# THE RELATIONSHIP BETWEEN ACCESS TO TOILETS AND SCHOOL ENROLLMENT IN PAKISTAN

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

By

Fatima Akram Hayat, B.Sc

Washington, DC
July 22, 2017

THE RELATIONSHIP BETWEEN ACCESS TO TOILETS AND SCHOOL ENROLLMENT IN PAKISTAN

Fatima Akram Hayat, B.Sc

Thesis Advisor: Adam Thomas, Ph.D.

## ABSTRACT

The Pakistani public-school system is struggling and has fallen substantially behind other developing countries in enrollment rates. The Government of Punjab, Pakistan's largest province, initiated a program to improve school infrastructure to increase enrollment. Some studies suggest that access to usable toilets can increase school enrollment, attendance rates, and educational outcomes. Using annual census data on government schools in Punjab, this paper contributes to existing literature by examining the relationship between changes in school enrollment and changes in the number of usable toilets, controlling for a number of variables that were ignored by previous studies. I find that the availability of usable toilets is positively and significantly associated with enrollment. This relationship is stronger for schools in rural areas, for female-only schools and for secondary schools. I find no evidence of a relationship between the availability of toilets and enrollment for boys-only schools.

I would like to thank my father, Akram Naeem Bharoka, for his unfaltering love and unfailing support:

‘Even after all this time the Sun never says to the Earth “You owe me”
Look what happens with a love like that, it lights the whole sky’ – Hafiz.

I would also like to express my most sincere gratitude to my thesis advisor, Dr. Adam Thomas for his many valuable inputs, patience and knowledge. For being supportive and encouraging just when I needed it, I would like to thank Dr. Jane Holahan, Professor Jeffrey Mayer, Nirmala Fernandes and Celeste Kelly.

Many thanks,
Fatima Akram Hayat

TABLE OF CONTENTS

LIST OF FIGURES



LIST OF TABLES

# CHAPTER 1
# INTRODUCTION

According to UNESCO's 2015 Annual Report, across the globe, 120 million children are out of school. The problems with low school enrollment are gravest for developing countries largely due to corruption in the public sector, cultural attitudes and various socio-economic factors. Governments are experimenting with various demand and supply side interventions to achieve the sustainable development goal of universal school enrollment (UNESCO, 2015).

In Pakistan, however, such efforts have yet to be fully successful. Even though education is a fundamental constitutional right of every Pakistani child and is mandatory according to article 25-A of the Constitution, primary school enrollment rate in 2015 for was about 70%, while middle school enrollment rate was a mere 34% (Alif Ailaan, 2015). Enrollment rates were dramatically lower for female students - only 33% in primary school and 19% in secondary school (Alif Ailaan, 2015). In 2014, there were 25 million children not attending schools, and most of them were in Punjab (Alif Ailaan, 2015). According to the World Bank, one-third of the country's population is 18 or younger, therefore, it can be argued that increasing school enrollment is the necessary next step to improving Pakistan's political and economic prospects.[1]

Punjab, Pakistan's largest province, is home to over 100 million people and has the largest share in the national budget (Alif Ailaan, 2015). It also fares better than other provinces on most social indicators (Jamal et al., 2007). Around ten million students are enrolled in the 53,373 public schools in the 36 districts of the province (Punjab School

[1] Taken from World Bank country profile statistics (http://data.worldbank.org/country/pakistan)

Census, 2015). Under the Pakistan Muslim League Noon, many new initiatives to improve welfare have been initiated (Punjab School Census, 2015). In 2010, the Government of Punjab implemented the School Infrastructure Project, which was implemented in 2012. Through this initiative, a total of $3 million was allocated for improving school infrastructure in the public schools of Punjab.[2]

Thus far, however, there has not been any research on the new initiative's impact on the fifty-three thousand government schools in the province.[3] The present study aims to fill this gap by elucidating whether the building of toilets is associated with changes in school enrollment.

This study's conclusions will have important policy implications because the Pakistani government has a limited fiscal window for spending on education, and it is crucial for policy makers to understand what strategies are most likely to solve problems in the education sector. If building toilets is correlated with increased enrollment, the government can direct resources to provide better toilets as well as expand the existing initiative that provides them. In a country where dropout rates are extremely high, the necessary first step to improving educational outcomes is to get children to school.

[2] Unless otherwise stated, all factual assertions in this paragraph are taken from the Punjab School Census (2015).
[3] Punjab has around fifty-three thousand schools in total (primary, secondary and high schools) but my empirical analysis consists of primary and secondary schools only.

# CHAPTER 2
# BACKGROUND

Historically, government schools in Pakistan have lacked basic infrastructure and facilities such as accessible and clean toilets, covered classrooms, electricity, sewage systems and clean drinking water, and this lack of infrastructure may be responsible for high student absenteeism rates (Jasper et al., 2012) and high dropout rates (Adukia, 2014) mainly because of diseases, for low teacher participation (Chaudhury et al., 2006), and for sub-optimal learning outcomes (Adukia, 2014).

In addressing this problem, the government of Punjab has taken its inspiration from neighboring India, where the government launched a healthcare initiative called the 'Clean India Campaign' through which clean toilets are being built for every Indian school. The intervention provided government funding to schools for the construction of toilets (Adukia, 2014). In general, the Indian experience has been encouraging. For example, Adukia (2014) found that improving access to clean drinking water and building toilets was positively related to school enrollment and learning outcomes.

Pakistan's education system has both public and private schools, with private education on the rise. Increased private school enrollment may be attributable in part to studies indicating a positive correlation between private schooling and improved educational outcomes (Chaudhury et al., 2006). While some of the country's schools are mixed gender, the substantial majority is gender-segregated. Budgets allocated to the building and maintaining these schools, and even standards for teacher hiring, are gender-biased. For instance, according to Alif Ailaan's report (2015), average annual budgets for

the nation's female-only schools were only $1300 per school as compared to $2800 for male-only schools in the same *tehsil* (sub-district).

Differential spending levels are related to attendance, for example, according to a World Bank report (1996), lack of basic infrastructure dissuades households in Pakistan from sending children to schools that do not have boundary walls and usable toilets - and this reluctance is more pronounced in the case of girls, primarily due to a concern for their safety, as well as to protect them from sexual harassment and health problems (UNICEF, 2014).

These differences in boys' and girls' educational experiences increase over time. There is a sharp increase in dropout rates among female students during middle school (Alif Ailaan, 2015), as this is the time when females start menstruating. This 'menstruation hypothesis' suggests that the onset of puberty for girls has a strong correlation with school absenteeism and dropout rates if there are no separate toilets available for them to use (Herz & Sperling, 2004). Due to a lack of toilets, female students in Pakistan miss a significant amount of school days, and students are more likely to drop out if they miss classwork because it may be difficult to catch up (Burgers, 2000).

Another potentially positive impact of toilet construction is the improvement in student health, which may decrease the probability of them dropping out of school due to elongated disease. Helminthic infections such as roundworms and hookworms affect millions of children of school going age due to poor sanitation (Burgers, 2000). Open defecation can also lead to diarrhea, dehydration and stunting in children, according to UNICEF (2014). Such health conditions may cause students to drop out of school. Thus,

building toilets may increase school attendance generally and may also reduce the differences between the educational opportunities afforded to male and female children.

Another issue specific to Pakistan is that, due to the lack of basic facilities in government schools, parents seek to enroll their children in private schools, which are expensive but also tend to have better educational outcomes (Alif Ailaan, 2015). According to Alif Ailaan (2015), 36% of all students in Punjab are enrolled in private schools. However, private schooling is a luxury that most households cannot afford consistently, and when they face economic trouble, female educational expenses are often among the first budgetary cuts that they make (Alif Ailaan, 2015). Improvements in public schools' infrastructure may allow families to feel more comfortable in enrolling their children in public school rather than in private school, thereby reducing the economic burdens that they must contend with. I concentrate my analyses on public schools only, where the problems with basic infrastructure are most pronounced.

# CHAPTER 3
# LITERATURE REVIEW

Since the 1990s, researchers and public health practitioners have sought to improve access to usable toilets and sanitation for children in developing countries as a means of improving students' health, participation, and their levels of achievement (Greene et al., 2012).

Glewwe et al. (1994) found that schools in Ghana with bad infrastructure also produced poor student learning outcomes. The infrastructure differences for which the study accounted included classrooms with leaking roofs, damaged furniture, and toilets that needed repair. However, the authors found no statistically significant relationship between learning outcomes and the addition of usable toilets. Investment in school infrastructure may also be correlated with teachers' attendance. Chaudhury et al. (2005) found that an increase in the infrastructure index (which was created using information on the availability of toilets, electricity, clean floors, libraries and covered classrooms) is negatively associated with teacher absence. This is important because improved teacher attendance may be positively related to student attendance and students' learning outcomes.

While some of this research is focused on children's academic outcomes, other research has focused on the effect of improved infrastructure on student enrollment and attendance. The central hypothesis driving education and public health research has been that improvements in sanitation may lead to improved health outcomes among children, which in turn results in improved student outcomes (World Bank, 2015 and UNICEF, 2014). For countries that have a large gender gap, students coming from households with

improved sanitation facilities also had lower dropout rates (Koolwal & van de Walle, 2010).[4] Similarly, Snel at al. (2002) found that the attendance of all students, especially female students, improves considerably when schools have proper sanitation facilities.

The converse is also true. According to a World Bank report (1996), parents in Pakistan are often hesitant in sending their children, especially their daughters, to schools without separate, usable toilets and boundary walls. Although this report lacks a statistically rigorous methodology, it makes several important recommendations – the most notable for the purposes of this study being that for all girls' schools, the construction of boundary walls and suitably designed toilets should be prioritized.

According to Dostie et al. (2006), school infrastructure, along with village and household characteristics, plays an important role in schooling decisions made by Indian households, especially for female students. On the other hand, Birdthistle et al. (2011) conducted an analysis of the literature to determine whether there is any effect of making single sex toilets available on a variety of outcomes including attendance, completion of primary and secondary education and enrollment levels. Mainly due to lack of segregated data, they found no evidence to either support or contradict the hypothesis that separate toilets for girls are associated with increased enrollment (Birdthistle et al., 2011). In Burkina Faso, Kazianga et al. (2013) found that the construction of well-resourced schools with separate toilets for girls was correlated with a 20% increase in female enrollment. Separate toilets were also positively related to test scores in female students (Kazianga et al., 2013).[5]

In India, under the School Sanitation and Hygiene Education initiative, funds were

[4] The authors studied countries like Yemen, Morocco, Nepal and Pakistan

allocated to schools to improve infrastructure facilities and 220,000 new toilets were constructed around the country for various schools using these funds. In her evaluation of this initiative, Adukia (2014) found that toilet construction was significantly correlated with enrollment levels. She concluded that toilet construction in the schools in her sample was associated with an enrollment increase of about 8% in secondary schools and about 12% in primary schools. Interestingly, the association between building toilets and enrollment stays positive and statistically significant even after three years (Adukia, 2014).

There is evidence that a lack of toilets in schools has a stronger relationship with academic outcomes for girls who are old enough to menstruate than for younger girls or for boys of the same age. Some of these studies look at enrollment and absenteeism, while other studies look at test scores and academic outcomes. But there is a consistent finding among all the studies that once girls are old enough to menstruate there is a particularly strong relationship between the availability of toilets in schools and positive academic outcomes. Adukia (2014) formed separate estimates by gender and age, and found that the availability of segregated toilets is especially largely correlated with improved learning outcomes and school attendance for pre-pubescent female students (Adukia, 2014).

In contrast, some studies cast doubt on the 'Menstruation Hypothesis'. For example, Oster and Thornton (2009) conducted a randomized control trial conducted in Nepal, distributing menstrual cups to girls at random. However, the authors found that this intervention did not have a significant impact on school enrollment for treatment group members (Oster and Thornton, 2009). Mensch & Lloyd (1998) reported that in Kenya, schools' sanitation and drinking water infrastructure was not correlated with the gap in

enrollment rates between male and female students.

Also relevant is the fact that, while discussing the limitations of her results, Adukia (2014) points out that schools could have chosen to make improvements other than constructing toilets with the funds allocated to them under the School Sanitation and Hygiene Education initiative. For example, if drinking water facilities were improved in addition to constructing toilets, then the increase in enrollment could have been due to the availability of clean drinking water and not to the increase in the availability of toilets. Thus, the author's estimate of the impact of toilet availability may have been upwardly biased.

Nonetheless, Adukia's work has been cited in numerous studies arguing for improved sanitation in schools. For example, Murnane et al. (2014a) recommend building toilets to increase school enrollment, based solely on the results of Adukia (2014). Similarly, based on Adukia's (2014) work, a book published by Brookings Institution (2015) argued that the construction of toilets may be associated with low dropout rates. This growing influence of Adukia's research demonstrates the necessity of replicating and building on her work using data from other countries.

To conclude, while the literature provides conflicting results, the bulk of the limited existing research on the effect of improved sanitation and school enrollment indicates that there may be a positive relationship between the two. Most studies that find no statistically significant relationships between these variables were not carried out in South Asia and therefore had different contextual backgrounds. Many also lacked updated data on segregated toilets, or the data were from the 1990s, when school infrastructure improvement interventions had not fully taken shape.

However, recent evidence suggests that such a relationship does exist in India, and a similar relationship could also exist in neighboring Pakistan because the two countries have similar demographic and socio-economic profiles.[6] The present study tests this hypothesis. In the existing stream of literature, school infrastructure and sanitation have not been studied for South Asian countries other than India. I extend Adukia's work to Punjab, the most populous and resource rich province in Pakistan, where in the last decade, the government has invested in improvements in school infrastructure in order to produce better educational outcomes.

[6] World Bank country profile statistics reveal that Pakistan and India are in the same bracket for Gross National Income, population density, life expectancy and forecasted Gross Domestic Product growth rates. The gross primary school enrollment ratio for Pakistan in 2015 was 92.7 and for India it was 108.6.

# CHAPTER 4
# CONCEPTUAL FRAMEWORK

The existing literature suggests that, in developing countries, students with access to improved school infrastructure and sanitation may be more inclined to attend school.[7] Therefore, I hypothesize that schools with improved access to toilets will experience higher enrollment.

My empirical model includes a rich set of control variables that literature suggests are plausibly associated with both the construction of toilets and school enrollment. These controls are shown in Figure 4.1 below and discussed in the following sub-sections:

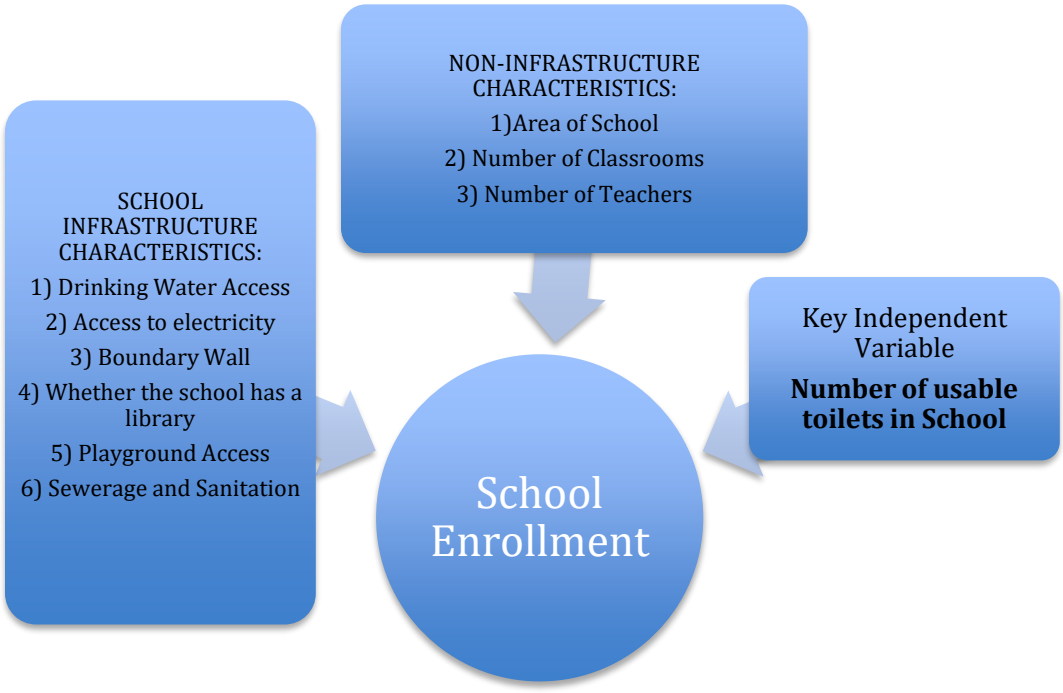


**Figure 4.1: Conceptual Framework**

[7] UNICEF (2015) defines 'improved sanitation facilities' as those that are not contaminated by fecal matter. I use the indicator for "usable toilets" in my analysis based on this definition.

My sample consists of 36,295 government schools – both gender-segregated and mixed gender - at primary and secondary levels (up to grade 8). I selected primary and secondary schools for this analysis because the steepest drop out rates are experienced during this time, with school populations shrinking by 48% (Alif Ailaan, 2015).

In addition to the presence or absence of toilets, changes in school enrollment may also be plausibly related to many other factors, and these factors must be controlled for, to allow me to isolate the relationship between the construction of toilets and school enrollment. These controls fall into two main categories: non-infrastructure characteristics and school infrastructure characteristics.

*1) NON-INFRASTRUCTURE CHARACTERISTICS:*

It is likely that school buildings that are constructed over relatively larger areas can attract more relatively students because there is more space available for the construction of additional classrooms. These schools may also have larger playgrounds or more laboratories and perhaps also more space for the construction of toilets (Adukia, 2014). Adukia controls for school area in her analysis, and I do the same. The literature suggests that two factors in particular are closely related to decreasing dropout rates: the number of teachers and the number of classrooms (Chaudhury et al., 2006). Evidence suggests that, with a reduced student-to-teacher ratio, children are less inclined to stop attending school (Burgers, 2000). Further, increases in the number of classrooms, even if they are 'open air classrooms' as in many government schools in Punjab (Alif Ailaan, 2012), can be

correlated with increased student enrollment (Murnane et al., 2014a). Adukia (2014) did not control for these two variables in her research, but I introduce them into my model because there is evidence to suggest that they might be important factors in school enrollment.

*2) SCHOOL INFRASTRUCTURE CHARACTERISTICS:*

Adukia (2014) found that schools with boundary walls had higher enrollment, as these schools are safer from terrorist activities. She also controls for whether or not the school had a boundary wall because, if the school had a boundary wall, it was more likely to have more toilets constructed. This is because, if the school already had basic facilities like a boundary wall, a library, a playground and sewerage facilities, resources could be spent on constructing more toilets. Following Adukia's model, I control for all these variables in my analysis. I also control for the availability of clean drinking water, Access to clean water is correlated with enrollment because many women in South Asian countries have the sole responsibility for fetching water from wells and other sources for their household and their schools (Murnane et al., 2014a). Thus, if water is not readily available in schools, they have the added responsibility of getting it from other sources, leaving less time for them to study. This dynamic could be related to the probability of lagging behind at school and eventually dropping out (Birdthistle et al., 2011).

In addition, I control for whether a school has electricity, because having electricity is potentially correlated with school enrollment as temperatures in the summer are extreme in Pakistan and, without running fans, it may be difficult for some students to

attend school. Moreover, a school without electricity would not have laboratories or computer labs, both of which may be negatively related to enrollment rates (Dostie et al., 2006).

# CHAPTER 5
# DATA AND METHODS

The provincial government of Punjab carries out an annual census of public schools.[8] A total of 53,372 schools were included in the government's census in 2014. The resulting census data include information on school enrollment, academic facilities like the number of classrooms and number of teachers, sports facilities, availability of basic school infrastructure, and sanitation.[9] This paper uses Punjab School Census data for 2012, 2013 and 2014.

I construct a balanced panel for my analysis in order to allow me to follow only schools that remained in existence throughout this period. Appending these data from different years to create panel data was made easier by the fact that each school is assigned a unique ID code in the school census.

My dependent variable is school enrollment. My key independent variable is the number of usable toilets in a school. This number does not include toilets that need repair, because toilets that are not in usable form are unlikely to impact enrollment.[10] As

[8] From 2012 to 2014, every public school in the province of Punjab was given an extensive questionnaire including questions about infrastructure, facilities and school enrollment. Of the 53,373 total public schools in the province, my empirical model includes only primary and secondary schools. These were 38,127 in number, but the number decreased over the years due to merging of different school levels. Upon constructing a balanced panel, I was left with 37,477 schools in my sample. I dropped 2122 observations because they had missing data on my key independent variable. Another 220 observations were dropped because they had missing data for school enrollment. Of the remaining observations, 1,204 more were dropped because they had missing values for at least one of my control variables. Thus, my final analytical sample consisted of 108,885 observations based on 36,295 schools.

[9] All the information presented in this paragraph is taken from Punjab School Census Data 2015

[10] My analysis uses UNICEF's (2015) guidelines to define the usability of toilets. Instead of looking at whether a school had toilets, I chose to look at the number of usable toilets in a school (this is my key independent variable). These were toilets not contaminated with fecal matter (UNICEF 2015), and those that had a working toilet seat and sewerage system. These data painted a clearer picture for my analysis, as damaged or unusable toilets are unlikely to contribute to better health outcomes and better learning outcomes.

previously discussed, I control for school infrastructure and non-infrastructure variables.[11]

Table 5.1 below describes each variable in my model:

**Table 5.1: Variable Definitions**

| **Variables** | **Definitions** |
|---|---|
| *Dependent Variable* | |
| **School Enrollment** | A continuous variable measuring total enrollment for a given year for a given school at the start of the academic year. |
| *Key Independent Variable* | |
| **Number of Usable Toilets in School** | A continuous variable measuring total number of usable toilets in a school in a given year. |
| *Control Variables* | |
| School Area (in Kanals) | A continuous variable measuring the total area the school for a given year. |
| Number of Classrooms in School | A continuous variable that measures the number of classrooms in a school for a given year |
| Number of Teachers in School | A continuous variable that measures the number of teachers in a school in a given year |
| Drinking Water in School | A dichotomous variable [12] that indicates whether clean drinking water was available in a school for a given year |

[11] For my control variable measuring school area, there were 151 missing values. I used interpolation to fill in those missing values that had data missing for the middle year (2013) where I could use the average of 2012 and 2014. I used the same strategy for variables that mostly remain fixed, for example – the presence of a boundary wall, playground and availability of electric power. For the missing values on my control variable defining the number of classrooms in a school, I used single imputation to redefine values. I then made an interpolated data dummy to test if imputed values looked different. A t-test, by comparing characteristics of observations for which I imputed data and characteristics for observations for which I did not impute data, showed that there was no significant difference between the two.
I carried out a sensitivity analysis using dummy variables for the controls where I used imputation to redefine values for missing observations. Running basic regressions with and without these dummies showed that there was a very slight difference between the coefficient on my key independent variable in both these regressions (2.0456 without dummies with imputed values and 2.067 with the dummies included). This is probably because the total number of missing observations in my data (after deleting the ones that had missing values for my key independent and dependent variables) were a mere 8.4% of my overall sample size.

[12] The Punjab School Census survey gave three answer choices for most of my dichotomous variables. For example, respondents would circle option 1 if there was no library in the school, option 2 if there was not one, and option 3 if the library existed but was not functional for any reason. I remodeled data for all my dichotomous variables such that options 1 and 3 would reflect the lack of a library and option 2 would reflect its presence.

**Table 5.1 (Cont.)**

| Variables | Definitions |
|---|---|
| Electricity in School | A dichotomous variable that indicates whether the school had electric power during a given year |
| Boundary Wall | A dichotomous variable that indicates whether the school had a boundary wall during a given year |
| Sewerage Access in School | A dichotomous variable that indicates whether the school had a functioning sewerage system during a given year |
| Playground | A dichotomous variable that indicates whether the school had a playground |
| Library | A dichotomous variable that indicates whether the school had a library |

My analysis uses a first differences empirical strategy to assess the relationship between toilets in schools and enrollment. This strategy automatically controls for fixed observable and unobservable differences (for variables such as cultural attitudes towards schooling, underlying demographic and labor market features and government policy features that could not be controlled for). The model is estimated as:

$$\Delta \mathbf{Y_{sdt}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1}\,(\Delta\,\mathbf{Toilets_{it}}) + \boldsymbol{\beta_2}\,(\Delta\,\mathbf{Classrooms_{it}}) + \boldsymbol{\beta_k}(\Delta\,\mathbf{School\ Infrastructure\ Characteristics})_{it} + \boldsymbol{\beta_j}(\Delta\,\mathbf{Non-Infrastructure\ Characteristics})_{it} \ldots + \boldsymbol{u_{it}}$$

I regress the outcome ΔY (change in school enrollment) in school *i*, and year *t* on a continuous variable indicating the changes in the number of usable toilets in a school for that year, and changes in a variety of control variables such as the number of classrooms in the school, the number of teachers in the school, school area, availability of drinking water, electricity, sewage, playground, library and boundary walls. By estimating a first

differences model, I am automatically controlling for all fixed characteristics of schools in my data that are fixed over time. For example, the geographic location of the schools, weather conditions, and norms related to gender roles that are all largely stable over my period of analysis.

My models also largely capture differences between schools in the cultural attitudes of the nearby communities, as variables of this sort are presumably largely unchanged in the three years of my analysis. Therefore, the estimation of a first differences model allows me to substantially reduce the extent of omitted variable bias in my estimate of the relationship between the availability of toilets and school enrollment. My coefficient of interest reflects the relationship between the change in enrollment and the change in toilets holding the other variables constant.

# CHAPTER 6
# DESCRIPTIVE STATISTICS

The table below provides descriptive statistics for my dependent variable, key independent variable and control variables. My analytical sample includes a total of 108,885 observations, which corresponds to 36,295 schools. The average enrollment level among the schools in my sample is 148 students per school. There is considerable variation in my sample; the standard deviation in enrollment is also quite large at 158 students. This could be because of the wide ranges in the relative sizes of the schools in different districts, along with other factors like the number of classrooms, and teachers, and basic school facilities provided. School area is also a determinant in this large standard deviation, because larger schools can accommodate more students.

**Table 6.1: Descriptive Statistics**

| Variable | Mean | Min | Max | SD |
|---|---|---|---|---|
| **Number of Usable Toilets** *(independent variable)* | 2.63 | 0 | 45 | 1.47 |
| **School Enrollment** *(Dependent Variable)* | 147.04 | 3.4 | 2753 | 157.8 |
| School Area (kanals) | 5.65 | 0.43 | 510 | 8.814 |
| Number of Classrooms | 3.98 | 1 | 40 | 2.57 |
| Number of Teachers | 6.43 | 1 | 50 | 1.39 |

**Table 6.1 (Cont.)**

| Variable | Mean | Min | Max | SD |
|---|---|---|---|---|
| Drinking Water | 0.28 | 0 | 1 | 0.14 |
| Electricity | 0.08 | 0 | 1 | 0.37 |
| Boundary Wall | 0.25 | 0 | 1 | 0.23 |
| Sewerage | 0.66 | 0 | 1 | 0.32 |
| Playground | 0.41 | 0 | 1 | 0.51 |
| Library | 0.64 | 0 | 1 | 0.60 |

The average number of toilets in my sample is little less than 3. However, there is a considerable range here as well. The minimum number of toilets is zero (in fact, twelve percent of my sample has no working toilets, although this is not included in the table) up to a maximum of 45. Also, although not shown in the table, I found that there was significant variation for the number of usable toilets in schools that the policy targeted to compared to those that were not part of the initiative.

There is also significant variation in my control variables. The average school in my sample is built on 5.6 *kanals* of land, with a standard deviation of 8.8 *kanals*, because government schools for the largest districts and *tehsils* are usually allotted a lot more space. There are 3.9 classrooms and about 6 teachers on average per school. Only 8% of the schools have electricity. About 28% of the schools have drinking water and 25% have a boundary wall, which is consistent with the government's initiative to provide basic facilities to schools. 66% of the schools have a boundary wall, 41% have a playground and about 64% have a library.

The average number of teachers and the number of classrooms are roughly the same, so there is about one teacher per classroom. Only a fifth have electricity and only a third have boundary walls, barely ten percent have sewerage. These schools are lacking in a lot of amenities, not just toilets. I control for these characteristics in my multiple regressions whose results are reviewed in the next section.

# CHAPTER 7
# REGRESSION RESULTS

Table 7.1 presents the results of my regression analyses. In all my regressions, the dependent variable is the change in school enrollment and the key independent variable is the change in the number of usable toilets in a school. The regression results table lists my key independent variable without any other controls first, and then introduces a set of controls divided into two categories: non-infrastructure school characteristics and school infrastructure characteristics. Robust standard errors are reported for all coefficients.

Regression (1) reports the raw correlation between the change in the number of usable toilets and the change in school enrollment without any control variables included; model (2) adds controls for changes in non-infrastructure school characteristics such as the total school area, the number of classrooms and number of teachers available; and model (3) adds to this regression a set of school infrastructure controls – such as the availability of drinking water, libraries, playgrounds, sewerage systems, boundary walls and electricity. Models (4) through (6) include interactions between the number of usable toilets and three important control variables: whether the school is located in a rural district (as opposed to an urban district), whether the school is a girls-only school (as opposed to boys-only), and whether the school is a primary school (as opposed to secondary). For models (4) through (6), F-test results are also reported at the bottom of the table assessing the joint significance of the number of usable toilets and the relevant interaction term.

Model (1), a simple OLS regression without any controls included, shows that an increase of one usable toilet is associated with a statistically significant increase of

approximately 4 students enrolling in the school. However, this coefficient is likely to be upwardly biased because of the exclusion of many variables that are plausibly correlated with both school enrollment and the number of usable toilets in a school in the same way (either positively or negatively correlated with both). For example, the number of classrooms is positively correlated with both school toilets and school enrollment.

In model (2), the inclusion of non-infrastructure controls produces a coefficient on usable toilets that is smaller in magnitude but is still statistically significant. This change is not surprising, because one would expect that the area of a school and the number of classrooms and teachers it has to be strongly predictive of its enrollment figures and perhaps also of the availability of toilets.

In model (3), school infrastructure controls are added. This model is least likely to be affected by omitted variable bias, given that all control variables are included in the regression. The results from this specification indicate a positive, statistically significant relationship between the number of usable toilets and school enrollment. The coefficient on usable toilets is almost the same in magnitude as in model (2), which is to say that my key estimate is robust to the inclusion of the controls for school infrastructure characteristics.

Some of the findings for the control variables are also worth noting. For example, the availability of drinking water and the availability of electricity and playgrounds in a school are consistently positively related with enrollment levels. Surprisingly, the availability of libraries and presence of a boundary wall are negatively related to enrollment. There is no evidence of a meaningful relationship between sewerage in a school and enrollment levels.

Adukia (2014) found that the correlation between toilets and enrollment was highest for primary school students, as children benefit most from toilet construction. The literature also elucidates that the correlation between school enrollment and access to sanitation and clean drinking water is different for male and female students. The next three models further explore these nuances in the relationship between the number of usable toilets and school enrollment by including interaction terms. I have not included these as controls in my regressions because they are fixed characteristics and therefore are automatically controlled for via my first differences specification.

Model (4) includes an interaction between the change in number of usable toilets and a dummy indicating whether a school is located in a rural district to test whether there is any difference in my relationship of interest between schools from rural and urban districts. I find that among urban schools, an increase of one usable toilet is associated with a statistically significant increase in enrollment of a little less than two students. For schools in rural districts, I find that an addition of one usable toilet is associated with about a 3.2 (1.742 + 1.438) student increase in enrollment. This relationship is also statistically significant, as reflected in the results of the F-test given at the bottom of the table. My relationship of interest is stronger for schools in rural districts, a finding that was expected, as rural districts are more likely to lack toilets both at home and in schools and that the addition of toilets in rural schools would have a stronger relationship with enrollment as compared with urban schools, where there were already more toilets both in schools and people's homes than in rural areas.

Model (5) consists of a dummy variable indicating whether the school is a primary school interacted with the change in number of usable toilets to test whether

there is any difference in my relationship of interest between primary schools and secondary schools. I find that the relationship between the change in usable toilets and enrollment for secondary schools is positive and statistically significant. An increase in one usable toilet for secondary schools is associated with an increase of roughly 3 children in enrollment. The relationship for primary schools is also positive, has a magnitude of 1.208 (3.302 – 2.094), and from the F-tests we can see that is it is statistically significant. This means that an increase of one usable toilet for a primary school is associated with an enrollment increase of a little more than one student. This finding is in line with the menstruation hypothesis, given that female students start menstruating during secondary school and are less inclined to drop out if usable toilets are available to them.

In model (6), a dummy indicating whether a school is a girls-only school is interacted with toilet construction to test whether my key relationship is different for girls' schools and boys' schools. There is no evidence of a relationship between change in usable toilets and enrollment for boys' schools. However, for girls' schools, the F-test given at the bottom of the table indicates that there is a statistically significant relationship. More specifically, a one toilet increase in girls' schools is associated with an increase of approximately four (0.323 +3.617) students. This finding corresponds directly to the menstruation hypothesis laid out in the relevant literature (Adukia, 2014), which theorizes that female students drop out of school when they start menstruating due to a lack of separate, usable toilets for them.

**Table 7.1: Regression Results**

***Dependent Variable: School Enrollment (Change in number of students enrolled per year)***

| VARIABLES | (1) Base Case | (2) Non-Infrastructure Characteristics | (3) School Infrastructure Characteristics | (4) Rural Interaction | (5) Primary School Interaction | (6) Female-only Interaction |
|---|---|---|---|---|---|---|
| Number of Usable Toilets | 4.278*** | 2.483*** | 2.486*** | 1.742*** | 3.302*** | 0.323 |
| | (0.296) | (0.219) | (0.219) | (0.316) | (0.258) | (0.298) |
| **NON-INFRASTRUCTURE CHARACTERISTICS** | | | | | | |
| School Area | | 0.230*** | 0.230*** | 0.228*** | 0.228*** | 0.221*** |
| | | (0.0663) | (0.0664) | (0.0662) | (0.0663) | (0.0650) |
| Classrooms | | 5.347*** | 5.346*** | 5.319*** | 5.290*** | 5.221*** |
| | | (0.212) | (0.212) | (0.210) | (0.211) | (0.212) |
| Teachers | | 2.879*** | 2.873*** | 2.856*** | 2.869*** | 2.841*** |
| | | (0.249) | (0.249) | (0.248) | (0.250) | (0.249) |
| **SCHOOL INFRASTRUCTURE CHARACTERISTICS** | | | | | | |
| Availability of Drinking Water | | | 0.906** | 0.914** | 0.888* | 0.871* |
| | | | (0.461) | (0.461) | (0.460) | (0.462) |
| Availability of Electricity in School | | | 0.779** | 0.764** | 0.729** | 0.742** |
| | | | (0.349) | (0.349) | (0.349) | (0.348) |

| **Table 7.1 (Cont.)** | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| VARIABLES | Base Case | Non-Infrastructure Characteristics | School Infrastructure Characteristics | Rural Interaction | Primary School Interaction | Female Interaction |
| Presence of Boundary Wall | | | -1.494*** | -1.471*** | -1.584*** | -1.191*** |
| | | | (0.435) | (0.435) | (0.436) | (0.434) |
| Availability of Sewerage | | | -0.104 | -0.116 | -0.0970 | -0.132 |
| | | | (0.252) | (0.252) | (0.252) | (0.252) |
| Availability of Playgrounds | | | 0.452** | 0.450** | 0.448** | 0.438** |
| | | | (0.191) | (0.191) | (0.191) | (0.190) |
| Availability of Libraries | | | -0.440** | -0.445** | -0.417* | -0.425** |
| | | | (0.214) | (0.214) | (0.214) | (0.214) |
| ***INTERACTIONS*** | | | | | | |
| NUMBER OF USABLE TOILETS*RURAL | | | | 1.438*** | | |
| | | | | (0.417) | | |
| NUMBER OF USABLE TOILETS*PRIMARY | | | | | -2.094*** | |
| | | | | | (0.451) | |
| NUMBER OF USABLE TOILETS*FEMALE | | | | | | 3.617*** |
| | | | | | | (0.454) |

**Table 7.1 (Cont.)**

| | | | | | | |
|---|---|---|---|---|---|---|
| Constant | 1.905*** | 1.495*** | 1.452*** | 1.441*** | 1.434*** | 1.445*** |
| | (0.115) | (0.106) | (0.109) | (0.108) | (0.108) | (0.110) |
| Observations | 85,908 | 85,908 | 85,908 | 85,908 | 85,908 | 85,908 |
| R-squared | 0.015 | 0.049 | 0.049 | 0.050 | 0.050 | 0.052 |
| F-test | | | | 83.47 | 87.25 | 65.02 |
| Number of Usable Toilets and interaction = 0 | | | | (0.00) | (0.00) | (0.00) |

Robust standard errors are reported in parentheses beneath each coefficient, and p-values are reported in parenthesis beneath each F-statistic.
*** p<0.01, ** p<0.05, * p<0.1

# CHAPTER 8
# DISCUSSION

This study explores the relationship between changes in the number of usable toilets and changes in enrollment for public schools in Punjab, Pakistan. In line with my hypothesis, I find evidence of a positive, statistically significant relationship between these two variables. Estimating a first differences model further reduces the extent of omitted variable bias in my regressions, and allows me to control for all fixed school-level characteristics. Moreover, estimates of my key relationship were robust to the inclusion of school infrastructure characteristics to the regression. This suggests that the first differencing strategy on its own substantially reduced omitted variable bias to a large extent.

For boys' schools, I find no evidence of a relationship between toilet availability and enrollment. However, for girls' schools, there is a positive, statistically significant relationship. These findings are consistent with the menstruation hypothesis because girls are inclined to stop attending school when they start menstruating if the school lacks usable, separate toilets (Herz & Sperling, 2004). I also find that the relationship between the availability of toilets and enrollment is stronger for secondary schools than for primary schools. This finding is also consistent with the menstruation hypothesis because it is usually during secondary school years that girls start menstruating. Moreover, as previously discussed, rural schools have much more limited access than urban schools to toilets so my finding that my relationship of interest is stronger for schools in rural areas than schools in urban areas is also consistent with my expectations.

Some of the findings for my control variables are also noteworthy. For example, the presence of a boundary wall and availability of libraries in schools is negatively related to enrollment. These unexpected findings suggest potentially fruitful avenues for future research.

My results differ from those of Glewwe et al. (1994), who find that there is no statistically significant relationship between school infrastructure and learning outcomes in Ghana. However, my methodology and sample more closely resemble those of Adukia (2014) than those of Glewwe et al., and my findings are very consistent with Adukia's for India. My findings are also broadly consistent with those of Kazianga et al. (2013), who find that access to sanitation is positively related to learning outcomes for schools in Burkina Faso, and with a broader literature showing that the onset of menstruation is negatively correlated with student performance (Birdthistle et al., 2011) and is positively correlated with dropout rates (Herz & Sperling, 2004) among female students.

There are also limitations associated with my analyses. It is possible that there was a random error in the measurement of my variables in the database, in which case my estimates would be biased towards zero. The basic assumption in my model was that there were no time-lag effects, and that toilets built in 2012 would lead to a change in enrollment in 2013. This is not always true, and there could be a time-lag associated with the correlation in building these toilets and changes in enrollment. Additionally, a lack of data precluded me from controlling for community level characteristics (such as economic conditions of the community, crime rates and incidence of terrorism in the community), and parent characteristics (such as parents' education levels and income).

The economic conditions of the community are likely to be related to both school enrollment and toilet construction, because it can influence the quality of schools in the area. A weak economy might affect school budgets negatively, making local governments less likely to invest in infrastructure. Additionally, parents might prefer to send their children to work instead of schooling them if the economy is doing badly. Since weak economic conditions are likely to be negatively related to both toilet construction and school enrollment (and vice versa), the omission of this variable from my regression likely results in upward bias in my estimates.

Similarly, the crime rate and incidence of terrorist activity in the area can negatively affect both the school infrastructure and the willingness of parents to send their children to school. Communities prone to terrorism and crime have fewer toilets built because construction workers may not want to go there to work and similarly, lower enrollment rates because it may be unsafe to attend school. This upwardly biases my estimates.

In the same way, parents with higher levels of income and education may be more invested in their children's education and this may be related to higher enrollment rates for these children. If parents have a high income, they do not need to have their children work and can afford to send them to school. To the extent that education and income are correlated with community characteristics (which are correlated with toilet construction), then education and income will also be positively correlated with toilet availability. So, the omission of these characteristics also asserts an upward bias on my estimates. In sum, the discussion above suggests that my estimates of the relationship between toilet availability and enrollment are likely to be upwardly biased.

Future studies might attempt to address some of these limitations by controlling for household incomes, crime rates, and other related factors. Another avenue for future research would be studies conducted on a different scale. My study focuses only on Punjab, which is the largest province of Pakistan but contains only a subset of the entire Pakistani population. A nationwide study would be a welcome addition to the literature, as would other studies in neighboring South Asian countries such as Sri Lanka, Nepal and Bangladesh where the population's socio-economic characteristics are very similar.[13]

From a programmatic perspective, my results provide suggestive evidence that, to the extent that Pakistani provincial governments are interested in raising enrollment rates among students in general and among girls, in particular, improvements in school infrastructure – and especially in the availability of toilets – are a potentially promising policy option.

[13] According to World Bank country profile statistics, South Asian countries like Pakistan, Sri Lanka, Nepal and Bangladesh are similar along a variety of dimensions including age profiles, macroeconomic characteristics like predicted GDP growth rates and gross national income and even educational attainment. South Asia as a region has had substantial increase in enrollment rates at the secondary level in the last decade according to UNICEF (2015), but overall enrollment numbers are still low for the region.

## WORKS CITED


Adukia, A. (2014). Sanitation and Education. Cambridge, MA: Harvard Graduate School of Education.

Alif Ailaan. (2015). 25 Million Broken Promises: The Crisis of Pakistan's Out-of- School Children. Islamabad: Alif Ailaan

Birdthistle, I., Dickson, K., Freeman, M., & Javidi, L. (2011). What Impact Does the Provision of Separate Toilets for Girls at Schools Have on Their Primary and Secondary School Enrollment, Attendance and Completion?: A Systematic Review of the Evidence. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Burgers, L. (2000). Background and rationale for School Sanitation and Hygiene Education. New York, UNICEF.

Chaudhury, Nazmul, et al (2006). "Missing in Action: Teacher and Health Worker Absence in Developing Countries." The Journal of Economic Perspectives 20.1 (2006): 91-116.

Dostie, B., & Jayaraman, R. (2006). Determinants of School Enrollment in Indian Villages. Economic Development and Cultural Change, 54(2), 405-421.

Glewwe, P., & Jacoby, H. (1994). Student Achievement and Schooling Choice in Low Income Countries: Evidence from Ghana. Journal of Human Resources, 843-864.

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

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

Herz, B. K., & Sperling, G. B. (2004). What Works in Girls' Education: Evidence and Policies from the Developing World. Council on Foreign Relations.

Jamal, H., & Khan, A. J. (2007). Trends in Regional Human Development Indices. Social Policy and Development Centre.

Jasper, Christian, Thanh-Tam Le, and Jamie Bartram (2012). "Water and Sanitation in Schools: A Systematic Review of The Health and Educational Outcomes." International Journal of Environmental Research and Public Health 9.8: 2772-2787.

Kazianga, H., Levy, D., Linden, L. L., & Sloan, M. (2013). The Effects of "Girl-friendly Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso. American Economic Journal: Applied Economics, 5(3), 41-62.

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

Mensch, B. S., & Lloyd, C. B. (1998, June). Gender Differences in the Schooling Experiences of Adolescents in Low-Income Countries: The Case of Kenya. Studies in Family Planning, 29(2), pp. 166-183.

Murnane, Richard J., and Alejandro J. Ganimian (2014a). Improving educational outcomes in developing countries: Lessons from rigorous evaluations. No. w20284. National Bureau of Economic Research, 2014.

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

Snel, Mariëlle, Sumita Ganguly, and Kathleen Shordt (2002). School Sanitation and Hygiene Education - India. Delft: Unicef India and IRC International Water and Sanitation Centre, 2002. Print

Sperling, Gene B., and Rebecca Winthrop. What Works in Girls' Education: Evidence for the World's Best Investment. Brookings Institution Press, 2015.

UNESCO (2015). A Growing Number of Children and Adolescents Are out of School as Aid Fails to Meet the Mark. <http://www.uis.unesco.org/Education/Documents/fs-31-out-of-school-children-en.pdf>.

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

UNICEF (2014) Lack of Toilets Dangerous for Everyone, UNICEF Says. N.p., 19 Nov. 2014. Web. <http://www.unicef.org/media/media_77952.html>.

World Bank (1996) Improving Basic Education in Pakistan,. Report 14960-PAK. Washington DC

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

World Bank (2017). Country Profiles. Retrieved from: http://www.worldbank.org/en/country