Senior Thesis

Department of Economics

"Access to Improved Sanitation Facilities and Primary Education: A Cross-Country Analysis of Developing Countries"

Michael Grund

May, 2016

Advisor: *Dr. Francesco Renna*

Abstract

This study examines the link between access to improved sanitation facilities and primary education performance. A 217 Country/Year panel data set from 40 *Low or Lower Middle* income countries are used from the World Bank's World Development Indicators. I estimate the impact of access to improved sanitation facilities on primary education completion and repetition rates using both pooled OLS and two-way fixed effects regression models. Results of the two-way fixed effects model imply at a 99 percent significance level that as access to improved sanitation facilities in a country increase by 1 percent, a country will see a 1.3 percent increase in primary education completion rates and a 0.37 percent decrease in primary repetition rates at the 5 percent significance level. This paper suggests the effect is caused by increases in a person's health which then leads to increases in primary education performance.

Acknowledgements

I would like to thank Dr. Francesco Renna and Dr. Michael Nelson for advising and providing guidance during the creation of this senior thesis. Dr. Renna is an Associate Professor of Economics at The University of Akron specializing in the fields of Applied Econometrics and Labor Economics. Dr. Nelson is a Professor of Economics and Department Chair at The University of Akron specializing in the Public Sector and Economic Development fields.

Table of Contents

Introduction

According to the World Health Organization (Key Facts JMP2015)), "About 2.6 billion people – half the developing world – lack even a simple 'improved' latrine and 1.1 billion people have no access to any type of improved drinking source of water." With half of the developing world lacking access to improved disposal of fecal waste, people in developing countries are at a much higher risk of catching fecal borne diseases. Many of the diseases transmitted through the fecal-oral route such as Cholera, Typhoid fever, tape worms, E. Coli, and diarrhea can cause debilitating or even life threatening illnesses because healthcare is limited and too expensive for those living in developing countries. According to the WHO, "low-income countries have ten times fewer physicians than high income countries" (Key Facts JMP2015).

People who do not have access to an improved sanitation facilities are subject to deteriorated health that can impact many aspects of their lives. Empirical studies have shown that certain diseases and medical conditions can lead to decreased school performance and decreases in cognitive ability. The removal of water-borne disease has been shown to increase years of schooling, enrollment, attendance, and literacy rates of both children and young adults (Barde 2014). Education in addition to health are two main components of human capital formation and human capital formation has been shown to be a significant factor of economic growth.

This study examines the link between access to improved sanitation facilities and education. The link between health and education has been researched by others including Thuilliez (2007) who studied the link between malaria and primary education. Primary education completion rates and repetition rates are used to represent education and the percent of a country that has access to an improved sanitation facility is used to represent sanitation access. This paper's main contribution to the literature is analyzing the relationship between improved

sanitation facility access and educational performance, specifically primary education. Other studies such as Barde (2014) examine the link between tap water access and cognitive ability and Spears (2013) who examines the link between access to improved sanitation facilities and height as a proxy for health.

An advantage of this study is that it focuses on "improved" sanitation facility access using 40 countries over a 16-year time span. Improved sanitation facilities are a category defined by the world bank that includes specific forms of facilities such as toilets. This paper provides a policy option for developing countries to increase human capital formation by building or providing access to improved sanitation facilities for people currently without access.

Literature Review

Past studies have examined the link between various determinants of health and educational attainment levels. Studies including Barde (2014) and Thuilliez (2007) examine the link to primary education specifically based on a student's access to tap water and their susceptibility to malaria. Spears (2013) focuses specifically on not having access to a sanitation facility and the effect on a person's height. These studies help to provide a foundation for the premise of this study.

Barde and Walkiewicz (2014) examine the relationship between piped water and test scores for students in Brazilian primary schools between the years of 1998 and 2005. The motivation behind their study was the high number of people in the world living without access to water and sanitation infrastructure. Without this infrastructure, many people's health becomes compromised due to disease. Barde (2014) wanted to further study the effect of the compromised health on human capital formation, specifically education. The data for this study was obtained from the *Sistema Nacional de Avaliação da Educação Básica,* a survey collected by Brazil's

department of education. Barde and Walkiewicz (2014) found a "positive and significant effect" that explains 14 percent of the standard deviation of test scores caused by differences in piped water access. After finding this result the authors went further in studying the relationship between the mother's education levels and the child's because a child's test scores could be affected by how well a mother teaches them. A mother with lower educational attainment is less likely to teach their child proper hygiene education. They concluded that a large emphasis should be placed on providing education programs on hygiene because children who have access to piped water but have mothers who do not have basic education perform similarly to children with educated mothers and no piped water. A further study into the differences between rural and urban areas is stated as needed and that this study focused primarily on urban areas.

Beach (2014) motivation is the large and documented effect of water infrastructure on short term health of children but lack of knowledge on the effect of this infrastructure on human capital 20- 30 years after. Beach finds that contaminated water will lead to decreased earnings of 1-9 percent and education of a person by 1-9 months after early life exposure. As a policy recommendation they state that the cost of investing in water infrastructure would be offset by future increased earnings. The data for this study was from 1900 and 1940. Typhoid fever rates from the 1900 and 1940 U.S censuses were used to represent water contamination at an early age and education and income were used for human capital later in a person's life. The analysis primarily focused on white males due because this time period was before the "great migration" of African Americans to cities. The data used are from 75 different cities and the majority of these cities were in the Northeast and Midwest. Beach found that eradication of typhoid fever alone would have led to an average increase of education of one to nine months per person and in increase of income of 1-9 percent.

Height has been used to represent human capital in many of the studies on the subject of sanitation and human capital in addition to education because children who face poor health conditions as children have stunted height according to Spears (2013). Dean Spears is an author of many articles on the subject. For example, "Essays in the Economics of Sanitation and Human Capital in Developing Countries" (2013) is based on free data from Demographic and Health Surveys. Hammer and Spears (2013) focused on a village sanitation program by the government of Maharashtra, India in 2004 and the effect on child height as a representation of human capital. They found an increase in child height of 1.3 centimeters in the villages where the sanitation program was put into place compared to the villages where the program did not take place. The authors state the many limitations to this study, one of which is that they could not actually prove latrine use because the questions in the survey pertained to ownership rather than usage. It is important to include latrine use because ownership of a latrine does not decrease fecal borne disease but usage of the latrine does.

Negative externalities are found to be caused by open defecation and as Spears suggests this as a reason to change public policy to focus on sanitation due to its high welfare gains. Hammer and Spears note that not all parts of the government projects to eliminate open defecation are the best use of limited funds because the effect on health is less in rural areas compared to urban. They specify the building of sanitation infrastructure as their policy recommendation in densely populated areas to increase levels of health, which they used height to represent.

A primary goal of the paper by Spears and Lamba (2011) was to find a relationship between the Total Sanitation Campaign and cognitive skills of Indian children. The study shows that a low cost latrine sanitation system put in by the government had successful results in

enhancing cognitive achievement of children in India. Data for this study were collected from the local district governments in India where the Total Sanitation campaign took place. They found that exposure to the Total Sanitation campaign at later points in one's life after the early devolvement showed little to no effect on their cognitive performance. In addition to cognitive performance, Spears and Lamba (2011) also compare the findings to height and find similar upward trends for both height and cognitive ability due to the sanitation efforts.

Spears (2013) claims to have made four contributions to the literature in attempting to explain the variation in child height between countries: (1) it is the first cross-country quantitative analysis between sanitation and height; (2) it finds a negative relationship between population density and height (based on sanitation) because children are exposed to higher amounts of waste in denser populations; (3) it gives a reasonable explanation of the "Asian enigma", i.e. Asians are short because they lack access to sanitation; and (4) it concludes that height is not only caused by malnutrition but also by lack of sanitation.

Data for this study was assembled from four different sources but the main source was Demographic and Health Surveys. The study shows that open defection in urban areas have a larger negative effect on height than in rural areas. Children in India where 55% of households openly defecated were two standard deviations shorter than the mean height. Finally, they find conclusive evidence that genetics does not play as large of role in explaining the difference in height for India and Africa as it does in the United States. (Ex If mother is short; their child will be short).

Thuilliez (2007) conducted a cross-country study of the effects of malaria on education to suggest that diseases such as malaria result in decreases in cognitive ability. This study finds that countries with intensive malaria have a 29 percent lower primary school completion rate

compared to countries without malaria. Thuilliez (2007) also finds that the countries with intensive malaria have 9 percent higher primary school repetition rates than countries without malaria. These results suggest that disease and illness can have a significant effect on education rates such as the significantly lower repetition and completion rates found by Thuilliez (2007). This study is pertinent to this paper because Thuilliez (2007) uses malaria as a proxy for health but a country's access to improved sanitation infrastructure can also be used as a health determinant of education based on the theoretical model presented in Thuilliez (2007).

A current gap in the literature exists for having access to improved sanitation infrastructure and its effect on educational attainment. Most studies focus on the direct effect of sanitation on health or the effect of water infrastructure (tap water, piped water) on cognitive ability. This paper takes the study further by attempting to explain the relationship between having an improved household sanitation facility and levels of educational attainment. Based on Thuilliez (2007) literature, I expect that educational attainment and having an "improved" sanitation facility will be positively related because open defecation leads to disease as found by Spears (2013). I base this hypothesis on Thuilliez (2007) theoretical model for the determinants of primary education completion and repetition rates. This paper's model will not include Malaria or other malaria specific variables but will include a country's percent access to improved sanitation facilities as shown in the conceptual framework section.

Conceptual Framework

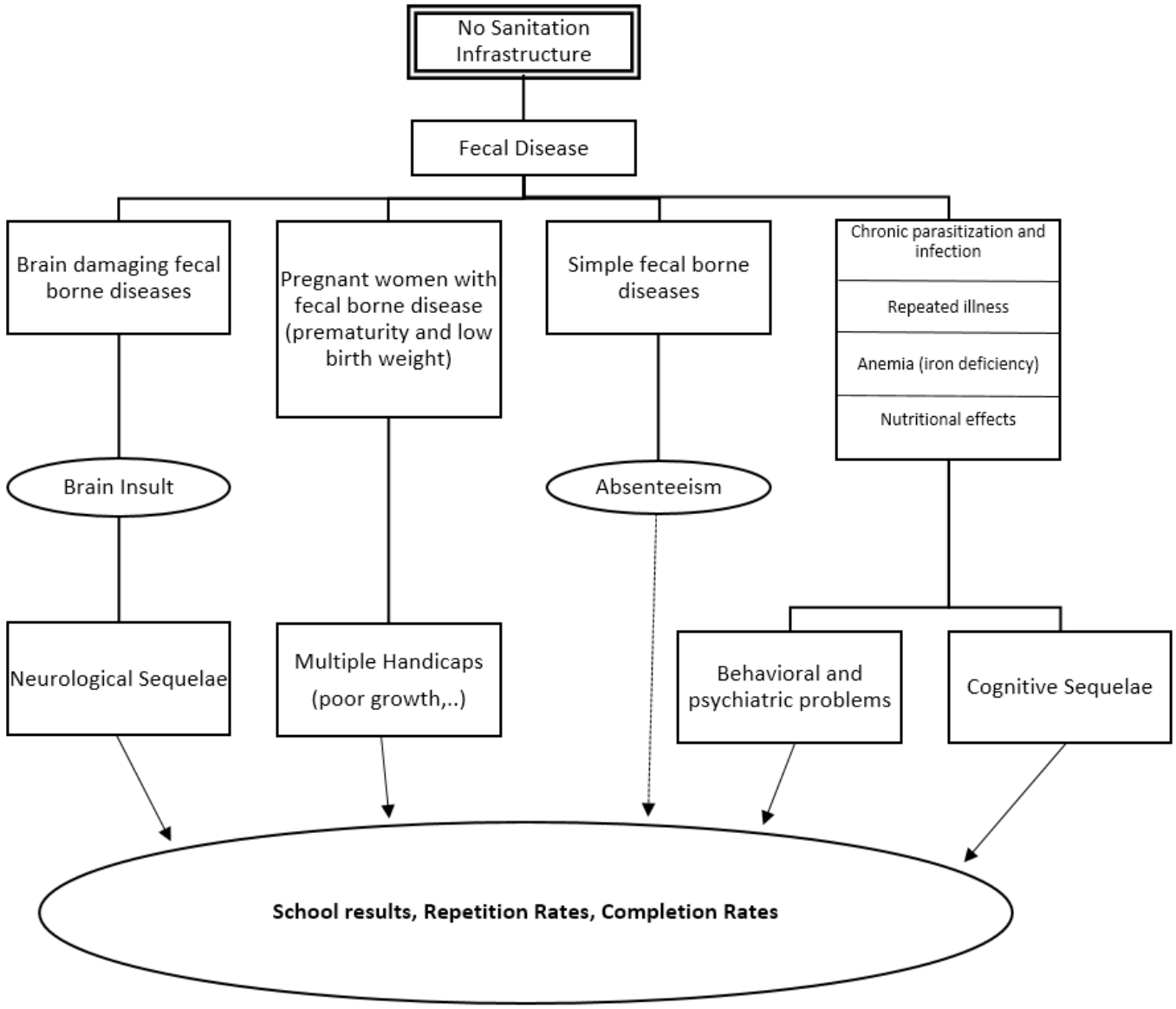
This study's conceptual framework is based on Thuilliez (2007). Below is the production function of education for this study where Primary Education is used as the dependent variable, specifically *Primary repetition and Completion Rates*:

$$Q = Q\,(F,R,H) + \varepsilon$$

In this model, Q represents schooling quality, F represents family factors, R represents public resources used for primary education, H represents health conditions, and ε represents unmeasured influencing factors. Thuilliez (2007) primary independent variable of interest is Malaria and its effect on primary education performance. I will use a country's percentage of people who have access to improved sanitation facility as part of H in the above production function in place of malaria but will maintain the variables used by Thuilliez (2007) for Q, F, and R.

In Figure 1 below I show how health and ultimately education can be effected by not having access to improved sanitation facilities. Examples of illnesses and health conditions caused by fecal borne diseases represented in Figure 1 are found in Thuilliez (2007), Brown (2013), and Heijnen (2014). Figure 1 shows that not having access to an improved sanitation facility can lead to fecal borne disease. This fecal borne disease can lead to sickness that results in school absenteeism and decreases in cognitive ability. These negative effects of not having access to an improved sanitation facility ultimately lead to decreased school completion and repetition rates.

Figure 1: Direct Impacts of Fecal Borne Disease on School Performance (*Source*: Thuilliez (2007))



This study will use *Low or Lower-middle income* countries as defined by the World Bank because high income OECD and non OECD countries have high percentage access to improved sanitation facilities.

Data and Methodology

The World Bank's World Development Indicators are used to create a panel dataset with eight variables drawn from 40 different countries over the years of 1998 through 2014. This creates a dataset with 217 Country/Year observations from 40 developing countries considered to have *Low* or *Lower Middle income*. Table 2 is located in the appendix showing which countries and years are used in this study because data are not available for every year for every country which limits the data set to 40 countries.

This study's dependent variable of interest is education. I will use two different dependent education variables based on Thuilliez (2007). The first dependent education variable being used is primary school repetition rates which are defined by the World Bank as being "calculated by dividing the sum of repeaters in all grades of primary school by the total number of students enrolled in primary school, and multiplying by 100". Thuilliez (2007) states that repetition rates are more directly affected by school results than other factors including different school and educational policies. Test scores are not directly used as a dependent variable because they are not comparable across countries due to different standards.

The second dependent variable used will be primary school completion rates. The World Bank defines Primary Completion Rates as being calculated by "dividing the number of new entrants (enrollment minus repeaters) in the last grade of primary education, regardless of age, by the population at the entrance age for the last grade of primary education and multiplying by 100". Primary education rates "can exceed 100% due to over-aged and under-aged children who enter primary school late/early and/or repeat grades" according to the World Bank. A second variable is being used because it has been shown that not all effects on education are shown by just repetition rates (Al Smarrai (2006)). Thuilliez (2007) uses primary repetition and primary

completion rates because they are often used in cross-country comparisons of school outcomes in international monitoring reports.

To represent improved sanitation facilities in this study I use the percentage of people in a country with access to an improved sanitation facility. Improved sanitation facilities include: flush/pour flush, ventilated improved pit latrines, pit latrines with slab, and composting toilets. The World Bank provides this data from the World Health Organization and the United Nations Children Fund and the original data was collected through household surveys.

Other independent variables include expenditure on education as percent of total government expenditure, pupil-teacher ratio in primary education (headcount basis), government expenditure per primary student as a percent of GDP per capita, and expenditure on primary education as a percentage of government expenditure on education as identified by Thuilliez (2007). GDP per capita is also included as an explanatory variable because I use 40 different countries and higher levels of GDP per capita have been shown to be a proxy for parents' income (Lee 2000).

Mortality rates under the age of 5 per 1000 people are used as an explanatory variable because countries with higher mortality rates for children have been shown to have generally overall lower levels of child health. Percent of a country considered to be Urban is used as an explanatory variable because there are lower transportation costs to get to school compared to rural areas that could affect school attendance. Another factor of urbanization is increased access to amenities such as electricity that can increase educational performance.

Econometric Model

This study uses both Pooled OLS models and two-way fixed effects models to examine the relationship between a country's access to improved sanitation facilities and primary education completion and repetition rates. The econometric models are represented below for primary completion rates and primary repetition rates respectively:

$$PrimaryCompRate_{it} = \alpha_0 + \alpha_1 SaniFacil_{it} + \alpha_2 GDPPCAP_{it} + \alpha_3 UrbPop_{it} - \alpha_4 MortRate5_{it} - \alpha_5 TeacherRatio_{it} + \alpha_6 GovEduExp_{it} + \mu_i + \gamma_t + \varepsilon_t$$

$$RepeaterPercent_{it} = \alpha_0 + \alpha_1 SaniFacil_{it} + \alpha_2 GDPPCAP_{it} + \alpha_3 UrbPop_{it} - \alpha_4 MortRate5_{it} - \alpha_5 TeacherRatio_{it} + \alpha_6 GovEduExp_{it} + \mu_i + \gamma_t + \varepsilon_t$$

The Two-Way fixed effects model is used in addition to the pooled OLS model to control for unobserved heterogeneity between countries and over time. Variable descriptions are provided in appendix Table 1. Multiple regressions will be run for each model with the different education variables included to compare as performed in previous studies.

Pooled OLS Results

Primary Completion Rate

Results for the Pooled OLS regressions are shown in Appendix Table 5. Three different models were run including different measures of government spending on education for each dependent variable. The results from the three different primary completion rate models are very similar to one another and none of the government education spending variables were statistically significant.

The main variable of interest of a country's percent access to improved sanitation facilities did not provide statistically significant results in any of the three models. Additionally, the results showed a 0.1 percent decrease in primary education completion rates for every 1 percent increase in improved sanitation facility access. This sign is opposite of the expected sign based on the literature and conceptual framework. The expected result is an increase in primary education completion rates as access to improved sanitation facilities in a country increases.

Four of the five control variables were statistically significant in all three models with the exception of the government education spending variables. The variable shown to have the highest change in primary education completion rates is GDP per capita. As GDP per capita increases by 1 percent, primary education completion rates increase between 4.7 and 5.1 percent in all three models. Results may be biased due to multicollinearity between a country's percent access to improved sanitation facilities and GDP per capita, mortality rate, and primary education completion rates as shown in Table 4.

Primary Repetition Rate

Results of the pooled OLS regression for primary repetition rates also have low explanatory value for access to improved sanitation. Two of the three regressions are statistically significant at the 10 percent level but all three estimates are below 0.06 of a percent change in primary repetition rate as improved sanitation facility access increases by 1 percent. The small change makes the variable's impact negligible. The positive sign of the estimate is not as expected based on the literature. All three estimates suggest that an increase in a country's access to improved sanitation results in an increase in primary repetition rates.

All the control variables used in this model are statistically significant with the exception of a country's average primary education student teacher ratio. As GDP per capita increases in a

country by 1 percent, there is a primary education repetition rate decrease between 1.96 and 2.13 percent across the three models. The education spending variables are all statistically significant in this model but show different signs. Overall, both of the pooled OLS models are inconsistent and do not correspond with the literature. Results for the two-way fixed effects model suggest that there is unobserved heterogeneity in the pooled OLS models.

Two-way Fixed Effects Results

Results for the two-way fixed effects models yield much different results than the pooled OLS models because they control for unobserved heterogeneity between countries and over time. The same observations and variables are used in the two-way fixed effects models as were used in the pooled OLS models. The results for the two-way fixed effects models can be found in Appendix Table 6.

Primary Completion Rate

The main variable of interest percent access to improved sanitation facilities in a country is statistically significant at the 1 percent level across all models. The estimates show that as access to improved sanitation facilities in a country increase by 1 percent, primary education completion rates increase between 1.2 and 1.3 percent across all three models. The mean value for improved sanitation facilities across the 40 countries used is 36 percent (Descriptive Statistics are located in Appendix Table 4). The sign of the estimate is positive as expected based on prior literature and the theoretical model. These results are much higher and more statistically significant than those from the pooled OLS model.

Two of the control variables are significant in the primary education completion rate two-way fixed effects models in addition to improved sanitation facility access. Mortality rate under the age of 5 in a country is significant at the 1 percent level but yields small parameter estimates.

The estimates show that as the mortality rate under the age of 5 in a country increase by 1 percent, primary education completion rates will decrease by 0.2 percent. The second variable of statistical significance is government education spending as a percentage of government spending.

Primary Repetition Rate

Results for the primary education repetition rate improve after using the two-way fixed effects model to control for unobserved heterogeneity compared to the pooled OLS. The main variable of interest shows that as access to improved sanitation facilities increases by 1 percent, there is a 0.31 percent decrease in primary education repetition rates. This result is significant at the 5 percent level and has the expected negative sign. The result is significant and the economic impact at 0.31 of a percent compared to the mean repetition rate across all 40 countries of 12.6 percent has a real world impact.

The control variables used in this model show little explanatory value due to either their level of statistical significance and their parameter estimate's real world impact. GDP per capita has the expected positive sign but is not statistically significant for any of the three regressions. Another issue with a control variable is Urban Population's negative sign. The theory shows that higher percentages of urban populations should lead to a decrease in repetition rates. The results show a statistical significance level of 10 percent for all three regressions that repetition rates will increase by approximately 0.65 percentage for every 1 percent increase in the urban population.

Another statistically significant variable is the pupil-teacher ratio but the results are opposite of the expected sign. As the pupil-teacher ratio increases by one student, primary repetition rates decrease by approximately 0.17 of a percent. The expected result is that smaller

classes would decrease repetition rates. The last set of education variables have no statistical significance and all have estimates of 0.10 percent or less. As education spending increases by 1 percent, repetition rates decrease by approximately 0.10 percent. This is the expected sign of education spending.

Conclusions

This study finds that both primary education completion and repetition rates are can be improved by increasing access to improved sanitation facilities. These results were expected based on Thuilliez (2007) study of malaria's impact on primary education completion and repetition rates. The results of this paper show that increases in access to improved sanitation facilities lead to an increase in primary education completion rates of 1.3 percent for every 1 percent increase in improved facility access. Countries with lower levels of improved sanitation facility access could see significant gains in completion rates if they improved their country's access. Repetition rates show lower gains due to improved sanitation facility access but with larger increases in access by a country repetition rates could decline by an impactful amount.

Increasing sanitation facility access as a policy campaign as done in India and studied by Spears (2013) could provide a tool for increasing developing economies levels of human capital. Human capital has been shown to be a key to economic development and increasing a country's access to improved sanitation facilities could show a return on investment. Further studies comparing this policy option to others such as improving tap water access as studied by Barde (2014) could provide insight to which investment would be best for a developing country to undertake in increasing human capital and spurring economic growth.

Limitations

The purpose of this study was to show that educational attainment can be increased if improved sanitation systems are installed. However, it is important to recognize that proper waste disposal's effect on education is indirect because proper disposal actually is increasing a person's health as shown by Spears (2013) and Thuilliez (2007). There are many other factors that affect health in addition to fecal-borne disease that could have large impacts on educational attainment. Other educational factors should be assessed as well and included in future research. While these health and education factors are not in the scope of this research study, they should be further researched and compared to these results.

References

Adukia, Anjali. "Sanitation and Education." Cambridge, MA: Harvard Graduate School of Education (2013).

Al-Samarrai, Samer. "Achieving education for all: how much does money matter?" Journal of international development 18, no. 2 (2006): 179-206.

Barde, Julia A., and Juliana Walkiewicz. 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, 2014.

Beach, Brian, Joseph Ferrie, Martin Saavedra, and Werner Troesken. Typhoid Fever, Water Quality, and Human Capital Formation. No. w20279. National Bureau of Economic Research, 2014.

Brown, Joe, Sandy Cairncross, and Jeroen HJ Ensink. "Water, sanitation, hygiene and enteric infections in children." Archives of disease in childhood 98, no. 8 (2013): 629-634.

Hammer, Jeffrey S., and Dean Spears. "Village sanitation and children's human capital: evidence from a randomized experiment by the Maharashtra government." World Bank Policy Research Working Paper 6580 (2013).

Heijnen, Marieke, Oliver Cumming, Rachel Peletz, Gabrielle Ka-Seen Chan, Joe Brown, Kelly Baker, and Thomas Clasen. "Shared sanitation versus individual household latrines: a systematic review of health outcomes." PLoS One 9, no. 4 (2014): e93300.

"Key Facts from JMP 2015 Report." World Health Organization. Accessed February 10, 2016. http://www.who.int/water_sanitation_health/monitoring/jmp-2015-key-facts/en/.

Lee, Jong–Wha, and Robert J. Barro. "Schooling quality in a cross–section of countries." *Economica* 68, no. 272 (2001): 465-488.

Spears, Dean, and Sneha Lamba. Effects of Early-Life Exposure to Sanitation on Childhood Cognitive Skills: Evidence from India's Total Sanitation Campaign. The World Bank, Policy Research Working Paper Series: 6659, 2013.

Spears, Dean. "How much international variation in child height can sanitation explain?" World bank policy research working paper 6351 (2013).

Thuilliez, Josselin. "Malaria and primary education: A cross-country analysis on primary repetition and completion rates." (2007).

Appendix

Table 1: Variable Descriptions and Sources

Variable	Description	Website
PrimaryCompRate	Primary completion rate, both sexes (%)	World Bank http://www.worldbank.org/
RepeaterPercent	Percentage of repeaters in primary education, all grades, both sexes (%)	World Bank http://www.worldbank.org/
SaniFacil	Improved sanitation facilities (% of population with access)	World Bank http://www.worldbank.org/
GDPPCAP	GDP per capita, PPP (constant 2011 international $)	World Bank http://www.worldbank.org/
UrbPop	Urban population (% of total)	World Bank http://www.worldbank.org/
MortRate5	Mortality rate, under-5 (per 1,000)	World Bank http://www.worldbank.org/
TeacherRatio	Pupil-teacher ratio in primary education (headcount basis)	World Bank http://www.worldbank.org/
GovEduExp	Expenditure on education as % of total government expenditure (%)	World Bank http://www.worldbank.org/
GDPPrimExp	Government expenditure per primary student as % of GDP per capita (%)	World Bank http://www.worldbank.org/
GovPrimeExp	Expenditure on primary as % of government expenditure on education (%)	World Bank http://www.worldbank.org/

Table 2: Countries used in Study

Country	Year																
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Bangladesh											X	X					
Benin				X	X	X	X				X		X		X	X	X
Bhutan								X			X	X	X	X		X	X
Bolivia		X	X	X	X	X											
Burkina Fas								X	X	X			X		X	X	
Burundi			X	X	X		X	X				X	X	X	X	X	
Cabo Verde										X	X	X		X			
Cambodia	X			X	X								X				
Cameroon									X	X	X	X	X	X	X		
Central Afr												X	X				
Comoros					X						X						
Congo, Dem.													X			X	
Congo, Rep.					X								X				
El Salvador						X		X	X	X	X	X					
Eritrea				X	X		X										
Gambia, The				X	X	X	X	X	X	X	X			X	X	X	
Ghana									X	X	X	X		X	X		
Guatemala									X	X	X		X	X	X	X	
Guinea											X	X	X		X	X	
India		X	X			X								X			
Indonesia										X	X	X	X	X	X	X	
Lao PDR					X			X									
Lesotho		X	X	X				X	X		X						
Madagascar	X							X	X	X	X	X		X	X		
Malawi													X	X			X
Mali											X	X	X	X		X	X
Moldova										X	X	X					
Morocco	X	X									X	X					
Mozambique									X						X	X	
Nepal				X	X	X					X						X
Niger									X	X	X	X	X	X	X	X	X
Pakistan																X	X
Philippines				X	X	X	X	X		X	X						
Rwanda			X								X		X		X	X	
Senegal	X							X	X		X	X	X				
Swaziland		X	X	X	X	X	X	X	X				X	X			
Tanzania							X	X	X	X		X					
Togo		X	X	X		X	X	X	X		X	X	X	X	X	X	X
Uganda							X					X	X	X			
Ukraine														X	X	X	
Zambia			X				X	X									

Table 3: Descriptive Statistics

Descriptive Statistics					
Variable	**Number of Observations**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
Year	217	2007.16	4.29	1998	2014
PrimaryCompRate	217	68.11	19.62	22.56	107.03
RepeaterPercent	217	12.84	8.19	0.07645	36.22
SaniFacil	217	35.98	20.49	8.20	75.00
GDPPCAP	217	2954	2318	596.32	9674.61
UrbPop	217	35.65	14.78	8.25	69.27
MortRate5	217	87.56	37.14	17.40	183.80
TeacherRatio	217	40.72	12.02	15.75	94.61
GovEduExp	217	17.35	5.05	4.77	37.69
GDPPrimExp	217	46.81	11.51	17.11	74.36
GovPrimeExp	217	12.61	6.51	3.14	41.72

Table 4: Correlation

Pearson Correlation Coefficients, N = 217											
Prob > |r| under H0: Rho=0											
	PrimaryCompRate	SaniFacil	GDPPCAP	UrbPop	MortRate5	GovEduExp	TeacherRatio	GDPPrimeExp	RepeaterPercent		GovPrimeExp
PrimaryCompRate	1	0.47	0.624	0.602	-0.765	0.161	-0.592	-0.198	-0.516		-0.111
Primary completion rate, both sexes (%)		<.0001	<.0001	<.0001	<.0001	0.0178	<.0001	0.0033	<.0001		0.1027
SaniFacil		1	0.63	0.415	-0.539	-0.16	-0.423	-0.193	-0.297		-0.096
Improved sanitation facilities (% of population with access)			<.0001	<.0001	<.0001	0.0181	<.0001	0.0044	<.0001		0.1596
GDPPCAP			1	0.479	-0.586	0.118	-0.625	-0.279	-0.385		-0.089
GDP per capita, PPP (constant 2011 international $)				<.0001	<.0001	0.082	<.0001	<.0001	<.0001		0.1899
UrbPop				1	-0.467	0.033	-0.423	-0.127	-0.465		-0.174
Urban population (% of total)					<.0001	0.6324	<.0001	0.0611	<.0001		0.0103
MortRate5					1	-0.0467	0.51585	0.18378	0.44568		0.08919
Mortality rate, under-5 (per 1,000)						0.4942	<.0001	0.0066	<.0001		0.1906
GovEduExp						1	-0.2718	-0.1149	0.08097		0.45858
Expenditure on education as % of total government expenditure (%)							<.0001	0.0914	0.2349	<.0001	
TeacherRatio							1	0.187	0.397		-0.237
Pupil-teacher ratio in primary education (headcount basis)								0.0057	<.0001		0.0004
GDPPrimeExp								1	-0.003		0.091
									0.9638		0.1824
RepeaterPercent									1		-0.142
Percentage of repeaters in primary education, all grades, both sexes (%)											0.0369
GovPrimeExp											1
Expenditure on primary as % of government expenditure on education (%)											

Table 5: Pooled OLS Results

OLS						
Variable	**Primary Completion Rate**			**Primary Repetition Rate**		
Intercept	49.54*** (13.81)	61.05*** (14.596)	53.65*** (14.8874)	16.63* (8.84)	34.41*** (9.25)	30.83*** (9.57)
SaniFacil	-0.0473 (0.048)	-0.073 (0.046)	-0.07064 (0.046)	0.054* (0.0309)	0.019* (0.02903)	0.0219 (0.0296)
LGDPPCAP	4.73*** (1.58)	4.69*** (1.581)	5.055*** (1.61)	-1.97* (1.01)	-2.13** (1.00225)	-2.083** (1.036)
UrbPop	0.299*** (0.0602)	0.28*** (0.0612)	0.295*** (0.06057)	-0.141*** (0.0386)	-0.17*** (0.03875)	-0.141** (0.039)
MortRate5	-0.26*** (0.026)	-0.26*** (0.0612)	-0.26*** (0.02632)	0.046*** (0.0168)	0.052*** (0.017)	0.046*** (0.017)
TeacherRatio	-0.19** (0.0824)	-0.28*** (0.087)	-0.23*** (0.07972)	0.122** (0.0527)	-0.0108 (0.0554)	0.076 (0.051)
GovEduExp	0.24 (0.157)			0.32*** (0.10078)		
GDPPrimExp		-0.18 (0.1236)			-0.296*** (0.078)	
GovPrimeExp			-0.00403 (0.06629)			-0.103** (0.043)
Number of Observations	217	217	217	217	217	217
R squared	0.7082	0.7080	0.7050	0.3383	0.3514	0.3259

1. Standard errors in parenthesis.
2. Statistical significance denoted by *** at 1% level, ** at 5% level, and * at 10% level.

Table 6: Two Way Fixed Effects Results

Two-Way Fixed Effects						
Variable	**Primary Completion Rate**			**Primary Repetition Rate**		
Intercept	89.04 (56.55)	121.44** (57.96)	127.43** (60.14)	29.16 (8.84)	34.70 (26.53)	30.13 (27.63)
SaniFacil	1.31*** (0.34)	1.20*** (0.34)	1.195*** (0.34)	-0.34** (0.16)	-0.32** (0.16)	-0.32** (0.16)
LGDPPCAP	-5.63 (8.25)	-7.43 (8.37)	-7.98 (8.51)	2.56 (3.81)	1.80 (3.83)	2.196 (3.91)
UrbPop	0.004 (0.45)	-0.10 (0.45)	-0.066 (0.45)	-0.65*** (0.21)	-0.67*** (0.21)	-0.65*** (0.21)
MortRate5	-0.22*** (0.049)	-0.21*** (0.05)	-0.21*** (0.05)	0.038* (0.022)	0.0398* (0.0224)	0.037 (0.023)
TeacherRatio	0.0606 (0.1396)	-0.030 (0.137)	0.015 (0.138)	-0.188*** (0.0644)	-0.175*** (0.063)	-0.16*** (0.06)
GovEduExp	0.309** (0.1428)			-0.084 (0.066)		
GDPPrimExp		-0.238* (0.143)			-0.101 (0.065)	
GovPrimeExp			-0.107 (0.079)			-0.103 (0.043)
Number of Observations	217	217	217	217	217	217
R squared	0.9440	0.9433	0.9429	0.9340	0.9343	0.9334
F Value	11.86	11.75	11.69	25.42	25.69	24.99

1. Standard errors in parenthesis.
2. Statistical significance denoted by *** at 1% level, ** at 5% level, and * at 10% level.

SAS Code

```
Title1 'Michael Grund';

Title2 'MAG138@zips.uakron.edu';

Title3 'Access to Improved Sanitation Facilities and Primary Education: A Cross-Country
Analysis of Developing Countries';


data Grund.Data1; /* Rename variables */

set grund.data9;

Rename

Pupil_teacher_ratio_in_primary_e = TeacherRatio

Government_expenditure_per_prima= GDPPrimExp

Expenditure_on_primary_as___of_g= GovPrimeExp

Percentage_of_repeaters_in_prim= RepeaterPercent

Expenditure_on_education_as___of= GovEduExp

Mortality_rate__under_5__per_1_= MortRate5

Urban_population____of_total___S= UrbPop

GDP_per_capita__PPP__constant_2= GDPPCAP

Primary_completion_rate__both_s= PrimaryCompRate

Improved_sanitation_facilities__= SaniFacil;

run;

data Grund.Data2; /* Drop missing values in data set, Drop Countries with only single
observations for Proc Panel, Drop values outside range*/

set grund.Data1;

/* Drop missing values in data set*/

if TeacherRatio="." then delete;

if GDPPrimExp="." then delete;

if GovPrimExp="." then delete;

if RepeaterPercent="." then delete;

if GovEduExp="." then delete;
```

```
if MortRate5="." then delete;
if UrbPop="." then delete;
if GDPPCAP="." then delete;
if PrimaryCompRate="." then delete;
if SaniFacil="." then delete;
if TeacherRatio="0" then delete;
if GDPPrimExp="0" then delete;
if GovPrimExp="0" then delete;
if RepeaterPercent="0" then delete;
if GovEduExp="0" then delete;
if MortRate5="0" then delete;
if UrbPop="0" then delete;
if GDPPCAP="0" then delete;
if PrimaryCompRate="0" then delete;
if SaniFacil="0" then delete;
if SaniFacil > "80" then delete; /* Drop values outside range*/
/* Drop Countries with only single observations for Proc Panel*/
if Country_Name="Vietnam" then delete;
if Country_Name="Timor-Leste" then delete;
if Country_Name="Sri Lanka" then delete;
if Country_Name="South Sudan" then delete;
if Country_Name="Sierra Leon" then delete;
if Country_Name="Nicaragua" then delete;
if Country_Name="Kenya" then delete;
if Country_Name="Georgia" then delete;
if Country_Name="Chad" then delete;
run;
data Grund.Data3; /* Create Log of Variables and add to new data set*/
set grund.Data2;
LPrimaryCompRate=Log(PrimaryCompRate);
```

```
LSaniFacil=Log(SaniFacil);
LGDPPCAP=Log(GDPPCAP);
LUrbPop=Log(UrbPop);
LMortRate5=Log(MortRate5);
LGovEduExp=Log(GovEduExp);
LTeacherRatio=Log(TeacherRatio);
LGDPPrimExp=Log(GDPPrimExp);
LRepeaterPercent=Log(RepeaterPercent);
LGovPrimExp=Log(GovPrimExp);
run;
data Grund.Data4; /* Create Log of Variables and add to new data set*/
set Grund.Data3;
Label
PrimaryCompRate="Primary completion rate, both sexes (%)"
SaniFacil="Improved sanitation facilities (% of population with access)"
GDPPCAP="GDP per capita, PPP (constant 2011 international $)"
UrbPop="Urban population (% of total)"
MortRate5="Mortality rate, under-5 (per 1,000)"
GovEduExp="Expenditure on education as % of total government expenditure (%)"
TeacherRatio="Pupil-teacher ratio in primary education (headcount basis)"
GDPPrimExp="Government expenditure per primary student as % of GDP per capita (%)"
RepeaterPercent="Percentage of repeaters in primary education, all grades, both sexes (%)"
GovPrimeExp="Expenditure on primary as % of government expenditure on education (%)"
LPrimaryCompRate="Log of Primary completion rate, both sexes (%)"
LSaniFacil="Log of Improved sanitation facilities (% of population with access)"
LGDPPCAP="Log of GDP per capita, PPP (constant 2011 international $)"
LUrbPop="Log of Urban population (% of total)"
LMortRate5="Log of Mortality rate, under-5 (per 1,000)"
LGovEduExp="Log of Expenditure on education as % of total government expenditure (%)"
LTeacherRatio="Log of Pupil-teacher ratio in primary education (headcount basis)"
```

```
LGDPrimExp="Log of Government expenditure per primary student as % of GDP per capita (%)"

LRepeaterPercent="Log of Percentage of repeaters in primary education, all grades, both sexes (%)"

LGovPrimeExp="Log of Expenditure on primary as % of government expenditure on education (%)"

run;

Proc sort data=Grund.Data4; /*Sort for Proc Panel*/

by Country_Name;

run;

proc corr data= Grund.Data4; /*Correlation test*/

run;

proc means data= Grund.Data4; /*Means for descriptive statistics*/

run;

ods pdf file="F:\WWW\Portfolios\Fall2014\226\mag138\SeniorProject\SASDraftResults.pdf";

Title4 'GovEduExp';

proc reg data= Grund.Data4; /*Pooled OLS Regression for Primary Completion Rate*/

model PrimaryCompRate = SaniFacil LGDPPCAP UrbPop MortRate5 GovEduExp TeacherRatio;

run;

proc reg data= Grund.Data4; /*Pooled OLS Regression for Primary Grade Repeat*/

model RepeaterPercent = SaniFacil LGDPPCAP UrbPop MortRate5 GovEduExp TeacherRatio;

run;

proc panel Data= Grund.Data4; /*Two Way Fixed Effects Regression for Primary Grade Completion*/

    id  Country_Name Year;

    model PrimaryCompRate = SaniFacil LGDPPCAP UrbPop MortRate5 GovEduExp TeacherRatio/fixtwo;

run;

proc panel Data= Grund.Data4; /*Two Way Fixed Effects Regression for Primary Grade Repeat*/

    id  Country_Name Year;
```

```
    model RepeaterPercent = SaniFacil LGDPPCAP UrbPop MortRate5 GovEduExp
TeacherRatio/fixtwo;

run;

Title4 'GDPPrimExp';

proc reg data= Grund.Data4; /*Pooled OLS Regression for Primary Completion Rate*/

model PrimaryCompRate = SaniFacil LGDPPCAP UrbPop MortRate5 GDPPrimExp
TeacherRatio;

run;

proc reg data= Grund.Data4; /*Pooled OLS Regression for Primary Grade Repeat*/

model RepeaterPercent = SaniFacil LGDPPCAP UrbPop MortRate5 GDPPrimExp
TeacherRatio;

run;

proc panel Data= Grund.Data4; /*Two Way Fixed Effects Regression for Primary Grade
Completion*/

    id  Country_Name Year;

    model PrimaryCompRate = SaniFacil LGDPPCAP UrbPop MortRate5 GDPPrimExp
TeacherRatio/fixtwo;

run;

proc panel Data= Grund.Data4; /*Two Way Fixed Effects Regression for Primary Grade
Repeat*/

    id  Country_Name Year;

    model RepeaterPercent = SaniFacil LGDPPCAP UrbPop MortRate5 GDPPrimExp
TeacherRatio/fixtwo;

run;

quit;

Title4 'GovPrimeExp';

proc reg data= Grund.Data4; /*Pooled OLS Regression for Primary Completion Rate*/


model PrimaryCompRate = SaniFacil LGDPPCAP UrbPop MortRate5 GovPrimeExp
TeacherRatio;

run;

proc reg data= Grund.Data4; /*Pooled OLS Regression for Primary Grade Repeat*/
```

```
model RepeaterPercent = SaniFacil LGDPPCAP UrbPop MortRate5 GovPrimeExp
TeacherRatio;

run;

proc panel Data= Grund.Data4; /*Two Way Fixed Effects Regression for Primary Grade
Completion*/

    id  Country_Name Year;

    model PrimaryCompRate = SaniFacil LGDPPCAP UrbPop MortRate5 GovPrimeExp
TeacherRatio/fixtwo;

run;

proc panel Data= Grund.Data4; /*Two Way Fixed Effects Regression for Primary Grade
Repeat*/

    id  Country_Name Year;

    model RepeaterPercent = SaniFacil LGDPPCAP UrbPop MortRate5 GovPrimeExp
TeacherRatio/fixtwo;

run;

ods pdf close;

quit;
```