (George et al., 2018; Kosek et al., 2013), stunting (George et al., 2015; Spears et al., 2013), as well as diarrhea (Cronin et al., 2016), and a host of other water-borne diseases among young children (Mertens et al., 1990; Daniels et al., 1990). Although Bangladesh has substantially reduced its open defecation rates, its continued lack of improved sanitation facilities[3] (World Bank, 2018) could adversely affect children's health, both early on and throughout their lives.

Bangladesh has made significant progress towards addressing the problem of poor sanitation. In line with the SDG goal of achieving access to adequate and equitable sanitation and hygiene for all and end open defecation, the country has reduced open defecation among its population from 35 percent to 2 percent by 2015 (World Bank, 2018). Rural households are more likely to openly defecate compared to urban households (DHS, 2015). Access to improved sanitation has also increased from 34 percent to 63 percent by 2015 (World Bank, 2018). Rural households are about twice as likely to have access to unimproved sanitation compared to their urban counterparts (DHS, 2014).[4]

Sixty-five percent of the households in the country have access to a private sanitation facility in 2014 (DHS, 2016). However, the figure dramatically declines if we consider whether the facility is improved or not. In 2014, only 45 percent of the households had a private improved sanitation facility – an increase from 39 percent in 2011 (DHS, 2016). Interestingly, rural households are more likely to use

[3] The World Bank (2018) estimates that 40 percent of Bangladeshis, or 58 million people, do not have access to a private improved sanitation facility in their home. This lack of access is associated with poverty: the vast majority of these 58 million people are in the bottom 50 percent of the national wealth distribution.

[4] https://dhsprogram.com/pubs/pdf/FR311/FR311.pdf

an unshared toilet facility compared to urban households (DHS, 2016). Within urban communities, slum dwellers are poorly served. Only 13 percent of households in the country's slums have access to their own sanitation facilities (World Bank, 2018). It is common for 10 households or more to share a single sanitation facility (World Bank, 2018). Shared sanitation facilities are considered more unhygienic compared to private ones. As a result, children living in urban slums are more likely to be stunted compared to those living in other urban areas (World Bank, 2018).

Poor sanitation could affect the timing of children's school enrollment. For example, repeated diarrhea and environmental enteric dysfunction (EED) could result in malnutrition, which could negatively affect early childhood cognitive development indicators such as learning abilities and motor skills. As a result, parents could be motivated to keep their children at home, waiting for them to grow and get stronger. Several studies have investigated the impact of water quality at the community level on education outcomes (Beach et al., 2016; Zhang & Xu, 2016). However, studies on the impact of sanitation on education outcomes are virtually non-existent.

In this study it is hypothesized that exposure to poor sanitation early in life, defined as the prevalence of unimproved sanitation facilities in the community,[5] adversely affects young children's school enrollment.[6] Due to poor sanitation, the child is likely to be enrolled in primary school late (or not at all). Conversely, a child who is exposed to improved sanitation is more likely to be enrolled in primary school on time. It should also be noted that delayed enrollment in primary school can have a serious impact on a child's cognitive development.[7] We use the children's residence (specifically communities known to have unimproved sanitation), and the primary school enrollment data at the time of the survey. The goal of the paper is to determine whether exposure to unimproved sanitation early in life negatively affects the likelihood of being enrolled in primary school between ages 6 and 9. Are children exposed to unimproved sanitation facilities early in life less likely to attend primary school?

Our findings indicate that unimproved sanitation has a negative and significant impact on primary school enrollment. A child's early exposure to unimproved sanitation decreases that child's likelihood of being enrolled in primary school by 6 to 9 percentage points on average compared to a child with access to improved sanitation. The effect is particularly strong – a difference of 8 to 10 percentage points – for children ages 6 to 7. It is also strong for children living in rural areas. We find that the results are statistically robust to errors due to potential omitted variable bias.

[5] We define "community" in the data set as a household cluster.
[6] We define school enrollment as primary school enrollment for children between 6 and 9 years.
[7] See, for example, Sclar et al. (2017) and Cole and Cole (1996), who reviewed findings from several studies conducted in low- and middle-income countries where schooling is not universal. For example, the memory performance of schooled children is comparable to their American counterparts, unlike unschooled children from the same village (Cole et al., 1971).

This paper proceeds as follows. In the next section we review the literature that motivates the analysis. In Section III we discuss the empirical framework and data while Section IV discusses the results. In Section V we check for robustness of the results and Section VI concludes.

II. Literature Review

Health and nutrition status in the early years of childhood has lasting impact on later-life outcomes including cognitive development, educational attainment, economic success and adult health. For example, Bengtsson and Lindström (2000) found that the disease load experienced during the early years of life has a strong impact on mortality in later life. Healthier children are more likely to become productive adults (Currie, 2009) because they are more likely to reach both their physical and cognitive potential (Case & Paxson, 2008).

Bozzoli et al. (2009) showed that in countries with relatively high infant mortality rates the height of the population on average is lower. De Oliveira and Quintana-Domeque (2014) found that GDP per capita during the birth year is positively associated with adult height in Brazil. Using cross-sectional data to analyze educational outcomes, Glewwe and Jacoby (1995) found that shorter and stunted children – indicative of childhood malnutrition – in Ghana were more likely to experience delayed primary school enrollment. Meanwhile for the Philippines, Glewwe et al. (2001) using longitudinal data found that better nourished children perform significantly better in school, partly because they enter school earlier and thus have more time to learn but mostly because of greater learning productivity per year of schooling. Vogl (2014) used adult height as a proxy for early life health and found that taller individuals in Mexico had higher earnings. Using the same proxy, Sohn (2015) found similar results for Indonesia. However, using height as a proxy for childhood nutrition does not account for specific aspects of childhood health conditions that have long-term impact (Currie & Vogl, 2013). Handa and Peterman (2007) using a longitudinal data set from South Africa found that children who are malnourished at age 3 or younger are less likely to be enrolled in school later in life. Similar results were found for Pakistan using longitudinal data (Alderman et al., 2001). Jamison (1986) for China and Moock and Leslie (1986) for Nepal found a significant negative relationship between stunting and school progression outcomes. Finally, for the state of Rhode Island in the United States, it was found that early life interventions significantly increased primary school enrollment and declines in social expenditures needed for such individuals (Chyn et al., 2019).

Studies have also been undertaken on the impact of specific conditions or shocks (such as nutritional deficiency, drought, famine, war, disease, and exposure to pollutants) early in a child's life as well as specific genetic factors (Zhu et al., 2018). Pathania (2009) found that external shocks such as a drought at birth were associated with a 0.3 centimeter (cm) drop in height among upper-caste women in rural India between the years 1950 and 1999, mostly driven in utero. Rong et al. (2017) found that early life exposure to famine in China between 1959 and 1961 contributed to a cognitive decline among adults. Almond et al. (2007), who studied the same famine in China, also found that those individuals who were most exposed to famine were more likely to be illiterate, unemployed, and disabled. Similar results were found by other studies (Kim et al., 2017, 2014; Chen and Zhou, 2007; Li and Yang, 2005). For Uganda it was shown that children in utero who were exposed to the 1980 famine were less likely to complete primary school and less likely to be literate (Umana-Aponte, 2011). Finally, for Peru, Leon (2012), found that among young children exposed to the Peruvian conflict in the 1980s and 1990s, the number of subsequent years in school declined for each additional year of civil war exposure.

The impact of nutrition interventions on academic achievement was also analyzed. It was found that female children below 7 years who were given a high-protein drink through a long-running randomized nutrition program in Guatemala attained a higher schooling level than their control group counterparts (Maluccio et al., 2009). Meanwhile, Hoddinott et al. (2008), found that nutritional supplements in the first two years of life increased male workers' wages by 46 percent for the same program. For Kenya, Miguel and Kremer (2004) found that a school-based mass treatment with deworming drugs resulted in reduced absenteeism by 25%. For the United States, the Head Start program significantly increased high school graduation rates (Garces et al., 2002). For India, Kramer et al. (2019) find that delivering double-fortified salt with iron and iodine improved test scores among students. Galasso and Wagstaff (2019) found a rate-of-return of 12%, and a benefit-cost ratio of 5:1-6:1 for a package of 10 nutrition interventions across different countries. Finally, for Honduras (Millan et al., 2020) it was found that exposure to the Conditional Cash Transfer program implemented by the government increases secondary school completion rates as well as the probability of increased university admissions.

Sanitation plays a major role on the health and nutrition status of the population. In developing countries, a crawling child is likely to encounter fecal matter and bacteria through contaminated water and soil. Ingested bacteria can lead to severe intestinal infections that might manifest in repeated bouts of diarrhea, as in the case of an *E. Coli* infection (Wanke, 2001), or even be without symptoms, as in the cases of environmental enteropathy (Korpe & Petri, 2012; Lin et al., 2013; Neto et al., 1994). For Nigeria, Oyemade et al. (1998) identified several risk factors for repeated bouts of diarrhea among children under age 6 such as the mother's hygiene practices after the child's defecation, poor disposal of garbage, as well as contaminated food and water. As a result of repeated exposure and infection, the children's ability to absorb nutrients would eventually be impaired, leading to occurrences of stunting. Mondal et al. (2012) found that infants who were stunted by 12 months of age were more likely to experience prolonged diarrhea and intestinal barrier dysfunction due to enteric bacteria. Alam et al. (2000) found that stunting among children under age 5 in Bangladesh was associated with diarrhea and dysentery episodes.

A number of studies focus on five main fecal-oral pathways for enteric infections: fluids, fingers, fields, flies, and food. However, recent studies have also found evidence that geophagy – defined as the consumption of soil, dirt, or mud – is highly common among very young children and thus is an additional risk factor for enteric infections (Ngure et al., 2014). Children can unknowingly consume soil in their play spaces that is contaminated with bacteria leaked from poor and unimproved sanitation facilities (George et al., 2018). In Kenya the practice of geophagy is significantly associated with diarrhea in children under 5 years (Shivoga and Moturi, 2009). Finally, for Bangladesh, George et al. (2015) showed that the practice of geophagy among children is associated with the prevalence of stunting and environmental enteropathy.

Some studies have also empirically established the link between sanitation and incidences of malnutrition among young children. Fink et al. (2011) found that access to improved sanitation is associated with lower risks of mild or severe stunting and even lower mortality rates, relative to a lack of access. Meanwhile, Dearden et al. (2017) showed that children in Ethiopia, India, Peru, and Vietnam who have access to improved sanitation are less likely to experience stunting. This is also true in Bangladesh, where Ahmed et al. (2012) found that the incidence of stunting among children is associated with living in households with unhygienic toilet facilities.

Until recently, most studies such as Akter (2019), Beach et al. (2016), Zhang and Xu (2016), as well as Barde and Walkiewicz (2014) have focused on the impact of water quality on children's education outcomes and not on the impact of sanitation. Furthermore, other studies have measured the immediate effects of water and sanitation on educational outcomes rather than their subsequent impacts. These studies often focus on water and sanitation facilities at the school level (Jasper et al., 2012) and not at the community level. The provision of treated drinking water in schools helped reduce absenteeism in Cambodia (Hunter et al., 2014) and in the Nyanza Province of Kenya (Freeman et al., 2012). Similar results have been found for Indonesia (Duflo, 2001) and a number of other low-income countries (Jasper et al., 2012). A few studies on the link between sanitation and schooling focused on issues other than students' health. For India, Adukia (2017) found that the construction of sex-specific school latrines increased the enrollment of teenage girls. The use of unimproved facilities combined with high population density can increase the risk of the transmission of fecal bacteria, which adversely affects children's health (Hathi et al. 2014).

Finally, in a recent study, Spears and Lamba (2015) studied the effects of childhood cognitive achievement of early-life exposure to India's Total Sanitation Campaign (TSC), a large government program that encouraged local governments to build toilets and promote their use. It was found that the TSC resulted in higher likelihood of six-year-olds recognizing numbers and letters. However, similar studies are virtually non-existent for other developing countries including Bangladesh. Thus, by focusing on unimproved sanitation at the community level, our study intends to fill the gap in the literature.

III. Empirical Framework and Data

In this section, we specify the empirical approach used in this study to analyze the impact of improved sanitation on primary school enrollment. We then describe the data set used.[8]

(i) Basic Specification and Logistic Regression

Our econometric specification is as follows:

$$ENROLL_t = \beta_0 + \beta_1 UNIMPSANRATIO_{t\text{-}x} + \beta_2 URBAN + \beta_3 TOP60 + \beta_4 HHSIZE + \beta_5 MPRM + \beta_6 MSEC + \beta_7 FEMHEAD + \beta_8 FEMCHILD + \beta_9 LOGCHILDAGE + + \beta_{10} LOGHILDAGESQ + \beta_{11} IMPSANACCESS + \beta_{12} IMPWATERACCESS + \beta_{13} PRIMSCHRATIO_{t\text{-}x} + \beta_{14} YEAR(2011) + \beta_{15} YEAR(2014) + \varepsilon$$

The dependent variable is the likelihood of enrollment (ENROLL) in primary school among children ages 6 to 9 for Bangladesh for the years 2007, 2011, and 2014. The primary variable of interest is the concentration of unimproved sanitation that the children are exposed to in their early years or at the baseline[9] within each household cluster[10] (UNIMPSANRATIO). Community unimproved sanitation is expected to have a negative impact. URBAN denotes households living in urban areas, and we have no particular expectation regarding this variable. A household in the top 60 (TOP60) percentile of the national wealth distribution is expected to have a positive impact on enrollment. Household size (HHSIZE) is expected to be negatively related. The higher the number of members, the lower the likelihood of primary school enrollment, since financial resources are being spread too thin. Mother's graduating from primary school (MPRM) is expected to have a positive impact on enrollment. The impact is expected to be similar for mothers who have completed secondary education (MSEC). We have no particular expectation regarding the impact of female-headed households (FEMHEAD). In some cases, female-headed households are poorer and hence their children are less likely to attend primary school. In other cases, female-headed households tend to emphasize education and hence children are more likely to attend school. A female child (FEMCHILD) is more likely to attend primary school than a male child. In recent years enrollment rates have increased sharply and are higher for girls than boys in Bangladesh's urban areas, according to UNICEF.[11] The log of a child's age (LOGCHILDAGE) is likely to have a positive impact, as higher the age, the higher the likelihood of attending primary school. We also included a squared term for the log of a child's age (LOGCHILDAGESQ).

[8] The data and their definitions are discussed in detail in the next section.

[9] As we demonstrate later, we define the baseline as seven years before the time of the survey. For example, the baseline year for children sampled in the 2007 DHS will be the year 2000.

[10] For simplicity, we will refer to this variable simply as community unimproved sanitation.

[11] See https://www.unicef.org/bangladesh/children_355.htm.

Household access to improved sanitation (IMPSANACCESS) is expected to be positively related to the likelihood of school enrollment. Meanwhile, household access to improved water (IMPWATERACCESS) is also expected to have a positive impact. The concentration of primary schools within each household cluster (PRIMSCHRATIO) is expected to have a positive impact, since the availability of primary schools in a community reduces the possibility of delayed enrollment. We do not have any particular expectations regarding the relationship between DHS years and the likelihood of enrollment (Year [2011], Year [2014]). The analysis uses a logistic regression model with division and year fixed effects to determine the likelihood of a child's being enrolled in primary school after being exposed to poor sanitation early on in life. We will also examine the impact for children living in urban and rural areas and across two age groups: children ages 6–7 and children ages 8–9.

(ii) Robustness Check – Alternative Specifications and Coefficient Stability

We also examine the robustness of the estimated results by addressing the potential omitted variable bias (Verbeek, 2008) which may exist within our preferred specification. We estimate (i) several alternative specifications by including district-level fixed effects as well as the addition of several covariates measured at the baseline and using (ii) the bounding method introduced in Oster (2019) and Altonji, et al. (2005).

One argument against the finding that exposure to unimproved sanitation early in childhood could reduce the likelihood of or delay the primary school enrollment of a child, is that our measured variable – the prevalence of unimproved sanitation in the community – generally reflects living conditions within the area. For example, inadequate community infrastructure such as the lack of roads can make schooling difficult (Yamauchi et al., 2011). Under these circumstances, our hypothesis will be invalid. We test this assumption by estimating separate logistic regressions. First, we replace the division-level fixed effects with district-level fixed effects to account for more local-level effects. Further, we estimate two subsequent logistic regressions by introducing two additional community-level covariates, measured at the baseline, that could reflect or approximate the general living conditions of each community. They are (i) the community-level electricity access at the baseline and (ii) the community-level wealth index at the baseline (computed as the median of household wealth within each household cluster). If the estimated effects of unimproved sanitation change sign or become insignificant by controlling for these two additional variables, our hypothesis would be rendered invalid.

Last, the bounding method examines the stability or robustness of the estimated coefficient to omitted variable bias under hypothetical circumstances, where the estimated model R-squared is increased and if the unobserved variables are more important than the included covariates in the estimated model. If the estimated effect of exposure to unimproved sanitation on school enrollment did not change sign or become zero, the result is considered to be robust to omitted variable bias.

(iii) Causal Inference through Regression on Matched Samples

To further estimate the causal effect of unimproved sanitation on school enrollment, we match similar children based on observed characteristics across two defined groups using the concentration of unimproved sanitation within each household cluster and re-estimate the preferred regressions on these matched samples. Children who are exposed to higher concentrations of unimproved sanitation within the household cluster (defined as being above the median value) during their early years are included in the treatment group, while children who are exposed to lower concentrations of unimproved sanitation within the household cluster (defined as being below the median value) during their early years are included in the comparison group.[12] Matching children with a similar set of observed characteristics would help reduce the bias of the estimated effect due to the confounding variables or underlying systematic differences between children who are exposed to higher and lower levels of unimproved sanitation (Rosenbaum and Rubin, 1983). In other words, matching aims to improve the sample balance which is the degree to which the treatment and comparison covariate distribution resemble each other.

We use a series of major matching methods to produce the matched samples of similar treatment and comparison children. They are (i) coarsened exact matching (Iacus et al., 2012; Blackwell et al., 2009) and (ii) propensity score matching (Rosenbaum & Rubin, 1983). In fact, for the latter, we estimate four different ways by which propensity score matching can be done: (i) caliper matching with the three nearest neighbors, (ii) caliper matching with the five nearest neighbors, (iii) radius matching (Dehejia & Wahba, 2002), and (iv) kernel matching (Heckman et al., 1997). We use six different matched samples to estimate and compare the causal effect of unimproved sanitation on primary school enrollment. We expect that each matched sample would yield estimates consistent with the others.

We approach the matching techniques as data pre-processing, where we will re-estimate separate logistic regressions on the resulting matched samples through the matching methods described above. While estimating a simple difference in means on the resulting matched sample could yield a causal estimate, following the suggestions in Ho et al. (2007), we further estimate a regression on the resulting matching specifications. By estimating another regression, we will be able to account for the remaining imbalances or differences between children in the treatment and comparison groups. Moreover, the two-step procedure is considered doubly robust (Robins & Rotnitzky, 2001). It requires only one of the matching or regression models to be specified correctly to produce consistent estimates. The process of estimating a simple difference in means for the outcome on matched samples would not otherwise possess this property of double robustness. Further, while it may seem to be counterintuitive to rely on smaller

[12] The median unimproved sanitation ratio is found to be 34.5 percent; 3,517 (55 percent) children under the median value are categorized as the comparison group while 2,882 (45 percent) children above the median value are included in the treatment group.

matched samples where unmatched observations will be inevitably dropped, Smith (1997) showed that matched samples can increase the efficiency of estimates.[13]

(iv) Data Sources

The sample in this analysis consists of 6,399 children ages 6–9 years from both urban and rural areas of Bangladesh. The data are put together from DHSs 2007, 2011, and 2014, which are nationally representative at the division level. The dependent variable is dichotomous and indicates whether a child aged 6–9 is enrolled (1) or not enrolled (0) in primary school at the time of the survey. The primary variable of interest is the concentration of unimproved sanitation at the baseline within each household cluster. We consider this variable to be reflective of the child's early exposure to unimproved sanitation. The variable is obtained by geomatching household clusters over a seven-year period.[14] Clusters from DHS 2000 were geomatched with DHS 2007 to estimate the exposure to unimproved sanitation when children in the 2007 sample were 1 or 2 years old.[15] Here we make two assumptions: (i) a household is likely to have unimproved sanitation if the majority of its neighboring households have it; and (ii) the families stayed in the same household and household cluster over the seven-year period. The percentage of children enrolled in primary school at the baseline was also obtained using the geomatching technique. This variable is defined as the proportion of primary schools in a cluster. Finally, the lagged electricity ratio was calculated using the same procedure.

The indicator "URBAN" takes the value of 1 if the household is located in an urban area and 0 otherwise. In our analysis, household wealth quintiles are constructed using a principal component analysis (Filmer & Pritchett, 1999) that accounts for housing conditions, ownership of durable goods, land, etc. Information on the type of water source and sanitation facility used by the household is not included in the calculation since these are considered separate variables in our analysis. The top three quintiles are combined to form the wealth variable, which indicates whether a household is in the top 60 percent of the population in terms of wealth. The analysis takes into account whether water and sanitation facilities are improved or not.[16]

---

[13] Through theoretical and simulation results in a wide range of scenarios, matched samples can result in substantial bias and variance reduction even if compared with random samples of the same size (Rubin & Thomas, 1992, 1996).

[14] The baseline for children in the 2007 DHS is the 2000 DHS, the baseline for children in the 2011 DHS is the 2004 DHS, and the baseline for children in the 2014 survey is the 2007 DHS.

[15] Of the 6,399 children in our sample, we matched 1,626 children from the 2007 DHS, 2,508 children from the 2011 DHS, and 2,265 children from the 2014 DHS. Figure 1 in the annex shows the geographic distribution of household clusters available before and after geomatching.

[16] According to the Joint Monitoring Programme (JMP), an improved sanitation facility is defined as "one that hygienically separates human excreta from human contact." The JMP considers the following categories as improved sanitation: flush to piped sewer system, flush to septic tank, flush to pit (latrine), flush to unknown place, ventilated improved pit latrine, pit latrine with slab, and composting toilet. Shared sanitation is excluded from the improved sanitation category. The following categories are considered as improved water: piped into dwelling; piped into compound, yard, or plot; piped to neighbor; public tap/standpipe; bottled water; tube well/borehole; protected spring; and protected well.

Household size is defined as the number of members living in the household. The mother's primary education takes the value of 1 if the mother has at least a primary education and 0 otherwise. Similarly, the mother's secondary education takes the value of 1 if the mother has at least a secondary education and 0 otherwise. Maternal education is included in this analysis due to its positive association with child educational outcomes as discussed in the literature. Given the evidence linking maternal education with child educational outcomes, it is reasonable to assume that it has a positive influence on a child's primary school enrollment as well.

Household heads take the value of 1 if the head is female and 0 otherwise. The child takes the value of 1 if the child is female and 0 otherwise. The child's age is the child's actual age measured in months. The household's access to improved sanitation takes the value of 1 if the household has improved sanitation and 0 otherwise.

Our data set is a pseudo-panel with children ages 6–9 sampled from three cohort pairs of independent repeated cross-sections of DHS 2000–07, 2004–11, and 2007–14, following Deaton (1985), Magadi (2016), and Ncube and Shimeles (2012), respectively. Unlike true longitudinal panel data, we replace cohort means with individual observations. If an additive individual fixed effect exists, a corresponding additive cohort fixed effect will also exist. The sample cohort means from the surveys are consistent. However, they are error-ridden estimates of the true cohort means. Such data are immune to attrition bias and can be used over long periods. For Bangladesh, the same individuals are not followed over time (unlike longitudinal studies), and histories of individuals are not available (Verbeek, 2008). The individuals sharing the year of birth are grouped into cohorts, and averages for each of the cohorts are treated as observations in a pseudo-panel. All tables and figures are presented in the Annex.

IV. Results

(i) Descriptive Statistics

Table 1 (Annex) presents the weighted mean values of the variables used in the logistic model. In our sample, 80 percent of all children between 6 and 9 years are enrolled in school at the time of the survey. About 11 percent of the children live in households headed by a female. About half of the children are female and the average age of the child in a household is 7.5 years. In terms of sanitation and water facilities, 41 percent of the children live in households with access to improved sanitation and almost everyone has access to improved water.

Data presented in Table 2 for children from rural areas indicate that 82 percent of all children ages 6–9 were enrolled in school at the time of the survey. Sixty-one percent of the children come from households that are in the top 60 percentile of the wealth distribution. Around 11 percent live in female-headed households. Half of the children are female, and the average age is 7.5 years. Regarding sanitation and water facilities, 39 percent of households have improved sanitation – which is less than in urban areas – and almost everyone has access to improved water.

Finally, Table 3 presents the corresponding figures for children from the urban areas. Seventy-eight percent of all children between 6 and 9 years were enrolled in primary school at the time of survey and 93 percent of them were in the top 60 percentile of the national wealth distribution. About 11 percent of the children belong to households headed by a female. Half of the children are female, and the average age is 7.5 years. Regarding sanitation and water facilities, 43.9 percent of the children belong to households that have improved sanitation, while almost all of them have access to improved water. Overall, in our sample although urban households are wealthier and more likely to have improved sanitation, school attendance is slightly lower in urban than in rural ones.

(ii) Regression Results

Table 4 presents the logistic regression results[17] for the total sample (column 1) followed by children ages 6–7 and children ages 8–9 (columns 3 and 5). Exposure to unimproved sanitation in the early years or at baseline[18] significantly decreases the likelihood of school enrollment by 5 percentage points. The effect seems to be stronger for children between 6 and 7 years, for whom unimproved sanitation exposure decreases the likelihood of school enrollment by 7 percentage points. The exposure does not seem to significantly affect the school enrollment of children between 8 and 9 years. Children in households with improved sanitation at the time of survey are about 4% more likely to be enrolled in school.

As expected, the mother's education has a positive and significant impact across all age groups. Children whose mothers completed primary school are 10 percentage points more likely to be enrolled in primary school compared with those whose mothers have no education. For children whose mothers completed secondary education or higher, primary school enrollment is 14 percentage points more likely. Again, this effect seems to be stronger among children aged 6–7 than among children aged 8–9. Female children are also more likely to be enrolled in school than male children. The older the child, the more likely the enrollment.

Table 5 presents a different set of results for children from rural areas across all age groups followed by children between 6 and 7 years and children between 8 and 9 years (columns 3 and 5). Exposure to unimproved sanitation in the early years significantly decreases the likelihood of school enrollment by 12 percentage points. Furthermore, similar to the findings discussed earlier, exposure to unimproved sanitation has a stronger effect on the school enrollment of children ages 6–7 years (14

[17] For each logistic regression, we also estimated a corresponding linear probability model to compare the estimated marginal effects. These results will also be used in examining the coefficient stability through the bounding method introduced by Oster (2019). They are presented in columns 2, 4, and 6.

[18] The phrases "at baseline" and "early years" are used interchangeably.

percentage points) than it does on children ages 8–9 (11 percentage points). Thus, children in rural areas are particularly vulnerable to the deleterious effect of unimproved sanitation on school enrollment.

Finally, Table 6 presents the regression results for children living in urban areas across all age groups as well as children between 6 and 7 years and children between 8 and 9 years. Unimproved sanitation exposure in the early years does not have a significant impact on the primary school enrollment of children in Bangladesh's urban areas. Children living in rural areas are more vulnerable than those in urban areas to the negative impact of unimproved sanitation exposure during their early years on later primary school enrollment.

Consistent with our findings earlier in Table 4, results from Tables 5 and 6 also show that maternal education, and the child's sex and age also significantly predict school enrollment across both rural and urban samples. The positive effect (on school enrollment) of having access to improved sanitation at the time of the survey seems to be significant only for children in the urban areas and not for those living in the rural areas of Bangladesh.

V. Robustness Check

(i) Alternative Specifications

As mentioned earlier, one argument against the finding that exposure to unimproved sanitation in the early years could reduce the likelihood of or delay the school enrollment of a child is that the measured variable generally reflects living conditions. Thus, it may actually be poor living conditions in general that negatively affect school enrollment, and not necessarily exposure to improved sanitation. We test this assumption by estimating alternate logistic regressions and the results are presented in Table 7. For each of the result matrixes, the estimated coefficient (marginal effects) of unimproved sanitation, standard error, as well as model R-squared (in square brackets) are reported.

Columns (1) and (2) report the estimated coefficients for unimproved sanitation from the logistic and linear probability model (LPM) regressions from our preferred specifications across the different samples by age group for comparison. Column (3) presents the same estimated coefficient for unimproved sanitation when we used district-level fixed effects rather than division-level fixed effects to account for more granular geographical differences in our analysis. The results are generally consistent with the results discussed in the previous section. In fact, under this specification, the estimated effects of unimproved sanitation on school enrollment is marginally stronger for the significant results. For example, exposure to unimproved sanitation at baseline decreases the likelihood of school enrollment by 13 percentage points for children across both age groups living in rural areas under this specification compared to 12 percentage points in the earlier estimations.

Column (4) presents the results for unimproved sanitation with baseline community-level electricity access included. The results are almost identical to those estimated and discussed in the earlier section. The same applies for column (5) when baseline community wealth – as measured by the median of the household wealth index – is added on top of the baseline electricity access variable. The results are consistent and in fact stronger. One noticeable change that can be observed in the results presented in columns (4) and (5) is that the effect of unimproved sanitation on school enrollment becomes significant for children ages 8–9 years after the inclusion of the two additional covariates. The effect was previously insignificant. For example, if both community electricity access and wealth at baseline are included in the model as additional covariates, the exposure to unimproved sanitation reduces the likelihood of school enrollment by 6 percentage points for all children ages 8–9 years, about twice the size of the estimated effect when community electricity access and wealth at baseline are not accounted for in our preferred specifications. The findings also suggest that the estimated coefficients for both community electricity access and community wealth variables are barely significant or insignificant, predicting school enrollment significantly for two to three subgroups but counterintuitively in the wrong direction – suggesting that the effect of unimproved sanitation on school enrollment is indeed more robust and meaningful than those two variables.

(ii) Coefficient Stability

While the discussion above addresses potential omitted variable biases that could arise due to general living conditions confounding our results, other omitted variable bias may underlie our preferred specifications. Our findings in Table 8 address these issues.

Table 8 further examines the robustness of the negative effects of children's exposure to unimproved sanitation at baseline on school enrollment at the time of the survey, found in Tables 4, 5, and 6. This follows the bounding method introduced in Oster (2019) as well as Altonji, et al. (2005), who developed a method to investigate the coefficients' sensitivity to potential omitted variable bias. The estimates reported here are calculated using the Stata module *psacalc* introduced in Oster (2019). Since the bounding method can only be implemented on linear models, we rely on the results from the LPM equivalents for each logistic regression in Tables 4, 5, and 6. The marginal effects, standard error, and model R-squared (in square brackets) for each of the LPM specifications are reported in column (3). Also, as we have shown earlier, the marginal effects estimated from the LPMs are very similar to those of the logistic regression models, reported in column (1). Column (2) reports the marginal effects from the LPM without controlling any covariates we included in our analysis.

First, as suggested by Oster (2019) we hypothetically increase each model's R-squared in column (5) by a factor of 2.2 and examine the changes on each of the model's coefficient. Will the negative impact of unimproved sanitation on school enrollment still hold if we include unobserved variables that would increase the current model R-squared by a factor of 2.2?[19]

The results are presented in column (5). For all the specifications, the changes in the coefficients are not large. For example, increasing the R-squared by a factor of 2.2 would change the effect of unimproved sanitation at baseline on school enrollment at the time of the survey from -5 percentage points to -6 percentage points for all children and from -13 percentage points to -9 percentage points for all children from rural areas. The changes were also similar for children between 6 and 7 years as well as rural children between 8 and 9 years. Since the coefficients estimated in column (5) do not change signs or become zero, they are robust to omitted variable bias. In fact, all fall within the 95 percent confidence interval estimated from their respective LPMs.

Column (7) presents the bias-adjusted coefficients for each specification under the assumption that the unobservable variables are 1.5 times as important as the covariates included in our estimated regressions – an unlikely situation given that we have extensively included variables that have been found to be important predictors of school enrollment and/or unimproved sanitation. Still, the results do not pose any concern given that the adjusted coefficients presented in column (7) under such an assumption do not change sign or are not reduced to zero, which further indicates that the coefficients are robust to omitted variable bias.

Second, following Altonji et al. (2005), we computed the ratios of the impact of omitted variable(s) relative to the included covariates that are needed to fully explain away the impact of unimproved sanitation exposure during children's early years on their later primary school enrollment. The results in column (8) show that for all specifications, the unobserved variables must be at least 1.8 to 1.9 times as important as the included control variables to fully explain away the impact of arsenic contamination on early childhood development. Our findings suggest that the exposure to unimproved sanitation earlier in life has a significant negative impact on school enrollment among young children and this estimated effect is also robust to potential omitted variable bias.

(iii) Causal Effects from Regressions Estimated on Matched Samples

The possibility of bias could still arise due to the potential underlying differences in characteristics between children who are exposed to high concentrations of unimproved sanitation and children who are not exposed. Such differences or selection bias could confound our estimates rendering them inaccurate. To address this issue, we apply several different matching methods to account for these differences by matching children who are exposed to a higher concentration of unimproved sanitation with children who are exposed to a lower concentration of unimproved sanitation. Both groups have otherwise similar

[19] See Hener et al. (2016) for an example of the application of Oster's (2019) bounding method.

characteristics. Once we obtain a sample of children who are matched (defined as the matched sample), we re-estimate the logistic regression on the matched sample. As mentioned earlier, in the context of this study, we consider children exposed to unimproved sanitation above the median value to be in the treated group while observations below the median value are included in the comparison group.[20] We also largely favor estimating regressions on the matched sample as opposed to simply estimating raw differences in the outcome value due to the double robustness discussed earlier.[21] Figure 2 presents the Propensity Score Balance between Treatment and Comparison Groups Prior to Matching. Figures 3 and 4 illustrate the propensity score balances before and after the matching specifications respectively. After matching, the estimated propensity scores in both the treatment and comparison group strongly align with each other. Similarly, as shown in Figure 4, the covariate balances across the different propensity score matching specifications have improved or narrowed.

Table 9 shows the results from our logistic regressions on the matched samples. Column (3) presents the regression results estimated on the coarsened exact matched sample,[22] and the results are very similar to those estimated from the nearest neighbor matched sample. For all age groups, it is estimated that unimproved sanitation exposure during the children's early years decreases the likelihood of school enrollment by 7 percentage points, 8 percentage points for children between 6 and 7 years and around 4 percentage points for children between 8 and 9 years.

Columns (4) to (7) present the regression results on various propensity score matched samples. Similar but marginally stronger results are evident. Column (4) presents the regression results estimated on a caliper matched sample with three nearest neighbors. Unimproved sanitation decreases the likelihood of school enrollment by 8 percentage points for all children, about 8 percentage points for children between 6 and 7 years, and 4 percentage points for children between 8 and 9 years. The results were very similar when specified for the five nearest neighbors, as shown in column (5), suggesting that there is little concern for biases due to matching quality or number of matches while increasing the sample of

[20] Again, the median unimproved sanitation ratio is found to be 34.5 percent; 3,517 children (55 percent) under the median value are categorized as the comparison group while 2,882 children (45 percent) above the median value are included in the treatment group.

[21] In addition, these regressions allow for the estimation of the effects of unimproved sanitation in its current configuration as a continuous variable rather than a dichotomous variable. We use the dichotomous version of the variable for the matching process while retaining its continuous form for the subsequent regression analyses on the matched samples. This will also enable comparison with the earlier estimated results.

[22] We specify covariates that are identified to be key correlates of school enrollment found in the earlier logistic regressions and then estimate the effect of unimproved sanitation on school enrollment controlling for the other covariates using a logistic regression (Blackwell et al., 2009). The resulting logistic regression will be able to control for the remaining imbalance resulting from the matching process. The variables included for the coarsened exact matching process are dummy indicators for urban households, households in the top 60 percentile of the wealth distribution, mother's education, child's sex and age, improved sanitation, as well as the Division. We matched the children within each DHS year, using the Stata user-written program *cem* introduced in Blackwell et al. (2009) to execute the matching. Of the 6,399 children, 2,519 treatment children are matched to 3,175 children in the comparison group.

matched children from the comparison group.[23] Exposure to unimproved sanitation in the early years decreases the likelihood of school enrollment by 8 percentage points for all children, 8 percentage points for children between 6 and 7 years, and about 4 percentage points for children between 8 and 9 years.

Columns (6) and (7) present the effect of unimproved sanitation on school enrollment when the regression is estimated on a radius matched sample and kernel matched sample, respectively.[24] While both matching methods are dissimilar, they share a desired property, ie., they maximize the utilization of observations in the comparison group – unlike the earlier matching specifications – and this results in a substantially larger matched sample with little difference in terms of the sample size compared to the unmatched sample. The results from both are similar. Unimproved sanitation decreases the likelihood of school enrollment by about 8 percentage points for all children, and those between 6 and 7 years, and by about 5 to 6 percentage points for children between 8 and 9 years. The stronger results estimated from the propensity score matched samples come as no surprise since they greatly reduce the average sample bias measured by the impact sizes (Cohen's D) of the covariates between the children in the treatment and comparison groups compared to the two preceding matching methods.[25]

Finally, it should be noted that the impact of unimproved sanitation estimated from all the matching specifications falls within the 95 percent confidence interval of the unimproved sanitation coefficients estimated from the corresponding logistic regressions. The effects were stronger for students ages 6–7 than for students ages 8–9, as in the case of earlier regressions.

[23] All the covariates included in our earlier regression models are used to estimate the propensity scores. We restrict the sample to an area of common support by removing children in the treatment group with propensity scores that are above the maximum propensity score and below the minimum propensity score of children in the comparison group. We specify that each child in the treatment group is to be matched to three similar children in the comparison group in one estimation and to five similar children in the comparison group in another estimation. As with the earlier two matching estimations, we matched the children within each DHS year. Also, we set the maximum propensity score distance between each matched pair of treatment and comparison children to be 0.25 times the standard deviation of the estimated propensity scores (Rosenbaum & Rubin, 1985). This is around 7, 6, and 5 percentage points for the 2007, 2011, and 2014 DHS, respectively. If two or more comparison children with the same propensity score are matched to the same child in the treatment group, only one of them is chosen. The combination of caliper restriction and matching with multiple neighbors balances the trade-off between bias and sample variance (Caliendo & Kopeinig, 2008; Lunt, 2013). Of the 6,399 children, 2,782 treatment children are matched to 2,065 comparison children in the first estimation (three nearest neighbors) and 2,782 treatment children are matched to 2,462 comparison children in the second estimation (five nearest neighbors).

[24] Dehejia and Wahba (2002) introduced a different version of caliper matching known as radius matching, which uses not only the specified amount of the number of nearest neighbors during the matching process but all nearest neighbors within the specified caliper size – which is 0.25 times the standard deviation of the estimated propensity scores as mentioned above. As Caliendo and Kopeinig (2008) have discussed, a benefit of using this approach is that it allows for additional matches when good matches – defined as matches within the caliper size – are available. Meanwhile, Kernel matching (Heckman et al., 1997) uses weighted averages of all individuals in the comparison group to construct the counterfactual outcome which would suggest lower variance since more information is used. As the matching results show in Table 7, both of these methods fully utilize all the available observations in the comparison group.

[25] Each of the propensity score matching specifications reduces the average sample bias from 0.17 to about 0.03 (82 percent reduction).

VI. Implications and Concluding Remarks

This paper estimates the impact of children's exposure to unimproved sanitation early in life (at ages 1–2) on their primary school enrollment later in life (at ages 6–9) in Bangladesh. Children's primary school enrollment is adversely affected by their earlier exposure to poor sanitary conditions. This is consistent with our hypothesis that children who suffer poor health due to their exposure to poor sanitation early in life are more likely to experience delayed school enrollment. Our results from the logistic regressions are robust to potential omitted variable biases, and placebo effects. They are further confirmed to be consistent through doubly robust causal estimations across multiple matched samples.

Two notable findings in our study merit further discussion. The first is that while exposure to poor sanitation early in life has a significant negative impact on primary school enrollment across the ages 6 to 9 years, the effect is stronger among children ages 6–7 than ages 8–9. This likely confirms our hypothesis that parents of children in poor health as a result of exposure to poor sanitation are motivated to delay enrolling their children in school until they are slightly older and healthier.

The second finding is that the effect of exposure to poor sanitation early in life is much stronger for children living in rural areas than for children in urban areas. In fact, the effect is statistically insignificant for children in urban areas. There are a few plausible explanations for this. One is the difference in the magnitude of health care coverage between the urban and rural areas. Because of a deficit of professional health care staff, it is estimated that close to 90 percent of the rural population in Bangladesh does not receive adequate health care. In urban areas, this share is estimated at 78 percent (Scheil-Adlung, 2015).[26] Fewer than 20 percent of the health care facilities across the country are providing services to more than 75 percent of the rural population (World Bank, 2015).[27] There appears to be a persistent "urban bias" in the government health care workforce (Ahmed et al., 2011). Moreover, there is also a significant gap in terms of the quality of water and sanitation facilities available between health care facilities in the urban and rural areas. Health care facilities in the rural areas, for example, are more susceptible to water shortages (World Bank, 2018). While urban health care facilities have almost universal sanitation coverage, the coverage for rural health care facilities is 85 percent. Thus, both the differences in availability and quality of health-care services across urban and rural areas could have confounded the effect of poor sanitation on school enrollment. The availability of health care services in urban areas, for example, would make it easier for parents to seek medical help, preventing their children from suffering longer-term health deficits.

[26] Based on data collected in 2011.

[27] It was also shown that the doctor-to-population ratio is 1:1,500 in urban areas, but it is 10 times worse in rural areas – 1:15,000 (Mabud 2005).

Our empirical results also indicate that the educational achievement of the mother matters in the enrollment of children. Female children are more likely to be enrolled than male children in primary schools, a result in line with recent UNICEF findings. Household size plays a major role in the enrollment rates of the 8- to 9-year-old age group but not the 6- to 7-year-old group, suggesting that children in larger households are less likely to attend school than those living in smaller households, probably because the resources are spread too thin or because children have to drop out to help at home. The impact of wealth on enrollment varied. Richer households were less likely to send their children to school in earlier years (6–7 years) and more likely to send their children to school in later years (8–9 years). This also may imply that poorer households are more likely to send their children to school in earlier years, but less likely to send their children to school in later years. Poorer households generally tend to have larger families and may not be able to afford to send children to school in later years.

Overall, providing more improved sanitation at the community level in Bangladesh could help increase enrollment in primary schools. Despite making progress toward the eradication of open defecation, Bangladesh still faces severe deficits in terms of the availability of improved sanitation. Construction of improved toilets in public places as well as incentives for households to construct improved toilets need to be emphasized. The challenge of expanding sanitation services and providing toilets will require the adoption of new and improved technologies, particularly in rural areas. Improved sanitation coverage at the household level needs to be intertwined with improved sanitation in public places, such as improvements in the coverage of sewerage systems, particularly in urban areas. Our findings suggest that the provision of good and proper sanitation is not only a public good by itself, but vital to improvements in health and nutrition, as well as education, particularly for increasing enrollment in Bangladesh's primary schools.

REFERENCES

Adukia, A. 2017. "Sanitation and Education." *American Economic Journal: Applied Economics*, 9:23-59.

Ahmed, T., Roy, S., Alam, N. and Hossain, M. I. 2012. "Determinants of Undernutrition in Children Under 2 Years of Age from Rural Bangladesh." *Indian Pediatrics*, 49:821-824.

Ahmed, S. M., Hossain, M. A., Raja Chowdhury, A. M. and Bhuiya, A. U. 2011. "The Health Workforce Crisis in Bangladesh: Shortage, Inappropriate Skill-Mix and Inequitable Distribution." *Human Resources for Health*, 9:3.

Akter, S. 2019. "Impact of Drinking Water Salinity on Children's Education: Empirical Evidence from Coastal Bangladesh." *Science of the Total Environment,* 690:1331-1341.

Alam, D. S., Marks, G. C., Baqui, A. H., Yunus, M. and Fuchs, G. J. 2000. "Association Between Clinical Type of Diarrhea and Growth of Children Under 5 Years in Rural Bangladesh." *International Journal of Epidemiology*, 29:916-921.

Alderman, H., Behrman, J. R., Lavy, V. and Menon, R. 2001. "Child Health and School Enrollment: A Longitudinal Analysis." *Journal of Human Resources*, 36:185-205.

Almond, D., Edlund, L., Li, H. and Zhang, J. 2007. *Long-term effects of the 1959-1961 China famine: Mainland China and Hong Kong* (No. w13384). Working paper, National Bureau of Economic Research.

Altonji, J. G., Elder, T. E. and Taber, C. R. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113:151-184.

Barde, J. A. and Walkiewicz, J. 2014. *Access to Piped Water and Human Capital Formation: Evidence from Brazilian Primary Schools* (No. 28). Discussion Paper Series, University of Freiburg, Department of International Economic Policy.

Beach, B., Ferrie, J., Saavedra, M. and Troesken, W. 2016. "Typhoid Fever, Water Quality, and Human Capital Formation." *The Journal of Economic History*, 76:41-75.

Bengtsson, T. and Lindström, M. 2000. "Childhood Misery and Disease in Later Life: The Effects on Mortality in Old Age of Hazards Experienced in Early Life, Southern Sweden, 1760-1894." *Population Studies*, 54:263-277.

Blackwell, M., Iacus, S. M., King, G. and Porro, G. 2009. "Coarsened Exact Matching in Stata." *The Stata Journal*, 9:524-546.

Bozzoli, C., Deaton, A. S. and Quintana-Domeque, C. 2009. "Adult Height and Childhood Disease." *Demography,* 46: 647-669.

Caliendo, M. and Kopeinig, S. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys*, 22:31-72.

Case, A. and Paxson, C. 2008. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy*, 116:499-532.

Case, A. and Paxson, C. 2010. "Causes and Consequences of Early-Life Health." *Demography*, 47:S65-S85.

Chen, Y. and Zhou, L-A. 2007. "The Long-Term Health and Economic Consequences of the 1959-1961 Famine in China." *Journal of Health Economics,* 26:659-681.

Chyn, E., Gold, S. and Hastings, J.S. 2019. *The Returns to Early-Life Interventions for Very Low Birth Weight Children.* (No. 25753), Working Paper, National Bureau of Economic Research.

Cole, M. and Cole, S.R. 1996. *The Development of Children,* 3rd ed. Freeman Publishing, New York.

Cole, M., Gay, J, Glick, J.A. and Sharp, D.W. 1971. *The Cultural Context of Learning and Thinking.* Basic Publishers, New York.

Cronin, A. A., Sebayang, S. K., Torlesse, H. and Nandy, R. 2016. "Association of Safe Disposal of Child Feces and Reported Diarrhea in Indonesia: Need for Stronger Focus on a Neglected Risk." *International Journal of Environmental Research and Public Health*, 13:310.

Currie, J. and Vogl, T. 2013. "Early-Life Health and Adult Circumstance in Developing Countries." *Annual Review of Economics*, 5:1-36.

Currie, J. 2009. "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development." *Journal of Economic Literature*, 47:87-122.

Cutler, D., Fung, W., Kremer, M., Singhal, M. and Vogl, T. 2010. "Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India." *American Economic Journal: Applied Economics,* 2:72-94.

Daniels, D. L., Cousens, S. N., Makoae, L. N. and Feachem, R. G. 1990. "A Case-Control Study of the Impact of Improved Sanitation on Diarrhea Morbidity in Lesotho." *Bulletin of the World Health Organization*, 68:455.

De Oliveira, V. H. and Quintana-Domeque, C. 2014. "Early-Life Environment and Adult Stature in Brazil: An Analysis for Cohorts Born Between 1950 and 1980." *Economics & Human Biology*, 15:67-80.

Dearden, K. A., Schott, W., Crookston, B. T., Humphries, D. L., Penny, M. E. and Behrman, J. R. 2017. "Children with Access to Improved Sanitation But not Improved Water are at Lower Risk of Stunting Compared to Children Without Access: A Cohort Study in Ethiopia, India, Peru, and Vietnam." *BMC Public Health*, 17:110.

Deaton, A. 1985. "Panel Data from Time-Series of Cross-Sections." *Journal of Econometrics*, 30:109-126.

Dehejia, R. H. and Wahba, S. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics*, 84:151-161.

Duflo, E. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91:795-813.

Filmer, D. and Pritchett, L. 1999. "The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries." *Population and Development Review*, 25:85-120.

Fink, G., Günther, I. and Hill, K. 2011. "The Effect of Water and Sanitation on Child Health: Evidence from the Demographic and Health Surveys 1986–2007." *International Journal of Epidemiology*, 40:1196-1204.

Freeman, M. C., Greene, L. E., Dreibelbis, R., Saboori, S., Muga, R., Brumback, B. and Rheingans, R. 2012. "Assessing the Impact of a School-Based Water Treatment, Hygiene and Sanitation Program on Pupil Absence in Nyanza Province, Kenya: A Cluster-Randomized Trial." *Tropical Medicine & International Health*, 17:380-391.

Galasso, E. and Wagstaff, A. 2019. "The Aggregate Income Losses from Childhood Stunting and Returns to Nutrition Intervention Aimed at Reducing Stunting." *Economics and Human Biology,* 34:225-238.

Garces, E., Thomas, D. and Currie, J. 2002. “Longer-Term Effects of Head Start.” *American Economic Review,* 92:999-1012.

George, C. M., Oldja, L., Biswas, S., Perin, J., Lee, G. O., Kosek, M., et al. 2015. “Geophagy is Associated with Environmental Enteropathy and Stunting in Children in Rural Bangladesh.” *The American Journal of Tropical Medicine and Hygiene*, 92:1117-1124.

George, C. M., Burrowes, V., Perin, J., Oldja, L., Biswas, S., Sack, D., et al. 2018. “Enteric Infections in Young Children are Associated with Environmental Enteropathy and Impaired Growth.” *Tropical Medicine & International Health*, 23:26-33.

Glewwe, P., Jacoby, H. G. and King, E. M. 2001. “Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis.” *Journal of Public Economics*, 81:345-368.

Glewwe, P. and Jacoby, H. G. 1995. “An Economic Analysis of Delayed Primary School Enrollment in a Low-Income Country: The Role of Early Childhood Nutrition.” *The Review of Economics and Statistics*, 77:156-169.

Handa, S., and Peterman, A. 2007. “Child Health and School Enrollment a Replication.” *Journal of Human Resources*, 42:863-880.

Hathi, P., Haque, S., Pant, L., Coffey, D. and Spears, D. 2014. “Place and Child Health: The Interaction of Population Density and Sanitation in Developing Countries”. *Demography*, 54:337–360.

Heckman, J. J., Ichimura, H. and Todd, P. E. 1997. “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program.” *The Review of Economic Studies*, 64:605-654.

Hener, T., Rainer, H. and Siedler, T. 2016. “Political Socialization in Flux? Linking Family Non-Intactness During Childhood to Adult Civic Engagement.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179:633-656.

Ho, D. E., Imai, K., King, G. and Stuart, E. A. 2007. “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference.” *Political Analysis*, 15:199-236.

Hoddinott, J., Maluccio, J. A., Behrman, J. R., Flores, R. and Martorell, R. 2008. “Effect of a Nutrition Intervention During Early Dhildhood on Economic Productivity in Guatemalan Adults.” *The Lancet*, 371:411-416.

Hunter, P. R., Risebro, H., Yen, M., Lefebvre, H., Lo, C., Hartemann, P., et al. 2014. “Impact of the Provision of Safe Drinking Water on School Absence Rates in Cambodia: A Quasi-Experimental Study.” *PloS one*, 9:e91847.

Iacus, S., King, G. and Porro, G. 2012. “Causal Inference Without Balance Checking: Coarsened Exact Matching.” *Political Analysis,* 20:1-24.

Jamison, D. T. 1986. “Child Malnutrition and School Performance in China.” *Journal of Development Economics*, 20:299-309.

Jasper, C., Le, T. T. and Bartram, J. 2012. “Water and Sanitation in Schools: A Systematic Review of the Health and Educational Outcomes.” *International Journal of Environmental Research and Public Health*, 9:2772-2787.

Kim, S., Fleisher, B. and Sun, J.Y. 2017. “The Long-Term Health Effects of Fetal Malnutrition from the 1959-1961 Great Leap Forward Famine.” *Health Economics,* 26:1264-1277

Kim, S., Deng, Q., Fleisher, B.M. and Li, S. 2014. "The Lasting Impact of Parental Early Life Malnutrition on the Offspring: Evidence from the China Great Leap Forward Famine." *World Development,* 54:232-242.

Korpe, P. S. and Petri, W. A. 2012. "Environmental Enteropathy: Critical Implications of a Poorly Understood Condition." *Trends in Molecular Medicine*, 18:328-336.

Kosek, M., Haque, R., Lima, A., Babji, S., Shrestha, S. and Qureshi, S., et al. 2013. "Fecal Markers of Intestinal Inflammation and Permeability Associated with the Subsequent Acquisition of Linear Growth Deficits in Infants." *The American Journal of Tropical Medicine and Hygiene*, 88:390-396.

Kramer, M., Kumar, S. and Vollmer, S. 2019. *Improving child health and cognition: evidence from a school-based nutrition intervention in India.* Papers and Proceedings, AEA Meetings, Atlanta, Georgia, USA.

Lawson, N. and Spears, D. 2016. "What Doesn't Kill You Makes You Poorer: Adult Wages and Early-Life mortality in India." *Economics & Human Biology*, 21:1-16.

Leon, G. 2012. "Civil Conflict and Human Capital Accumulation the Long-Term Effects of Political Violence in Peru." *Journal of Human Resources*, 47:991-1022.

Li, W. and Yang, D.T. 2005. "The Great Leap Forward: Anatomy of a Central Planning Disaster." *Journal of Political Economy,* 113:840-847.

Lin, A., Arnold, B. F., Afreen, S., Goto, R., Huda, T. M. N., Haque, R., et al. 2013. "Household Environmental Conditions are Associated with Enteropathy and Impaired Growth in Rural Bangladesh." *The American Journal of Tropical Medicine and Hygiene*, 89:130-137.

Lunt, M. 2013. "Selecting an Appropriate Caliper Can be Essential for Achieving Good Balance with Propensity Score Matching." *American Journal of Epidemiology*, 179:226-235.

Mabud, M. A. 2005. *Demographic Implications for Health and Human Resources for Bangladesh*. Centre for Health, Population and Development, Independent University, Bangladesh.

Magadi, M. A. 2016. *Application of "Pseudo Panels" to Investigate Causal Link Between HIV and Fertility in Sub-Saharan Africa*. Conference Paper, International Conference on Sequence Analysis and Related Methods, Lausanne, June 8-10, 2016.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. 2009. "The Impact of Improving Nutrition During Early Childhood on Education Among Guatemalan Adults." *The Economic Journal*, 119:734-763.

Mertens, T. E., Fernando, M. A., Cousens, S. N., Kirkwood, B. R., Marshall, T. F., and Feachem, R. G. 1990. "Childhood Diarroea in Sri Lanka: A Case-control Study of the Impact of Improved Water Sources." *Tropical Medicine and Parasitology*, 41:98-104.

Miguel, E., and Kremer, M. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72:159-217.

Millan, T.M., Macours, K., Maluccio, J.A. Tejerina, L. 2020. "Experimental Long-Term Effects of Early-Childhood and School-Age Exposure to a Conditional Cash Transfer Program." *Journal of Development Economics,* 143:102385.

Mondal, D., Minak, J., Alam, M., Liu, Y., Dai, J., Korpe, P., et al. 2012. "Contribution of Enteric Infection, Altered Intestinal Barrier Function, and Maternal Malnutrition to Infant Malnutrition in Bangladesh." *Clinical Infectious Diseases*, 54:185-192.

Moock, P. R., and Leslie, J. 1986. “Childhood Malnutrition and Schooling in the Terai Region of Nepal.” *Journal of Development Economics*, 20:33-52.

Ncube, M. and Shimeles, A. 2012. *The Making of the Middle Class in Africa*. African Development Bank. Available at: http://www.afdb.org/fileadmin/uploads/afdb/Documents/Knowledge/AEC

Neto, U. F., Martins, M. C., Lima, F. L., Patricio, F. R. and Toledo, M. R. 1994. “Asymptomatic Environmental Enteropathy Among Slum-Dwelling Infants.” *Journal of the American College of Nutrition*, 13:51-56.

Ngure, F. M., Reid, B. M., Humphrey, J. H., Mbuya, M. N., Pelto, G. and Stoltzfus, R. J. 2014. “Water, Sanitation, and Hygiene (WASH), Environmental Enteropathy, Nutrition, and Early Child Development: Making the Links.” *Annals of the New York Academy of Sciences*, 1308:118-128.

Oster, E. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics*, 37:187-204.

Oyemade, A., Omokhodion, F. O., Olawuyi, J. F., Sridhar, M. K., and Olaseha, I. O. 1998. “Environmental and Personal Hygiene Practices: Risk Factors for Diarrhea Among Children of Nigerian Market Women.” *Journal of Diarrhoeal Diseases Research*, 16:241-247.

Pathania, V. 2009. *The Long Run Impact of Drought at Birth on Height of Women in Rural India*. Working paper, UC Berkeley.

Robins, J. M. and Rotnitzky, A. 2001. “Comments on the Bickel and Kwon Article, ‘Inference for Semiparametric Models: Some Questions and an Answer.” *Statistica Sinica*, 11: 920-936.

Rong, H., Xi, Y., An, Y., Tao, L., Zhang, X., Yu, H., et al. 2017. “The Correlation Between Early Stages of Life Exposed to Chinese Famine and Cognitive Decline in Adulthood: Nutrition of Adulthood Plays an Important Role in the Link?” *Frontiers in Aging Neuroscience*, *9*:444.

Rosenbaum, P.R. and Rubin, D.B. 1985. “Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score.” *The American Statistician,* 39:33-38.

Rosenbaum, P.R. and Rubin, D.B. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika,* 70:41-55.

Rubin, D. B. and Thomas, N. 1996. “Matching Using Estimated Propensity Scores: Relating Theory to Practice.” *Biometrics*, 52:249-264.

Rubin, D. B. and Thomas, N. 1992. “Characterizing the Effect of Matching Using Linear Propensity Score Methods with Normal Distributions.” *Biometrika*, 79:797-809.

Scheil-Adlung, X. 2015. *Global Evidence on Inequities in Rural Health Protection. New Data on Rural Deficits in Health Coverage for 174 Countries* (No. ESS 47). ESS Paper Series (SECSOC), International Labor Organization.

Sclar, G.D., Garn, J.V., Penakalapati, G., Alexander, K.T., Krauss, J., Freeman, M.C., et al. 2017. “Effects of Sanitation on Cognitive Development and School Absence: A Systematic Review.” *International Journal of Hygiene and Environmental Health,* 220:917-927.

Shivoga, W. A. and Moturi, W. N. 2009. “Geophagia as a Risk Factor for Diarrhea.” *The Journal of Infection in Developing Countries*, 3:094-098.

Smith, H. L. 1997. “Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies.” *Sociological Methodology*, 27:325-353.

Sohn, K. 2015. “The Height Premium in Indonesia.” *Economics & Human Biology*, 16:1-15.

Spears, D. and Lamba, S. 2015. "Effects of Early-Life Exposure to Sanitation on Childhood Cognitive Skills: Evidence from India's Total Sanitation Campaign." *Journal of Human Resources,* 51:298-327.

Spears, D., Ghosh, A. and Cumming, O. 2013. "Open Defecation and Childhood Stunting in India: An Ecological Analysis of New Data from 112 Districts." *PLoS One*, 8:e73784.

Umana-Aponte, M. 2011. *Long-Term Effects of a Nutritional Shock: The 1980 Famine of Karamoja, Uganda* (No.11/258). Working paper, Centre for Market and Public Organization, Bristol Institute of Public Affairs, University of Bristol.

Verbeek, M. 2008. "Pseudo-Panels and Repeated Cross-Sections," in *The Econometrics of Panel Data: Fundamental and Recent Developments in Theory and Practice, 3rd ed.* (pp. 369-383). Springer, Berlin, Heidelberg.

Vogl, T. S. 2014. "Height, Skills, and Labor Market Outcomes in Mexico." *Journal of Development Economics*, 107:84-96.

Walker, S. P., Wachs, T. D., Gardner, J. M., Lozoff, B., Wasserman, G. A., Pollitt, E., et al. 2007. "Child Development: Risk Factors for Adverse Outcomes in Developing Countries." *The Lancet*, 369:145-157.

Wanke, C. A. 2001. "To Know Escherichia Coli is to Know Bacterial Diarrheal Disease." *Clinical Infectious Diseases*, 32:1710-1712.

World Bank. 2018. *Promising Progress: A Diagnostic of Water Supply, Sanitation, Hygiene, and Poverty in Bangladesh.* World Bank, Washington, DC.

World Bank. 2015. *The Path to Universal Health Coverage in Bangladesh: Bridging the Gap of Human Resources for Health.* World Bank, Washington, DC.

Yamauchi, F., Muto, M., Chowdhury, S., Dewina, R. and Sumaryanto, S. 2011. "Are Schools and Roads Complimentary? Evidence from Income Dynamics in Rural Indonesia." *World Development,* 39: 2232-2244.

Zhang, J., & Xu, L. C. 2016. "The Long-run Effects of Treated Water on Education: The Rural Drinking Water Program in China." *Journal of Development Economics*, 122:1-15.

Zhu, C., Zhang, X., Zhao, Q. and Chen, Q. 2018. "Hybrid Marriages and Phenotypic Heterosis in Offspring: Evidence from China." *Economics and Human Biology,* 29: 102-114.

ANNEX

Table 1: Weighted Sample Mean Values of Variables in Specification (1) for Bangladesh

Variable	**Obs**	**Mean**	**Std. Dev.**
Child Enrolled in School	6,399	0.804	0.397
Unimproved Sanitation Concentration (lagged)	6,399	0.385	0.279
Urban Households (=1)	6,399	0.463	0.499
Top 60 of Wealth Index (=1)	6,399	0.763	0.425
Household Size	6,399	6.774	3.200
Mother's Education			
No Education	6,399	0.266	0.442
Primary Education	6,399	0.291	0.454
Secondary Education and Above	6,399	0.443	0.497
Female Head of Household (=1)	6,399	0.107	0.309
Female Child (=1)	6,399	0.497	0.500
Child's Age	6,399	7.448	1.092
Improved Sanitation (=1)	6,399	0.414	0.493
Improved Water Access (=1)	6,399	0.983	0.129
Percent of Children in Primary School	6,399	0.822	0.113
Sample Size by Age Group 6–7	3,288		
Sample Size by Age Group 8–9	3,111		

Table 2: Weighted Sample Mean Values of Variables in the Rural Areas of Bangladesh

Variable	**Obs**	**Mean**	**Std. Dev.**
Child Enrolled in School	2,674	0.821	0.383
Unimproved Sanitation Concentration (lagged)	2,674	0.434	0.258
Top 60 of Wealth Index (=1)	2,674	0.614	0.487
Household Size	2,674	6.991	3.218
Mother's Education			
No Education	2,674	0.312	0.463
Primary Education	2,674	0.317	0.466
Secondary Education and Above	2,674	0.370	0.483
Female Head of Household (=1)	2,674	0.107	0.309
Female Child (=1)	2,674	0.500	0.500
Child's Age	2,674	7.444	1.090
Improved Sanitation (=1)	2,674	0.392	0.488
Improved Water Access (=1)	2,674	0.974	0.160
Percent of Children in Primary School	2,674	0.826	0.118
Sample Size by Age Group 6–7	1,392		
Sample Size by Age Group 8–9	1,282		

Table 3: Weighted Sample Mean Values of Variables in the Urban Areas of Bangladesh

Variable	Obs	Mean	Std. Dev.
Child Enrolled in School	3,725	0.784	0.412
Unimproved Sanitation Concentration (lagged)	3,725	0.329	0.293
Top 60 of Wealth Index (=1)	3,725	0.936	0.245
Household Size	3,725	6.521	3.160
Mother's Education			
No Education	3,725	0.212	0.409
Primary Education	3,725	0.260	0.439
Secondary Education and Above	3,725	0.527	0.499
Female Head of Household (=1)	3,725	0.107	0.309
Female Child (=1)	3,725	0.494	0.500
Child's Age	3,725	7.453	1.095
Improved Sanitation (=1)	3,725	0.439	0.496
Improved Water Access (=1)	3,725	0.994	0.078
Percent of Children in Primary School	3,725	0.818	0.108
Sample Size by Age Group 6–7	1,906		
Sample Size by Age Group 8–9	1,819		

Table 4: Logistic and LPM Regressions including Division and Year Fixed Effects—Correlates of School Enrollment with Early Exposure to Unimproved Sanitation—by Age Groups

VARIABLES	(1) Full Sample Logit All Age	(2) Full Sample LPM All Age	(3) Full Sample Logit 6 – 7	(4) Full Sample LPM 6 - 7	(5) Full Sample Logit 8 - 9	(6) Full Sample LPM 8 - 9
Unimproved Sanitation Ratio	-0.050***	-0.052***	-0.072**	-0.073**	-0.031	-0.031
	(0.019)	(0.020)	(0.030)	(0.031)	(0.020)	(0.021)
Urban	-0.051***	-0.051***	-0.052***	-0.051***	-0.051***	-0.050***
	(0.011)	(0.011)	(0.017)	(0.017)	(0.012)	(0.012)
Top 60 of Wealth Index (=1)	-0.004	-0.007	-0.041**	-0.044**	0.036**	0.031**
	(0.013)	(0.013)	(0.020)	(0.020)	(0.016)	(0.015)
Household Size	-0.002	-0.002	0.001	0.001	-0.006***	-0.006**
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Mother's Education						
Primary (=1)	0.101***	0.098***	0.120***	0.118***	0.076***	0.082***
	(0.014)	(0.014)	(0.022)	(0.022)	(0.015)	(0.016)
Secondary and Above (=1)	0.140***	0.138***	0.184***	0.181***	0.095***	0.099***
	(0.014)	(0.014)	(0.022)	(0.022)	(0.015)	(0.016)
Female Head of Household (=1)	-0.002	-0.001	0.028	0.028	-0.032*	-0.030*
	(0.015)	(0.015)	(0.024)	(0.024)	(0.017)	(0.016)
Female Child (=1)	0.045***	0.045***	0.038***	0.040***	0.054***	0.053***
	(0.009)	(0.009)	(0.015)	(0.015)	(0.010)	(0.010)
Child's Age	0.127***	0.905***	0.221***	0.229***	0.017*	0.017*
	(0.005)	(0.069)	(0.013)	(0.015)	(0.010)	(0.010)
Improved Sanitation (=1)	0.036***	0.036***	0.043**	0.042**	0.029***	0.028***
	(0.010)	(0.010)	(0.017)	(0.017)	(0.011)	(0.010)
Improved Water Access (=1)	-0.013	-0.013	-0.040	-0.050	0.025	0.022
	(0.047)	(0.041)	(0.069)	(0.062)	(0.055)	(0.043)
Percent of Children in Primary School	-0.027	-0.020	-0.075	-0.070	0.010	0.018
	(0.043)	(0.043)	(0.071)	(0.071)	(0.046)	(0.049)
Observations	6,399	6,399	3,288	3,288	3,111	3,111
Pseudo R2/ R-squared	0.155	0.150	0.122	0.137	0.083	0.051

Robust clustered standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

Table 5: Logistic and LPM Regressions including Division and Year Fixed Effects—Correlates of School Enrollment with Early Exposure to Unimproved Sanitation—Rural Samples by Age Groups

VARIABLES	(1) Rural Sample Logit All Age	(2) Rural Sample LPM All Age	(3) Rural Sample Logit 6 - 7	(4) Rural Sample LPM 6 - 7	(5) Rural Sample Logit 8 - 9	(6) Rural Sample LPM 8 - 9
Unimproved Sanitation Ratio	-0.121***	-0.130***	-0.139***	-0.147***	-0.110***	-0.115***
	(0.029)	(0.030)	(0.047)	(0.048)	(0.030)	(0.031)
Top 60 of Wealth Index (=1)	0.008	0.009	-0.007	-0.007	0.026	0.024
	(0.015)	(0.016)	(0.025)	(0.026)	(0.016)	(0.017)
Household Size	-0.001	-0.001	-0.000	0.000	-0.003	-0.002
	(0.003)	(0.003)	(0.006)	(0.005)	(0.003)	(0.004)
<u>Mother's Education</u>						
Primary (=1)	0.083***	0.090***	0.106***	0.118***	0.055***	0.065***
	(0.016)	(0.018)	(0.027)	(0.031)	(0.017)	(0.020)
Secondary and Above (=1)	0.108***	0.111***	0.156***	0.161***	0.059***	0.066***
	(0.019)	(0.020)	(0.031)	(0.033)	(0.021)	(0.021)
Female Head of Household (=1)	-0.017	-0.016	0.011	0.012	-0.038*	-0.036
	(0.022)	(0.022)	(0.039)	(0.037)	(0.021)	(0.025)
Female Child (=1)	0.048***	0.048***	0.025	0.025	0.076***	0.071***
	(0.013)	(0.013)	(0.022)	(0.022)	(0.016)	(0.015)
Child's Age	0.128***	0.858***	0.224***	0.231***	0.028**	0.029**
	(0.009)	(0.104)	(0.021)	(0.022)	(0.014)	(0.014)
Improved Sanitation (=1)	0.012	0.013	0.016	0.015	0.008	0.007
	(0.016)	(0.015)	(0.026)	(0.026)	(0.017)	(0.015)
Improved Water Access (=1)	-0.004	-0.005	-0.033	-0.048	0.041	0.032
	(0.054)	(0.048)	(0.079)	(0.072)	(0.045)	(0.049)
Primary School Ratio	-0.036	-0.020	-0.084	-0.077	0.014	0.032
	(0.059)	(0.060)	(0.097)	(0.098)	(0.066)	(0.072)
Observations	2,674	2,674	1,382	1,382	1,292	1,292
Pseudo R2/R-Squared	0.171	0.158	0.130	0.143	0.117	0.063

Robust clustered standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

c

Table 6: Logistic and LPM Regressions including Division and Year Fixed Effects—Correlates of School Enrollment with Early Exposure to Unimproved Sanitation—Urban Samples by Age Groups

VARIABLES	(1) Urban Sample Logit all age	(2) Urban Sample LPM all age	(3) Urban Sample Logit 6 – 7	(4) Urban Sample LPM 6 - 7	(5) Urban Sample Logit 8 – 9	(6) Urban Sample LPM 8 – 9
Unimproved Sanitation Ratio	0.005	0.004	-0.019	-0.015	0.032	0.029
	(0.026)	(0.026)	(0.040)	(0.041)	(0.029)	(0.028)
Top 60 of Wealth Index (=1)	-0.024	-0.030	-0.113***	-0.110***	0.039*	0.045
	(0.023)	(0.023)	(0.038)	(0.034)	(0.021)	(0.029)
Log of Household Size	-0.004	-0.003	-0.000	-0.001	-0.008***	-0.008***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Mother's Education						
Primary (=1)	0.091***	0.103***	0.107***	0.115***	0.066***	0.094***
	(0.017)	(0.020)	(0.028)	(0.031)	(0.018)	(0.024)
Secondary and Above (=1)	0.135***	0.146***	0.175***	0.183***	0.092***	0.113***
	(0.017)	(0.019)	(0.027)	(0.029)	(0.018)	(0.023)
Female Head of Household (=1)	0.006	0.009	0.038	0.039	-0.021	-0.020
	(0.020)	(0.020)	(0.033)	(0.033)	(0.019)	(0.022)
Female Child (=1)	0.039***	0.040***	0.040**	0.043**	0.041***	0.039***
	(0.012)	(0.012)	(0.019)	(0.019)	(0.014)	(0.014)
Log of Child's Age	0.127***	0.942***	0.224***	0.231***	0.008	0.009
	(0.007)	(0.092)	(0.018)	(0.020)	(0.014)	(0.014)
Improved Sanitation (=1)	0.053***	0.049***	0.061***	0.059***	0.045***	0.039***
	(0.014)	(0.013)	(0.022)	(0.022)	(0.017)	(0.014)
Improved Water Access (=1)	-0.077	-0.067	-0.175	-0.142	-0.000	-0.009
	(0.105)	(0.078)	(0.181)	(0.100)	(0.091)	(0.083)
Primary School Ratio	0.007	0.009	-0.039	-0.023	0.041	0.036
	(0.064)	(0.062)	(0.103)	(0.104)	(0.067)	(0.067)
Observations	3,725	3,725	1,906	1,906	1,819	1,819
Pseudo R2/R-Squared	0.156	0.154	0.133	0.150	0.087	0.058

Robust clustered standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

Table 7: Comparison of Estimated Coefficients for Unimproved Sanitation across Alternative Specifications

		(1)	(2)	(3)	(4)	(5)
Age Group	**Sample**	**Preferred Specification**		**With District-Level Fixed Effect**	**With Community Electricity Access**	**With Community Electricity Access & Wealth**
		Logit	**LPM**	**Logit**	**Logit**	**Logit**
All Age	Full Sample	-0.050***	-0.052***	-0.054***	-0.054**	-0.074***
		(0.019)	(0.020)	(0.021)	(0.022)	(0.024)
		[0.155]	[0.151]	[0.178]	[0.155]	[0.156]
	Rural	-0.121***	-0.130***	-0.132***	-0.113***	-0.123***
		(0.029)	(0.030)	(0.034)	(0.037)	(0.038)
		[0.171]	[0.158]	[0.214]	[0.171]	[0.172]
	Urban	0.005	0.004	-0.014	-0.019	-0.034
		(0.026)	(0.026)	(0.029)	(0.029)	(0.033)
		[0.156]	[0.154]	[0.179]	[0.157]	[0.157]
6–7 Years Old	Full Sample	-0.072**	-0.073**	-0.088***	-0.069**	-0.092**
		(0.030)	(0.031)	(0.032)	(0.035)	(0.038)
		[0.122]	[0.137]	[0.145]	[0.122]	[0.122]
	Rural	-0.139***	-0.147***	-0.148***	-0.130**	-0.153**
		(0.047)	(0.048)	(0.056)	(0.061)	(0.062)
		[0.130]	[0.143]	[0.158]	[0.130]	[0.132]
	Urban	-0.019	-0.015	-0.059	-0.035	-0.041
		(0.040)	(0.041)	(0.044)	(0.045)	(0.051)
		[0.133]	[0.150]	[0.159]	[0.133]	[0.133]
8–9 Years Old	Full Sample	-0.031	-0.031	-0.020	-0.042*	-0.062**
		(0.020)	(0.021)	(0.023)	(0.023)	(0.025)
		[0.083]	[0.051]	[0.116]	[0.084]	[0.086]
	Rural	-0.110***	-0.115***	-0.144***	-0.106***	-0.104**
		(0.030)	(0.031)	(0.042)	(0.040)	(0.040)
		[0.117]	[0.063]	[0.183]	[0.117]	[0.117]
	Urban	0.032	0.029	0.042	0.002	-0.024
		(0.029)	(0.028)	(0.037)	(0.032)	(0.037)
		[0.087]	[0.058]	[0.111]	[0.092]	[0.093]

Table 8: Robustness to Omitted Variable Bias—Coefficient Stability

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Logit		LPM		Bias-Adjusted Coefficient			Delta
Age Group	Sample	Controlled Marginal Effect	Uncontrolled Effect	Controlled Effect	95% CI	Rmax = 2.2*(R2)	Identified Set	Delta = 1.5	Beta = 0
All Age	All Children	-0.050***	-0.049***	-0.052***	[-0.091, -0.014]	-0.061	[-0.061, -**0.052**]	-0.071	3.709
		(0.019)	(0.019)	(0.020)					
		[0.155]	[0.001]	[0.151]					
	Rural	-0.121***	-0.147***	-0.130***	[-0.188, -0.071]	-0.09	[**-0.130**, -0.090]	-0.055	1.941
		(0.029)	(0.029)	(0.030)					
		[0.171]	[0.010]	[0.158]					
6–7 Years Old	All Children	-0.072**	-0.057**	-0.073**	[-0.134, -0.012]	-0.113	[-0.113, **-0.073**]	-0.159	28.699
		(0.030)	(0.029)	(0.031)					
		[0.122]	[0.001]	[0.137]					
	Rural	-0.139***	-0.158***	-0.147***	[-0.241, -0.054]	-0.121	[**-0.147**, -0.121]	-0.096	2.207
		(0.047)	(0.046)	(0.048)					
		[0.130]	[0.009]	[0.143]					
8–9 Years Old	Rural	-0.110***	-0.125***	-0.115***	[-0.176, -0.053]	-0.084	[**-0.115**, -0.084]	-0.052	1.852
		(0.030)	(0.030)	(0.031)					
		[0.117]	[0.015}	[0.063]					

Table 9: Comparison of Results from Matched Samples and Logit and LPM Neighbor including Division and Year Fixed Effects

	(1)	**(2)**		**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	(8)
	Unmatched Sample								
	LPM	**Logit**		**Coarsened Exact Matched Sample**	**Propensity Score Matched Samples**				**Estimated Values**
Age Group		**Marginal Effect**	**95% CI**		**Caliper Matching 3 Nearest Neighbors**	**Caliper Matching 5 Nearest Neighbors**	**Radius Matching**	**Kernel Matching**	**(4) – (8)**
All Age	-0.052***	-0.050***	[-0.091, -0.014]	-0.069***	-0.080***	-0.083***	-0.075***	-0.077***	**[-0.083, -0.069]**
	(0.020)	(0.019)		(0.022)	(0.023)	(0.022)	(0.021)	(0.021)	
Observations		6,399		5,694	4,885	5,295	6,298	6,298	
Treatment | Control		2,882 | 3,517		2,519 | 3,175	2,781 | 2,104	2,781 | 2,514	2,781 | 3,517	2,781 | 3,517	
Off Support				705	1,514	1,104	101	101	
R-squared		0.150 / 0.155		0.161	0.158	0.151	0.155	0.155	
6–7 Years Old	-0.073**	-0.072**	[-0.134, -0.012]	-0.084**	-0.078**	-0.079**	-0.075**	-0.078**	**[-0.084, -0.075]**
	(0.031)	(0.030)		(0.037)	(0.039)	(0.038)	(0.036)	(0.036)	
Observations		3,288		3,075	2,498	2,712	3,242	3,242	
Treatment | Control		1,482 | 1,806		1,350 | 1,725	1,436 | 1,062	1,436 | 1,276	1,436 | 1,806	1,436 | 1,806	
Off Support				213	790	576	46	46	
R-squared		0.137 / 0.122		0.119	0.109	0.105	0.111	0.109	
8–9 Years Old	-0.031	-0.031	[-0.069,0.008]	-0.041**	-0.037	-0.041*	-0.055***	-0.054***	**[-0.054, -0.037]**
	(0.021)	(0.020)		(0.020)	(0.023)	(0.023)	(0.020)	(0.021)	
Observations		3,111		2,897	2,346	2,535	3,042	3,043	
Treatment | Control		1,400 | 1,711		1,267 | 1,630	1,332 | 1,014	1,332 | 1,203	1,332 | 1,710	1,332 | 1,711	
Off Support				214	765	576	69	68	
R-squared		0.051 / 0.083		0.086	0.105	0.106	0.104	0.103	
Average Cohen's D		0.169		0.097	0.031	0.033	0.028	0.028	

Robust clustered standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Geographic Distribution of Household Clusters Prior and After Geomatching

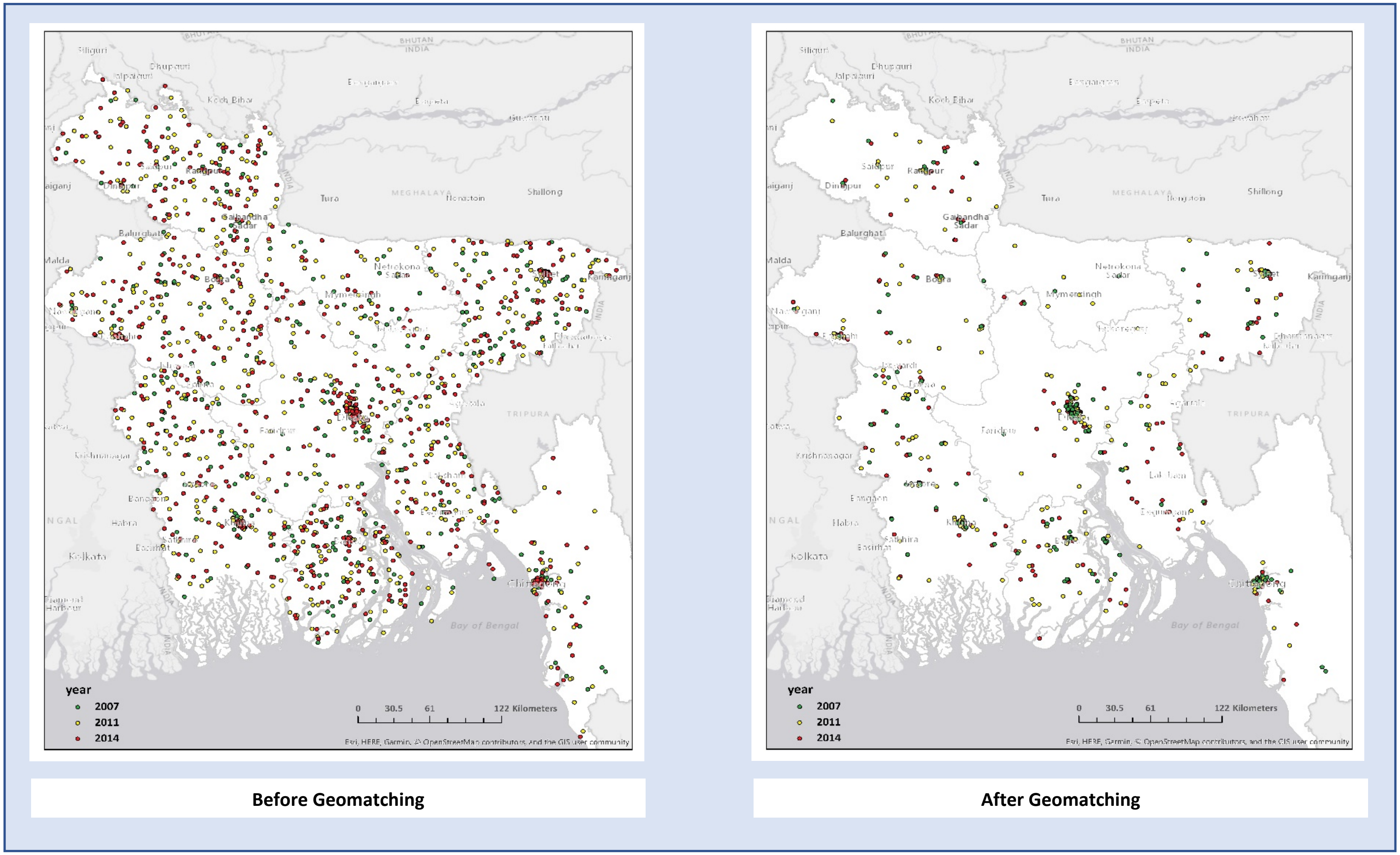

Figure 2: Propensity Score Balance between Treatment and Comparison Groups Prior to Matching

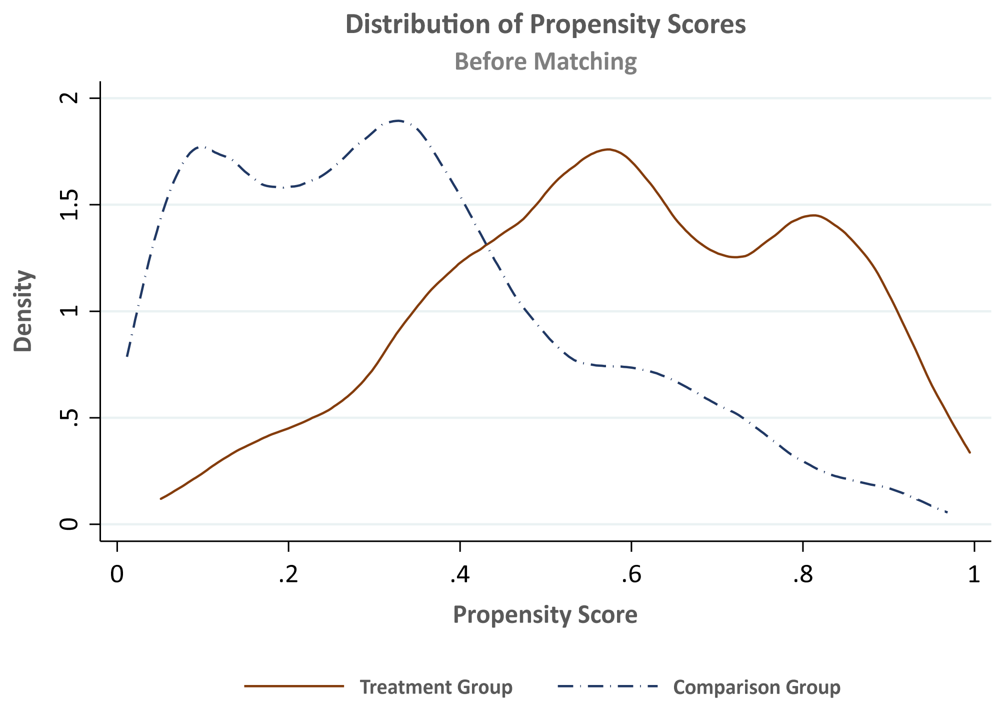

Figure 3: Propensity Score Balance between Treatment and Comparison Groups for Each Propensity Score Matched Sample

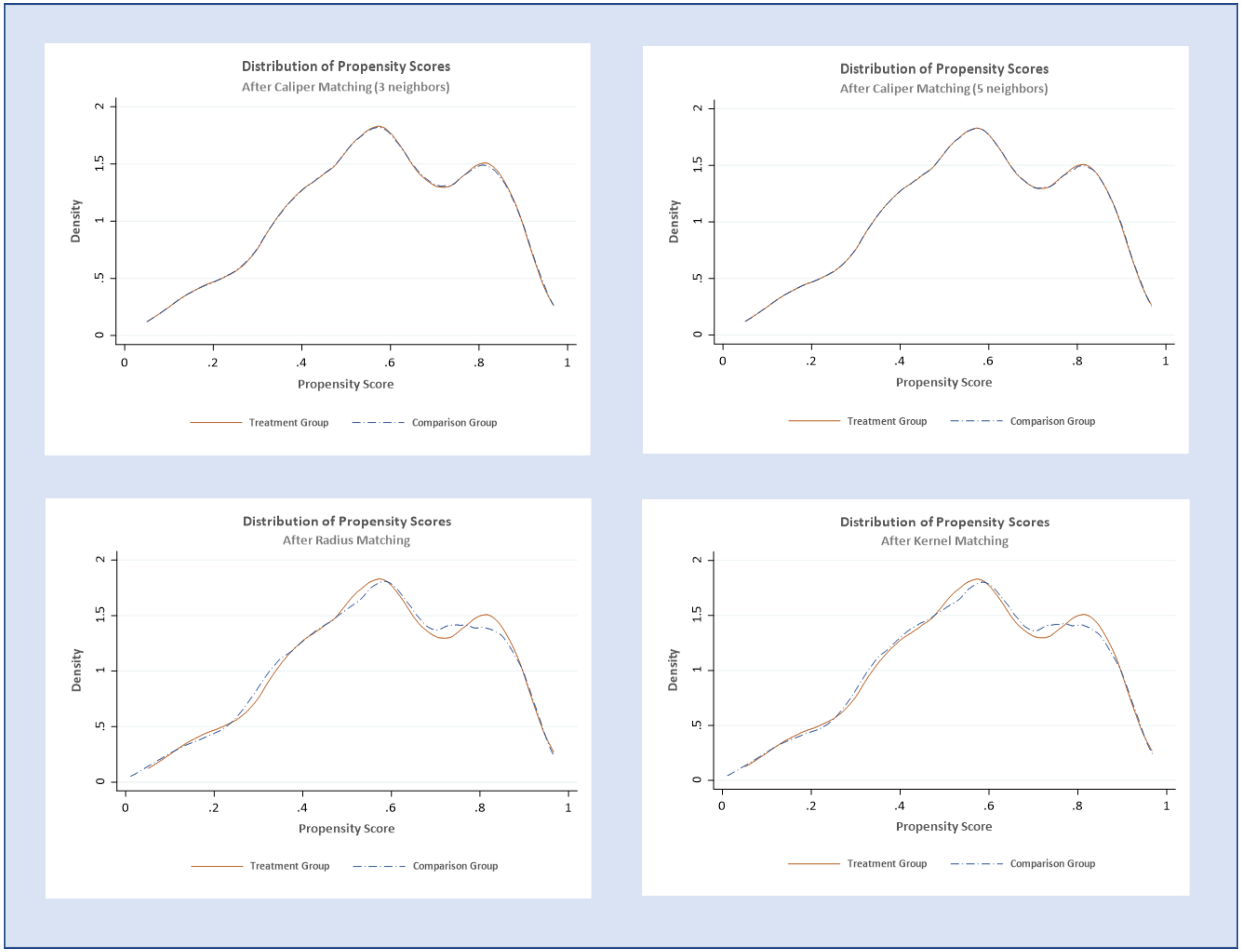

Figure 4: Comparison of Covariates Balance between Treatment and Control Group across Each Matched Sample