Are private schools delivering better education? An empirical analysis of the differences in academic achievement between children in private and public schools in Pakistan.

Henna Akram

620122

MSc Development Economics

Supervised by Professor Massoud Karshenas

Word count: 7,942 words

This dissertation is submitted in partial fulfillment of the requirements for the degree of MSc Development Economics of the School of Oriental and African Studies (University of London).

Submission date: 14/09/16

Student Declaration:

I have read and understood regulations 17.9 of the Regulations for students of the School of Oriental and African Studies concerning plagiarism. I undertake that all material presented for examination is my own work and has not been written for me, in whole or in part by any other person. I undertake that any quotation or paraphrase from the published or unpublished work of another person has been duly acknowledged in the work which I now present for examination. I understand that I must have prior approval before incorporating any coursework I have completed for any previous qualification either at SOAS or elsewhere, and that any such previous coursework is duly acknowledged in the dissertation now submitted. I give permission for a copy of my dissertation to be held for reference, at the School's discretion.

Student Name: Henna Akram
Student ID Number: 620122
Date: 14/09/16

Contents

Abstract:

The private sector has played an increasing role in providing education in Pakistan. This study uses econometric analysis to estimate the differences in academic achievement between children attending private and public schools to ascertain whether private schools are providing better education. It uses a large scale primary data set produced by Annual Status of Education Report (ASER) Pakistan for 2015 and following the literature, uses test scores to estimate the effect of private schools whilst controlling for individual characteristics. It overcomes the selection bias by including parental education variables to proxy for ability to control for more able children going to private schools. This study identifies and addresses the shortcomings of existing empirical literature by looking at school quality and conducting analysis of the private school effect at the regional level. It finds that by allowing the model to change flexibly across regions, the private school effect varies significantly, as does school quality, and this dilutes the private school effect found for the whole country. It concludes that national level analysis is not representative of all regions in Pakistan and instead future research should focus on school quality at the regional level.

Chapter 1: Introduction

Pakistan's population in 2014 was 183.57 million which represents an increase of over 300% from 44.91 million in 1960 (Country Economy , 2016). With a rapidly growing population, the need to have an education system able to cater for all is an important task. The strains of meeting the educational requirements have been highlighted much in the literature (see Mohiuddin, 2007; Mahmood, 1999) and even back in 1972 when the population was 65 million, it was noted that there was not enough schools to cater for all children in the country (The Family Planning Association of Pakistan , 1973).

The private sector has played an increasing role in providing education in Pakistan (Mahmood, 1999). Private education institutions are "those controlled and managed by non-governmental organizations, such as a religious body, trade union or business enterprise" (Lynd, 2007, p. 20) and public education institutions are defined as "institutions controlled and managed by a public education authority or a government agency" (ibid.). In 2006, 30% of primary school students were schooled in private institutions (UNESCO Institute for Statistics, 2005, quoted in Lynd, 2007, p.23). This is significantly larger when compared with counties Pakistan shares borders with where 17% of primary school children in India, in 2003, and 5% of primary school children in Iran, in 2005, were in private schools.

Thus, given the growing importance of the private sector in providing education, it is important to assess whether these private schools are offering better education than their public sector counterparts. This study carries out econometric analysis to

estimate the differences in academic achievement between children attending private and public schools in Pakistan, to ascertain whether private schools are providing better education. It uses a large scale primary data set produced by Annual Status of Education Report (ASER) Pakistan for 2015 and following the literature, uses test scores to estimate the effect of private schools. In doing so, this study acknowledges methodological issues such as biases in estimation due to unobservable variables that could affect the performance of children in private schools, for example a child's innate ability. It approaches these issues by controlling for individual characteristics and adding parental characteristics, where the latter are used as proxies for a child's ability.

Chapter two of this paper begins by outlining the education system in Pakistan. It then looks at the theoretical issues and the two common methodological approaches used to assess the effect of schools, either on academic achievements of students or on the wages earned by the students in the labour market. It then critically appraises the empirical literature specifically for Pakistan, as this is the approach this study takes, identifying shortcomings in the existing literature. In doing so, it finds that the Pakistani literature either do not adequately account for, or fail to factor in, school quality in the analysis.

Chapter three presents the data and methodology this paper uses to investigate whether private schools provide better education over public schools. Following the literature on Pakistan, test scores are used as the dependent variable and institution type as the key independent variable of interest, along with a number of control variables. Different model specifications are outlined including the way in which this

paper attempts to address the defects found in the existing empirical literature by looking at school quality. Limitations of the study are noted before Chapter four presents the findings from econometric analysis of the effect of attending a private school on test scores and analysis of school quality. The results from each model are critically reviewed and discussed, and finally, this paper ends with concluding remarks in Chapter five.

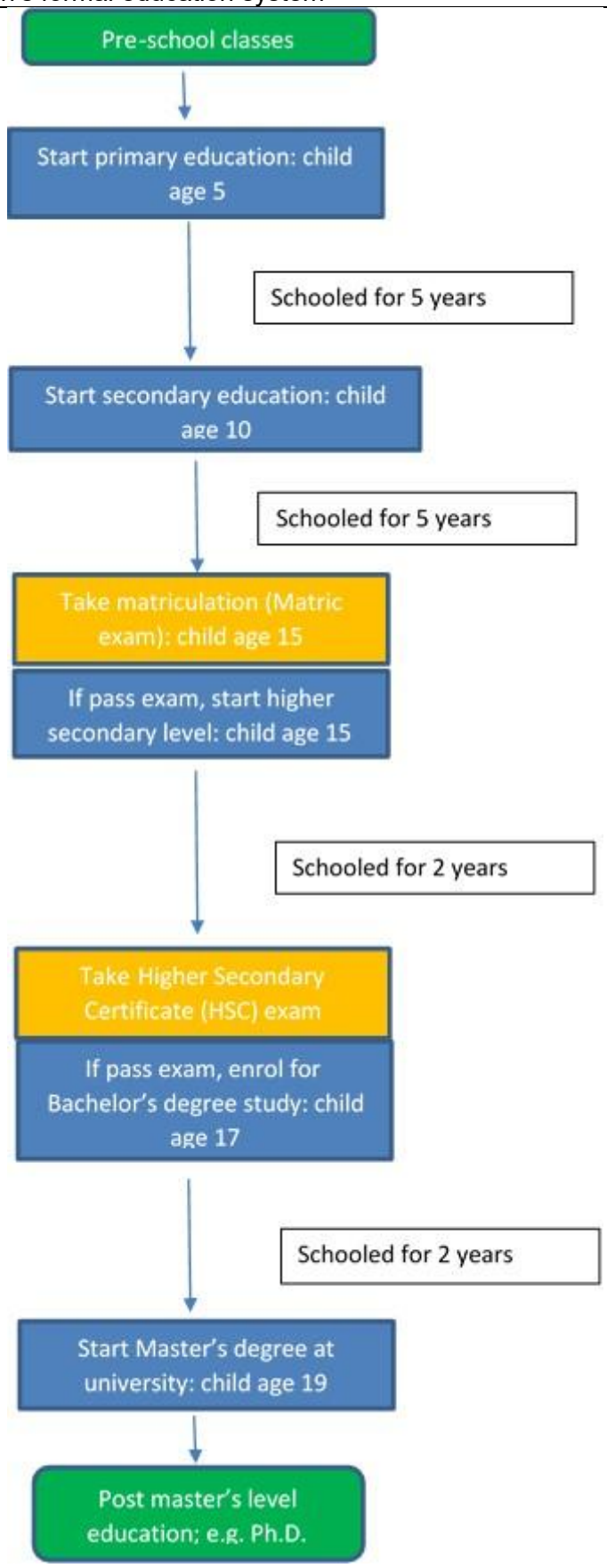
Chapter 2: Background

This chapter begins by outlining the education system in Pakistan. It then looks at the theoretical issues and the two common methodological approaches used to assess the effect of schools which are either on academic achievements of students or on the wages earned by the students in the labour market. Lastly, it critically appraises the empirical literature specifically for Pakistan as this is the approach this study takes, identifying shortcomings in the existing literature.

2.1 The Pakistani school system

Figure 2.1 represents Pakistan's formal education system diagrammatically. As an education system largely based on that of the UK due to colonial history, it begins with primary education and goes through to master's level degrees and beyond. Children enter primary school at age 5 where they are schooled for 5 years before entering secondary school, (sometimes split into middle secondary schools), where they are schooled for a further 5 years. At the end of this, the student, now aged 15, takes a matriculation exam (UK equivalent to GCSEs). After passing this, they can then enrol to study for the Higher Secondary Certificate (UK equivalent to A-Levels) which takes 2 years, before going onto complete a bachelor's degree. Unlike other countries, for example in the UK and the US, universities in Pakistan only offer education at the master's level and above (Mohiuddin, 2007, p. 267).

Figure 2.1: Pakistan's formal education system



Notes: In addition to formal examinations, annual examinations take place with the

purpose of assessing whether a child progresses to the next class or has to repeat the year.
Sources: Adapted from (Mohiuddin, 2007), (Mahmood, 1999)

2.2 Theory and methodology

Selection bias

One of the key concerns when estimating the differences between public and private schools is what is called the selection bias (Angrist & Pischke, 2009). It is plausible that innately intelligent students attend private schools, perhaps because parents want to get the best return on investment by sending their most able child to a private school. This can mean that the variable measuring the effect of attending a private school would be over estimating this effect, if it does not control for this bias, as any estimates would capture a child's innate ability plus any effects of schooling.

In an ideal world, to assess how effective different types of schools are, one would conduct a randomised control trial, whereby participants would be randomly assigned to attend a private or a public school to ensure that any differences in performance found can solely be attributed to the quality of education provided at that institution, rather than due to personal characteristics of the students in each school (Angrist & Pischke, 2009, p. 15). In principal, this would seem possible to carry out and one experiment, in the state of Tennessee in the US, known as the Student Teacher Achievement Ratio (STAR) experiment, allowed for the selection bias, when estimating the effect of class size on student performance, as students and teachers were randomly assigned to classes of different sizes (see Krueger, 1999) for an econometrics analysis of this data.

However, in practice, it is rare to find such studies. This may be because it could be difficult to persuade parents to allow their children to take part in such an experiment, due to the nature of education as a long term investment, and the need for such a study to last at least one year to be able to compare results.

Other ways of controlling for this selection bias comes from the inclusion of control variables and is the approach generally taken in the literature. These are useful for two reasons: first, if the control variables (such as age, gender, parental background) correlate with the variable of interest for example student/teacher ratios or institution type within a regression equation, their inclusion is likely to reduce the selectivity bias and enable better estimates of the coefficient of interest (Angrist & Pischke, 2009). Secondly, if these variables are quite similar for the variable of interest (i.e. parental education levels or income levels are the same for children in public and in private schools), their inclusion instead reduces the variance of the residuals in econometric estimation which in turn allows better estimation of individual coefficients (ibid).

Methodological approaches

There are generally two approaches taken to test the effect of schools whilst controlling for factors such as age, gender, and family background variables.
The first is to look at labour market outcomes of students who completed their studies at different schools. The effect of school quality (measured by student-teacher ratios, for example) in public and private schools would be measured

through differentials in the value placed on the skills and knowledge of the individual given by the labour market. For example, Card & Krueger (1992) conducted an empirical exercise to assess whether school quality matters in public schools in the United States where they used earnings to measure the effect of school quality. School quality was measured by pupil/teacher ratio, average term length, and relative teacher pay and they found that men who went to states with higher school quality, gained more from an additional year of schooling in terms of earning. Similarly, Barro & Lee (1996) use pupil/teacher ratio, spending per pupil and teacher salary as measures of school quality.

The second way to assess the quality of education provided is to look at the academic outcomes of students. This involves measuring differentials in test scores obtained by students in public and private schools. The Coleman report (1966) did this for different groups of children in schools in the US to assess the effect of school quality on their performance on test scores (Coleman, et al., 1966). See also Ehrenberg & Brewer (1995) where an econometric model was used to look at teacher characteristics (as a measure of quality) and their influence on student test scores, and Heyneman & Loxley (1983) for a cross country study looking at effects of school quality on test scores for 29 countries.

2.3 Empirical studies

This section considers literature, mostly on Pakistan, on the impact of private schools on outcomes but some studies on India have also been included to make up

for lacking research in this area and can be considered relevant due to India's geographical proximity to Pakistan.

There seems to be a common methodology used in studies that have estimated the private school effect in India and Pakistan. This tends to be to look at the impact of attending a private school on children's test scores, whereby the independent variable of interest takes the form of a dummy variable, where a value of 1 denotes that the school is a private school and 0 if it is a government school, see Desai, et al. (2008) for India and Muralidharan & Kremer (2006) specifically for rural areas in India, and Javaid, et al. (2012) for Pakistan. Javaid, et al. (2012) used Annual Status of Education Report (ASER) data for Pakistan for the years 2010 and 2011 with a total 104,328 observations to estimate the differential in the learning outcome of children in public and private schools. The results show that children attending private schools tend to do better in tests than their public school counterpart after controlling for individual and household factors, and when using regional and village fixed effects.

These studies are based on household surveys that collect information on child and household characteristics, and separately on school characteristics of public and private institutions. As part of the survey, a test is administered by the surveyor to each child within a household and this forms the data for test scores, which enables specific child and household characteristics to be controlled for within the analysis. However, what it does not allow to be controlled for is school quality. Although previous research conducted on South Asia provide information on how large the "private school effect" is (i.e. what the gains in terms of test scores are from

attending a private school after certain characteristics are controlled for), it does not provide an assessment of how quality differs across private and public schools, and if or how this can help provide explanations for the difference in test scores/academic achievement of children attending such schools.

Where work does include school quality, it has either been on a specific area of Pakistan or been a small scale study, which would not be representative of the whole country. For example, Alderman, et al. (1996) look at school quality measured by student teacher ratio and instructional expenditures, and find that it has mixed effects on academic outcomes. However, this study only looks at one district of Punjab, Lahore, and focuses on low income neighbourhoods, as the main aim of the study was to identify the factors that affect the decision of poor households as to which school to send their children to. Khan & Kiefer (2007) did a similar study for rural areas of Pakistan and surveyed schools rather than households, but the sample was biased as it was comprised of 43 Non-Governmental Organisations (NGO) schools and then 43 of the nearest, private and public schools to these each. They find that students in NGO schools perform the best followed by private and then public school students, but the authors emphasised caution in interpreting these findings as they acknowledged the shortcomings of the biased sample. Differently, Aslam (2007) estimates equations separately for public and private schools however, this study was restricted as it looked only at Grade 8 students in Lahore, a district in Punjab and did not explicitly account for quality differences across schools.

As outlined in the methodological approaches section above, literature in the US consider school quality and its impact on earnings or test scores, however, this is

defective in the literature on Pakistan because they do not allow for school quality and where the few studies that do, they focus on a single district or use an unrepresentative sample. This study intends to carry out econometric analysis to estimate the differences in test scores between children attending private and public schools to ascertain whether private schools are providing better education. It follows the work of Javiad et al (2012), by reproducing some of their results with the latest available data (2015) but goes further to consider the role of school quality. The data and methodology used to carry out this analysis are discussed in the next chapter.

Chapter 3: Data and methodology

This chapter presents the data and methodology this study uses to investigate whether private schools provide better education over public schools (as measured by student test scores), and whether the quality of education provided in private schools is different to those in public schools. The chapter begins by describing the primary dataset used and the variables available within this. It then moves onto outlining the regression models that will be used to analyse this issue and ends by noting limitations of the study.

3.1 Data

The analysis in this study is based on the latest available cross sectional data from the Annual Status of Education Report (ASER) Pakistan for 2015. ASER is a household survey covering the characteristics of children aged from 3 to 16, and schools, with a focus on all rural and most urban districts of Pakistan (ASER Pakistan, 2014). It is the largest available data set collecting individual micro data and provides reliable information on children and school characteristics with good coverage of all areas in Pakistan. 600 households are surveyed in each district with a focus on capturing information in rural areas (ibid.).

Pakistan has four provinces Balochistan, Khyber-Pakhtunkhwa, Punjab and Sindh. It also has several federally administered territories (Islamabad Capital Territory and Federally Administrated Tribal Areas (FATA)) and autonomous territories (Azad Kashmir and Gilgit Baltistan) (Pakistan Student Association, n.d.).

Figure 3.1: Map of Pakistan

Source: (Pakistan Student Association, n.d.)

As the survey tests all children in a household, regardless of whether they are in school or not, this study only looks at children who are currently enrolled in a private or public school to accurately determine the effect of type of school on test scores. Furthermore, observations were only included for children who were aged from 5 to 16, as this is considered the school going age, and where there was no information available for a child's reading, English or math score these observations were dropped. After these considerations, the sample size equated to 169,786 household observations. Geographically, the data have been broken down into 13 provinces/administrative areas including an account for urban/rural areas where available.

Information is available both for public and private schools across Pakistan as at least one school from each village/block is included in the sample. There are a total of 4,302 public and 1,819 private schools in the sample. There are more public schools as the survey focuses on collecting data from rural areas where as previous literature suggests these areas are better covered by public schools, see for example, Mohiuddin (2007); Mahmood (1999, p. 19). Information available includes the number of teachers employed, total child enrolments for each school and school infrastructure characteristics.

3.2 Variables used in econometrics analysis

Dependent and independent variable

The dependent variable in this study is test scores as the aim is to measure whether private school children do better than public school children in terms of their academic achievement. Table 3.1 provides information on the type of test and grading level used to assess children's ability in reading (native language and English) and maths.

Table 3.1: Description of tests administered to children in the survey

Level	Reading (Urdu, Pashto and Sindhi)	Reading English	Maths
1	Recognises alphabets	Recognises capital letters	Recognises numbers 1-9
2	Recognises words	Recognises lower case alphabets	Recognises numbers 10-99
3	Reads sentence/paragraph	Reads simple words	Can do subtraction
4	Reads story	Reads sentence	Can do division

Source: (ASER Pakistan, 2014).

Following Javaid, et al. (2012) a z-score has been created for each child, for each of the three tests. This involves the subtraction of the mean from the sample and division by the standard deviation, for each test and allows for easy comparisons to be made. An average across the three tests was calculated to produce a single standardised score for each child. As the scores have been standardised it allows the following interpretation to be made. A child with a z-score of 0 implies that their score is exactly the same as the average of the sample whereas a score of 1 would mean that it is one standard deviation above the average and the opposite for a score of -1.

The independent variable of interest is the type of school a child attends. As this study seeks to estimate the difference in test scores of children in private and public schools, a private dummy variable is included, which takes the value of 1 if the child goes to a private school and 0 otherwise, following the literature. The value of the coefficient on this variable will estimate the effect of attending a private school on test scores.

Control variables

The literature review provided an indication of the types of factors one must control for when assessing the impact of type of school on test scores. The following variables, as in Javaid, et al. (2012), will be included with the aim of stripping out any effects of these confounding variables to enable better estimation of the independent variable of interest:

Child level:

1. Female (dummy variable =1 if female)
2. Age of child
3. Tuition (dummy variable =1 if child receives tuition)

Household level:

4. Number of siblings the student has
5. Father's age
6. Father's age squared
7. Father attended school (dummy variable =1 if father went to school)
8. Mother's age
9. Mother's age squared
10. Mother attended school (dummy variable =1 if mother went to school)
11. Wealth index (a weighted average of the following dummy variables (pucca* + house owned + electricity available + TV available + mobile available + smartphone available)

*pucca is a house made using concrete materials such as stones, concrete, burnt bricks and timber.

The expected biases of these variables and rationale for inclusion are discussed below when outlining the methodology.

3.3 Methodology

This study uses the Ordinary Least Squares (OLS) method to estimate the private school effect due to its simplicity and common use in the literature. Model

specifications are summarised mathematically in Table 3.2 with explanatory notes and rationale for each presented below this.

Table 3.2: Model specification	
Model no.	Specification
1	$Test\ score_i = \alpha_0 + \alpha_1 Private_i + \alpha_2 Female_i + \beta_1 Age_i + \varepsilon_i$ Where Private, Female=1 or 0 and ε_i is the idiosyncratic error term.
2	$Test\ score_i = \alpha_0 + \alpha_1 Private_i + \alpha_2 Female_i + \beta_1 Age_i + \beta_3 Father's\ age_i + \beta_4 Father's\ age^2{}_i + \alpha_4 Father\ school_i + \beta_5 Mother's\ age_i + \beta_6 Mother's\ age^2{}_i + \alpha_5 Mother\ school_i + \varepsilon_i$ Where Private, Female, father school, mother school, =1 or 0 and ε_i is the idiosyncratic error term.
3	$Test\ score_i = \alpha_0 + \alpha_1 Private_i + \alpha_2 Female_i + \beta_1 Age_i + \alpha_3 Tuition_i + \beta_2 Siblings_i + \beta_3 Father's\ age_i + \beta_4 Father's\ age^2{}_i + \alpha_4 Father\ school_i + \beta_5 Mother's\ age_i + \beta_6 Mother's\ age^2{}_i + \alpha_5 Mother\ school_i + \beta_7 wealth_i + \varepsilon_i$ Where Private, Female, tuition, father school, mother school, =1 or 0 and ε_i is the idiosyncratic error term.
4	$Test\ score_i = \alpha_0 + \alpha_1 Private_i + \alpha_2 Female_i + \beta_1 Age_i + \alpha_3 Tuition_i + \beta_2 Siblings_i + \beta_3 Father's\ age_i + \beta_4 Father's\ age^2{}_i + \alpha_4 Father\ school_i + \beta_5 Mother's\ age_i + \beta_6 Mother's\ age^2{}_i + \alpha_5 Mother\ school_i + \beta_7 wealth_i + \varepsilon_i$ Where age= children aged from 5 to 10, Private, Female, tuition, father school, mother school, =1 or 0 and ε_i is the idiosyncratic error term.
5	For each region: $Test\ score_i = \alpha_0 + \alpha_1 Private_i + \alpha_2 Female_i + \beta_1 Age_i + \alpha_3 Tuition_i + \beta_2 Siblings_i + \beta_3 Father's\ age_i + \beta_4 Father's\ age^2{}_i + \alpha_4 Father\ school_i + \beta_5 Mother's\ age_i + \beta_6 Mother's\ age^2{}_i + \alpha_5 Mother\ school_i + \beta_7 wealth_i + \varepsilon_i$ Where Private, Female, tuition, father school, mother school, =1 or 0 and ε_i is the idiosyncratic error term.

This study uses a step by step approach to take into account any biases when measuring the effect of attending a private school on test scores. The first model looks at this effect whilst controlling for gender and age. At the child level, the dummy for female is expected to be negative as females tend to perform less well in developing countries compared with males (see Dickerson, et al. (2013); Buchmann, et al. (2008). A positive effect of age is expected as the older a child is, the better their ability to answer questions on the test.

However, after controlling for gender and age, it could be that a more able child is sent to a private school as discussed in Chapter 2, this is known as the selection bias. Thus, Model 1 is said to suffer from endogeneity bias, more specifically omitted variable bias, whereby the ability of a child is uncontrolled for in the model, causing the error term to be correlated with the private school variable. As the ability of a child is unobservable, Model 2 adds proxies for a child's ability by including information about parent's education. It is expected that parental education has a positive impact on whether a child goes to school and their test scores (see Dickson, et al. (2013), Ermisch & Pronzato, (2010)). Parental age and age squared are included to capture any positive effects of a parent's age on a child's education and then any subsequent impacts as age increases.

Model 3 includes variables to account for other aspects of heterogeneity that may affect the allocation of students to private and public schools and thus their performance in tests. Where children receive extra help with their studies, for example in the form of paid tuition outside of school, it is expected that this would positively impact their performance on the tests, thus it is expected that the sign on tuition variable will be positive. The impact of number of siblings could be positive or negative. For example, Roscigno & Ainsworth-Darnell (1999, p. 166) found that children with more siblings tend to have more resources. This could lead to higher educational outcomes as they have more resources available. However, others such as Downey (2001) found that having lots of siblings also limits the child's individual access to such resources and thus the returns to their education. Finally, variables have been included to proxy for the wealth of a family as wealthier families are better

placed to afford private schools. It is expected that the more assets a family has, the more likely that the child performs better at school (see for example (Rumberger, 1983).

On closer inspection of the data, it was identified that the tests that were administered to all children were derived from the syllabus of children in Grades 2 and 3. In order to account for this, Model 4 restricts Model 3 to cover only the children for whom the test is most applicable to, children aged 5-10. It is anticipated that this should improve the estimation of the private school effect as this regression excludes children over the age of 10 whom may more easily be able to complete the test, thus biasing the results of private schools downwards.

Finally, the main contribution of this research, and that which is lacking in previous work, including that of Javaid et al. (2012), is to look at whether the quality of education provided in private schools is different to those in public schools in order to understand how quality of school affects test scores. Ideally, this would be done by including school quality variables such as student teacher ratios into the econometric model to ascertain whether it has a significant effect. However, although this was the intention of this work, it was not possible to match the school characteristics dataset to the household data because there was no unique identifier in each dataset. Attempts were made to match these data using other indicators such as village ID and school type as this information was available for both schools and households, however, due to a discrepancy between how villages were coded for public schools and households, it meant that only a small proportion of the data could be matched.

This resulted in a merged sample which predominantly consisted of only private schools and as this was a biased sample, these data were not used.

Instead, another approach will be taken to analyse how important quality of school is in explaining test scores. School infrastructure, student teacher ratios and teacher absentee rates by institution will be analysed at the regional level to identify any variations in school quality. Model 5, which is equivalent to Model 3, runs the regression equation separately for each region to investigate whether there are any differences between the private school effect on test scores across regions. This is another way to identify any relationship between test scores and quality of schools in each region and it is acknowledged that although these results may not be as robust as those in an econometric model, it is still a useful exercise.

3.4 Limitations of this research

Firstly, although this study attempts to control for omitted variable bias as a child's innate ability is unobservable, it could be that parental education may not capture this effectively and thus could result in estimated coefficients being biased. Furthermore, due to the subjectivity of the creation of the wealth variable and some of the regional school quality indicators, largely based on the available data, it is important to note that this may not fully represent the factor accurately but serve only as a proxy and it is acknowledged that other indicators may estimate this better.

Also, following the literature, standardised test scores have been used as a measure of student outcomes for a child's performance over three tests. Although it may be

expected that a child would on average perform similarly across different tests, it could be that there are variations between the subjects. Thus a combined score would not pick up on these nuances, should they exist. Moreover, the test score data are from tests designed using the curriculum for Grade 2 and 3 children and have been administered to all children aged 5 to 16. To overcome this, age is controlled for in the Models, as well as a separate model created to analyse the extent of this issue.

Although the data used covers the whole of Pakistan, it's coverage of regions vary particularly as the survey seeks to cover rural areas more thoroughly. However, there still remains a good number of observations for each area to draw conclusions from.

As with all surveys, they are subject to some shortcomings. For example, there could be issues with the respondents willingness to provide information, especially personal, sensitive information as is the case with this survey in terms of household characteristics (mothers and fathers education level), which could lead to misreporting, and furthermore there could be data entry problems, especially with large datasets. The next chapter presents the results from the empirical analysis.

Chapter 4: Empirical analysis

This chapter presents the findings from preliminary analysis and output from econometric analysis on the effect of attending a private school on test scores following the methodology as discussed in the previous chapter. The results from each model will be critically reviewed in terms of whether variables are statistically significant, particularly focussing on the private school effect. This chapter also presents analysis of school quality variables at the regional level and compares these with results of regional regressions of test scores on individual and household characteristics. Finally, this chapter ends with a discussion of the results.

4.1 Preliminary analysis

Expected sign of coefficients

Following the discussion in the previous chapter, a summary of the expected sign of coefficients for each variable are presented in Table 4.1. A positive sign on a variable indicates that as it increases, on average, it results in an increase in test scores and vice versa for negative coefficients.

Table 4.1: Expected sign of coefficients for each variable	
Variable	Expected sign of coefficient
Private	+
Child level	
Female (dummy variable =1 if female)	-
Age of child	+
Tuition (dummy variable =1 if child receives tuition)	+
Household level	
Number of siblings the student has	+/-
Father's age	+
Father's age squared	-
Father attended school (dummy variable =1 if father went to school)	+
Mother's age	+
Mother's age squared	-
Mother attended school (dummy variable =1 if mother went to school)	+
Wealth	+

Descriptive statistics

Descriptive statistics of the sample are shown in Table 4.2 where the mean, standard deviation, minimum and maximum value for each variable are given. This helps to look at the main features of the data set before conducting econometric analysis. The average age of children in the sample is 9.75 and females account for 35% of the sample. Mothers, on average, were five years younger than fathers and were less likely to have gone to school.

Table 4.2: Descriptive statistics					
VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Child age	169,786	9.750	3.185	5	16
Female	169,786	0.352	0.478	0	1
Tuition	169,786	0.134	0.340	0	1
Z-score	169,786	0.152	0.903	-1.50	1.20
Mother age	168,275	36.08	7.156	19	79
Mother age square	169,786	1,341	572.6	0	6,241
Sibling	169,786	2.635	1.666	-1	9
Mother gone school	169,786	0.334	0.472	0	1
Father age	168,493	41.09	8.005	20	85
Father age square	169,786	1,739	721.5	0	7,225
Father gone school	169,786	0.578	0.494	0	1
Wealth	147,761	0.657	0.236	0	1

Source: (ASER, 2015)

It is useful to look at differences between the mean values of each variable by public and private schools (Table 4.3) as it can highlight any differences between the two, before conducting econometric analysis which will determine if these are significant. The average z-score (standardised child test score) is much lower for a public school child compared with a private school child, which provides a raw indication that private schools do better. Econometric analysis will tell if this finding is robust once other factors are controlled for. Whilst child and parent age tend to be similar for students across school types, it is interesting to note that nearly 34% of private school children received paid tuition compared with 6% of public school students. This may suggest that tuition is an important factor explaining a child's performance on tests, particularly in private schools, and would need to be controlled for to accurately measure the effect of attending a private school.

Also, it was more likely that children in private schools had parents who attended school as 76% of fathers and 54% of mothers had gone to school compared with

52% and 26%, respectively for parents of public school children. There is a 15% difference in the wealth index of children in private schools vs. those in public schools, indicating that children at private schools come from wealthier families.

Table 4.3: Descriptive statistics by school type

Variable	Private	Public	Difference
Z-score	0.322	0.093	0.229
Female	0.390	0.339	0.051
Child Age	9.632	9.791	-0.160
Tuition	0.342	0.060	0.281
No. of Siblings	2.300	2.753	-0.453
Father Age	40.294	41.369	-1.074
Father school	0.757	0.515	0.242
Mother Age	35.273	36.362	-1.089
Mother school	0.544	0.260	0.284
Wealth	0.769	0.618	0.151

Source: (ASER, 2015)

4.2 Econometric analysis

Results from the regression Models

As the variable test score has been standardised, the coefficients on the independent variables refer to how many standard deviations test scores will change as a result of an increase of one standard deviation in the predictor variable. Table 4.4 presents the results from regression Models 1 to 4, as outlined in the data and methodology chapter. Column 1 presents results from Model 1 and shows that on average, when controlling for age and gender, a child that goes to a private school does 0.261 standard deviation times better than a child at a public school. A hypothesis test can be conducted to determine whether a coefficient is statistically different from 0 or if this finding happened by chance. Statistical software

automatically produce results from hypothesis tests and these are represented by asterisks by the standard errors of the coefficients in Table 4.4. The process by which statistical significance are calculated is demonstrated for the private school variable to test if the coefficient on private dummy is statistically significant at the 1% level, and can be viewed in Appendix A.

However, as discussed in the previous chapter, this model does not control for other factors such as innate ability, thus it is prone to omitted variable bias which makes the OLS estimators biased and inconsistent, i.e. it would suggest that attending a private school has a greater impact on test scores than the true value of an unbiased coefficient.

Column 2 of Table 4.4 presents the results from Model 2 which attempts to control for this bias and is the regression of test scores on child characteristics (age and gender) and parental characteristics (age and whether they attended school), where the latter are used to proxy for a child's ability. The results show that the private school effect drops from on average 0.261 standard deviations to 0.198 standard deviations, although it remains statistically significant at the 1% level. The variable controlling for gender differences is negative and statistically significant which means that females on average perform worse on tests than males, in line with the general literature. The coefficient on age is large and statistically significant, meaning that the age of a child is important in determining how well they do on a test.

Parent's education has a positive effect on test scores. If a child's father had gone to school, the child does on average 0.112 standard deviation times better than a child

whose father had not attended school and similar results are found for when a child's mother went to school. Mothers and fathers age both affect test scores and have a quadratic relationship whereby as parents get older, the effect on their child's test scores becomes negative. This could be because when a parent is younger they are more able and interested in their child's education but as they get older, this interest and ability to help reduces. The inclusion of parental variables have a significant effect both in explaining test scores but also in explaining some of the private school effect.

Model 3 incorporates other variables to control for heterogeneity, for example whether a child receives paid tuition outside of school or if wealth of the family has an impact on test scores and what is found is that the pure effect of attending a private school reduces to 0.141 and remains statistically significant at the 1% level (see column 3 of Table 4.4). This suggests that even after controlling for child and family characteristics, there seems to be an advantage of attending a private school over a public school.

The effect of including gender, age and parent's age were similar to those in Model 2, although the coefficient on whether a child's parents had attended school reduced slightly. From the descriptive statistics it was found that 34% of students in private school received paid tuition outside of school. When controlling for this in the model, it was found to be statistically significant that on average, where a child received tuition, their test scores increased by 0.13 standard deviations. The number of siblings a child has had a negative, albeit small, effect on their test score. This implies that a child with many siblings could limit their individual access to resources

and thus the returns to their education which supports the findings in Downey (2001). Finally, wealth, as expected has a large positive effect on test scores. Children of families that are well off tend to do on average 0.227 standard deviations better in tests.

Column 4 of Table 4.4 contains the results from Model 4. Model 3 is the same as Model 4 but for children aged 5 to 10 only, which reduced the sample size to around 92,000. As expected, the coefficient on the private school variable increases in magnitude to 0.163 standard deviations. This is because the test was designed based on the curriculum of grade 2 and 3 children, thus including the results from children older than these grades, whom should on average do better on the test will mean that part of the private school effect could be reduced because older students took the test, underestimating the effect of attending a private school. Restricting the sample to children 5 to 10 years of age eliminates this bias. Parental education and receiving private tuition had a slightly larger effect on test scores than what was found in Model 3, however, the results for the other coefficients were not drastically different, apart from the age of a child's mother becoming statistically insignificant.

Table 4.4: Estimated coefficients of the private school effect and various control variables on test scores, 2015

	(1)	(2)	(3)	(4)
Private	0.261	0.198	0.141	0.163
	(72.99)**	(53.54)**	(33.70)**	(29.43)**
Female	-0.000	-0.019	-0.027	-0.026
	(0.10)	(5.83)**	(7.88)**	(5.78)**
Age	0.196	0.193	0.190	0.270
	(397.26)**	(373.61)**	(347.97)**	(210.02)**
Fathers age		0.013	0.013	0.013
		(5.50)**	(4.93)**	(3.83)**
Fathers age squared		-0.000	-0.000	-0.000
		(6.69)**	(5.61)**	(4.42)**
Father gone school		0.112	0.101	0.126
		(31.18)**	(26.37)**	(25.27)**
Mother age		0.007	0.013	0.004
		(2.54)*	(4.63)**	(1.10)
Mother age squared		-0.000	-0.000	0.000
		(0.11)	(2.39)*	(1.16)
Mother gone school		0.133	0.088	0.111
		(34.99)**	(21.51)**	(20.47)**
tuition			0.130	0.184
			(24.56)**	(25.35)**
sibling			-0.023	-0.036
			(21.68)**	(25.10)**
wealth			0.227	0.183
			(29.56)**	(18.48)**
_cons	-1.824	-2.342	-2.506	-2.936
	(341.14)**	(77.66)**	(78.60)**	(71.89)**
R^2	0.49	0.50	0.52	0.38
N	169,786	168,275	146,702	92,098

Notes: * indicates that the coefficient is statistically significant at the 5% level. ** indicates that the coefficient is statistically significant at the 1% level.
Source: (ASER, 2015)

School quality:

After allowing for child and parental characteristics, the private school effect is still positive and statistically significant at the 1% level. This suggests that a child that attends a private school has an advantage in scoring higher in tests than a public school child. The R^2 value however, which measures how well the data fits the

model, is not greater than 52% for any of the models presented in Table 4.4 which suggests that there may be other variables that are responsible for explaining the variation in test scores. As outlined in the previous chapter, due to constraints of the way the data are coded, it was not possible to control for quality directly in the regression equation. Instead, this section of the chapter looks at school quality indicators by region, and then compares these with the results from the regression of Model 3 conducted separately for each region.

Information about the physical characteristics of schools and other school quality indicators were available for both public and private schools in the ASER Pakistan dataset. Table 4.5 shows analysis of a school infrastructure index (0-10 scale), student teacher ratios and teacher absentee rates for private and public schools in each region/province. The school infrastructure index that has been created is on a scale of 0 to 10 and measures the quality of infrastructure in a given school whereby a higher score on this index implies that the school has good quality infrastructure (see notes of Table 4.5 for further details on the components of the index). On average, infrastructure tended to be better in private schools and private schools tended to have both lower student teacher ratios and teacher absentee rates.

However, there is a stark difference in the quality of schools in different regions of Pakistan in terms of their infrastructure: public schools in rural parts of Balochistan scored the lowest on the infrastructure index at 1.7 and public schools in urban parts of Islamabad scored the highest at 8.1. Both public and private schools in Islamabad scored the highest in terms of their infrastructure. Student teacher ratios were lowest

in Gilgit-Baltistan whereby the average student teacher ratio in private schools was 19 students per teacher and for a public school it was 20 (see Table 4.5).

Private schools in urban Balochistan had the lowest teacher absentee rates at 5% and for public schools, the lowest teacher absentee rates were found in rural Punjab at 7% when compared with other areas. As these variations in school quality across regions are large, it is important to test whether the variations across the regions between public and private schools are much greater than the differences between public and private schools within regions, as it could mean that regional variations are important and that results presented earlier may in fact be capturing differences between regions more than the private/public school difference.

Table 4.5: Quality of school score based on school characteristics by area and institution, 2015

Area	School Infrastructure Index*		Student teacher ratio		Teacher absentee rate	
	Private school	Public school	Private school	Public school	Private school	Public school
Punjab rural	5.6	5.9	22	30	7	7
Sindh rural	5.6	3.1	26	31	12	14
Balochistan rural	3.8	1.7	21	28	7	15
Khyber Pakhtunkhwa rural	6.2	4.2	32	29	6	14
Gilgit-Baltistan rural	4.6	3.8	19	20	8	15
Azad Jammu and Kashmir	3.6	2.9	19	23	7	13
Islamabad – ICT	8.0	6.7	27	23	64	15
Balochistan Urban	5.4	5.7	21	25	5	10
Federally Administrated Tribal Areas	5.6	2.7	35	38	9	10
Punjab Urban	6.3	6.7	20	31	7	8
Sindh Urban	7.0	4.2	19	21	9	14
Khyber Pakhtunkhwa Urban	6.2	5.0	26	34	8	15
Islamabad Urban	-	8.1	-	22	-	12
Average for all areas	5.5	3.9	23	29	8	12

Notes: *This index is based on school infrastructure characteristics where a scale of 0 to 10 measures the quality of infrastructure in a given school: a score of 1 is given to each of the following questions if answered with a yes and 0 otherwise.

a. Is there drinking water facility available in the school for children?
b. Have the school complete Boundary wall?
c. Is there toilet available in the school for children?
d. Are there library books available?
e. Is there a library in the school?
f. Have the school playground wall?
g. Is there electricity connection available?
h. Is there laboratory available?
i. Is there computer lab available?
j. Is there internet facility available?

Thus a higher score on this index implies that the school has good quality infrastructure.
- Data unavailable.
Source: (ASER, 2015)

In order to test this in light of the differences in quality indicators across regions, results from Model 5 are presented in Table 4.6 for the coefficient of interest, private school, for each region. The full results by region are available to view at Appendix B.

There are vast differences in the effect of a child attending a private school on their test scores across different regions of Pakistan. The biggest effect of attending a private school once household characteristics have been controlled for was found in rural Balochistan where on average a child that went to a private school performed 0.26 standard deviations better than a child in a public school in the same area, statistically significant at the 1% level. There were some areas where attending a private school had a negative, albeit small, impact on test scores. Attending a private school in Azad Jammu and Kashmir meant that children did on average -0.07 standard deviations worse than children in public schools.

Other notable trends include that of the 4 provinces split by rural and urban areas. Private school effects tends to be lower in urban areas compared with rural areas and this is the case for Sindh, Balochistan and Khyber Pakhtunkhwa. Research shows that private schools tend be located in urban areas, whereas public schools tend to provide education in rural areas see, for example, Mohiuddin (2007); Mahmood (1999, p. 19). Thus, it could be that when a private school is located in a rural area, it has better quality of teaching than the public schools in the area.

However, Punjab shows a different trend, whereby the private school effect is much lower in rural areas compared with urban areas, where it was found that on average

there was a 0.06 standard deviation improvement in test scores for children in urban areas of Punjab compared with 0.137 for those in rural areas, both statistically significant at the 1% level. Results for Islamabad, the capital city, were found to be statistically insignificant which could be because private schools might not be much different to public schools in the city.

Table 4.6: Estimated coefficients of the private school effect and various control variables on test scores by region, 2015

Province ID	Province name	Private	R^2	N
2	Punjab rural	0.0651	0.49	27,993
		(0.0086202)**		
3	Sindh rural	0.1928	0.48	20,143
		(0.0168791)**		
4	Balochistan rural	0.2604	0.59	29,723
		(0.0178078)**		
5	Khyber Pakhtunkhwa rural	0.1023	0.49	23,937
		(0.0103736)**		
6	Gilgit-Baltistan	0.0351	0.54	7,139
		(0.0153307)*		
7	Azad Jammu and Kashmir	-0.0712	0.50	11,962
		(0.0118294)**		
8	Islamabad – ICT	0.0627	0.53	421
		(0.0733)		
9	Balochistan-Urban	0.0279	0.51	12,402
		(0.0155)		
11	Federally Administrated Tribal Areas	0.0402	0.54	4,050
		(0.0194409)*		
12	Punjab-urban	0.1370	0.57	5,485
		(0.0186853)**		
13	Sindh-Urban	0.1277	0.79	852
		(0.028565)**		
14	Khyber Pakhtunkhwa-Urban	0.0584	0.54	2,146
		(0.029026)*		
15	Islamabad-Urban	-0.0636	0.53	449
		(0.0400)		

Notes: * indicates that the coefficient is statistically significant at the 5% level. ** indicates that the coefficient is statistically significant at the 1% level.
Source: (ASER, 2015)

4.3 Discussion

This findings from this paper are similar to those of Javaid et al. (2012) when looking at the effect of private schools on test scores. There seems to be a private school effect whereby after controlling for individual and parental characteristics, a child that

goes to a private school, on average does better than their public sector counterpart. Javaid et al. (2012 p.17) found that this effect was 0.123 standard deviations using pooled ASER data for 2010 and 2011. The empirical analysis in this paper estimates the private school effect at 0.141 standard deviations using the latest available data for 2015 (column 3 of Table 4.4). This could mean that over time, there has been either an improvement in the quality of education in private schools or a deterioration in the quality of public school education, or some other variable, which could explain the increase in the private school effect over time.

However, Javaid et al. (2012) fail to take into account school quality, most likely due to the way the data are coded in studies that used ASER Pakistan data. This paper adds to the literature by looking at the private school effect by region especially relevant in light of the regional variations in school quality indicators. Comparing regional differences of the private school effect to quality indicators of private and public schools at the province level can allow some relationships to be explored. For example, quality of public schools in rural Balochistan was found to be poor both on the infrastructure score and for teacher absentee rates when compared with private schools in the same area. This could explain why there is a big difference in test scores of children at private schools, as these quality differences are large. However, to assess the significance of quality indicators in Balochistan, and other regions of Pakistan, it is important to include quality within econometric models in future studies.

Although previous studies such as Javaid, et al. (2012) and Khan & Kiefer (2007) use regional fixed effects models, this paper has shown that these are inadequate

for the case of Pakistan, as although such models attempt to control for unobserved heterogeneity across regions, it assumes that the regression equation remains the same for the whole country, i.e. all coefficients on the independent variables are the same. In reality, as this study finds is that by allowing the model to change flexibly across regions, the private school effect varies significantly, as does school quality, and this dilutes the private school effect found for the whole country. This suggests that national level analysis is not appropriate for Pakistan as it is not representative of all regions.

Chapter 5: Conclusion

The aim of this study was to estimate the differences in academic achievement of children attending private and public schools in Pakistan to ascertain whether private schools are providing better education. It began by critically appraising the existing empirical literature on Pakistan and in doing so, it found major defects in the current works whereby the quality of schools are not taken into account in the analysis. This is particularly important as it may plausibly be that differences in school quality are important in explaining differences in test scores between public and private schools.

The analysis in this study started by reproducing some of the results from Javaid et al. (2012) using data for 2015, and found that there seems to be a private school effect whereby after controlling for individual and parental characteristics, a child that goes to a private school, on average does better than their public sector counterpart. However, when school quality was introduced and found to vary significantly across regions, a re-run of the model by region found that the private school effect also varied significantly across regions. In some regions the impact of a child attending a private school on their test score was small and statistically insignificant and in others it was larger than the result found at the national level.

Although this research looks at quality, a limitation of the study, due to data constraints, is that the significance of quality on test scores is yet to be explored in econometric analysis. This research has however, conducted school quality analysis at the regional level and shed light on the importance of including school quality

indicators and conducting analysis at the regional level given the large variations across the different regions in Pakistan.

Thus, this study concludes that conducting analysis using a dummy variable to distinguish between institution types is inadequate. Ideally, what should be done is to collect data on children, their test scores and their characteristics along with information about the quality of the school they attend. This would allow matching of children to schools and being able to account for school quality into econometric modelling to determine their significance. Due to the large nature of the regional variation across Pakistan, future studies should adequately address this heterogeneity either by modelling at regional level or below, or by taking a qualitative case study approach to investigate what exactly is going on in each region.

References and Bibliography

References

Alderman, H., Orazem, P. & Pate, E., 1996. *School quality, School Cost, and the Public/Private School choices of Low Income Household in Pakistan,* s.l.: World Bank. Impact Evaluation of Education Reforms. Working Paper 2.

Angrist, J. D. & Pischke, J.-S., 2009. *Mostly harmless econometrics : an empiricist's companion.* 1st ed. Princeton, NJ: Princeton University Press.

ASER Pakistan, 2014. *Standards Manual.* [Online]
Available at:
http://aserpakistan.org/document/aser/ASER%20STANDARD%20MANUAL.pdf
[Accessed 14 August 2016].

ASER, 2015. *Household and school characteristics.* [Online]
Available at: http://aserpakistan.org/index.php?func=data_statistics
[Accessed 10 June 2016].

Aslam, M., 2007. *The quality of school provision in Pakistan: Are girls worse off? ,* Oxford: GPRG working paper no. 066. University of Oxford..

Barro, R. & Lee, J. W., 1996. International Measures of Schooling Years and Schooling Quality. *The American Economic Review,* 8(2), pp. 218-223.

Buchmann, C., DiPrete, T. & McDaniel, A., 2008. Gender Inequalities in Education. *Annual Review of Sociology,* Volume 34, pp. 319-337.

Card, D. & Krueger, A. B., 1992. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *The Journal of Political Economy,* 100(1), pp. 1-40.

Coleman, J. et al., 1966. *Equalty of Educational Opportunity,* Washington: Government Printing Office .

Country Economy , 2016. *Country comparison Pakistan vs India.* [Online]
Available at:
http://countryeconomy.com/countries/compare/pakistan/india?sector=Population&sc=XE23#tbl
[Accessed 4 August 2016].

Desai, S., Dubey, A., Vanneman, R. & Banerji, R., 2008. *Private schooling in India,* s.l.: Brookings-NCAER India Policy Forum .

Dickerson, A., McIntosh, S. & Valente, C., 2013. *Do the Maths: An Analysis of the Gender Gap in Mathematics in Africa,* Bonn: The Institute for the Study of Labor (IZA): Discussion Paper No. 7174.

Dickson, M., Gregg, P. & Robinson, H., 2013. *Early, late or never? When does parental education impact child outcomes?,* Bristol: Centre for Market and Public Organisation: Working Paper No. 13/298.

Downey, D. B., 2001. Number of siblings and intellectual development: The resource dilution explanation. *American Psychologist,* 56(6-7), pp. 497-504.

Ehrenberg, R. & Brewer, D., 1995. Did teachers' verbal ability and race matter in the 1960s? Coleman revisited.. *Economics of Education Review,* Volume 14, p. 1–21.

Ermisch, J. & Pronzato, C., 2010. *Causal Effects of Parents' Education on Children's Education,* Essex: Institute for social and economic research: Working paper No. 2010-16.

Heyneman, S. & Loxley, W., 1983. The Effect of Primary-School Quality on Academic Achievement Across Twenty-nine High- and Low-Income Countries. *American Journal of Sociology,* 88(6), pp. 1162-1194.

Javaid, K., Musaddiq, T. & Sultan, A., 2012. *Prying the Private School Effect: An Empirical Analysis of Learning Outcomes of Public and Private Schools in Pakistan,* LUMS, Department of Economics: Lahore.

Khan, S. R. & Kiefer, D., 2007. Educational Production Functions for Rural Pakistan: A Comparative Institutional Analysis. *Education Economics ,* 15(3), p. 327–342.

Krueger, A., 1999. EXPERIMENTAL ESTIMATES OF EDUCATION PRODUCTION FUNCTIONS. *The Quarterly Journal of Economics,* 2(114), pp. 497-532.

Lynd, D., 2007. *The Education System in Pakistan: Assessment of the National Education Census,* Islamabad: UNESCO.

Mahmood, N., 1999. *Educational Development in Pakistan: trends, issues, and policy concerns,* Islamabad: Pakistan Institute of Development.

Mohiuddin, Y. N., 2007. *Pakistan: a global studies handbook.* California: ABC-CLIO, Inc.

Muralidharan, K. & Kremer, M., 2006. Public and private schools in rural India. In: R. Chakrabarti & P. Peterson, eds. *School.* s.l.:s.n.

Pakistan Student Association, n.d.. *Map of Pakistan.* [Online]
Available at: https://www.tuhh.de/psa/map.htm
[Accessed 14 August 2016].

Roscigno, V. J. & Ainsworth-Darnell, J. W., 1999. Race, Cultural Capital, and Educational Resources: Persistent Inequalities and Achievement Returns. *Sociology of Education,* 72(3), pp. 158-178.

Rumberger, R. W., 1983. The Influence of Family Background on Education, Earnings, and Wealth. *Social Forces,* 61(3), pp. 755-773.

The Family Planning Association of Pakistan , 1973. *Pakistan - a profile.* Lahore : Communication & Producation Unit TREC, PPPC.

Bibliography

Andrabi, T., Das, J. & Khwaja, I., 2008. A Dime a Day: The Possibilities and Limits of Private Schooling in Pakistan. *Comparative Education Review,* 52(3), pp. 329-355.

Angrist, J. D. & Pischke, J.-S., 2009. *Mostly harmless econometrics : an empiricist's companion.* 1st ed. Princeton, NJ: Princeton University Press.

ASER Pakistan, 2014. *Standards Manual.* [Online]
Available at:
http://aserpakistan.org/document/aser/ASER%20STANDARD%20MANUAL.pdf
[Accessed 14 August 2016].

ASER, 2015. *Household and school characteristics.* [Online]
Available at: http://aserpakistan.org/index.php?func=data_statistics
[Accessed 10 June 2016].

Ashley, L. et al., 2014. *The role and impact of private schools in developing countries: a rigorous review of the evidence. Final report,* London: Department for International Development.

Aslam, M., Jami, B. R. & Rawal, S., 2011. *Teachers and School Quality: Some Policy Pointers from Rural Punjab.* [Online]
Available at:
http://www.aserpakistan.org/document/aser_policy_briefs/2011/Policy_Brief_Teachers_and_School_Quality.pdf
[Accessed 09 August 2016].

Baker, E. et al., 2010. *Problems with the Use of Student Test Scores to Evaluate Teachers.* [Online]
Available at: http://www.epi.org/files/page/-/pdf/bp278.pdf
[Accessed 8 August 2016].

Barro, R. & Lee, J. W., 1996. International Measures of Schooling Years and Schooling Quality. *The American Economic Review,* 8(2), pp. 218-223.

Bold, T., Kimenyi, M., Mwabu, G. & Sandefur, J., 2013. *The high return to low-cost private schooling in a developing country,* London: International Growth Centre, LSE, Working Paper.

Card, D. & Krueger, A. B., 1992. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *The Journal of Political Economy,* 100(1), pp. 1-40.

Chaudhry, H., 2016. *Why do so many children drop out of Pakistani schools?.* [Online]
Available at: http://www.dawn.com/news/1241630
[Accessed 09 August 2016].

Coleman, J. et al., 1966. *Equalty of Educational Opportunity,* Washington: Government Printing Office .

Country Economy , 2016a. *Country comparison Pakistan vs India.* [Online]
Available at:
http://countryeconomy.com/countries/compare/pakistan/india?sector=Population&sc

=XE23#tbl
[Accessed 4 August 2016].

Country Economy, 2016b. *United Kingdom - Population.* [Online]
Available at: http://countryeconomy.com/demography/population/uk
[Accessed 4 August 2016].

Desai, S., Dubey, A., Vanneman, R. & Banerji, R., 2008. *Private schooling in India,* s.l.: Brookings-NCAER India Policy Forum .

Gul, R., Gulshan & Ali, A., 2013. Causes of Dropout Rate in Government High Schools (Male). *American International Journal of Research in Humanities, Arts and Social Sciences,* 13(223), pp. 120-125.

Hussain, A., Salfi, N. A. & Khan, T. M., 2011. Causes of Students' Dropout at Primary Level in Pakistan: An Empirical Study. *International Journal of Humanities and Social Science,* 1(12).

Hussain, R., Rana, J. S., Waqar, K. & Shah, N., 2012. Effects of teaching and learning in large primary classes in Peshawar primary schools . In: A. Bashiruddin, Z. Bana & A. K. Afridi, eds. *Education in Pakistan: Learning from Research Partnerships.* Karachi: Oxford University Press, pp. 27-58.

Interface, 2010. *Quality education in Pakistan.* [Online]
Available at: http://www.interface.edu.pk/students/Aug-10/Quality-education-in-Pakistan.asp
[Accessed 8 August 2016].

Javaid, K., Musaddiq, T. & Sultan, A., 2012. *Prying the Private School Effect: An Empirical Analysis of Learning Outcomes of Public and Private Schools in Pakistan,* LUMS, Department of Economics: Lahore.

Khan, A. I., 1994. *Education, concept and process: a critical analysis of the education system in Pakistan.* First ed. Lahore: Polymer Publications.

Khan, S. R., 2005. *Basic Education in Rural Pakistan.* First ed. Karachi: Oxford University Press .

Ladd, H. & Loeb, S., 2013. The Challenges of Measuring School Quality: Implications for Educational Equity. In: D. Allen & R. Reich, eds. *Education, Justice, and Democracy .* Chicago, IL: University of Chicago Press , pp. 22-25.

Lynd, D., 2007. *The Education System in Pakistan: Assessment of the National Education Census,* Islamabad: UNESCO.

Mahmood, N., 1999. *Educational Development in Pakistan: trends, issues, and policy concerns,* Islamabad: Pakistan Institute of Development.

Mohiuddin, Y. N., 2007. *Pakistan: a global studies handbook.* California: ABC-CLIO, Inc.

Mujiaba, A. D., 2015. Comparative Study of Drop-out Rates among Government and Private Primary Schools. *EUROPEAN ACADEMIC RESEARCH,* 3(8), pp. 9470-9480.

Orr, D. et al., 2013. *What are the impacts and cost-effectiveness of strategies to improve performance of untrained and under-trained teachers in the classroom in developing countries?,* London: EPPI-Centre, Social Science Research Centre, Institute of Education, University of London..

Pakistan Bureau of Statistics, 2014. *Pakistan Statistical Year Book 2014.* Islamabad: Government of Pakistan .

Pakistan Student Association, n.d.. *Map of Pakistan.* [Online]
Available at: https://www.tuhh.de/psa/map.htm
[Accessed 14 August 2016].

Popham, J. W., 1999. Using Standards and Assessments. *Eductaional Leadership,* 56(6), pp. 8-15.

Rumberger, R. W., 1983. The Influence of Family Background on Education, Earnings, and Wealth. *Social Forces,* 61(3), pp. 755-773.

The Family Planning Association of Pakistan , 1973. *Pakistan - a profile.* Lahore : Communication & Producation Unit TREC, PPPC.

World Bank , 2000. *Beyond Economic Growth Student Book.* [Online]
Available at: http://www.worldbank.org/depweb/english/beyond/global/chapter4.html
[Accessed 4 August 2016].

World Development Indicators, World Bank, 2016. *Government expenditure on education, total (% of GDP).* [Online]
Available at:
http://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS?name_desc=false
[Accessed 07 August 2016].

Zarif, T., 2012. Reasons of High Dropout Rate of Students in Grades 5-6 at Public Schools of District Kashmore (Sindh-Pakistan) in 2010-11. *Journal of Research and Reflections in Education,* 6(2), pp. pp145 -153.

Appendix A

A hypothesis test can be conducted to determine whether a coefficient is statistically different from 0 or if this finding happened by chance. Statistical software automatically produce results from hypothesis tests and these are generally represented by asterisks next to the standard errors of the coefficients. The process by which statistical significance are calculated is demonstrated for the private school variable in Model 1 below to test if the coefficient on private dummy (0.261) is statistically significant at the 1% level.

In the case of the private dummy, the null and alternative hypothesis is stated below.

H_0: α1=0; the coefficient on private dummy equals 0

H_1: α1≠0; the coefficient on private dummy is different from 0

The t statistic is used to perform the test due to the large sample size which means that the distribution approximates to that of a normal distribution. The relevant critical value for this two-tailed test at the 1% significant level is $t_{0.005}$= 2.67.
The test statistic is calculated as $t = \frac{\widehat{\alpha_1}}{se\widehat{\alpha}_1}$, which in this case equals 72.99; where $se\hat{\alpha}_1$ is the standard error of the private dummy coefficient.

As 72.99 is greater than 2.67, the null hypothesis is rejected and it can be concluded that the coefficient on private dummy is statistically significant at the 1% level. This means that children at private schools on average obtain higher marks on tests.

Appendix B

This appendix presents the full results of the regression:

$$Test\ score_i = \alpha_0 + \alpha_1 Private_i + \alpha_2 Female_i + \beta_1 Age_i + \alpha_3 Tuition_i + \beta_2 Siblings_i + \beta_3 Father's\ age_i + \beta_4 Father's\ age^2{}_i + \alpha_4 Father\ school_i + \beta_5 Mother's\ age_i + \beta_6 Mother's\ age^2{}_i + \alpha_5 Mother\ school_i + \beta_7 wealth_i + \varepsilon_i$$

Where Private, Female, tuition, father school, mother school, =1 or 0 and ε_i is the idiosyncratic error term, for each region.

pid = 2	Punjab rural
Number of obs	27993
R-squared	0.4857
Adj R-squared	0.4855

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.065	.0086202	7.56	0	0.04824	0.082037
Female	0.024	.007927	3.07	0.002	0.00878	0.039851
Age	0.188	.0013021	144.29	0	0.18533	0.190433
Tuition	0.089	.0094887	9.34	0	0.07005	0.107244
Sibling	0.000	.0025883	0.16	0.869	-0.0046	0.005499
Fathers age	0.000	.0059246	0.05	0.962	-0.0113	0.011895
Fathers age squared	0.000	.0000655	-1.11	0.268	-0.0002	5.58E-05
Father gone school	0.046	.0091347	5.05	0	0.02825	0.064054
Mother age	0.033	.0067576	4.81	0	0.01929	0.045776
Mother age squared	0.000	.000086	-3.17	0.002	-0.0004	-0.0001
Mother gone school	0.092	.0089319	10.28	0	0.07427	0.109283
wealth	0.201	.0217109	9.25	0	0.15831	0.243421
_cons	-2.514	.0826423	-30.41	0	-2.6755	-2.35153

pid = 3	Sindh rural
Number of obs	20143
R-squared	0.4823
Adj R-squared	0.482

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.193	.0168791	11.42	0	0.15974	0.225904
Female	-0.026	.0095729	-2.67	0.008	-0.0443	-0.00678
Age	0.184	.0014992	122.78	0	0.18113	0.187006
Tuition	0.219	.0200252	10.94	0	0.17974	0.258238
Sibling	-0.020	.0028145	-7.02	0	-0.0253	-0.01424
Fathers age	0.021	.0067009	3.21	0.001	0.00835	0.034618
Fathers age squared	0.000	.0000727	-4.31	0	-0.0005	-0.00017
Father gone school	0.065	.0101563	6.37	0	0.04481	0.084621
Mother age	0.004	.0073219	0.59	0.555	-0.01	0.01867
Mother age squared	0.000	.0000905	0.99	0.324	-9E-05	0.000267
Mother gone school	0.082	.0119915	6.87	0	0.05885	0.105859
wealth	0.176	.0200104	8.81	0	0.13715	0.21559
_cons	-2.653	.0869217	-30.52	0	-2.8231	-2.48236

pid = 4	Balochistan rural
Number of obs	29723
R-squared	0.588
Adj R-squared	0.5878

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.260	.0178078	14.62	0	0.2255	0.295306
Female	-0.085	.0074906	-11.31	0	-0.0994	-0.07002
Age	0.216	.0011867	182.36	0	0.21408	0.218736
Tuition	0.245	.0245837	9.98	0	0.19728	0.293648
Sibling	-0.018	.0019683	-9.32	0	-0.0222	-0.01449
Fathers age	0.005	.0050914	0.90	0.366	-0.0054	0.014578
Fathers age squared	0.000	.000057	-0.19	0.848	-0.0001	0.000101
Father gone school	0.084	.0079367	10.60	0	0.06858	0.099688
Mother age	-0.006	.0051491	-1.14	0.254	-0.016	0.004224
Mother age squared	0.000	.0000634	1.45	0.148	-3E-05	0.000216
Mother gone school	0.040	.0114553	3.49	0	0.01751	0.062418
wealth	0.219	.0132211	16.55	0	0.19289	0.24472
_cons	-2.383	.0594243	-40.10	0	-2.4993	-2.26633

pid = 5 Khyber Pakhtunkhwa rural

Number of obs 23937

R-squared 0.4874

Adj R-squared 0.4872

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.102	.0103736	9.86	0	0.08194	0.122602
Female	-0.068	.0084331	-8.06	0	-0.0845	-0.05145
Age	0.174	.0012702	136.96	0	0.17147	0.176454
Tuition	0.077	.0135188	5.69	0	0.05042	0.10342
Sibling	-0.017	.0026642	-6.25	0	-0.0219	-0.01143
Fathers age	-0.016	.0069733	-2.36	0.018	-0.0301	-0.00281
Fathers age squared	0.000	.0000814	1.30	0.192	-5E-05	0.000266
Father gone school	0.066	.0093337	7.03	0	0.04729	0.08388
Mother age	0.024	.0071666	3.34	0.001	0.00991	0.038003
Mother age squared	0.000	.0000953	-2.15	0.032	-0.0004	-1.8E-05
Mother gone school	0.039	.0093225	4.20	0	0.02092	0.057466
wealth	0.200	.0193116	10.37	0	0.1625	0.238208
_cons	-1.707	.0832666	-20.50	0	-1.8702	-1.54378

pid = 6 Gilgit-Baltistan

Number of obs 7139

R-squared 0.5423

Adj R-squared 0.5416

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.035	.0153307	2.29	0.022	0.0051	0.065201
Female	-0.029	.0142467	-2.05	0.041	-0.0571	-0.00125
Age	0.183	.0022389 81.51		0	0.17811	0.186889
Tuition	0.087	.0248448	3.48	0	0.03788	0.135283
Sibling	-0.014	.0047443	-2.97	0.003	-0.0234	-0.00479
Fathers age	0.034	.0094906	3.61	0	0.01565	0.052857
Fathers age squared	0.000	.0001008	-3.96	0	-0.0006	-0.0002
Father gone school	0.094	.0159117	5.93	0	0.06309	0.12547
Mother age	0.023	.0107971	2.16	0.031	0.00215	0.044485
Mother age squared	0.000	.0001359	-1.40	0.163	-0.0005	7.66E-05
Mother gone school	0.026	.0181157	1.46	0.144	-0.0091	0.061972
wealth	0.291	.0375108	7.76	0	0.21742	0.364489
_cons	-3.031	.1483012 -20.44		0	-3.3218	-2.74038

pid = 7 Azad Jammu and Kashmir

Number of obs	11962
R-squared	0.4965
Adj R-squared	0.496

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	-0.071	.0118294	-6.02	0	-0.0944	-0.04802
Female	0.000	.01097	-0.02	0.987	-0.0217	0.021326
Age	0.177	.0018371	96.48	0	0.17364	0.180841
Tuition	0.069	.0209389	3.27	0.001	0.02751	0.109595
Sibling	-0.010	.00436	-2.31	0.021	-0.0186	-0.00152
Fathers age	-0.024	.01014	-2.41	0.016	-0.0443	-0.0046
Fathers age squared	0.000	.0001108	2.88	0.004	0.0001	0.000537
Father gone school	-0.001	.0147092	-0.04	0.971	-0.0294	0.028303
Mother age	0.061	.0106918	5.75	0	0.04052	0.082437
Mother age squared	-0.001	.0001316	-6.22	0	-0.0011	-0.00056
Mother gone school	0.002	.0132779	0.16	0.872	-0.0239	0.028167
wealth	0.126	.0301427	4.18	0	0.06681	0.184979
_cons	-2.057	.1216685	-16.90	0	-2.2952	-1.81826

pid = 8	Islamabad – ICT
Number of obs	421
R-squared	0.5298
Adj R-squared	0.516

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.063	.0732754	0.86	0.393	-0.0814	0.206715
Female	0.032	.0622484	0.52	0.605	-0.0901	0.154604
Age	0.194	.0111836	17.31	0	0.17159	0.215564
Tuition	-0.010	.0702302	-0.14	0.889	-0.1478	0.128267
Sibling	-0.015	.0256533	-0.59	0.555	-0.0656	0.035256
Fathers age	0.074	.0500781	1.48	0.14	-0.0243	0.172553
Fathers age squared	-0.001	.0005539	-1.33	0.184	-0.0018	0.000352
Father gone school	0.026	.1021575	0.26	0.797	-0.1746	0.227057
Mother age	-0.055	.0609283	-0.90	0.37	-0.1745	0.065084
Mother age squared	0.001	.0007891	0.93	0.353	-0.0008	0.002285
Mother gone school	0.294	.0796683	3.69	0	0.13698	0.450204
wealth	0.137	.1814829	0.76	0.45	-0.2195	0.494031
_cons	-2.655	.869609	-3.05	0.002	-4.3642	-0.94529

pid = 9 Balochistan-Urban

Number of obs	12402
R-squared	0.5122
Adj R-squared	0.5117

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.028	.0155308	1.80	0.072	-0.0025	0.058389
Female	-0.133	.0127279	-10.47	0	-0.1583	-0.10836
Age	0.186	.001835	101.44	0	0.18255	0.189739
Tuition	0.287	.0251094	11.43	0	0.23782	0.336254
Sibling	-0.020	.0035581	-5.75	0	-0.0274	-0.01348
Fathers age	0.031	.0112949	2.77	0.006	0.0091	0.053378
Fathers age squared	0.000	.0001362	-1.99	0.047	-0.0005	-4.18E-06
Father gone school	0.098	.0117131	8.33	0	0.07457	0.120489
Mother age	-0.007	.0116482	-0.61	0.541	-0.03	0.015714
Mother age squared	0.000	.0001548	0.09	0.932	-0.0003	0.000317
Mother gone school	0.093	.0156285	5.98	0	0.06282	0.124093
wealth	0.126	.0257066	4.88	0	0.07515	0.175929
_cons	-2.213	.1102519	-20.07	0	-2.4291	-1.99688

pid = 11	Federally Administrated Tribal Areas
Number of obs	4050
R-squared	0.5424
Adj R-squared	0.541

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.040	.0194409	2.07	0.039	0.00213	0.07836
Female	0.032	.0182183	1.78	0.075	-0.0032	0.068187
Age	0.178	.0030334	58.84	0	0.17253	0.184425
Tuition	0.111	.0191732	5.78	0	0.07323	0.148409
Sibling	-0.017	.0071789	-2.30	0.021	-0.0306	-0.00245
Fathers age	-0.003	.0139764	-0.23	0.818	-0.0306	0.024178
Fathers age squared	0.000	.0001513	0.21	0.834	-0.0003	0.000328
Father gone school	0.028	.025474	1.12	0.264	-0.0215	0.07839
Mother age	0.069	.0145516	4.72	0	0.04014	0.097203
Mother age squared	-0.001	.0001777	-4.27	0	-0.0011	-0.00041
Mother gone school	0.084	.0233554	3.58	0	0.03789	0.12947
wealth	0.138	.0609388	2.27	0.024	0.01858	0.257529
_cons	-2.984	.2177379	-13.71	0	-3.4111	-2.55737

pid = 12	Punjab-urban
Number of obs	5485
R-squared	0.565
Adj R-squared	0.564

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.137	.0186853	7.33	0	0.10032	0.173581
Female	0.001	.0148147	0.09	0.929	-0.0277	0.030366
Age	0.172	.0024401	70.30	0	0.16675	0.176322
Tuition	0.202	.0152291	13.24	0	0.17175	0.231464
Sibling	-0.015	.0052914	-2.82	0.005	-0.0253	-0.00455
Fathers age	0.032	.0136347	2.35	0.019	0.00526	0.058723
Fathers age squared	0.000	.0001529	-2.55	0.011	-0.0007	-9.1E-05
Father gone school	-0.053	.0231003	-2.31	0.021	-0.0986	-0.00802
Mother age	0.034	.0152872	2.19	0.028	0.00358	0.063515
Mother age squared	0.000	.0001948	-1.79	0.074	-0.0007	3.36E-05
Mother gone school	0.072	.0197025	3.67	0	0.03367	0.110915
wealth	0.244	.0481438	5.06	0	0.14937	0.338134
_cons	-3.070	.1801906	-17.04	0	-3.423	-2.71655

pid = 13	Sindh-Urban
Number of obs	852
R-squared	0.7896
Adj R-squared	0.7866

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.128	.028565	4.47	0	0.07161	0.18374
Female	-0.022	.0310027	-0.70	0.486	-0.0825	0.039246
Age	0.231	.0051305	44.94	0	0.2205	0.240644
Tuition	0.134	.0489728	2.74	0.006	0.03806	0.230311
Sibling	0.056	.0101664	5.50	0	0.036	0.075913
Fathers age	-0.025	.0220833	-1.11	0.265	-0.068	0.018737
Fathers age squared	0.000	.0002504	0.69	0.487	-0.0003	0.000666
Father gone school	0.109	.0318462	3.42	0.001	0.04637	0.171387
Mother age	0.079	.0229978	3.43	0.001	0.03364	0.123915
Mother age squared	-0.001	.0002821	-2.62	0.009	-0.0013	-0.00019
Mother gone school	0.062	.0356254	1.75	0.08	-0.0075	0.132339
wealth	-0.150	.0816244	-1.84	0.066	-0.3102	0.010228
_cons	-3.532	.2888706	-12.23	0	-4.0992	-2.96519

pid = 14	Khyber Pakhtunkhwa-Urban
Number of obs	2146
R-squared	0.5376

Adj R-squared 0.5349

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	0.058	.029026	2.01	0.045	0.00144	0.11528
Female	-0.051	.0267723	-1.92	0.055	-0.1039	0.001075
Age	0.188	.0043002	43.76	0	0.17974	0.196602
Tuition	0.031	.0338502	0.91	0.362	-0.0355	0.097234
Sibling	-0.004	.0089568	-0.40	0.687	-0.0212	0.013961
Fathers age	-0.001	.0267429	-0.02	0.984	-0.053	0.051893
Fathers age squared	0.000	.0003078	0.18	0.857	-0.0005	0.000659
Father gone school	-0.017	.0335567	-0.51	0.611	-0.0829	0.048719
Mother age	0.014	.0263685	0.53	0.596	-0.0377	0.065692
Mother age squared	0.000	.0003451	-0.74	0.46	-0.0009	0.000422
Mother gone school	0.057	.0294855	1.95	0.052	-0.0004	0.115218
wealth	-0.430	.0662095	-6.50	0	-0.56	-0.30027
_cons	-1.510	.3102724	-4.87	0	-2.1181	-0.90112

pid = 15 Islamabad-Urban
Number of obs 449
R-squared 0.5288
Adj R-squared 0.5158

Test Scores	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Private	-0.064	.0400479	-1.59	0.113	-0.1423	0.015161
Female	0.017	.0346794	0.50	0.616	-0.0508	0.085558
Age	0.106	.0060372	17.59	0	0.09432	0.118055
Tuition	0.050	.0481995	1.03	0.302	-0.0449	0.144534
Sibling	-0.021	.0154807	-1.33	0.183	-0.0511	0.009781
Fathers age	-0.001	.0433358	-0.02	0.984	-0.086	0.084297
Fathers age squared	0.000	.0005207	0.05	0.958	-0.001	0.001051
Father gone school	-0.001	.1223406	-0.01	0.994	-0.2414	0.239542
Mother age	0.157	.0570017	2.75	0.006	0.04465	0.268716
Mother age squared	-0.002	.0007779	-2.71	0.007	-0.0036	-0.00058
Mother gone school	0.088	.0911666	0.97	0.332	-0.0907	0.26765
wealth	0.148	.2334884	0.63	0.527	-0.3111	0.606674
_cons	-3.187	.671337	-4.75	0	-4.5064	-1.86748