DETERMINANTS OF PUPILS' EXAM PERFORMANCE IN THE UK AND PAKISTAN

By:

Uzma Ahmad

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Economics
Faculty of Social Sciences
The University of Sheffield

June 2018

Abstract

This thesis consists of three chapters exploring the determinants of pupil performance in high-stakes exams in UK and Pakistan.

The first empirical chapter looks at the effect of parents' education on their children's education adopting Instrumental Variable (IV) methodology. I exploit the 1972 UK RoSLA (raising of school leaving age) as a source of exogenous variation using the Longitudinal Study of Young People in England (LSYPE) dataset linked to the National Pupil Database (NPD). The results show that parental education has a significant and positive impact on their children's educational outcomes, as measured by performance in GCSE examinations taken at age 16.

The subsequent chapter explores key issues relating to teacher quality. Firstly, it investigates the impact of being taught by high quality or low quality teachers on students' performance in exams, and secondly, which teacher characteristics are associated with student performance. It uses the survey data of 611 pupils from one region of Pakistan studying in Year 9 which is also linked to the administrative data on the students' exam scores.

The unique feature of the data tracks students' scores across multiple subjects at a single point in time. Teacher fixed effects and pupil fixed effects methods are used. First stage results show that there are significant variations in teacher fixed effects within the schools, suggesting important unobservable differences among teachers. A good teacher, defined as being at the 75th percentile of the teacher fixed effects distribution, is related to an increase in score by 0.15 standard deviations relative to the omitted teacher, while a bad teacher decreases the score by 0.77 standard deviations. Therefore, a pupil having been taught by a good teacher scores 0.92 standard deviations more than the pupil who is taught by a bad teacher (25th percentile teacher) leaving a significant effect on pupil performance. Second stage results suggest that teacher observed characteristics do not explain the variation in teacher quality.

The final chapter studies the effect of distance on participation in post-compulsory education in the Pakistan context, and whether socioeconomic characteristics have an effect on achievement in a value-added model taking selection into secondary education into account using a survey dataset on pupils studying in the post-compulsory grade (Year 12) in 2011-2012 from one district of the Punjab province of Pakistan. In this chapter I used two variables as instruments: one is the log of distance to nearest post-compulsory education institution, as a measure of proximity to an educational institution and the second is urban location. The results show that participation and performance in post-compulsory education are two different processes, with participation being driven by availability of post-compulsory institutions within travel distance, while performance once in post-compulsory education is determined by ability.

Acknowledgements

While my name may be alone on the front cover of this thesis, there are a number of people behind this piece of work who deserve to be both acknowledged and thanked here.

Firstly, I am grateful to Allah Almighty, who gave me an opportunity to take up this challenging task and gave strength to write this dissertation.

After that, I would like to express my most sincere, special gratitude to my principal supervisor, Steven McIntosh for his constant kindness, honesty, inspiration, patience and support throughout this long journey of my PhD. Without his support this work would not have been possible. He also helped me grow my skills as an independent researcher, academic writing, good research practise and gave me the space and freedom I needed from day one. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. He is the best ever supervisor. I am forever truly grateful to him for being a tremendous mentor for me. His advice and help on both research as well as on my career has been priceless.

I must express a great thanks to my other supervisor, Gurleen Popli for her valued comments, brilliant and constructive guidance, and her deep understanding in this field that led to a substantial improvement to my work. I would like to thank her for always encouraging my research and improving the writing of thesis. She also deserves thanks for helping me keep things in perspective.

In addition, I would like to say thanks to all other administrative and academic staff members within department of economics and University of Sheffield, especially Arne Risa Hole. I am also thankful to Karl Taylor and the Department of Economics for financial assistance to present my research work in a number of conferences in Germany, Spain, York, Leeds, and Sheffield. I am also grateful to those faculty members who contributed effectively in departmental conferences and seminars.

I would like to show appreciation to Dr. Bilal A Khan, Vice Chancellor, The Islamia University Bahawalpur Pakistan for providing the scholarship and also to Syed Wasim Hashmi in Higher Education Commission, Pakistan for resolving issues relating to funding.

I am also eternally obliged to my teacher and my MPhil supervisor Dr. Karamat Ali, who always persuaded me to pursue higher studies. He inspired me with his strong vision. Completing the rigorous work of survey data collection would have been significantly more difficult were it not for his advice and support.

Finally, words cannot express how grateful I am to my beloved husband for providing support and encouragement throughout this process. Thank you for your part in my journey. I would like to give my greatest heartfelt thanks to him for his enormous patience, understanding and complete support in every way possible.

A big special thanks to all of my lovely three children for their maturity, patience, love and support through the ups and downs of my research. Having a third baby in the 3rd year of my PhD and looking after three young children whilst undertaking my PhD needed equal engagement and placed a great deal of strain on me. Despite this, they brought lots of smiles and laughter to my worries and provided a much needed form of escape from my studies. Thanks to all of you for your contributions and immense understanding throughout this stressful process.

Lastly, but by no means least, I would like to express my appreciation to my wonderful parents for their constant and selfless support. I would also like to thank my supportive younger sister for all of the sacrifices that she has made on my behalf. Their prayers for me were what sustained me thus far. I am very grateful to my uncle and his family for their contributions and help throughout. I am also thankful to my mother-in-law and father-in law for their prayers and support.

Data Acknowledgements

The analysis of chapter 2 is based using data from the Longitudinal Study of Young Peoples in England (LSYPE) and National Pupil Database (NPD) provided by the Department of Education (DfE), UK Data Service. Chapter 3 and 4 uses the administrative data on pupils' exam sores provided by the Board of Intermediate and Secondary education (BISE) Bahawalpur, Pakistan.

List of Contents

List of Tables

List of Figures

Chapter 1:

Introduction

1.1 Aims and Motivations

Human capital is the stock of knowledge, skills and abilities that enables individuals to be productive and thus earn income. It can be accumulated through investment in education, training, and other factors that increase productivity and earnings. The significance of the role of human capital cannot be denied in economic development during the last half century, for economists and international development organizations (Mankiw et al., 1992). Simultaneously, it has also been proved that people-oriented human resources provide any organization an edge by creating superior share-holder value, so the education component of human capital is a key motivational force for economic growth (UNDP, 2003).

Economics has provided great insights into a multitude of topics surrounding education; most notably, human capital theory pioneered by Becker (1964) assisted in explaining why individuals choose to invest in education and training, while Mincer (1974) developed the earnings function which has been widely used to estimate returns to education.

Recent growth in the economics of education research has come in the form of improved data availability and development of new methodological approaches (Machin, 2008). This has improved understanding of issues relating to intergenerational mobility, the impact of school choice and competition on pupil outcomes, and the link between teacher quality and student performance (Dickson et al., 2016; Gibbons et al., 2008; Azam and Kingdon, 2015; Bohlmark and Lindahl, 2007). In addition, the economics of education has expanded to include policy evaluation research, which has made the field more relevant for informing policy makers.

In particular, major developments have been made in explaining the determinants of pupil outcomes. Several studies have explored the various determinants of educational attainment or performance, which are mainly family characteristics, teachers and schools (Hanushek and Rivkin, 2012; Vignoles and Meschi, 2010). These studies consider the primary, secondary and post-compulsory level of education in different countries. Pupil outcomes are likely to play a significant role in an individual's life chances and adult outcomes (Blanden et al., 2012; Field, 2010; OECD, 2007, ONS, 2011). It is therefore crucial to understand the determinants of pupil outcomes such as individual, family, teacher, and school characteristics. Equally, it is important to understand how policies and interventions may be used to improve outcomes.

Each chapter of this thesis looks at different determinants of performance in high stakes exams. The GCSEs exams in England at age 16 (chapter 2) and the Year 9 exams in Pakistan (chapter 3) which are held at the age of 14-15 are high stakes tests and are the key gateway exams into post-compulsory education and also very important for labour market outcomes. Further, performance in these high stakes exams determines the sixth form pupils go to, the qualifications they take next, the eligibility for a university course, the universities they can apply to and future prospects. The grade 12 exams in Pakistan (chapter 4) taken in the age group of 17 to 18 are also crucial as they provide access to university and renowned fields such as medicine, engineering, etc. In Pakistan, Year 10 at age 15-16 is equivalent to O-Level exams in UK, and Year 9, the exam preceding GCSE and taken at age 14 or later. Years 11 and 12 are equivalent to A-Levels exams usually sat at age 18 or later.

Every chapter is a micro-econometric investigation of the determinants of pupil performance in high stakes exams. Chapter 2 uses the Longitudinal Study of Young People in England (LSYPE) and National Pupil Database (NPD) of the UK adopting an Instrumental Variable (IV) methodology. Chapter 3 uses survey panel data for Pakistan collected by the author in order to discuss the issue of teacher quality using teacher fixed effects and pupil fixed effects approaches. Chapter 4 uses the same dataset for Pakistan looking at the determinants of performance and participation in higher grades using Probit and Heckman Selection Models. Overall, the thesis uses a range of econometric methodologies and number of different datasets to study the pupil outcomes in high stakes exams in a developed (UK) and developing country (Pakistan).

1.2 Structure and content of the thesis:

The thesis consists of three different chapters: The first chapter is based on the intergenerational mobility of education in England, whereas the other two chapters address the issue of teachers' quality and determinants of sustained academic achievements in Pakistan. Thus, the present study has serious implications for both developed (UK) and developing (Pakistan) countries particularly in an educational context.

1.2.1 Brief overview of chapter 2

Recent studies provide evidence that an intergenerational correlation exists between the education of parents and their children. There are two possible explanations of a positive intergenerational correlation: one is direct and the other is indirect. The first, and direct, cause states that it could be the result of the genetic transmission of ability such that talented parents have more able children. If this is the sole reason for the intergenerational relationship, then the issue of higher achievements amongst future generations can be ignored when evaluating educational policy aimed at raising present education levels, since the inherited genes will not have been affected by the existing situation.

The second and indirect cause of the intergenerational correlation works through two routes. One indirect route functions via the direct transfer of knowledge: for example, the more motivated and educated parents are in a better position to help and push their children as they have experienced the benefits of education themselves. The other indirect route works through income and lifestyle. It argues that more educated parents have higher incomes which can buy many things like private schooling, books, tutors and an affluent neighbourhood.

Chapter 2 investigates the causal mechanism underpinning this intergenerational correlation – vis-à-vis the established link between parents' and children's educational outcomes – using the Longitudinal Study of Young People in England (LSYPE) dataset. Using the raising of the school leave age (RoSLA) in 1972 as an instrument for parental education allows us to isolate the exogenous variation in parents' education.

The results confirm that parental education is positively related to their children's educational outcomes, as measured by performance in GCSE examinations taken at age 16, and suggest that the effect of nurture (upbringing) is mainly responsible for the intergenerational relationship. Further results suggest that controlling for both parents' education, mothers' education is positively related to children's education while the effects of fathers' education disappear. These findings are robust across boys and girls samples.

1.2.2 Brief overview of chapter 3

Teachers differ significantly in how they teach their students, but little is known about which teacher attributes account for this. Improving weak teaching may be one of the most effective means of raising pupil achievement. Teacher quality is a question of interest for education policy-makers and researchers.

The majority of studies estimate teacher effects at elementary/primary school level. However studies addressing the issue of teacher quality at high school level are limited. Slater et al. (2012) estimate teacher effects at senior secondary level, Aaronson et al. (2007); Azam and Kingdon (2015); and Kingdon (2006) study teacher quality at high school level. Chapter 3 also discusses teacher quality at higher secondary level, as one of the concerns about estimating teacher effects using school data at primary grades is that all students are taught by the same teacher in primary grades. This implies that one cannot estimate the effect of multiple teachers on the same student in different subjects.

Chapter 3 uses the survey data of 611 pupils, from one of the districts of Punjab, (Pakistan), studying in Year 9 (age 14-15) in 2008-09, from both private and public schools; the survey data is linked with the administrative data of the student exam scores, to address the issue of teacher quality using a teacher fixed effects approach. The present study takes the direct approach, linking teacher characteristics to student outcomes in an achievement production function, but with two innovations.

Firstly, it estimates the effect of teachers on student achievement (exam score) using teacher fixed effects models. Secondly, it relates the estimated teacher effects to the observable characteristics of the teachers.

First stage results show that there are significant variations in teacher fixed effects within the schools, suggesting important unobservable differences among teachers. A good teacher, defined as being at the 75th percentile of the teacher fixed effects distribution, is related to an increase in score by 0.15 standard deviations relative to the omitted teacher, while a bad teacher decreases the score by 0.77 standard deviations. Therefore, a pupil having been taught by a good teacher scores 0.92 standard deviations more than the pupil who is taught by a bad teacher (25th percentile teacher) leaving a significant effect on pupil performance. Second stage results suggest that teacher characteristics do not explain the variation in teacher quality.

1.2.3 Brief overview of chapter 4

The fourth chapter is an investigation of the determinants of participation and performance in post-compulsory education, controlling for the selection into post-compulsory education and prior ability, using a unique primary dataset on pupils studying in the post -compulsory grade (Year 12) in 2011-2012 from one district of the Punjab province of Pakistan.

The results show that participation and performance in post-compulsory education are two different processes, with participation being driven by availability of post-compulsory institutions within travel distance, while performance once in post-compulsory education is determined by ability.

This is an important and interesting result with obvious policy relevance - those participating in post-compulsory education are not necessarily the most able, but rather those with the best access to post-compulsory education.

To provide context for Chapters 3 and 4, the following section will discuss the structure of compulsory schooling in Pakistan.

1.2.4 Brief Overview of the Education System in Pakistan

The education system in Pakistan is divided into five different levels.

1) Primary Level (grades 1 to 5)
2) Middle Level (grades 6 to 8)
3) High/Secondary Level (grades 9 and 10, leading to the Secondary School Certificate or SSC)
4) Intermediate/Higher/Higher Secondary Level (grades 11 and 12, leading to the Higher Secondary Certificate or HSC)
5) University Level (leading to undergraduate and graduate degree and research).

The Secondary School Certificate (SSC) includes the grades 9 and 10 examinations, and is equivalent to GCSE/O level in England, while the Higher Secondary School

Certificate includes grades 11 and 12 examinations, also known as HSC, is equivalent to GCE A Level in England. Both SSC and HSC exams are conducted through the Board of Intermediate and Secondary Education (BISE). There are also alternative qualifications available in Pakistan, where SSC and HSC are replaced by Ordinary Level (O Level) and Advanced Level (A Level) which are managed by the British examination boards of Cambridge University, however the present study does not include these alternatives.

There is also vocational and technical education available for the development of the skilled workforce. Technical education comprises of three years of education after matriculation/Grade10. Vocational training consists of 6 months to two years duration after Grade 8 or Grade 10. Analysis done here focuses only on formal academic education, and does not consider technical or vocational training. According to Article 25-A of the Constitution of Pakistan, the state must provide free and compulsory education from age 5 to 16 years, i.e. for Grades 1 to 10.

There are two main types of schools that exist in Pakistan, private and public schools, different to England, where mainly school types are community schools, foundation schools, independent schools and voluntary schools. Although similar in terms of their education structures they both differ in terms of finances and regulations. Most of the cost of operating the public school system is borne by the public exchequer.

Private schools are owned by sole proprietors, trusts or non-governmental organizations (NGOs) or other forms of management, work for profit and follow either the national curriculum or a curriculum approved by foreign educational institutions and are fee-charging schools. In public schools, female students are always taught by female teachers and male students by male teachers, while in private school, it could be male or female teachers depending upon the availability.

Public schools follow the national curriculum, charge no tuition fee and are registered with the Education Department. Although every private school is required to be registered with the Education Department, most of the schools remain unregistered and consequently the size of the private school sector is unmeasured in government statistics (Aslam, 2009). In this way such schools avoid large taxation and other costs.

All schools, private or public, which are registered with the Education Department, are linked with the administrative data, from which the test scores of the pupils can be obtained (similar to the National Pupil Database (NPD) in the UK). If a private school is not registered with the Education Department then they cannot be linked with the administrative data.

In my sample I have pupils from both the private and public schools, and all the private schools are registered with the Education Department, so I can link them up to the administrative data. It is generally believed that private schools provide a better education as compared to public schools. Also previous research has shown that pupils from private schools perform better than those from public schools at both primary and middle levels of education (Alderman et al., 2001; and Aslam, 2009).

As previously discussed above, the private and public schools are the main education providers in Pakistan. All Public schools are free in providing education from Grade 1 to Grade 10, including free books and uniforms, while private schools charge fees ranging from 500 to 12,000 rupees per month depending on their standard and location etc. Also there is a separate charge for examination fees, books and uniform.

Turning to post-compulsory education, both private and public colleges charge fees. Public colleges charge fees up to 8,000 rupees, while private college fees vary a lot between 15,000 to 150,000 Rs**.** Mostly, private colleges offer A-levels with traditional academic tracks in a variety of subjects while some public colleges offer vocational education with other traditional academic routes in a variety of subjects. Most of the private colleges have co-education while all public colleges are single sex. Some private colleges involve entry tests at the time of admission but public colleges do not.

Chapter 2:

Measuring the Nature and Nurture Effects in Intergenerational Transmission of Human Capital in England

2.1 Introduction

There is an established link between parents' and children's educational outcomes as parents always play an important role in their children's education regarding their human capital investments. The children of highly educated parents have higher educational levels and better labour market outcomes as compared to those children who grow up in less educated families. Why does this happen - "Is it that more able parents have more able children"? This chapter considers the ideas of selection and causation. The selection theory states that the parents, who are highly educated, have children with higher educational levels, regardless. The story of causation works through another way, parents with more education are in an improved position to assist their children by not only giving them motivation and encouragement but also providing resources. So, it is very important to distinguish between these scenarios particularly from a policy perspective.

There is a vast literature examining educational choices and the determinants of children's educational attainment. Recent research has shown an intergenerational correlation with respect to education, income and occupational status between present and previous generations. Typically, the studies in the US (Solon, 1999) and in the UK (Dearden et al., 1997) have found an intergenerational correlation between the earnings of fathers and sons of 0.40 and 0.60 respectively.

The present study is conducted in order to investigate the cause of this intergenerational correlation i.e. the established link between parents' and children's educational outcomes, as it is not only important for the evaluation of educational policy but for also designing policies reducing educational inequality. Particularly in Britain it is a very important issue as the recent government have planned to lower the number of children leaving school at 18. Young people now must remain in education or training until age 18. Since education is a main priority of governments, these results are important for policy-makers in order to design policies for those children who are at risk of under-achievement. Also findings that educational investment on present generations has a positive effect on future generations in terms of higher productivity plus non-economic social benefits for the society, indicating higher social returns to education, has a very important role in the cost-benefit analysis of educational investment.

There are two possible explanations of a positive intergenerational correlation, one is direct and the other is indirect. The first and direct cause states that it could be the result of the genetic transmission of ability such that talented parents have more able children. If this is the sole reason for the intergenerational relationship, then the issue of higher achievements amongst the future generations can be ignored when evaluating the educational policy of raising the present education levels since the inherited genes have not been affected in this situation. It might be the field of a genetic engineer.

The indirect cause of intergenerational correlation works through two routes. The first indirect route functions via the direct transfer of knowledge for example the more motivated and educated parents are in a better position to help and push their children as they have experienced the benefits of education themselves. The second indirect way works through income and lifestyle. It argues that more educated parents have higher incomes which can buy many things such as private schooling, books, tutors and an affluent neighbourhood.

Recent research has investigated whether the intergenerational link is causal and whether the link is due to nature (inherited genes) or nurture (upbringing). The main difficulty in sorting the intergenerational link into nature and nurture is separating the genetic effects and other characteristics explaining the educational outcome that might be transmitted from parents to children.

The present study investigates the causal mechanism underpinning this intergenerational correlation – vis-à-vis the established link between parents' and children's educational outcomes using the Longitudinal Study of Young People in England (LSYPE) dataset linked with the National Pupil Database (NPD). Using the raising of the school leaving age (RoSLA) implemented in 1972 as an instrument for parental education allows us to isolate the exogenous variation in parents' education. The OLS and IV results confirm that parental education is positively related to their children's educational outcomes, as measured by performance in GCSE examinations taken at age 16, and suggest that the effect of nurture (upbringing) is mainly responsible for the intergenerational relationship. Further results suggest that when controlling for both parents' education, mothers' education is positively related to children's education while the effects of fathers' education disappears after applying IV. These findings are robust across the boys and girls samples.

This chapter proceeds as follows: section 2.2 reviews previous literature, section 2.3 and 2.4 describe the data and methodology, section 2.5 presents the results and 2.6 presents robustness checks and section 2.7 then concludes.

2.2 Literature Review

2.2.1 Intergenerational Mobility in Education

Fewer studies have analyzed the intergenerational mobility in education, compared to income. The estimated elasticity for intergenerational mobility in education varies widely from 0.14 to 0.45 in the US (Mulligan, 1999) and 0.25 to 0.40 in the UK (Dearden et al., 1997). The estimates of the elasticity vary based on the data, methods used, and the specific outcomes being looked at.

The nature versus nurture argument is about the relative influence of an individual's innate attributes as opposed to the acquired attributes from social and environmental factors in which one is brought up. For example nature counts as the physical and personality traits determined by your genes which remain the same irrespective of where you were born and raised, while nurture is how you were brought up. The main identification issue is decomposing the total effect of parental schooling into nature and nurture effects. Three identification strategies exist in the literature: twin parents, adoptees children or instrumental variable (IV) studies.

(1) The twins approach holds genetic effects constant between twin parents as the genes passed on should be identical, so any observed difference in the relationship between parents' and children's schooling within twin pairs can be attributed to upbringing

effects. Behrman and Rosenzweig (2002) first applied this approach using Minnesota data on female and male twin pairs to difference out any intergenerational correlation attributable to genetics. Results from ordinary least square estimates, even after controlling for fathers' schooling and earnings, reveal large effects: an additional year of maternal schooling causes an increase in the children's years of education by 13% while the effect of fathers' schooling was approximately double (25%) that of mothers' schooling. However when they look within female identical twin pairs, thereby eliminating mothers' unobservable characteristics that are shared by twins and holding genetic ability constant, they find no impact of mothers' education on children's educational attainment , although the effect of fathers' education is still positive and significant.

One critical issue with such studies is that some part of the influence of mothers' education is still transmitted via her partner's genes due to assortative mating effects[1]. Obviously, this would not be an issue if the parents would have randomly met and married as in this case inclusion of the partners' schooling would have no impact on mobility estimates. But in the case of inclusion of the partner's schooling the mobility estimates measure the impact of increased parents' schooling on the children's schooling, net of the assortative mating effects.

However the above specification (assortative mating) depends upon the nature of the analyzed policy. If, for example, the policy makers are interested in raising the education of parents, they should not worry how it works either through assortative mating or not. But if they are interested in exploring the consequences of gender specific programs such as School Age Mothers[2] (SAM) Programme that are aiming to increase the schooling of mothers' but not fathers', they need to control for assortative mating effects and should include the fathers' and mothers' schooling simultaneously. In the context of the current study, there is another issue: in countries such as Pakistan and Bangladesh marriages/partnerships are not formed by the choice of women. They are 'arranged' by families, based on a wider set of characteristics which may/may not include education. So, the assortative mating argument is still applicable, as arranged

[1] Assortative mating occurs when individuals select partners non-randomly from within their population, on the basis of a trait that both they and their partners express, for example, more educated women in almost all societies marry more educated men, given own ability-schooling correlations.

[2] The aim of the School Age Mothers Programme is to support young women of school age, who are pregnant or parenting, to continue in compulsory education and beyond if this is appropriate.

marriages make assortative mating more likely. There is at least a chance marriage is random if based on attraction and love but not if arranged on the specific basis of observed characteristics. So assortative mating is more of an issue in countries like Pakistan, India and Bangladesh. Although the research in this chapter is not based on these countries, but still it controls for the assortative mating effect.

The finding of the Behrman and Rosenzweig (2002) study, that the impact of father's education via the upbringing route is more than that of the mother's education has been replicated in the literature of twin studies, in Scandinavian countries, by Holmlund et al. (2011) for Sweden and Pronzato (2012) for Norway, using both monozygotic (identical) and dizygotic (non-identical) twins. The study of Behrman and Rosenzweig (2002) is only about monozygotic (identical) twins, as it has to be, otherwise it would not be identical genes. Another study by Antonovics and Goldberger (2005), however, calls into question the results of the study of Behrman and Rosenzweig (2002), and suggests that these results are sensitive to educational measurement issues and coding of data. Behrman et al. (2005) replicate the original study with a larger Chinese dataset and find the same results as the previous Minnesota analysis. Moreover Bingley et al.'s (2009) study shows no correlation between mothers' schooling and children's educational attainment in the case of the identical twins.

Overall, there are numerous problems with twin studies. Firstly, they make measurement issues worse, since the method relies on differences between twins, many of which will be zero, and so measurement error will be a higher proportion of the differences than it would have been of the levels. Secondly, there will be a small sample size of twins and finally, there is concern of non random occurrence of different educational levels in two twins perhaps being due to twins' unobservable characteristics. The possibility that twins have different educational levels not by random but because of the difference between unobservable characteristics of twins again produces an unobserved ability bias.

(2) The second identification strategy to account for genetic effects compares natural born and adopted children, who share the same family environment but not their parents' genetic inheritance; therefore, any differences in educational attainment among children in the same family are driven by nature effects not by nurture. Sacerdote (2007) uses data on Korean American Adoptees, and reports a positive impact of

mothers' education on children's outcomes, after controlling for ability and assortative mating. In the economics literature, many researchers (Dearden et al., 1997; Bjorklund et al., 2006; Plug, 2004 & 2006; and Sacerdote, 2007) have estimated the intergenerational schooling effects using data on parents and their adoptees.

Bjorklund et al. (2006) uses a large sample of adoptees born between 1962 and 1966 in Sweden containing also information on adopted children's new siblings in adopted families and their biological parents as well. By using such data they are able to separate the genetic components through biological parents, and upbringing effects through adopted parents. The results depict that in the case of fathers' education, education works equally through genes and upbringing, while a genetic effect dominates in the case of maternal education. To put it simply, the education of both the natural father and adoptee father matter, but for the mother it is only the biological mother's education that matters.

Another study of adoptees undertaken by Plug (2004) using data from the American Wisconsin Longitudinal Study finds that parental education effects are not high but remain significant in the case of adoptees. Therefore the studies of Plug and Sacerdote highlight that nurture effects outweigh the nature effects. The adoptee studies have their own limitations, for example, small sample size; the methodology of these studies assumes that children are randomly allocated to new families and the adoption takes place at birth so that means adoptees spent no time with natural parents, which are not necessarily the case.

The adoptees studies (Wisconsin, Sweden, UK and other US states) always find positive and significant schooling effects of fathers' and mothers' schooling, provided that mothers' and fathers' schooling are included as separate regressors. Provided that these models are correctly specified, after allowing for the assortative mating effects, when fathers' and mothers' schooling are included simultaneously, they find that the father's schooling effect is bigger than that of mothers. Therefore, mostly the evidence in the literature gives support in favour of the argument: nurture effects are certainly important for a child's educational outcome. However these studies also indicate that the contribution of paternal schooling is bigger than maternal schooling. Despite the fact that adoptees studies take a different route for the effect to take by eliminating the genetic link between both parents and child, whereas, the twin approach differences out the genetic effect for only one parent, another angle of this debate suggests that, there is still a non-genetic effect transmitted from parents to children via parenting style, which leads to the correlation of parents' education with children's education. More to the point, small sample size even in a registry dataset and non-random placement of adoptees children limit the usefulness of this approach.

(3) The third and final identification strategy, IV methods, is based on natural experiments, and the one adopted in this study, uses educational policy reforms (e.g. raising of school leaving age, RoSLA) as an instrument, in order to isolate the exogenous variation in parents' schooling, without directly affecting the children. This approach separates the nature and nurture transmission factors since the variation exploited in parental education is orthogonal to the unobservables. Therefore, any association of parents' education on children's education remaining will be attributed to nurture effects only; as such variation will be orthogonal to genes. Chevalier (2004); Brown et al. (2011); Chevalier et al. (2013) for UK; Black et al. (2005) for Norway; Holmlund et al. (2011) for Sweden; Oreopoulous et al. (2006) for U.S., are the followers of this strategy.

Other researchers used different instruments depending upon the nature of study and data availability, such as: Brown et al. (2011) use age at which NCDS respondents start full time schooling determined by the Local Education Authority (LEA) policy, and Carneiro et al. (2007) use exogenous changes in the cost of education. All of these instruments find a positive correlation among parental education and children's education.

Brown et al. (2011) use the British National Child Development Study to contribute to the intergenerational literature by investigating the relationship between the ability test scores (literacy and numeracy) of parents and children. They find that parents' performance in reading and mathematics test scores is positively associated with the corresponding test scores of their children at a similar age. Further, the results of the study suggest that nurture effects are mainly responsible for the intergenerational correlation in literacy while inheritance is more important with respect to numeracy.

Black et al. (2005) found high correlations between parents' and children's schooling mainly because of selection not causation. In order to generate exogenous variation in parents' education that is independent of endowments, they use changes in compulsory schooling laws as an instrument, introduced in different Norwegian municipalities in the 1960s in which compulsory schooling increased from 7 to 9 years. Due to this reform some parents experienced two extra years of schooling who wanted to leave school at their first opportunity. They found a small but significant relationship between mothers' and sons' schooling and no significant relationship between mothers' and daughters' schooling or fathers' and sons' schooling.

Chevalier (2004) investigates the causal relationship between parental education and children's education using a change in SLA in Britain (British Family Resource Survey) over a period of 1994-2002 on the probability of staying in full time education at age 16-18 using a probit model. He initially finds a positive impact of parental schooling on child's educational outcomes: an additional year of parent's education increases the probability of staying on by 4 to 8 percentage points. However the study is limited because there is no cross-sectional variation in the British compulsory schooling law, as legislation was implemented nationwide. The larger the variation in compulsory schooling reforms, the more precise the estimates. For example, Black et al. using Norwegian reforms exploit a larger variation across municipalities (700 municipalities). So, using such a big source of municipality - variation, they arrive at more precise estimates. While considering the nationwide implementation of the law, it is possible that the changes in the law intermingle with the trend changes in parental income.

In order to investigate the discrepancies across methods Holmlund et al. (2008) applied three identification strategies: identical twins; adoptees; and instrumental variables to one particular Swedish data set. The results of their study are consistent with the results of previous studies. They found that the maternal effect is half the paternal effect in twin samples. On the contrary the opposite holds in the case of adoptees samples. Instrumental variable estimates give no significant paternal schooling effect but a quite large maternal effect. In addition to the above, they find non linearity in the effect of education indicating larger parental education effects at higher levels of education.

Finally a recent study by Dickson et al. (2016) studies the intergenerational mechanism between parents' and children's educational outcomes using the 1972 reform of school leaving age as an instrument in England and Wales using data from The Avon

Longitudinal Study of Parents and Children (ALSPAC) in the Avon[3] area. Parents' education is positively related to child educational outcomes at age 4 and continues until high stake exams taken at age 16. The parents who are affected by the reform gain results 0.1 standard deviations more than those parents who remain unaffected. The impact is stronger for the parents who are at the bottom of the educational distribution. He finds no difference across numeracy and literacy test scores.

The main criticism of using changes in the minimum school leaving age as an instrument is that it only provides the LATE[4] (Local Average Treatment Effect) estimates, certainly not comparable to OLS estimates, as the effect of this instrument on the population lying at the bottom of the schooling distribution is likely to be larger than at the top. It is due to the fact, that reform of school leaving age induce certain cohorts (with lower prior educational attainment) to increase their schooling as compared to previous cohorts. These changes are most likely to affect the proportion of people already at the margin of deciding whether to stay on or not.

On the other hand such estimates are worth noting for the educational economists who are particularly interested about early school leavers.

2.2.2 Intergenerational Mobility of Income

Researchers have also looked at the effect of parents' income on children's educational outcomes as well when studying determinants of educational outcomes. Many studies do find an important effect for family income. A broad literature is based on the intergenerational transmission of income in the United States. Solon (1999) finds the more compressed income distribution leads to smaller correlations between parents' and children's outcomes. Despite the fact that children who are nurtured in less favourable circumstances achieve lower qualifications, Carnerio and Heckman (2003) find parental income does not affect children's educational decisions, while parental education has a positive effect on children's outcomes. Similarly, Mayer (1997) finds

[3] This is former administrative area in South West of England, which includes the city of Bristol.

[4] The **average treatment effect (ATE)** is a measure used to compare treatments (or 'interventions) in randomized experiments, evaluation of policy interventions, and medical trials. The ATE measures the average causal difference in outcomes under the treatment and under the control. The Local Average Treatment Effect (LATE) is the average of the unit level causal effect for the compliers. The use of RoSLA as an instrument identifies a local average treatment effect, since it has only relevance for those who are affected by the RoSLA.

only modest and a sometimes negligible effect of parents' long-run income on children's educational attainment in Norway.

Chevalier (2004) finds that when fathers' income is added to the schooling choice equation it shows no significant impact on the parental educational estimates, even though income on its own has a positive and significant effect. Blanden and Gregg (2004) using UK data find a positive relationship between parental income and the child's educational outcome, although the study does not simultaneously provide the estimates for parental education. Loken (2010) using oil shocks as an instrument, finds no causal relationship between family income and children's education in Norway.

Chevalier et al. (2013) is notable for being one of the studies that control for both income and education as it distinguishes between the causal effects of parental income and parental education levels. The outcome variable in this study is a dummy variable defined as participation in post-compulsory education taken at age 16-18 estimated using a probit model. They use the Labour Force Survey (LFS) which is a quarterly survey of households in U.K. pooled it over the period 1993-2012. They use the 1972 reform of school leaving age and month of birth to instrument the parents' education. Least squares estimates reveal the same results consistent with previous literature, using IV methodology: larger effects of maternal schooling than paternal schooling. Income has a strong and significant impact on the children's educational attainment. Further, controlling for parental income, IV results using a dummy variable for reform of school leaving age as an instrument reinforce the role of mothers' education, particularly for daughters, whilst fathers' education has no significant impact on sons' or daughters' schooling.

In summary, there is consensus in the literature that these studies (twin, adoptees and IV) find that: more educated parents have more educated children because of higher education. So, there is not intergenerational mobility. If there is intergenerational mobility then one generation's outcomes do not depend on the pervious generation's outcomes. However it is uncertain, whether it is the education of mother, the education of fathers or the education of both parents that is crucial. In the same way it is unclear while measuring the total effect of parents' education whether it is the nature or nurture that is the decisive factor. The present study is conducted in order to find out the cause in the intergenerational literature mainly focusing on the nature and nurture effects of this mobility.

It is well documented that compulsory schooling laws are good instruments as natural experiments providing the involuntary increases in schooling for those cohorts who want to leave school at their first opportunity. They are therefore frequently used as instruments as they are exogenously driven irrespective of gender, ethnicity, income, education, location and timing.

Within the UK compared to Chevalier (2004) and Chevalier et al. (2013) the current study is using child education outcomes as attainment. Both the above studies are limited in their analysis of child education measures i.e. both studies use the decision to participate in post compulsory education as an outcome measure at age 16-18 which is just a yes or no pupil outcome variable. Compared to all other studies on intergenerational mobility the current study uses the parents having no qualifications as a measure of parents' education which is more relevant since such parents are the ones most likely to be affected by the raising of the school leaving age. Also no previous study uses the LSYPE (Longitudinal Study of Young People in England) linked with the National Pupil Database (NPD) to answer this question. There are the benefits of using this dataset relative to the other data sets that have already been used in the literature. For example, ALSPAC is limited to one specific region of the country; LFS and FRS are ongoing surveys, and cover the whole population, but they are not cohort surveys, and are unlikely to have sufficient numbers of observations on just young people. LSYPE is a cohort survey and represent the young people in England. Also due to its feature of matching the NPD it allows us to use richer dataset of pupils' education and schools variables which Chevalier (2004) and Chevalier et al. (2013) do not have.

Similar to Chevalier (2004); Chevalier et al. (2013); Dickson et al. (2016), the current study uses the RoSLA of 1972 in England and Wales as a source of exogenous variation in the parents' education.

2.3 Data

The analysis is carried out using data from the LSYPE (Longitudinal Study of Young People in England). The LSYPE is a cohort study of young people (born between 1, September, 1989 and 31, August, 1990) first observed in 2004 (wave 1) when they were aged between 13 and 14 in Year 9 (or equivalent) of schools in England. Using multi-stage stratified sampling LSYPE gathered information on 15,770 households in wave one (2004), 13,539 in wave two (2005), 12,439, in wave three (2006), 11,449 in wave four (2007), 10,430 in wave 5 (2008) and 9,779 in wave 6 (2009).

In LSYPE, every wave until four contains three types of questionnaires: family background, parental attitudes and young person. In the first four waves at least one parent/guardian was interviewed where possible along with the young person while from the fifth wave onwards interviews only took place with the young person. The analysis of this chapter uses the first wave of this study by merging three types of files, which includes the information about parental socio-economic status, personal characteristics, attitudes, experiences and behaviour, attainment in education, income and family environment and deprivation, the school(s) the young person attends and the young person's future plans. The unique feature of this data set is that we are able to match it with the NPD, which enables us to access information about the school level variables and exam results throughout schools at Key Stage4[5].

In the UK educational context the statistic of interest is the proportion of early school leavers which is considered a problem. Raising the school leaving age has been a priority of recent governments. For example, in England and Wales it has been increased numerous times since the introduction of compulsory Education Act in 1870. The RoSLA that occurred in 1972 extended the minimum school leaving age in England and Wales from 15 to 16. Further it was increased to 17 years in 2013 and to 18 years in 2015. Basically the RoSLA aimed to generate more skilled labour by providing an additional year of schooling to gain additional qualifications and skills.
Due to the Education Act of 1972 individuals born before September 1957 could leave school at age 15, on the other hand those born after this date had to stay for an additional year of schooling[6]. This RoSLA brings a discontinuity in the education attained by the parents of the LSYPE sample, so the RoSLA behaves as a regression

[5] The National Curriculum in all schools in England is divided into the following key stages:
Key Stage, Year, Age
KS 1, Foundation 1 2 , 5-7
KS 2, 3 4 5 6, 8-11
KS 3, 7 8 9, 12-14
KS 4, 10 11, 15-16
At the end of key Stage 4, all students sit for GCSE (General Certificate of Secondary Education) in a variety of subjects.

[6] It is worth noting that there was a strict compliance of the RoSLA (Harmon and Walker, 1995).

discontinuity and picks up the effects of the policy change. It can be seen from Figure 2.1 and Figure 2.2. There is a noticeable jump in the education of fathers' and mothers' born after the RoSLA was implemented (a lower rate of no qualifications).

Since not all parents are affected by ROSLA as the very specific sample used is of a cohort of 13-14 year olds in 2004, and their parents. I.e. the sample of parents is restricted to those who have a 13/14 year old child in year 9 in 2004. The average age of fathers and mothers in current sample is 44 and 41 years respectively. Parents with an age more than 46 could leave school at 15 while those aged 46 or younger had to stay school for another year, until they turn at least 16. So, the younger group will have an exogenous increase in their schooling due the act in 1972. In the current sample 65% of fathers and 81% mothers are affected by the reform, i.e. those who were born after 1957.

Table 2.1: **Frequency distributions of the instrumental variable for mothers' education**

RoSLA mother	Frequency	Percent
0	2,043	18.04
1	9,284	81.96
Total	11,327	100.00

Table 2.2: Frequency distributions of the instrumental variable for fathers' education

RoSLA father	Frequency	Percent
0	3,893	34.37
1	7,434	65.63
Total	11,327	100.00

The frequency distribution of the two instrumental variables used in this study as separate instruments for fathers' and mothers' education are given in Table 2.1 and Table 2.2. In our sample 81% of mothers are affected by this reform and 65% of fathers are affected. Due to this reform these fathers and mothers have to stay at schools for an additional year. The proportions of sampled individuals allowed to leave school at 15 are 18% for mothers and 34% for fathers i.e. those who were born before 1957. These tables show that the proportion of mothers affected by the raising of school leaving age (RoSLA), is greater than the proportion of fathers. The average age of fathers and mothers in the current dataset is 44 and 41 years respectively. Given the three years age difference in fathers' and mothers' age, it makes sense that the number of mothers affected by RoSLA is more than that of fathers.

In the current dataset we have a mixture of variables indicating parents such as: main parent/father/mother. In order to define them consistently, we use all 'father' and 'mother' variables, rather than the 'main parent' variables. We also run the equations separately for mothers and fathers, rather than using their respective instruments in the same equation, so as to make sure that the instrument is applying to the correct person (i.e. RoSLAd on father's education and RoSLAm on mother's education).

The formal qualifications for each parent are categorized into the following levels: 1 for no qualifications, 2 for less than five GCSEs or equivalent, 3 for five or more GCSEs or equivalent, 4 for A levels, 5 for Higher education below degree and 6 for Degree/higher degree.

The analysis uses two dependent variables to measure the educational outcomes of pupils. One is the total GCSE/GNVQ total point score, having a maximum value in the data set of 886 (see Table 2.4), GCSE points score awards 8 points for an A*, 7 points for an A, 6 points for a B etc. It is a way of summing total performance across all the GCSEs taken. The second, dependent variable is pass_ac, total number of GCSE/GNVQ qualifications at grades A*-C having a maximum at 19. Both are the qualifications taken at the end of compulsory schooling at the age of 16 and measured at the individual level in high stake exams.

Chevalier et al. (2013) use an indicator of attaining five or more GCSEs graded A to C (a standard measure of educational achievement in the UK). Achieving five or more GCSEs at grade C or above as an important cut off point. Most schools will not allow pupils to stay on to do A levels if they have not achieved at least five GCSEs at least grade C. It is not considered here as the main indicator of pupil success, since it is a simple yes/no binary indicator and discards a lot of information that is available from the NPD on performance in each exam (as Chevalier did not have this).

Turning now to the control variables another important indicator of socio-economic background, being in receipt of free school meals (FSM) provided to the children from low income families was tried initially, but found to be strongly correlated with parents' education measured by having no qualifications. Therefore, the log of annual family income is used.

Schools in England are mainly classified as: community schools, independent schools, foundation schools and voluntary schools. Community schools are run by the local authority that sets the admission criteria such as the catchment area, employs the staff in school and owns the land and building of schools. Pupils attending these schools have to follow the national curriculum. The independent schools are private schools that charge the fees towards the cost of running the school. The pupils do not necessarily follow the national curriculum and schools are maintained by a governing body instead of a local education authority and admission criteria and school policies are governed by the head teacher and governing body that supports the head teacher.

The description of variables used in the study is given in Table 2.3. Descriptive statistics for these variables used in equation (1) are shown in Table 2.4.

Table 2.3: Description of Variables

Variables	Description
Total GCSE point score (Dependent variable)	Total GCSE/GNVQ point score
Total number of GCSE passed (Dependent variable)	Total number of GCSE/GNVQ qualifications at grades A*-C
Female	1= female 0 = male
Ethnicity [Reference category = White]	1= White 2= Mixed 3= Indian 4= Pakistani 5= Bangladeshi 6= Otherasian 7= Black
Health problems	1 = individual has no disability 0 = otherwise
Household income	Log of household income
Household income missing	Household income missing
Parents meeting	Attend meeting of parents and teachers 1 = yes 0 = no
Father's education [Reference category = Father edu: degree]	1= Father edu: no qualifications 2= Father edu: less than five GCSEs or equivalent 3= Father edu: five or more GCSEs or equivalent 4= Father edu: A levels 5= Father edu: Higher education below degree level 6= Father edu: higher degree/degree
Mother's education [Reference category = Mother edu: degree]	1= Mother edu: no qualifications 2= Mother edu : less than five GCSEs or equivalent 3= Mother edu: five or more GCSEs or equivalent 4= Mother edu : A levels 5= Mother edu: Higher education below degree level 6= Mother edu: higher degree/degree

Average parental age	Average of father's age and mother's age
School pass rate	Proportion of pupils getting 5 or more GCSEs
Working status of parents [Reference category = Neither parents working]	1= single parent working 2= both parents working 3 = neither parent working
Help at home in studying	1= yes 0= no
Future aspirations	1= continuing education 0= otherwise
Number of siblings	Number of siblings of young person in household
School type [Reference category = community schools]	1= community schools 2= independent schools 3= foundation schools 4 = voluntary schools
House type [Reference category = rented house]	1= owned house 2 = rented house 3= others
Computer at home	Computer at home 1= yes 0 = no

Notes: The variable household income missing is to allow for the lower number of observations on this variable.

Table 2.4: Descriptive Statistics

Variable	Mean	S.D	Min.	Max.
Total GCSE point score	394.63	144.25	0	886
Number of GCSE passed	6.69	4.09	0	19
Female	0.49	0.49	0	1
Mixed	0.04	0.21	0	1
Indian	0.08	0.27	0	1
Pakistani	0.06	0.23	0	1
Bangladeshi	0.04	0.21	0	1
Other Asians	0.01	0.11	0	1
Black	0.05	0.21	0	1
Whites (base category)	0.72	0.45	0	1
Health status	0.12	0.33	0	1
Log of household annual income	7.52	4.43	0	13.12
Household Income missing	0.25	0.43	0	1
Parents meeting	0.28	0.45	0	1
Father no qualifications	0.24	0.42	0	1
Father less than five GCSEs	0.08	0.27	0	1
Father five or more GCSEs	0.15	0.36	0	1
Father A levels	0.27	0.44	0	1
Father Higher education	0.11	0.31	0	1
Father degree (base category)	0.15	0.36	0	1
Mother no qualifications	0.24	0.42	0	1
Mother less than five GCSEs	0.10	0.30	0	1
Mother five or more GCSEs	0.26	0.44	0	1
Mother A levels	0.15	0.36	0	1
Mother Higher education	0.13	0.33	0	1
Mother degree (base category)	0.11	0.32	0	1
Average parental age	43.22	5.70	28.5	65.5
Working status of single parent	0.25	0.43	0	1
Working status of both parent	0.67	0.47	0	1
Neither parent working (base category)	0.08	0.28	0	1
Help at home in studying	0.83	0.38	0	1
Future aspiration	0.80	0.40	0	1
No of siblings	1.67	1.19	0	11
Independent schools	0.05	0.21	0	1
Foundation schools	0.17	0.37	0	1
Voluntary schools	0.14	0.34	0	1
Community schools (base category)	0.65	0.48	0	1
Owned house	0.81	0.39	0	1
Rented house (base category)	0.18	0.38	0	1
Others	0.01	0.10	0	1
Computer at home	0.56	0.34	0	1
School pass rate	51.71	21.04	0	100
Number of Observations fathers' sample	8380			
Number of Observations mothers' sample	8387			

Notes: The data were cleaned for measurement error in fathers' age and mothers' age variables since there were some extreme ages, for-example there were a few parents having recorded age as 18 years and 97 years, who are unlikely to have 13/14 years old children. These parents are dropped from the analysis. They are probably data entry mistakes. Further, due to missing values on some of other control variables and having a common sample across all three specifications within fathers' and mothers' equations, the current analysis arrive at 8380 observations for fathers and 8387 for mothers.

2.4 Methodology

This chapter is concerned with answering the question: What is the effect of parents' education on their children's education using OLS and IV methods, the latter accounting for the endogeneity bias? Initially, an OLS methodology is adopted. The OLS procedure includes all the controls for child, family and school characteristics.

The specific issue in this chapter is that the parental education variable could 'pick up' the effect of both parents' genes and parents' parenting style. IV is used purely to distinguish between these two 'causes' of the intergenerational correlation. So the IV is used to identify variation in parental education that could not possibly be correlated with genes, so if it still affects children's outcomes, then it must be working via the upbringing route. So if the analysis finds an exogenous increase in parents' education, it can only be correlated with (and so pick up) the effect of things that happen after that education (upbringing of their children) and not things that happen before that education (genetic make-up of the parents).

The RoSLA of the school leaving age of 1972, which raised the school leaving age from 15 to 16 years, serves as a source of exogenous variation that is used as an instrument. Our empirical model is summarized as following:

$$\mathbf{Y}_i^{\mathbf{child}} = \beta_0 + \beta_1 X_i^{\mathbf{child}} + \beta_2 X_i^{\mathbf{family}} + \beta_3 X_i^{\mathbf{school}} + \beta_4 X_i^{\mathbf{peer}} + \beta_5 \, \mathrm{FEdu}_i / \beta_6 \, \mathrm{MEdu}_i + \varepsilon_i \qquad (1)$$

Where $i = 1.... n$ denotes the child.

$Y_i^{\mathbf{child}}$ = Student academic achievement in school (Total GCSE points score)
$X_i^{\mathbf{child}}$ = Vector describing characteristics of the child
$X_i^{\mathbf{family}}$ = Vector describing characteristics of the families
$X_i^{\mathbf{school}}$ = Vector describing characteristics of the school
$X_i^{\mathbf{peer}}$ = Vector capturing peer effects
ε_i = Error term

In equation (1) $\mathbf{Y}_i^{\mathbf{child}}$, is the children's educational outcomes, as measured by performance in GCSE examinations taken at age 16,. $X_i^{\mathbf{child}}$ is a vector containing variables gender, ethnicity, health status and future aspirations, $X_i^{\mathbf{family}}$ indicates log of household income, whether parents attend school meetings, parental age, working status of parents (i.e. whether one, both or neither parents are working), home owned / rented, help at home in studying and number of siblings, X_i^{school} includes school type (independent / foundation / voluntary), X_i^{peer} indicates the proportion of pupils in the respondent's school who get 5 or more GCSEs. So it is a school level variable, and so can be used as an explanatory variable to pick up peer effects. $FEdu_i$ and $MEdu_i$ record the father's and mother's qualifications respectively. The ε_i is the error term which represents the effects of all other determinants of performance including the unobservable attributes of the child.

Our interest lies in ascertaining whether there exists a positive relationship between children's educational outcomes and parents' qualifications, that is, whether β_5 and $\beta_6 > 0$ and also to know the cause of any such positive relationship. Three possible reasons will be considered. The children who have parents with higher qualifications could benefit through genes, through higher income or through other factors related to high skill level via upbringing. Three versions of equation (1) are estimated. In the first, the only control variables are gender, ethnicity, health status, and numbers of siblings as all of these variables are the most exogenous variables in the model. This specification therefore estimates the raw (least conditioned) intergenerational coefficients between parents' and children's education.

The second specification adds different school control variables namely school type, and proportion of children getting 5 or more GCSEs.

The third specification includes family income variables such as whether the family owns their house or lives in rented accommodation, log of household income and working status of parents. In this way comparing the intergenerational coefficients before and after the inclusion of school related and family income variables will indicate whether this intergenerational relationship exists through the latter variables or there is no effect on intergenerational coefficients after controlling for these schools and family income attributes.

The final specification is estimated using two stage least square (2SLS), where equation (2) and (3) serve as a first stage in which the RoSLA of school leaving age is used as an instrumental variable, and then replacing fathers' and mothers' education in the second-stage regression (based on equation 1) with their predicted values, $\widehat{F}Edu_i$ and $\widehat{M}Edu_i$ based on estimating equations (2) and (3).

$$FEdu_i = \alpha 0 + \alpha_1 RoSLA^{D} + \alpha_2 X_i^{c} + \alpha_3 X_i^{f} + \alpha_4 X_i^{s} + \alpha_5 X_i^{p} + \varepsilon_{i2} \quad (2)$$

$$MEdu_i = \gamma 0 + \gamma_1 RoSLA^{M} + \gamma_2 X_i^{c} + \gamma_3 X_i^{f} + \gamma_4 X_i^{s} + \gamma_5 X_i^{p} + \varepsilon_{i3} \quad (3)$$

The estimate of β_5 and β_6 therefore estimates, conditional on covariates, the effect of parents' education on child schooling using only the part of the variation in parent's education caused by the RoSLA. This strategy is the one we apply in this study and similar to other studies (Oreopoulos et al., 2003; 2006).

In equation (2) and (3) RoSLA is the dummy variable, which takes the value of one if the individuals were affected by the RoSLA, and zero otherwise. $RoSLA^{\boldsymbol{D}}$ and $RoSLA^{\boldsymbol{M}}$ are the dummies for the fathers' and mothers' respectively. In our dataset due to the RoSLA of 1972, parents born before September 1957 could leave school at 15, while the parents who were born after September 1957, are affected by this RoSLA as a result of which, they had to stay at school for an additional year. In this way the education act of 1972 brings a discontinuity in the education obtained by the parents. The use of this RoSLA can isolate the exogenous variation in parents' education.

With the instrumental variable (IV) technique, the identification depends on the quality of the instrument. In order to obtain consistent estimates of β_5 using two stage least square (2SLS) on equation (1) and (2), two assumptions must be fulfilled by the instrument: 1) RoSLA has to be correlated with parents' education and 2) RoSLA has to be uncorrelated with the error term (ε_{i1}).

So in order to identify the nurture effect, an instrumental variables approach is used. The selection process of an instrumental variable is not straightforward as, especially in small samples, IV estimation might produce biased estimates. Also, there is an additional problem of choosing a weak instrument even when benefitting from a large sample. The first step followed in order to choose a suitable instrumental variable was to test the validity of the first assumption, known as the test of relevance. In the first stage regression results it can be seen that in all models and samples the coefficient on the instrument is significant.

In order to test whether there is a weak instrument problem or not, two approaches are considered. Baum et al. (2007) and Stock et al. (2002) suggest that an F statistic in the first stage regression that exceeds 10 may be deemed reliable when one endogenous regressor exists. In addition the Cragg-Donald F-statistic must exceed the critical values, which were tabulated by Stock and Yogo (2005) for the first-stage F-statistic to test whether instruments are weak. From Table 2.5, it may be seen that for the sample of mothers, the F-statistic is continuously over 10 for each of the outcome measures. For these samples, the instruments are seen as reliable and valid.

It should be noted that the F-statistic to test for joint significance of the coefficients in the first stage regression is always found to be significant thus the instruments have significant explanatory power for fathers' education and mothers' education once controlling for other exogenous variables. It may be argued that although the instrument performs well for mothers' and fathers' case. However, for the fathers' sample, only in specification 3, the F-statistic is just the borderline at 10.74.This will be considered when evaluating the results.

To summarise, the instrument seems to perform well under the testing procedure and indicates validity, relevance and in most cases does not show any signs of the weak instrument problem except in model 3 in the fathers' sample. The instrument does not seems to indicate complete weakness for the fathers' sample so it is still worth comparing the IV results with the OLS results for this sample.

Table 2.5: Instrumental variable testing

	Fathers' education	**Mothers' education**
Specification 1	F-value = 23.09	F-value = 62.08
Specification 2	F-value = 22.87	F-value = 55.39
Specification 3	F-value = 10.74	F-value = 38.84

Notes: Specification 1: Child characteristics.
Specification 2: Child characteristics and school characteristics.
Specification 3: Child characteristics, school characteristics and family characteristics.

In order to measure fathers' and mothers' education, dummy variables were used indicating if the father and mother have no qualifications. It can be seen from Figures 2.1 and 2.2 given below that the RoSLA brings a shift in the scatter plot for both fathers' education and mothers' education when measured by no qualifications. Other measures of fathers' and mothers' education have been tried, but did not appear to be as closely related to parents' age with a break in the series around the time of the RoSLA (see appendix, Figure A2.1 to A2.6).

Figure 2.1: Relationship between Instrument (RoSLA) and Proportion of Fathers' Education (with no qualifications) by Age

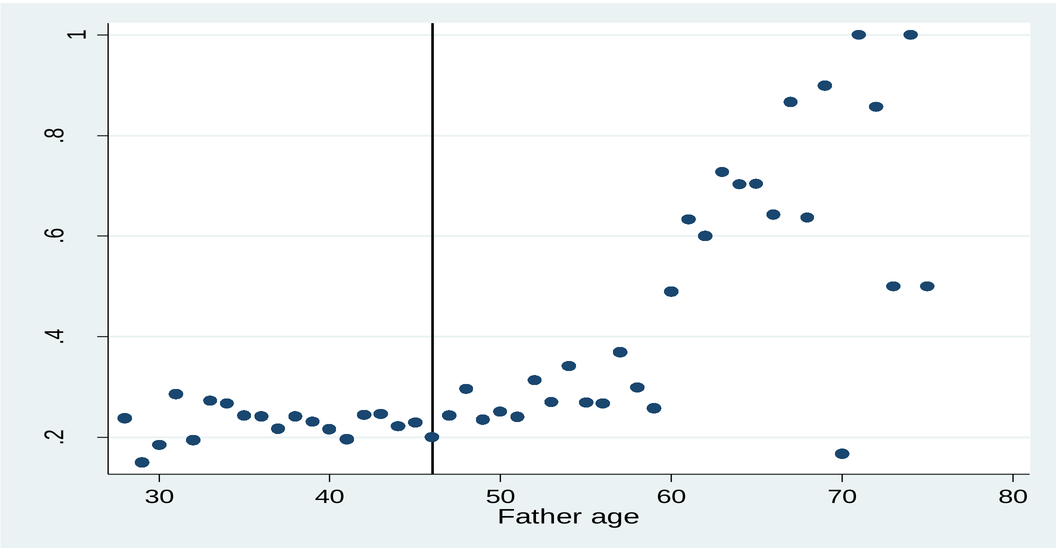
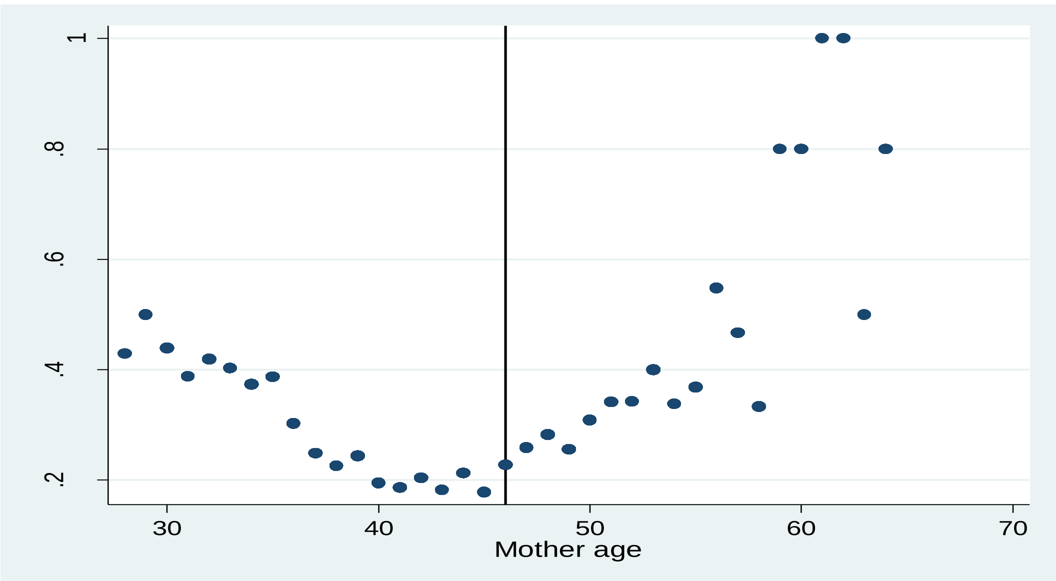


Figure 2.2: Relationship between Instrument (RoSLA) and Proportion of Mothers' Education (with no qualifications) by Age



However the effect of the RoSLA is not identical for fathers and mothers. Figure 2.1 suggests that for fathers the relationship is more flat around the RoSLA time relative to mothers. The cut-off age is 46 years for fathers and mothers as the RoSLA affected anyone born after 1st September, 1957. LSYPE was first undertaken in 2004. So by 2004, those affected would be 46 (or younger) as the survey was earlier in the year than September. So the cut-off age is 46. Any parent aged 46 or younger will have been affected by the RoSLA in 1972 and will have had to stay in school until at least 16. Any parent aged above 46 could have left school at 15. So the younger group will have an exogenous increase in their education, due to the RoSLA. We focused on the group of parents who are most likely to be affected by the RoSLA, these are the people who have no qualifications (Black et al, 2003).

Figures 2.1 and 2.2 show the relationship between the instrument and the proportions of fathers and mothers with no qualifications, by age. The figures show that those individuals who left school before the RoSLA are more likely to have no qualifications. It is obvious from both graphs that the RoSLA of 1972 has created a discontinuity (break in the series around the time of the RoSLA) and captures the effect of the policy change. It is near the current age of those parents affected by RoSLA (age 46). This makes sense, since the RoSLA is likely to affect people at the bottom of education distribution the most.

All the parents are not affected by the RoSLA since current sample is restricted to only those parents having children of age 13/14 years in 2004. Therefore, this is a very specific sample with some unusual patterns emerge in these graphs, Figures 2.1 and 2.2. Further, these graphs seem to suggest that for the fathers the proportion with no qualifications does start increasing around the mid-forties, though it is quite slowly at first, and it only really takes off after the mid-50s. For mothers, there is a clearer rise in the proportion with no qualifications after the mid forties. It is true that the proportion with no qualifications rises again amongst the younger mothers aged less than mid-30s, but this is not surprising as if they are a mother of a 13/14 year old at this age, then they had their child around the age of 20. Girls who have children this young are likely to have dropped out early from education.

Regarding the other graphs in the appendix, the story across all of these other graphs remains that there does not seem to be as clear a break at age 46 for these other education categories, which is why the current study focuses on no qualifications as the main variable in the analysis. The fact that in current study we have a very particular sample of parents (parents of 13/14 year olds in 2004), this can also explain the unusual patterns on some of the other figures e.g. Figure A2.2 - the one for whether they hold a degree or not, in the same way as explained here the high proportion with no qualifications amongst young mothers.

All three specifications estimated for fathers and mothers control for average parental age as parents' age could have a differential impact while nurturing their children. Trillingsgaard and Sommer (2016) explore the association between maternal age and children's socioeconomic development measured as behavioural, social and emotional difficulties at age 7, 11, and 15 and found that being an older mother is positively related to her child's socioeconomic development. Results of this study highlight that patience, maturity, and mental flexibility that come along with age give a better parenting style. So it is important to control for parents' age. In the current study the effect of average parental age is positive and strongly significant on children's total points score in all specifications. The effect of this variable is strong to an extent that if models are estimated without these, the effects of parents' education on children's education are peculiar i.e., become negatively significant. This is due the correlation between parent education and parent age.

While estimating fathers' and mothers' specifications, the possibility of controlling for their respective age was considered but the resulting IV estimates were meaningless due to the high correlation as parents' age is strongly related to the RoSLA instrument, since whether they are affected by the RoSLA is defined by their age. The correlation between the RoSLA instrument for the mother and mother's age and between the RoSLA instrument for the father and father's age is 0.71 and 0.76 respectively. Therefore the average age of both parents is used since this picks up the general age effect of parents without being too strongly correlated with the RoSLA variable for one particular parent.

Similar to other studies on intergenerational mobility in education, the current study faces the issue of high standard errors on coefficients on parents' education. Although in the current study the standard errors on the instrumented parental education are high but still most of the IV results are significant. Silles (2011) finds very high standard errors on parents' education leading to the IV estimates becoming insignificant on parents' education.

2.5 Results

2.5.1 OLS Results

Beyond this point, we use fathers' education and mothers' education measured by fathers and mothers having no qualifications respectively, remembering that the RoSLA has bite at the bottom of the education distribution. Throughout the chapter the coefficient on father's and mother's educations only are presented in all tables in the interest of brevity.

The raw intergenerational education coefficient, controlling for child characteristics only, is -74.85 for fathers' education and -76.61 for mothers' education as shown in Table 2.6. This means that if the father and mother have no qualifications, it will lead to a 74 point and 76 point decrease in children's total points score respectively.

Table 2.6: Intergenerational coefficients on fathers' and mothers' education: OLS Dependent variable is children's total points score

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification l Pupil characteristics only	-74.850*** (3.803)	-76.607*** (4.072)
No. of observations R^2	8380 0.12	8387 0.12
Specification 2 Pupil and school characteristics	-54.475*** (3.626)	-51.859*** (3.910)
No. of observations R^2	8380 0.23	8387 0.22
Specification 3 Pupil, school & family characteristics	-33.122*** (3.490)	-27.442*** (3.825)
No. of observations R^2	8380 0.33	8387 0.33

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Specification 2 includes numerous control variables for the school of the child, and the values of the intergenerational education coefficients of -54.47 and -51.86 show that these control variables have a clear impact on the intergenerational coefficients for fathers' and mothers' education. Thus, the effect of fathers' and mothers' education on children's performance in GCSE exams is partly being transmitted through these school control variables, that is better educated parents sending their children to better schools.

The third specification adds various family control variables, for example, household income, working status of parents, house tenure, parents' meeting at schools etc. The results from this specification (with coefficients of -33.12 for fathers' education and -27.44 for mothers' education) show that the effect of fathers' and mothers' education also exists through these latter variables. Thus, the source of the correlation between parents' education and child performance occurs at least in part through the school and family variables i.e. parents with no qualifications have lower-achieving children, partly due to a lower income and other adverse upbringing effects.

2.5.2 IV Results

The IV method isolates random variation in fathers' education and mothers' education due to the RoSLA, raising of the school leaving age, as this cannot be transmitted genetically. Two- Stage Least Squares is used since a linear equation is estimated in the second stage. Again three specifications are estimated. The first one is just controlling for child characteristics only, the second adds numerous control variables for the school of the child, and the third specification includes various family control variables. Results from the IV method are higher than the corresponding OLS estimates which could be downward biased due to measurement error as measurement error biases towards zero. The coefficients on fathers' education and mothers' education are reported in Table 2.7.

Table 2.7: Intergenerational Coefficients on parents' education: IV Results Dependent variable is children's total points score.

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 4	-223.740***	-161.464***
Pupil characteristics only	(78.799)	(48.650)
No. of observations	8380	8387
F	23.09	62.08
Specification 5	-197.383***	-121.216***
Pupil and school characteristics	(75.539)	(49.077)
No. of observations	8380	8387
F	22.87	55.39
Specification 6	-128.426	-103.912*
Pupil, school & family characteristics	(101.501)	(57.497)
No. of observations	8380	8387
F	10.74	38.40

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

As found in the existing literature, the IV estimates of fathers' education and mothers' education are larger than the corresponding OLS estimates. The IV strategy isolates the exogenous variation in fathers' and mothers' education due to the effect of the 1972 RoSLA and as such could not be passed on genetically to the children because such variation will be orthogonal to genes so any established link between parents and children education can be attributed to nurture effects only.

The IV results, therefore, shows that having removed any genetic effect, there is an intergenerational relationship between parents' education and children's performance, suggesting that the source of relationship is not a genetic effect. The results further suggest that the effect of nurture (upbringing) is mainly responsible for the intergenerational relationship and rejects the idea that genetic effects are the dominant source of the intergenerational relationship. One possible reason for the fact that the IV coefficients not only lose none of their value compared to the OLS coefficients but are actually considerably larger is that our instrument is correcting for measurement error, that is, measurement error in the education variable causes the OLS estimates to be downward biased, but this is corrected by the IV procedure. However, it is also consistent with a LATE interpretation, as the parents who are affected by the RoSLA are most likely to have a lower level of educational qualifications as compared to the average parents and an increase in the education for parents at the lower end of the distribution may have more effect on their children then a further increase in education for parents already at the top of the education distribution. These results are consistent with the study of Chevalier (2004).

Within this chapter the standard errors on IV estimates[7] are found to be higher, as studies by Silles (2011), Black et al. (2005) and Chevalier et al. (2013) also found higher standard errors on IV estimates. As a consequence of this they found insignificant IV estimates for fathers' and mothers' education. Compared to them, even though the standard errors are high, the current study finds most of the IV estimates are significant.

From marriage market effects it is known that men and women tend to marry assortatively by qualification i.e. individuals who hold a CSE/O-level are much more likely to be married to someone with a qualification than is an unqualified individual. They also tend to marry with a couple of years' age gap, with the man being a bit older.

This raises the question of whether the positive correlation in partners' qualifications is partially a causal relationship: does holding a qualification make one more likely to marry someone with a qualification. The evidence suggests yes (Anderberg and Zhu, 2014). That would then make the partner's qualification dependent on your own qualification which is a problem when we run regressions on e.g. only the mother or only the father.

It is also not clear how this works over the RoSLA: consider the women born in 1957 who have more qualifications due to the RoSLA. These women are now more likely to marry more qualified males. By the typical husband wife age gap, these first-affected women would tend to marry pre-RoSLA men. Hence their higher qualification may not have been associated with a higher rate of qualification among their husbands (but possibly better "quality" on other dimensions – ability, social background etc.). Hence the effect of a woman being RoSLA affected may be low since her qualification was not associated with higher qualifications for her husband.

[7] Results are checked after dropping these variables (parental aspiration and study support); the results remain qualitatively the same as the previous results.

In contrast, consider the first RoSLA affected men: they would typically marry women born around 1960 (that is, into the post-RoSLA period). For these first-RoSLA affected men, a qualification may well have ensured a qualified wife. Hence being RoSLA affected may have increased the spouse qualification more for men than for women.

These effects may affect the validity of the IV approach, and may help account for the asymmetric findings for mother/father.

In the present study, the standard errors in the fathers' education equation are bigger than in the mothers' education. It suggests that there is more noise in the IV estimates for fathers' education relative to mothers' education. It could be due to one of these things explained above.

However, as far as the interpretation of the IV estimates is concerned, caution should be taken to account for the LATE (Local Average Treatment Effect) nature of these estimates, due to the fact that the IV estimate is specific to those affected by the instrument, RoSLA. This implies that the effects of education reform is not homogenous for all the parents, instead it is relevant to those parents who have a lower level of qualifications and lower preference for education (parents who wished to leave school at their earliest opportunity i.e. at 15), hence lying at the bottom of the education distribution. There is no effect for parents who are higher in the education distribution. This is consistent with (Chevalier, 2004; Imbens and Angrist, 1994)

2.6 Robustness checks

The identification strategy assumes that due to the RoSLA, there will be an increase in the amount of schooling of those parents who are affected by the RoSLA. In this section a number of modifications are made to the estimated relationships in the previous section, in order to assess the robustness of the results.

2.6.1 Restricted Sample around the RoSLA

The results for robustness checks are presented in the table below. In this study our identification strategy assumes that the RoSLA has increased the parents' schooling. However, there is a possibility that the RoSLA has no identifying power and results might be caused by the unobservable differences between those affected and unaffected by the RoSLA and cohort effects, although we have controlled for parents' age. For example, current 60 years old and 30 years old parents may have had large differences in parenting their children as compared with the differences between the current 40 years and 50 years old parents. For this reason, we restrict the sample to those parents in the close vicinity of the RoSLA (born five years before and after the RoSLA) to capture the treatment effect as tightly as possible. The frequency distribution of mothers and fathers affected by RoSLA for the restricted sample is given below in Table 2.8 and 2.9.

Table 2.8: Frequency distributions of the instrumental variable for mothers' education in restricted sample

RoSLA mother	Frequency	Percent
0	1,381	28.14
1	3,527	71.86
Total	4,908	100.00

Table 2.9: Frequency distributions of the instrumental variable for fathers' education in restricted sample

RoSLA dad	Frequency	Percent
0	2,009	34.91
1	3,745	65.09
Total	5,754	100.00

In choosing the size of window there is a trade off between comparing parents born just after and before the RoSLA and increasing the sample size by allowing a wider window on each side of the discontinuity. The former reduces any bias resulting from the treatment effect when moving further away from the time of the policy change and the latter improves the precision of the estimates. I chose the 5 years before and after the RoSLA to get a reasonable sample size so that precision of estimates is not sacrificed. This restricted sample therefore controls for these differences in unobserved characteristics by comparing people of similar age who were affected and unaffected by the RoSLA, thus enabling us to make a fairer comparison of parents.

Comparing the results in Table 2.10 to Table 2.6 shows that estimating the sample in this way has not greatly affected the OLS results. Note that to have a common sample across fathers' and mothers' specifications, both parents would need to be in the 5 year age window. For any couple with a larger age difference, one parent could be outside the window, and so both would be dropped when using a consistent sample. A consistent sample for the 5 year window robustness check and in other results is therefore not used.

Table 2.10: **OLS Results for 5 years restricted sample around the RoSLA Dependent variable is children's total points score.**

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 7 Pupil characteristics only	-83.955*** (5.275)	-79.185*** (6.372)
No. of observations R^2	4362 0.11	3784 0.09
Specification 8 Pupil and school characteristics	-63.422*** (5.047)	-51.008*** (6.105)
No. of observations R^2	4362 0.21	3784 0.21
Specification 9 Pupil, school & family characteristics	-39.989*** (4.880)	-25.263*** (6.00)
No. of observations R^2	4362 0.32	3784 0.31

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Looking at the IV results for the full set of controls in Table 2.11, and comparing them to the IV results in Table 2.7, the restricted results are similar, for fathers.

Table 2.11: **IV Results for 5 years Restricted Sample around the RoSLA Dependent variable is children's total points score.**

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification l0 Pupil characteristics only	-201.655** (87.762)	25.236 (117.305)
No. of observations F	4362 17.54	3784 11.92
Specification 11 Pupil and school characteristics	-199.114*** (84.711)	-3.407 (104.106)
No. of observations F	4362 18.00	3784 13.15
Specification 12 Pupil, school & family characteristics	-129.777 (110.907)	-71.316 (86.690)
No. of observations F	4362 9.02	3784 18.26

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

However, the results for mothers in the restricted sample are radically different as they become insignificant, but similar results have been found before in the literature - e.g. a famous study by Behrman and Rosenzweig (2002) who found a significant effect of mothers' education in their OLS equation, but insignificant for their within-twin pair estimates (the latter blocking off the genetic effect and so isolating the upbringing effect).

Another study by Silles (2011) found the instrumental variable estimates are not sufficiently precise to find that either parent's schooling has a beneficial effect on children's cognitive and non- cognitive development. In Britain Chevalier et al. (2013) found an insignificant IV effect of maternal education on their children's decision to continue in education, using the 1972 RoSLA.

These results confirm the validity of the identification strategy in that differences in the ages of parents affected by the RoSLA do not seem to be driving the results in case of fathers. Thus, the RoSLA has created an exogenous variation (increase) in fathers' education.

However, caution should be taken in the interpretation of the IV models. As in the present study the results for mothers are not robust for the restricted sample for five years around the RoSLA. It could be possible that the true upbringing effect for mothers is zero, and this analysis is accurately estimating that. But it should be noted that there is still some intergenerational effect coming through the other parent (father).
However the instrumental variable estimates can only be interpreted as Local Average Treatment Effect (LATE) – as the RoSLA has not a homogenous effect on the post-RoSLA cohorts (as shown in the figures A2.1 to A2.6 in appendices). My results are conditional for being of stable homes and both parents.

2.6.2 Using number of GCSEs passed as a dependent variable

In this part again three specifications are estimated by OLS and IV, using number of GCSEs passed as another dependent variable. Although the total points score is a better measure of student outcome as it has a lot of variation, but an employer may be interested in knowing the number of GCSEs passed. The results given in Tables 2.12 and 2.13 remain qualitatively the same as in the previous case when using total points score as a dependent variable.

Table 2.12: Intergenerational coefficients on fathers' and mothers' education using OLS.
Dependent variable is children's number of GCSE passed.

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 13 Pupil characteristics only	-2.069*** (0.108)	-2.261*** (0.115)
No. of observations R^2	8380 0.11	8387 0.12
Specification 14 Pupil and school characteristics	-1.517*** (0.104)	-1.597*** (0.111)
No. of observations R^2	8380 0.21	8387 0.21
Specification 15 Pupil, school & family characteristics	-0.902*** (0.098)	-0.911*** (0.109)
No. of observations R^2	8380 0.32	8387 0.32

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Table 2.13: Intergenerational Coefficients on parents' education: IV Results
Dependent variable is children's number of GCSE passed.

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 16 Pupil characteristics only	-6.656*** (2.274)	-4.091*** (1.366)
No. of observations F	8380 23.09	8387 62.08
Specification 17 Pupil and school characteristics	-6.015*** (2.201)	-3.043*** (1.392)
No. of observations F	8380 22.87	8387 55.39
Specification 18 Pupil, school & family characteristics	-4.236 (2.959)	-2.473* (1.621)
No. of observations F	8380 10.74	8387 38.84

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

2.6.3 Controlling for the education of both parents

So far we have looked at the effect of father's education and mother's education on children's education independently using their respective instruments. However one potential criticism could be that these estimates measure the direct effects of each parent's education including the indirect effects coming through the assortative mating.

The current data reveals that 24% of mothers and 24% of fathers have education measured as no qualifications. Given that parents have identical schooling levels and due to potential correlations[8] between each other's endowments and schooling due to non-random marital sorting, then this may lead to upward biased estimates.

Therefore it is important to consider the intergenerational effect of the partner's schooling. Due to the endogenous nature of father's education and mother's education, we used the appropriate gender instrument for each.

When including both parents' education to isolate the direct effect of each parent's education from indirect effects coming through assortative mating effects, we found an interesting results: the mother's education effect outweighs the father's education effect. The coefficients on mother's education are significant in all three IV specifications but father's education has insignificant coefficients, once controlling for mother's education in all three specifications. The IV results are given in Table 2.14.

Table 2.14: Intergenerational Coefficients on parents' education: IV Results Dependent variable is children's total points score. Controlling for both parents' education: full sample

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 19 Pupil characteristics only	-96.817 (74.230)	-176.464*** (46.768)
No. of observations F	8367 12.61	
Specification 20 Pupil and school characteristics	-94.484 (68.624)	-143.352*** (46.650)
No. of observations F	8367 13.25	
Specification 21 Pupil, school & family characteristics	-62.290 (85.877)	-135.962** (63.755)
No. of observations F	8367 6.15	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

[8] These correlations arise from the fact that women with better levels of schooling tend to have children with better educated men who may also be better endowed.

In all specifications, the impact of mothers' education is larger in terms of magnitudes and significance than fathers' education (the latter being insignificant in all three IV specifications) as also found by Black et al. (2005). It could be due to the reason that children spend more time with their mother than their fathers.

Another interesting point in the first stage regressions, when both of the RoSLA instruments are used, the appropriate gender one (i.e. RoSLAm for mothers and RoSLAf for fathers) is the one with the largest coefficient, and often is the significant one while the inappropriate-gender instrument gets an insignificant coefficient. This is re-assuring, and suggests the instruments used in this study contain real information.

Another concern while using both parents' education simultaneously is a potential high correlation between instruments used together in the same specification. Contrary to Chevalier et al.'s (2013) study who found a high correlation of 0.67 between the RoSLA instrument of father and mother, the correlation between the instruments for father and mother in the current study is 0.51, which is not as high compared to their study. So in this particular model while using both parents' education simultaneously, it seems less likely that results are driven by the high correlation between instruments.

Further, it is interesting to explore further the effect of both parents' education on sons and daughters separately. And further, these IV results remain robust across specifications in most cases when splitting the sample for boys and girls. Results are given in Tables 2.15 and 2.16. These results confirm that mothers' education has a bigger impact for both sons and daughters. This finding is contrary to Chevalier et al. (2013) who found strong IV effects of paternal education for daughters.

Table 2.15: Intergenerational Coefficients on parents' education: IV Results
Dependent variable is children's total points score.
Controlling for both parents' education: Girls sample

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 22 Pupil characteristics only	-188.343 (146.253)	-206.707*** (74.362)
No. of observations F	4114 3.37	
Specification 23 Pupil and school characteristics	-162.125 (135.351)	-186.617*** (75.350)
No. of observations F	4114 3.50	
Specification 24 Pupil, school & family characteristics	-189.637 (260.025)	-216.912 (165.483)
No. of observations F	4114 0.86	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Table 2.16: Intergenerational Coefficients on parents' education: IV Results
Dependent variable is children's total points score.
Controlling for both parents' education: Boys sample

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 25 Pupil characteristics only	-46.962 (86.638)	-178.579*** (68.355)
No. of observations F	4253 9.76	
Specification 26 Pupil and school characteristics	-63.129 (78.804)	-125.567*** (65.877)
No. of observations F	4253 10.16	
Specification 27 Pupil, school & family characteristics	-25.979 (83.194)	-113.122 (79.699)
No. of observations F	4253 6.95	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

2.6.4 Highly educated

The 1972 education act increased the compulsory schooling from age 15 to age 16. As previously explained in the chapter, as a result of this RoSLA, one should expect a small effect of the RoSLA on education attainment of those parents who are highly educated.

To verify this, we estimated the IV models on the restricted sample of those individuals who have highest qualifications as degree/high degree.

The results given in Table 2.17 show that the first stage has no predictive power and the IV results are peculiar, with the coefficients on parents' education being insignificant. This also supports that the argument that no qualifications used as in the main analysis as a measure of parents' education is appropriate. The huge standard errors are a sign of miss-specification.

The results confirm the previous finding that the effect of an exogenous increase in education via the raising of the school leaving age is confined to the lower part of education distribution, and that there is no effect of the RoSLA further up the education distribution (Imbens and Angrist, 1994).

Table 2.17: Intergenerational Coefficients on parents' education: IV Results Dependent variable is children's total points score.

	Fathers' Education (Degree)	**Mothers' Education (Degree)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 28 Pupil characteristics only	-4507.704 (17559.840)	-824.309 (562.948)
No. of observations F	8380 0.07	8387 3.22
Specification 29 Pupil and school characteristics	-2726.956 (7296.008)	-499.371 (331.595)
No. of observations F	8380 0.15	8387 4.41
Specification 30 Pupil, school & family characteristics	-402.220 (495.107)	-281.402 (196.096)
No. of observations F	8380 1.32	8387 7.24

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

2.6.5 Fake RoSLA

There may be concern that generating the RoSLA variable as an instrument using the age of the father and mother (the cut off age is 46) may not be correctly modelling the discontinuity in education attainment and only pick up a general upward trend in parental education (Chevalier et al., 2013).

In order to check this we created a fake RoSLA – to simulate the effect by miscoding the RoSLA to its actual introduction. For example, the RoSLA is set to have apparently happened in 1976 (so the cut off age for fathers and mothers is 42 to be affected by this fake RoSLA, i.e. they must be four years younger since the cut-off age was 46 for those affected by the 1972 RoSLA).

The results are given in the Table 2.18 below and comparing them to the actual RoSLA results in Table 2.7, they suggest that the fake RoSLA has no identifying power in estimating the casual impact of father's education and mother's education on children's education. There are positive and insignificant effects of fake RoSLA on both parents' education with very low F- value in all specifications. Also, very high standard errors give an indication of an incorrect modelling. Hence it reassures that the actual RoSLA is picking up the discontinuity in parents' education and is correctly modelled.

Table 2.18: Intergenerational Coefficients on parents' education: IV Results using fake RoSLA, Dependent variable is children's total points score.

	Fathers' Education (no qualifications)	**Mothers' Education (no qualifications)**
	Coefficients Standard errors	Coefficients Standard errors
Specification 31 Pupil characteristics only	-26132.04* (455783.7)	2727.95 (6593.915)
No. of observations F	8438 0.00	8445 0.18
Specification 32 Pupil and school characteristics	847.858 (925.476)	764.397 (887.924)
No. of observations F	8438 1.08	8445 1.01
Specification 33 Pupil, school & family characteristics	241.002 (190.725)	165.354 (190.475)
No. of observations F	8438 4.85	8445 4.39

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

2.7 Conclusions

In this study the key objective is to explore whether the established link between parents' education and children's educational outcomes is due to genes or upbringing. In order to identify the effect of parents' education in two components, it is important to have a source of exogenous variation in parents' education, i.e. the source must be correlated with parents' educational choice and uncorrelated with the parent's ability and other factors.

Therefore, the chapter presents new evidence on the effect of fathers' education and mothers' education on children's education using the Longitudinal Study of Young People in England (LSYPE) which is matched to the National Pupil Database (NPD).

The initial OLS results, similar to other studies, suggest that parental education has a significant and positive impact on their children's educational outcomes, as measured by performance in GCSE examinations taken at age 16. These results are consistent with the evidence found in the previous literature, as documented as a positive intergenerational correlation between parents' education and their children's educational outcomes.

To identify the exogenous variation in parents' education, we used the 1972 RoSLA (raising of school leaving age) as an instrument. The IV results, therefore, show that having removed any genetic effect, there is an intergenerational relationship between parents' education and children's performance, suggesting that the source of the relationship is not a genetic effect. The results further suggest that the effect of nurture (upbringing) is mainly responsible for the intergenerational relationship. However this identification strategy estimates a (LATE) local average treatment effect with no ripple effect further up (Chevalier, 2004; Imbens and Angrist, 1994), as only parents who wished to leave school at 15, those who have either a lower level of qualifications, a lower taste for education, a lower ability or poor resources (financial constraints), were affected by the RoSLA. The IV estimates are therefore not directly comparable to the initial estimates.

The estimates of fathers' and mothers' education are only convincing for those who have lower education level and these are relevant for the population targeted by the recent policies on school leaving age introduced in Britain. Those parents who already possess higher human capital, and have higher ability as well, for those an additional increase in their education will have less of an effect on their children's performance, while, on the other hand, an additional increase in parents' education to those with lower education will certainly increase their awareness, hence have a positive effect on their children's performance. Increasing the education of the present generation has a positive impact on the future generation.

Focusing on the characteristics of those who drop out from the sample, they are of a low level of education and this is to be expected due to less educated people/parents being less likely to respond to surveys. This is a worry, given the focus is on the parents with no qualifications.

Appendices to Chapter 2

Figure A2.1: Relationship between Instrument (RoSLA) and Proportion of Fathers' Education (Degree) by Age

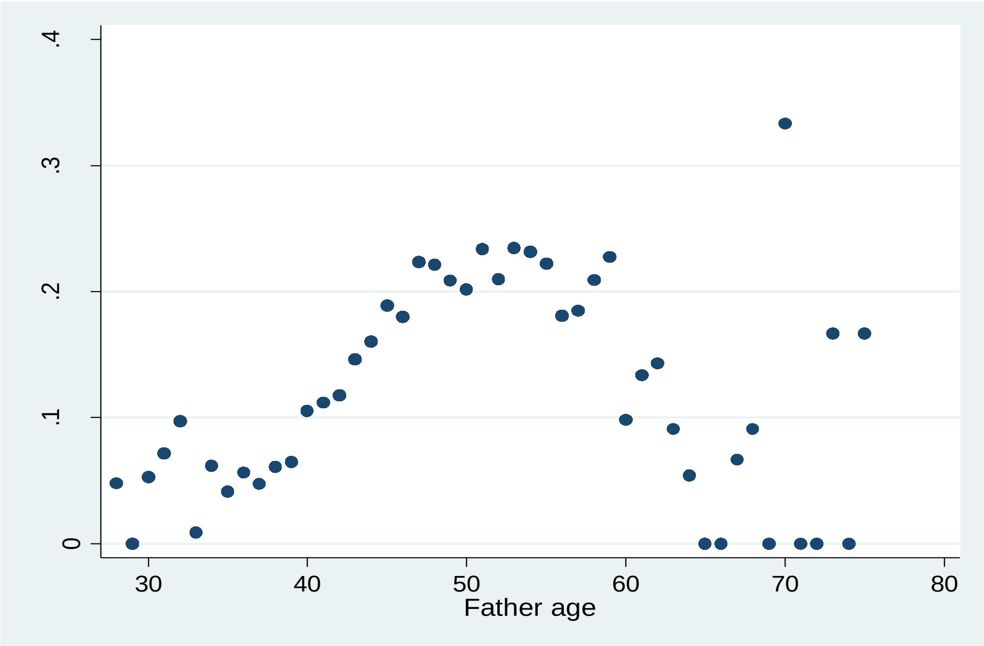


Figure A2.2: Relationship between Instrument (RoSLA) and Proportion of Mothers' Education (Degree) by Age

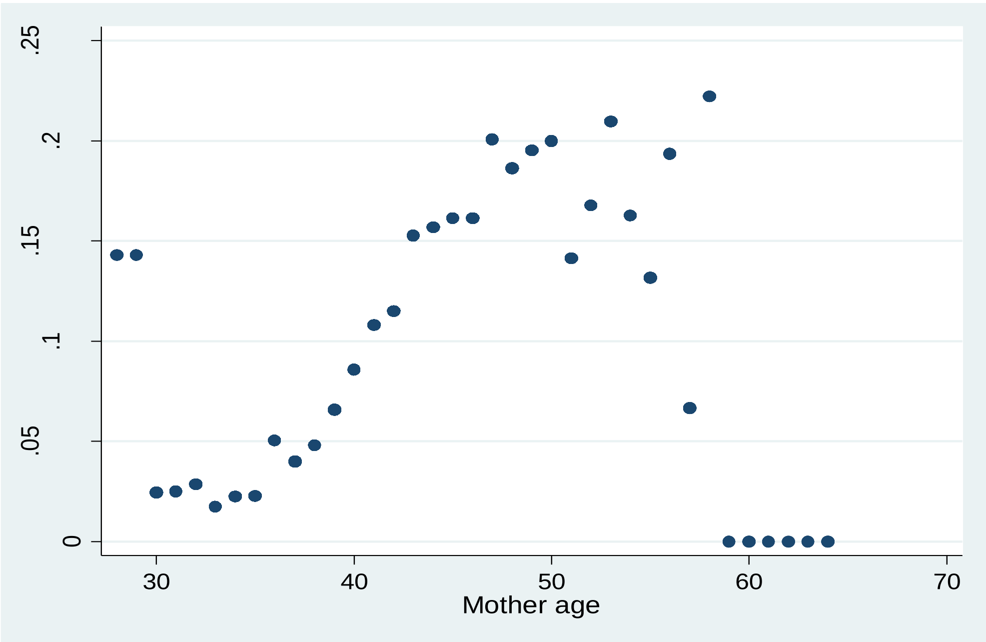


Figure A2.3: Relationship between Instrument (RoSLA) and Proportion of Fathers' Education (Best GCSE) by Age

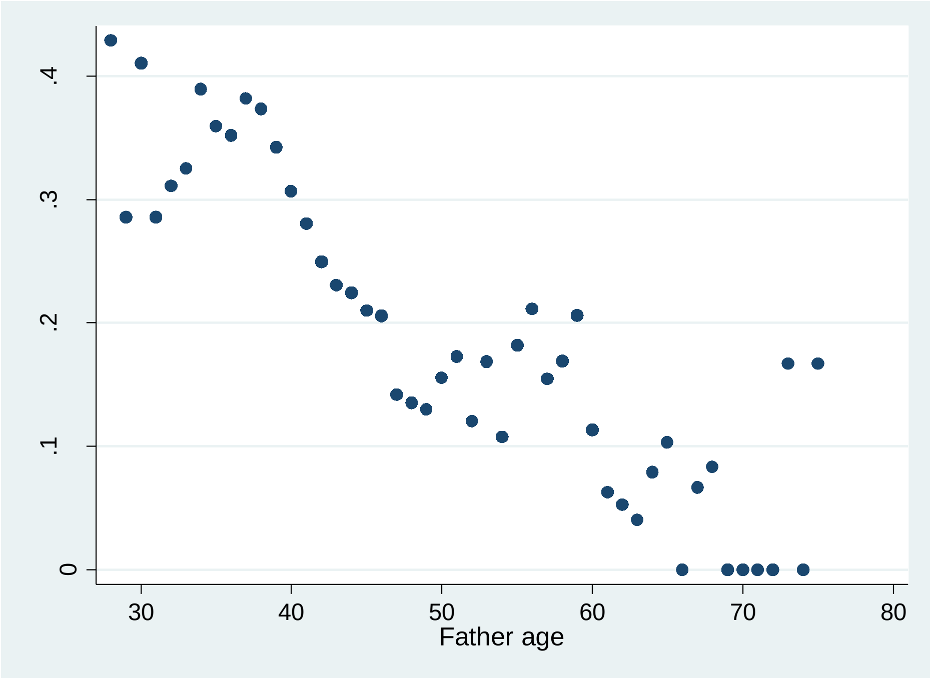


Figure A2.4: Relationship between Instrument (RoSLA) and Proportion of Mothers' Education (Best GCSE) by Age

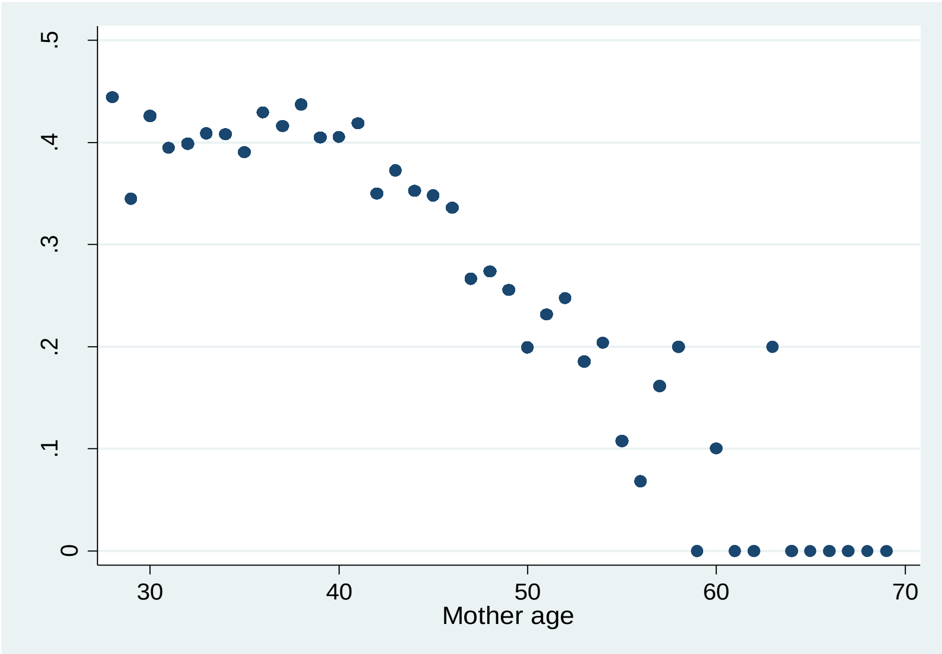


Figure A2.5: Relationship between Instrument (RoSLA) and Proportion of Fathers' Education (Post compulsory Education) by Age

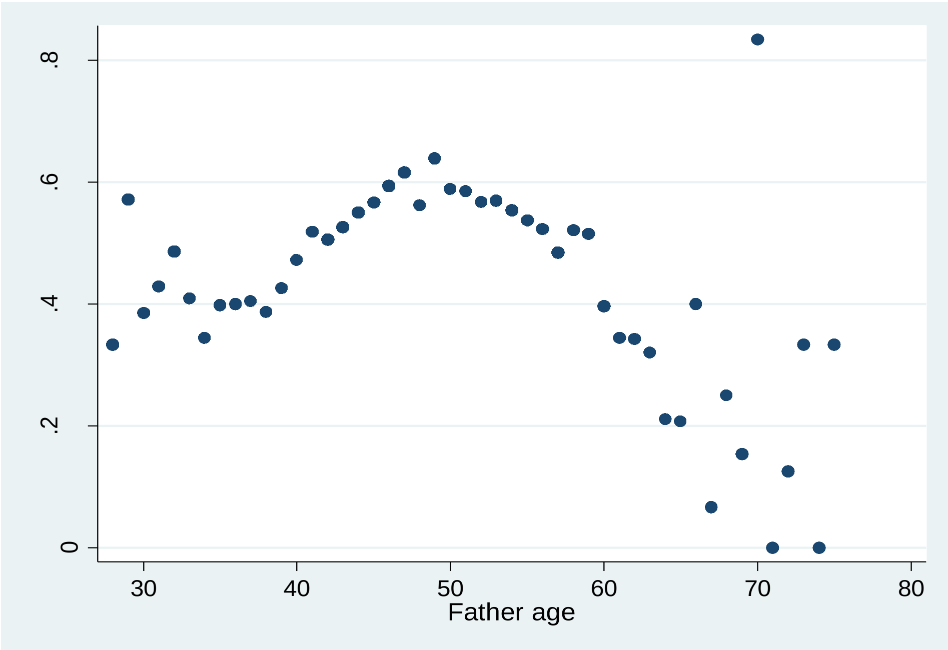


Figure A2.6: Relationship between Instrument (RoSLA) and Proportion of Mothers' Education (Post compulsory Education) by Age

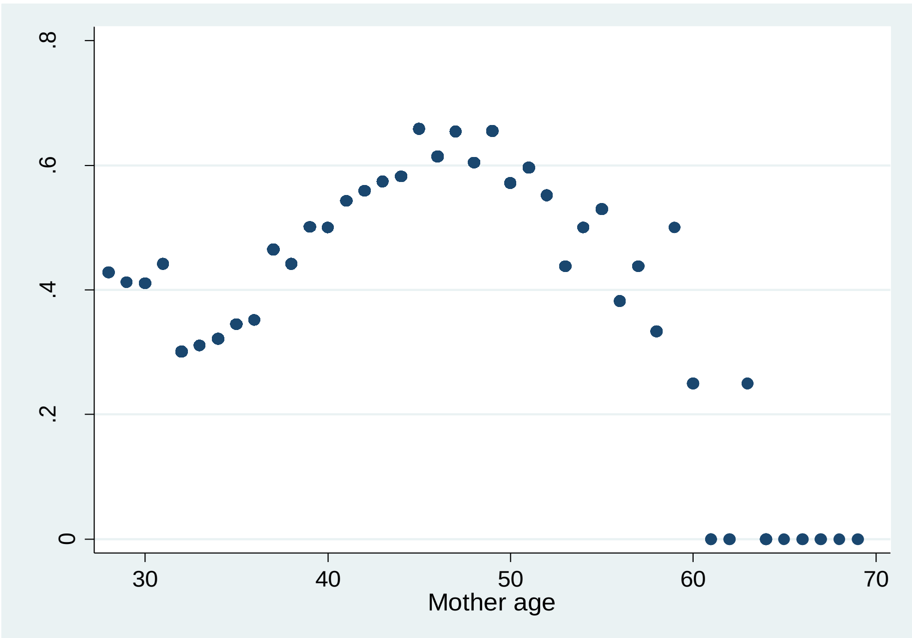


Table A2.1: OLS Results: Effects of Mothers' Education on Children's Education

Variables	Coefficients	S.E.
Female	21.187***	(2.625)
Mixed	16.937***	(6.345)
Indian	33.692***	(5.062)
Pakistani	10.328	(6.454)
Bangladeshi	52.512***	(7.518)
Other Asians	28.867***	(11.719)
Black	10.823*	(6.409)
Health status	-32.064***	(4.017)
Log of household annual income	3.382**	(1.634)
Household Income missing	29.140*	(16.583)
Parents meeting	-36.222***	(2.939)
Mother no qualifications	-27.442***	(3.825)
Average parental age	1.538***	(0.250)
Working status of single parent	18.102***	(5.623)
Working status of both parent	22.777***	(5.891)
Help at home in studying	9.662***	(3.483)
Future aspiration	104.157***	(3.460)
No of siblings	-8.305***	(1.244)
Independent schools	4.721	(6.953)
Foundation schools	7.922**	(3.698)
Voluntary schools	13.308***	(4.013)
Owned house	35.068***	(3.935)
Others	16.023	(13.421)
Computer at home	-55.654***	(5.477)
School pass rate	1.520***	90.076)
Number of observations	8387	
R^2	0.33	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Table A2.2: OLS Results: Effects of Fathers' Education on Children's Education

Variables	Coefficients	S.E.
Female	21.361***	(2.619)
Mixed	16.540***	(6.323)
Indian	33.333***	(5.062)
Pakistani	9.780	(6.330)
Bangladeshi	51.317***	(7.462)
Other Asians	25.134**	(11.620)
Black	9.772	(6.381)
Health status	-31.230***	(3.999)
Log of household annual income	3.361**	(1.630)
Household Income missing	29.018*	(16.545)
Parents meeting	-36.577***	(2.931)
Father no qualifications	-33.122***	(3.490)
Average parental age	1.636***	(0.249)
Working status of single parent	16.114***	(5.623)
Working status of both parent	22.188***	(5.858)
Help at home in studying	9.193***	(3.476)
Future aspiration	102.944***	(3.457)
No of siblings	-8.370***	(1.240)
Independent schools	-5.096	(6.933)
Foundation schools	8.656***	(3.691)
Voluntary schools	12.978***	(4.000)
Owned house	34.196***	(3.919)
Others	13.648	(13.372)
Computer at home	-55.833***	(5.448)
School pass rate	1.512	(0.076)
Number of observations	8380	
R^2	0.33	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Table A2.3: IV Results: Effects of Mothers' Education on Children's Education (First stage)

Variables	Coefficients	S.E.
Female	0.001	0.008
Mixed	0.072***	0.018
Indian	0.282***	(0.014)
Pakistani	0.427***	(0.018)
Bangladeshi	0.446***	(0.021)
Other Asians	0.362***	(0.331)
Black	0.045**	(0.018)
Health status	-0.043***	(0.011)
Log of household annual income	-0.017***	(0.005)
Household Income missing	-0.127***	(0.047)
Parents meeting	-0.029***	(0.008)
ReformM	-0.079***	(0.013)
Average parental age	0.001***	(0.000)
Working status of single parent	-0.111***	(0.016)
Working status of both parent	-0.221***	(0.017)
Help at home in studying	-0.047***	(0.009)
Future aspiration	-0.030***	(0.002)
No of siblings	0.017***	(0.003)
Independent schools	0.024	(0.020)
Foundation schools	-0.010	(0.010)
Voluntary schools	-0.007	(0.011)
Owned house	-0.132***	(0.111)
Others	-0.135***	(0.382)
Computer at home	-0.123***	(0.016)
School pass rate	-0.003***	(0.000)
Number of observations	8387	
F	38.84	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Table A2.4: IV Results: Effects of Fathers' Education on Children's Education (First stage)

Variables	Coefficients	S.E.
Female	-0.003	(0.008)
Mixed	0.051***	(0.020)
Indian	0.227***	(0.016)
Pakistani	0.308***	(.0120)
Bangladeshi	0.392***	(0.023)
Other Asians	0.182***	(0.363)
Black	0.001	(0.200)
Health status	-0.014	(0.013)
Log of household annual income	-0.018***	(0.005)
Household Income missing	-0.139***	(0.052)
Parents meeting	-0.024***	(0.009)
ReformF	-0.042***	(0.012)
Average parental age	0.003***	(0.001)
Working status of single parent	-0.121***	(0.018)
Working status of both parent	-0.174***	(0.018)
Help at home in studying	-0.057***	(0.011)
Future aspiration	-0.053***	(0.011)
No of siblings	0.012***	(0.004)
Independent schools	0.009	(0.021)
Foundation schools	0.007	(0.012)
Voluntary schools	-0.009	(0.013)
Owned house	-0.130***	(0.012)
Others	-0.125***	(0.042)
Computer at home	0.101***	(0.017)
School pass rate	-0.002***	(0.000)
Number of observations	8380	
F	10.74	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Table A2.5: IV Results: Effects of Mothers' Education on Children's Education (Second stage)

Variables	Coefficients	S.E.
Female	21.744***	(2.690)
Mixed	22.289***	(7.776)
Indian	55.052***	(16.957)
Pakistani	42.818*	(25.335)
Bangladeshi	87.387***	(26.502)
Other Asians	56.175***	(24.012)
Black	13.619**	(6.981)
Health status	-35.741***	(4.795)
Log of household annual income	2.249	(1.959)
Household Income missing	21.596	(18.540)
Parents meeting	-38.114***	(3.431)
Mother no qualifications	-103.912*	(57.497)
Average parental age	1.902***	(0.363)
Working status of single parent	9.967	(8.547)
Working status of both parent	6.264	(14.147)
Help at home in studying	5.554	(4.537)
Future aspiration	101.888***	(3.943)
No of siblings	-7.212***	(1.562)
Independent schools	-2.852	(7.309)
Foundation schools	7.296**	(3.839)
Voluntary schools	12.953 ***	(4.130)
Owned house	24.807***	(8.685)
Others	3.803	(15.830)
Computer at home	-46.151***	(9.108)
School pass rate	1.307***	(0.167)
Number of observations	8387	
F	38.84	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Table A2.6: IV Results: Effects of Fathers' Education on Children's Education (Second stage)

Variables	Coefficients	S.E.
Female	21.631***	(2.746)
Mixed	21.182***	(8.428)
Indian	55.157***	(23.767)
Pakistani	39.562	(32.161)
Bangladeshi	90.192***	(40.590)
Other Asians	42.102**	(21.939)
Black	9.651*	(6.677)
Health status	-32.940***	(4.398)
Log of household annual income	1.813	(2.527)
Household Income missing	17.732	(22.463)
Parents meeting	-38.527***	(3.891)
Father no qualifications	-128.426	(101.500)
Average parental age	2.265***	(0.691)
Working status of single parent	4.760	(13.532)
Working status of both parent	5.990	(0.32)
Help at home in studying	3.291	(6.830)
Future aspiration	97.781***	(6.531)
No of siblings	-7.273***	(1.820)
Independent schools	-4.497	(7.302)
Foundation schools	9.616***	(3.927)
Voluntary schools	12.292***	(4.295)
Owned house	21.609	(14.005)
Others	1.364	(19.107)
Computer at home	-46.025***	(11.829)
School pass rate	1.264***	(0.257)
Number of observations	8380	
F	10.74	

*** p<001, ** p<0.05, * p<0.1
Standard errors in parentheses

Chapter 3:

Teacher Characteristics and Pupil Performance in Pakistan: A Teacher Fixed Effects Approach

3.1 Introduction

The most important subject in education policy is how to improve the educational outcome within schools. Various measures such as increasing school inputs, lowering class size, incentive based policies and teacher qualities have been discussed in this regard. Policymakers, scholars, parents and school administrators have profound belief that teacher quality is one of the important factors in raising pupils' test scores performance. Improving weak teaching potentially improves the school quality (Glewwe and Kremer, 2006).

Most of the research on teacher quality comes from the US (Aaronson et al., 2007; Chetty et al., 2014; Hanushek and Rivkin, 2012; Rivkin et al., 2005; Rockoff, 2004) while outside it, teacher quality is under researched as compared to other characteristics of schools. Further, recent research on teacher quality among developed and developing countries (Azam and Kingdon, 2015; Slater et al., 2012) focuses on high school exams, as discussing teacher effectiveness/quality at high school (high stakes) exams yields additional benefits: first, in order to participate in the political and economic life of knowledge–based economies, the minimum requirement is a high school diploma. Second, there is the existing link between poor performance in core high school exams and failure in graduating (Allenswoth and Easton, 2007). It is therefore, important to look at teacher quality in high school exams.

So far research on teacher quality both in developed and developing countries has shown that there exists a significant variability in teacher quality, meaning that teachers differ to a great extent in what, or how, they teach their pupils (Azam and Kingdon, 2015; Rockoff, 2004; Hanushek and Rivikin, 2006; Aaronson et al., 2007; Slater et al., 2012; Leigh, 2010; Clotfelter et al., 2006). Assessing the relative effectiveness of teachers has an obvious policy implication for both developed and developing counties. However, a lack of evidence on the causal relationship between teacher characteristics and teacher quality, has led researchers to conclude that it is hard to discover which particular observable characteristics explain the differences in teacher quality, so that what makes a teacher good or bad remains a puzzle in education policy (Metzler and Woessmann, 2012; Rivkin et al., 2005; Aaronson et al., 2007).

The vagueness in assessing teacher quality arises due to different methodological challenges involved. First, is the data requirement, as Rockoff (2004) mentions that most issues in the field of teacher quality are due to data quality. To identify teacher quality, researchers have to focus on pupil-teacher data with rich information on pupil, family and school variables which is not easily available. Second, is the non-random matching of students and teachers to classrooms within school (Rothstein, 2009), leading to a possible correlation between observed teacher characteristics and unobserved student characteristics. Third, omitted variable bias in the form of unobserved teacher characteristics such as motivation and intelligence may exacerbate the results (Aslam and Kingdon, 2011). Fourth, it is often difficult to have detailed information on teacher credentials and teaching methods in classrooms simultaneously in the dataset affecting student performance.

In this study I try to overcome these issues: such as non-random sorting of teachers and students to classes, non-random attrition and omitted variable bias. The current methodology uses controls for both teacher-fixed effects and pupil-fixed effects due to available variation within teachers and within pupils and presents new evidence about teacher quality. However, this study faces the similar challenge as other studies on teacher quality, it is not possible to isolate the pure teacher effect from the school effect, nevertheless, the present study can control for at least some observable school characteristics.

The study uses unique survey data, designed and collected by me, which is linked to the administrative exams scores data of Year 9 pupils (taken at age 14-15) in private and public schools from one of the districts of the Punjab province in Pakistan in 2008-2009. The survey data I collected has detailed information on pupils' individual, family, peers, schools and teachers characteristics. The current dataset complements the previous leading datasets in the field of teacher effectiveness. For example, similar to Azam and Kingdon (2015), Aaronson et al. (2007), and Rockoff (2004), I can match pupils to their actual teachers rather than an average of teachers within a school;

dissimilar to Aaronson et al. (2007); Rockoff (2004); Kane et al. (2008); Rivkin et al. (2005); Clotfelter (2006) the current dataset examines pupils who are taking high stakes exams which are important for both schools and them. Despite the fact that my dataset carries advantages, there are shortcomings as well which I explain at the end.

3.1.1 Objectives of the Study

This chapter is concerned with answering two main research questions:

1) Does teacher quality matter? By this question I mean measuring the variation in teacher quality, evaluating teachers in terms of exam results of their pupils i.e. what is the impact on pupil scores of being taught by a low or high quality teacher?

2) What determines teacher quality? Under this question I am aiming to explore whether observed characteristics explain the variation in teacher quality.

In addition, the study further aims to explore the following:

3) What is more important in explaining pupil performance, teacher characteristics or process variables[9]?

The present study works through two steps. In the first step, it estimates the model of pupils' performance using an educational production function linking a substantial set of covariates of pupils' own, their family and schools characteristics, and incorporating teacher fixed effects. This step yields estimates of a fixed effect for each teacher, which gives an indication of the teachers' contributions towards student performance. Second, it relates the estimated teacher fixed effects to teachers' own credentials by regressing measured teacher fixed effects onto different teacher own characteristics and classroom practices.

First stage results show that there are significant variations in teacher fixed effects within the schools, suggesting important unobservable differences among teachers. A good teacher, defined as being at the 75th percentile of the teacher fixed effects distribution, is related to an increase in the student test score of 0.15 standard deviations relative to the omitted teacher, while a bad teacher decreases the score by 0.77 standard

[9]In this study "process variables" defines the variables that make the process of teaching more effective such as more time spent on teaching in a classroom, lower workload of teacher, planning lessons in advance, being more kind and helpful to students, behaving in a democratic way. They are included to capture the effect of teacher unobservables such as motivation, enthusiasm and ability (albeit imperfectly).

deviations. Therefore, a pupil having been taught by a good teacher (75th percentile teacher) scores 0.92 standard deviations more than the pupil who is taught by a bad teacher (25th percentile teacher) leaving a significant effect on pupil performance.

To address non random sorting of pupils and teachers to classrooms a pupil fixed effects model is also estimated along with the teacher fixed effects approach as the dataset used in this study has variation in test scores for multiple subjects for the same student. Results from the pupil fixed effects specification are theoretically more relevant, given that they control for individual pupils' characteristics (at least their fixed characteristics that are constant across teachers). Pupil fixed effect results highlight that teachers are important in explaining pupils' performance even after controlling for pupils' prior ability and home background effects.

Turning to the second stage, the results reveal that teacher observed characteristics explain very little of the variation in estimated teacher quality, showing a negative relationship between training and teacher performance, at the 10 % level, which corroborates previous findings in the literature (Rivkin et al., 2005; Aaronson et al., 2007).

The findings of this chapter question the emphasis put on the teacher characteristics (qualification, teaching experience and teacher training) in developed and developing countries[10]. In the status quo teachers are being rewarded for the possession of characteristics that have nothing to do with pupils' performance. It may be hard to identify good teachers ex ante but administrative data can be used to identify them ex post (Slater et al., 2012). It implies that improving teaching quality is important but it is very difficult to measure the characteristics that are responsible for bringing variation in teacher quality or at least good teachers cannot be identified on the basis of teachers'

[10] The importance of hiring more qualified teachers can be seen in public policy, for example, in the US, No Child Left Behind, 2002-2015, by law required highly qualified teachers within school to teach disadvantaged pupils. The program is aimed to be implemented on the basis of teacher qualification, as opposed to teacher effectiveness in classrooms. However, individual states were left to define what a 'highly qualified' teacher is. According to Hanushek and Rivkin (2012) most states simply picked up the variants of the existing requirement of teacher certification. In India, Minister of State for Human Development, Dr. Shashi Tharoor, on 20th August, 2013, introduced a three-points based strategy to improve the quality of teachers in school: 1) Strengthen the teacher education institutions; 2) The revision of teacher curriculum; and 3) Setting the minimum qualifications for Teacher Educators.

observable characteristics. However ex post evaluations, assessing teachers on the basis of their contributions towards pupils' performance is more appropriate. But the ex post evaluations need rich matched student-teacher data that are often not available to researchers or do not exist in developing countries.

Generally, teacher hiring or firing depends upon teacher certification. Results suggest that policy makers should move focus from policies rewarding teachers on the basis of credentials to performance based.

This study contributes to the literature of teacher effectiveness in the following ways. (1) This is the first study (to my knowledge) investigating teacher quality, using a survey dataset, at high school level/senior secondary level in Pakistan.[11] Aaronson et al. (2007) argued that, although it is important to study teacher effects at all levels of the education process, studying at high school has an additional advantage as classrooms are subject specific and the teacher-student match really corresponds to what one thinks of as the teacher effect. Furthermore, the exams at senior secondary and secondary school levels are high stakes exams. (2) In terms of methodology, this study is only one of a very small number (including Slater et al., 2012; Metzler and Woessmann, 2010; and Azam and Kingdon, 2015) using both teacher fixed effects across pupils, rather than over time, and pupil fixed effects approaches to examine teacher quality. (3) The present study increases the confidence of the reader, by corroborating the findings of developed countries, such as the US and UK, irrespective of the context.

The chapter is structured as follows: section 3.2 presents the review of previous literature regarding the effectiveness of teachers, section 3.3 elaborates on data used, section 3.4 discusses the methods and variables used in this study, section 3.5 interprets the results from the analysis and the final section 3.6 presents the conclusions and policy implications for policies related to teacher effectiveness.

[11]The study by Aslam and Kingdon (2011) is the only study for Pakistan on teacher effectiveness using a cross-sectional pupil fixed effects approach at the middle level of education (Year 8).

3.2 Literature Review

3.2.1 Teacher Effectiveness

Teacher quality is potentially a key determinant of student performance and teachers differ in terms of quality. Data requirements are complex to study teacher quality. Many previous studies in this field, surveyed by Hanushek (2002), do not allow matching students with their respective teachers. Also due to the nature of such data, these studies are not able to control for pupils' characteristics. The literature probing teacher effectiveness has adopted two approaches.

3.2.2 Teacher fixed effects approach

In the first approach teacher quality is measured as a teacher fixed effect in which different groups of students are taught by the same teacher in a given year or over time, while the performance of the same students with different teachers is also observed. This enables the researcher to calculate the total teacher effect on the basis that the better teachers would be the ones who produce higher marks for the students than other teachers with the same students. Several papers have used this approach (Aaronson et al., 2007; Azam and Kingdon, 2015; Hanushek et al., 2005; Hanushek and Rivkin, 2006; Leigh, 2010; Rockoff, 2004; Rivkin et al., 2005; Slater et al., 2012).

Rockoff (2004) identifies teacher effects using panel data covering over a decade (1989/90 to 2000/01) of students' test scores in grades 2 to 6 and teacher assignments from two contiguous districts in New Jersey. The data on teachers in multiple classrooms and test scores in multiple years allow the author to measure teacher fixed effects and student fixed effects. Generally in primary schools, students are taught by a single teacher. The shortcoming of using such data is that one cannot identify the impact of different teachers on particular students in different subjects at the same time. Measuring teacher quality using the teacher fixed effects method, he finds significant differences among teachers i.e., a one standard deviation increase in teacher quality leads to a 0.20 and 0.24 standard deviation increase in reading and maths test scores, respectively. Controlling for teacher fixed quality, teaching experience has a positive effect on reading test scores, while the other teacher observable characteristics such as gender, ethnicity and education are found to have no significant impact.

Other studies using the US data and the teacher fixed effects approach are as follows. Rivkin et al. (2005) use a unique matched panel dataset which spans grades 3 to 7 for three cohorts of students across 3000 schools in the state of Texas in the mid -1990s. The limitation of their data was that it does not match individual students to their teachers, rather only to a set of teachers in a grade within schools which is likely to attenuate estimated teacher effects. Their findings from the semi-parametric lower bound estimates of the variance in teacher quality imply that teacher observable characteristics such as education and experience have little impact on student educational outcomes. Nevertheless teachers have a powerful impact on reading and mathematics achievement. The results also suggest that reduction in class size by a costly ten students has fewer effects than the positive effects of moving one standard deviation up the teacher quality distribution.

Another study by Hanushek et al. (2005) estimates the variation in teacher quality using matched student and teacher data from grade 4 to 8 in the same schools as Rivkin et al. (2005) for the school years from 1995 to 2000 in a large district of Texas. They measured teacher quality using a semi-parametric approach based on value added student achievement. The findings of their study confirm the variation in teacher effectiveness within schools not between schools. Further, their findings confirm that teacher certification and experience explain little of the variation in teacher quality.

Aaronson et al. (2007) use the administrative dataset of 9th grade students over three years from 1997 to 1999 in one school district in Chicago. In contrast to many other studies, the key advantages of their study are the ability to match teachers to their students in a particular classroom and the availability of data on prior achievement of students which they assume addresses the issue of student heterogeneity. They find an increase of 0.15 standard deviations in student maths scores, if there is an increase of one standard deviation in teacher quality. Also they find that traditional human capital measures explain very little variation in estimated teacher quality.

Outside the United States, very little work has been done on the estimation of teacher quality using the teacher fixed effects approach. The only study in the UK by Slater et al. (2012) estimates the effect of individual teachers on student educational outcomes using the schools and teacher primary dataset for the England[12]. They also estimate the variability in teacher quality which they measured by the impact on test scores. The data used in this study relates 7305 pupils to their 740 teachers across 33 schools in England in each of the compulsory subjects; Maths, Science and English in GCSE exams taken at the age of 16. The model they use is the point-in-time fixed effects; to control for pupil heterogeneity, prior attainment is used. They find considerable variability in teacher effectiveness: a teacher being one-standard deviation better increases the student outcomes by 25% of a standard deviation which is a little higher than the estimates found in US studies. In addition, the results support the findings that observed teacher characteristics explain little of the differences in estimated effectiveness.

Similarly, Leigh (2010) is the only study using data from Australia; he estimates the effectiveness of teachers in raising students' test scores using a dataset covering 90,000 students in primary grade 3 to 7 and 10,000 school teachers in one of the states of Australia, Queensland. Teacher fixed effects are jointly significant and highly dispersed, even after adjusting for measurement error. The study finds a strong positive relation between teachers' gains in literacy and numeracy. Moving a teacher from the 25^{th} percentile to the 75th percentile on the teacher quality distribution will increase student test scores by one-seventh of a standard deviation. Teacher experience is positively related to teacher effectiveness; however, there is no positive effect of teacher qualifications on pupils' test scores. In literacy, female teachers do better than male teachers.

[12] The dataset for this study was collected by CMPO, for the evaluation of project "Performance Threshold" for teachers. The description of the project is given in Atkinson et al., (2009). They match this dataset with school level variables from the National Pupil Database (NPD).

The teacher fixed effects approach does not require identification of certain characteristics that generate student achievement, instead it estimates the overall effect of each teacher that captures in a very general way the influence of a certain teacher relative to other teachers in the sample. However this approach assumes that a particular teacher is equally effective for all pupils, which may not be the case necessarily.

In general, findings of all studies using the teacher fixed effects methodology suggest that teachers have a significant impact on student achievement that means teachers' assignment is a relevant issue for education policy. However, when they regress estimated teacher fixed effects on teacher observable characteristics, such as education, experience, gender and training, they hardly find variation in teacher quality due to these characteristics.

3.2.3 Direct approach

The second approach in the literature considers an educational production model linking teacher characteristics to student performance, controlling for student attributes. This approach estimates the direct relationship between teacher characteristics and student achievement. Different studies use different methodologies under this approach. The most commonly used methodologies vary from a standard cross-sectional achievement production function (Aslam and Kingdon, 2011; and Kingdon, 2006); or panel data approach (Clotfelter et al., 2006 and 2010; Azam and Kingdon, 2015); to the instrumental variable method (Hoxby, 1996; Kingdon and Teal, 2010); and experimental methods measuring the impact of teacher incentives (Lavy, 2002; Glewwe et al., 2010; Muralidharan and Sundararaman, 2011).

Clotfelter et al. (2006, 2007a, 2007b) explore the relationship between different teacher credentials and student achievement by directly regressing student achievement on teacher characteristics using a detailed panel dataset from North Carolina, US. They use data of primary grades from 3 to 5 and use student fixed effects to address non-random matching of students and teachers. They conclude that the effect of teaching experience on student achievement is positive and larger for maths than for reading. Clotfelter, et al. (2010), study the above relationship in the context of high school end of course exams.

Using a direct approach, few papers in developing countries have studied teacher effectiveness. Azam and Kingdon (2015), using administrative panel data, estimate a direct teacher value added approach of 8319 pupils studying in grade 12 in private schools from one of the districts in the state of Uttar Pradesh, India. They use a pupil fixed effects approach controlling for prior ability using grade 10 exam scores. They find a considerable variation in teacher quality. The results suggest that being taught over a time period of two years by a good teacher, defined as being at the 75th percentile, is related to an increase in student achievement by 0.47 standard deviations relative to a low quality teacher, defined as being at the 25th percentile. However they find that observable teacher characteristics explain very little variability in teacher quality.

Kingdon (2006) analyses the effect of teacher characteristics upon students' performance using a dataset from 186 schools in India under the control of the Council for Indian Secondary Certificate Examinations (CISCE) which permits the matching of students' subject test scores to the teachers who teach those subjects. The methodology uses the standard cross-sectional achievement production function allowing for pupil fixed effects. She assigns the average characteristics of all teachers in school to all students in grade 10 as she does not know the exact teacher who taught the student in a given subject. She finds teacher training and teacher qualifications (having a masters level or higher degree) would raise pupil performance by 0.09 standard deviations. She suggests that the above findings are upper bound estimates. However her study does not control for previous academic achievements of the pupils, hence she estimates the education production function in levels, not value added.

Using the same data and an estimation methodology similar to Kingdon (2006), Kingdon and Teal (2010) study the relationship between teachers having union membership and student achievement. They find that a teacher with union membership is negatively related to student performance. Further results from a school fixed effects model suggest that a teacher having union membership has a positive impact on teacher pay. Like Kingdon (2006), the study does not control for previous attainment.

For Pakistan, Aslam and Kingdon (2011) is the only study on teacher effectiveness using a standard production function (cross-section) looking at pupils studying at the middle level of education, grade 8. They uncover teacher characteristics, from their CVs, such as education, teaching experience and teacher training and teaching methods that affect pupil achievement the most. The study uses data from government and private schools in Pakistan, 2002-2003. Instead of using variation across time they use the pupil fixed effect approach across different subjects. Their pupil fixed effect results highlight the importance of teaching practices in classrooms while the teacher credentials do not significantly matter to pupil achievement.

The findings from the literature using the direct approach are not different to the ones measuring teacher quality using a teacher fixed effects approach, highlighting that teacher observable characteristics explain little of the variation in teacher quality. Use of the mentioned techniques, IV and panel data approach estimation, depends on availability of certain types of data. However statistical methodologies used in these studies are not above criticism. For example, instrumental variable methodology depends on a convincing instrument which is hard to find in reality. In most developing countries policies having exogenous variation, for example education policy regarding minimum school leaving age or maximum class size, exist in statutes, but are hardly ever adhered to in practice. Experimental methodology provides a good solution for the endogeneity issue, however, problems occur while researchers try to generalize the results outside the experiment (Todd and Wolpin, 2003). Similarly, even if one keeps aside the issue of measurement error[13] in panel data approaches, still it is a difficult task as data are hard to obtain and almost non-existent in developing countries.

Looking at the literature on teacher quality from the very rich US literature to outside the US, though under researched, using any teacher fixed effects or the direct approach, it is clear that teachers are not the same in what they teach their students and in how they teach, which indicates that it is important to study the effects of teacher effectiveness for student achievement. For developing counties Azam and Kingdon (2015) is the only study that discusses teacher quality at high school exams. Therefore,

[13] There is missing information on variables of interest, and always, it is not possible to match who teaches which subject. Consequently, it is likely to induce error in calculating the percentage of teachers, and number of teachers per grade or the ratio of two, hence this may contain a nontrivial amount of noise (Rivkin et al., 2005).

research is still needed to see, just like in US studies, whether teachers matter to student achievement at higher grades in the context of developing countries. Also Aslam and Kingdon (2011) is the only paper in the field of teacher quality, looking at whether teacher characteristics are more important or the teaching ways chosen by a teacher (the process variables) in the classroom.

The present study contributes to the literature on teacher quality studying the relationship between teacher characteristics and student achievement at higher levels of education in developing countries. Different from Aaronson et al. (2007); Rockoff (2004); Kane et al. (2008); Rivkin et al. (2005) and Clotfelter (2006) but similar to Clotfelter et al. (2010) and Azam and Kingdon (2015), the current study examines teacher quality for pupils taking high stakes exams. Similar to Aaronson et al. (2007), the present study is able to match all pupils to their subject specific teachers, though while they attain a 75% student-teacher match, this study has a 100 % match. Further, like Kingdon (2006) and Kingdon and Teal (2010), the present study does not control for prior ability, though it does include pupil fixed effects.

3.3 Data

One of the contributions of this study is that it utilizes the survey dataset collected by myself involving rigorous field work, interviewing 611 pupils studying in Year 9, using a random sampling technique, as well as interviewing their family, subject teachers and the school head teachers and also tracking the examination scores information in Year 9 of all pupils by linking them with administrative data. Thus, it is a comprehensive dataset containing rich information on all aspects of selected students which might affect their academic performance in a particular grade.

In total there were 14 schools included in the sample, including 6 Government and 8 Private schools in the Bahawalpur district, in the Punjab province of Pakistan. The choice of the Punjab and particularly the Bahawalpur district for the field survey was based on the following grounds: (1) Recent literature on private schooling in Pakistan has noticed that much of the expansion in private schooling has been particularly prominent in the province of the Punjab. (2) Bahawalpur is the largest district of the Punjab in terms of area; it has a high number of reputable educational institutions, and a large number of public schools is also available in the city. All schools are affiliated to the Board of Intermediate and Secondary Education (BISE) which is responsible for conducting exams in Year 9, 10, 11 and 12 in all private and public schools.

Data were collected from 611 pupils, among them 396 in science and 215 in an arts field of study in Year 9, by using pupil questionnaires filled in by the pupils themselves. Pupil questionnaires also provided information on personal characteristics such as gender, date of birth, physical activity, educational aspiration, hours of study at home, health status, and mode of transport, food frequency, and weekly eating pattern. In addition to this, each child's weight and height was measured to calculate the body mass index. A family questionnaire was sent to the home of each child and returned to the school authorities the next day completed by either parent (or the child completed the questionnaire by asking the parent questions if the parent was illiterate), containing information on parents' education, occupation, family size, family income, number of rooms, house tenure, neighbourhood, siblings' educational record etc.

The teacher dataset captured information not only on teachers' traditional human capital measures (education, teacher salary, experience, and training) but also on the process variables (teaching methods used in the classroom and questions related to teacher behaviour) adopted by teachers in the classroom. This was done by interviewing all subject teachers who taught the pupils in Year 9 in sampled schools by asking a series of questions, for example, daily workload[14], class size, lesson planning, behaviour with students (authoritarian or democratic behaviour), surprise tests or quizzes given to the students, class duration[15] and subject taught to the sampled Year 9 students. These process variables included in the teacher questionnaire are aimed to capture the effect of teacher "unobservables" to some extent. The intuition is that more efficient and motivated teachers spend more time in the classroom, have more involvement with students encouraging them to ask questions, give surprise tests and quizzes, and are more helpful to students behaving in a democratic way. This further suggests that teachers are more helpful to students who behave in a democratic way (Coe et al., 2014). In addition to the above, the comprehensive teachers' data permits us to separate the effects of observed teacher characteristics from unobserved aspects of

[14] Daily workload is measured as the number of classes a teacher teaches per day such as 4 classes or 5 classes.

[15] Class duration is measured in minutes spent in a class, average class duration was less than an hour.

teacher quality. The data for this study provides us with a unique advantage: the ability to relate teachers with students in certain classrooms in a specific grade (Year 9). Many other studies are missing this advantage, as they have data which only allow them to match students to the average characteristics of all teachers in a specific grade.

Further to making the dataset more useful, the school head of each sampled school was interviewed. School head questionnaires brought forth information on school head teacher characteristics such as education, gender, experience, training and characteristics of the school, such as enrolment in Year 9, number of teachers teaching to Year 9, school facilities and school resources such as number of books in the school library, number of computers, and area of playground etc. Overall, the private schools were better in responding relative to public schools.

The Year 9 and 12 examinations are examinations conducted by the Board of Intermediate and Secondary Education, Bahawalpur (BISE). Marks obtained by the students in the central board examination are a reliable measure of students' academic performance as the examination system is uniform all over the Punjab province. The only difference is that the medium of instruction is in the Urdu language in public schools and the English language in private schools.[16] The exams are held at the same time, scheduled by the Ministry of Education. The exam papers are the same except that for the private school students' exam papers are in English while they are in Urdu for the public school students. The exams are nationally set, so marking standards are the same and exams are marked outside the schools by the same set of external examiners, leaving little chance of systematic manipulations, tempering results, leakages of exam papers and minimizing the incidence of cheating, which is a big worrying factor at the level of high school exams (Kingdon, 1996).

The exam scores used in this study matter a great deal as good performance in these exams is considered by the universities and colleges in Pakistan at the time of giving admissions to higher education, hence they are a key gateway to further education for students having more scope and relevance for policy implications. Clotfelter et al.

[16] In Pakistan private schools could be English medium or Urdu medium, but all private schools in the current dataset are English medium.

(2010) argue that tests/exams being external to school as compared to within school means students are also more likely to take them seriously.

The examination results data were provided by BISE Bahawalpur, Punjab, using a unique pupil identifier code (called a board roll number or board registration number) of each individual student, provided by their schools and also for consistency double checked by asking the students for their board registration number. The mark sheets of students are made available online at the time of announcement of results on the BISE website and latter published in a gazette. As the data set contained information on student registration numbers, so I could search each sampled student score record online manually, by inputting the board number in the board number box and double checked this with the gazette marks record. In this way the information on examination score in this dataset is measured with little to no error.

The marks sheet (for Year 9) produced by the BISE is based on all the seven different subjects' marks and also the total marks obtained by a student which is the aggregate of all the subject marks. All science and arts field of study students have the same four compulsory subjects: Language[17], English, Mathematics and Religion regardless of any field, while the other three subjects are different for science and arts fields of study. These three subjects which are different in both fields are dropped from the model (for example, Physics, Chemistry and Biology for the science group) and also teachers who teach these subjects are dropped, effectively leaving the total number of teachers as 56 (down from 98) who teach Language, English, Mathematics and Religion, shared by both disciplines.

I restrict the analysis in chapter 3 to the four common subjects, in order to include the pupil fixed effects, and then essentially look at whether some teachers are doing better than the other teachers with the same pupils. This avoids the problem that field of study is an endogenous variable itself, chosen by the individual student, and is subject to many of the same influences as performance. Including pupils from all compulsory subjects avoids the results being affected by any possible selection issues where when faced with a choice, pupils of different quality are choosing different subjects (so some

[17] Language subject is another language called "Urdu" which is the National language of Pakistan, and is different to the English subject.

subjects and hence teachers may have more of the good pupils while other subjects have more of the lower quality pupils).

The final dataset for chapter 3 constitutes 2444 observations which records 4 observations for each pupil matched with their respective subject teacher characteristics. Table 3.1 describes the proportion of pupils in the dataset on the basis of school type, gender, field of study and location. The percentage of pupils in private schools in the dataset is 32%. Male pupils are 49% of the total. The pupils having science as a field of study are 64% and 35% for arts. Therefore, the present sample is a combination of private and public schools in Punjab and also in terms of gender and field of study.

It is very important to have a representative sample of all fields in the dataset for example as it is generally believed that people having science as a field of study are more able and motivated hence have better academic performance at schools relative to their counterparts. So if all pupils in the dataset are from one particular field this might create a sample selection problem. There were 14 schools in sample, and on average 43 observations (43 students) are from each school.

Table 3.1: Proportion of Pupils in Data by School Type, Gender, Field of Study and location in Year 9

	School Type	
School type	**Frequency**	**Percent**
Private schools Y9	197	32.30
Public schools Y9	413	67.70
Total	610	100.00
	Gender	
Gender	**Frequency**	**Percent**
Male	300	49.10
Female	311	50.90
Total	611	100.00
	Field of Study	
Field of Study	**Frequency**	**Percent**
Science	394	64.48
Arts	217	35.52
Total	611	100.00
	Location	
Location	**Frequency**	**Percent**
Urban	344	56.39
Rural	266	43.61
Total	611	100.00

3.3.1 Strengths and Weaknesses of the Data

1) The dataset used is primary data collected by myself and is a fair mix of pupils regarding the proportion of private and public school, rural and urban areas, males and females, and science and arts field of study.
2) It covers a wide range of characteristics of pupils and their families, which may affect their academic progress, such as: individual attributes, family attributes, subject teacher attributes, school head attributes.
3) The data are designed in a way that makes it possible to match with the administrative dataset of BISE, academic scores, through an unique pupil identifier code, which further enhances its strength creating a richer dataset of academic score variables.
4) Exam scores in Year 9 are high school exams, hence high stake exams.
5) The timing of data collection is also important, pupils are chosen randomly at the beginning of Year 9, at that time schools and teachers are not aware of their potential performance in board exams. In other cases, picking a sample after the declaration of official results, there is risk of bias associated with teachers putting their most able students forward. However, one could argue against this as teachers are likely to have some inkling of student potential. As there is possibility that they may not know the performance of students in the board exams but school and teachers do know the past ability of the student (which I do not) and in so far as the past ability is a predictor of future ability there could be a selection here. Therefore, to address the issue of non availability of past ability in the dataset, I estimate the pupil fixed effects model to control for unobserved pupil characteristics. Another key feature of data is that it has variation within student across subjects that enable me to estimate a pupil fixed effects model due to observing pupils studying in multiple subjects/classrooms within the same school.

Despite all the above mentioned strengths, it is true that sample size is small and so one has to be cautious when making claims for the entire population**.** In addition, data come from a specific region of Punjab, Pakistan, so generalisability of the findings outside this region should be done with caution.

3.4 Methodology

The objectives of the study are twofold; one, is there any variation in teacher quality[18]; and two, is the variation in teacher quality explained by teacher observable characteristics? Therefore, the methodology proceeds in two steps: measuring variation in teacher quality and examining what explains teacher quality.

3.4.1 Measuring variation in teacher quality

The major tool of analysis, which measures the relationship between the school inputs, like teacher quality, school facilities, child attributes, family factors, and pupil achievement, is the educational production function (EPF). The objective of this study is to estimate the standard EPF in a consistent and unbiased manner. The model underlying this approach is straightforward and postulates that the output of the educational process is related to a series of measureable inputs. The model of the standard educational production function is specified as follows:

$$Y_{ijsk} = \beta_0 + \beta_1 X_i + \beta_2 S_k + \varepsilon_i \qquad \text{(Eq 1)}$$

Where Y_{ijsk} is the academic performance, measured as the score obtained by the i^{th} student in the class of the j^{th} teacher in subject s in the k^{th} school and is determined by the vector of his/her personal and family characteristics (X) and by school characteristics (S); β's are parameters of interest; and ε_i is the error term.

The dependent variable in the model is marks achieved by a student in the four subjects: Language, English, Mathematics and Religion. The most common use of fixed effects is when we have panel data, i.e. data on the same individual at different points in time. In that case, we can include a fixed effect to pick up unobserved characteristics of teachers that remain constant over time. In this study, we do not have panel data, but still have data on the same teacher across different pupils. We can see that the principle is the same - in panel data we have more than one observation on each individual spread across different points in time, and in this case we have more than one observation on each teacher spread across different pupils. So including fixed effects

[18] Throughout this chapter teacher quality means the impact of a teacher on students' performance in exams.

for teachers will allow us to measure that part of teacher performance that is constant across pupils with different characteristics/abilities etc. In order to do this, we need to include a dummy variable for each teacher in the data set, which pick up the impact of all invariant (across pupils) teacher characteristics, whether observed or unobserved. The key point in the analysis is that I observe different teachers teaching the same pupils in the same schools. Therefore, pupil and school characteristics are in effect being held constant, and any variation in pupil performance across subjects for the same pupil is attributed to the effects of different teachers. Thus, the teacher fixed–effects achievement function can be estimated as follows:

$$Y_{ijsk} = \beta_0 + \beta_1 X_i + \beta_2 S_k + \alpha_j + \delta_s + \varepsilon_{ijsk} \qquad \text{(Eq 2)}$$

α_j is treated as a set of fixed parameters, named as teacher fixed effects and δ_s is the subject fixed effects to control for systematic variation in marks by subjects (for example if Mathematics systematically receives higher marks than English). Now the error term is ε_{ijsk} capturing the effect of individual, teacher, subject and school level unobservables.

This step (first stage) yields estimates of a fixed effect for each teacher, which gives an indication of teachers' contributions towards pupil performance. Therefore, the coefficients estimated, due to the inclusion of teacher fixed effects, are looking at how each teacher does with the pupils (within a school) relative to other teachers with the same pupils.

However, there are issues to be aware of. The main issue faced by every study in estimating the causal effect of teacher characteristics and student achievements is the potential bias due to non-random sorting of teachers and students to classrooms. For example, teachers with strong credentials may be assigned to the classes with unobservably more able and motivated students or indeed, vice versa.

This is particularly an important issue for any study taking into account the high school exams, as high school level students have more options and opportunities to choose their field of study, and ability-tracking[19] (ability grouping) is more prevalent than at the primary education level. However, ability grouping is not a common phenomenon in South Asia in general and particularly in Pakistan (Aslam and Kingdon, 2011). If the estimated teacher effect does not fully allow for non-random matching of pupils to teachers then it may be contaminated by pupils' unobservable characteristics such as ability and motivation. The estimates of teacher effects would be biased upward if more motivated and greater ability students are assigned to more effective teachers, and vice versa if school policy or education administrators deliberately assigned lower ability pupils to the more effective teachers.

The standard way to deal with this problem is to use longitudinal data. Longitudinal data provide the researcher with multiple measures, for instance, test score in mathematics for each student for multiple years. This property of data will allow a researcher to include in the model student fixed effects and thereby to control statistically for unobservable time invariant characteristics of individuals, such as ability and motivation, which could possibly be correlated with teacher characteristics (Kane et al., 2008).

The pupil fixed effects approach, which is basically within student estimation, addresses the problem of non-random assignments of teacher to students by identifying the effects of teacher attributes through exploiting within student variation in the data. However, this approach is less suited for high school level as at that level multiple outcome measures for the same subject are not available over time and also it is costly and time consuming to collect such data. Alternatively, similar analysis can be possible when test scores are available for multiple subjects for the same student. It will produce similar benefits of multiple measures of score for a single pupil. The dataset used in this study has variation in test scores for multiple subjects for the same student. For example, each student has test scores recorded for four subjects: Language, English, Mathematics and Religion that enable me to incorporate student fixed effects to control

[19] The educational practice of ability grouping emerged around the turn of the 20th century as a way to prepare students for their "appropriate" place in the workforce. Students with high abilities and skills were given intense, rigorous academic training while students with lower abilities were given a vocational education. Evidence suggests this is prevalent in UK.

for non-random assignments of students and teachers and as a control for prior ability, which would otherwise be a limitation of the present study.

However, one of the consequences of using the pupil fixed effects model is one cannot use any characteristics that do not vary across subject, such as gender, thus most of the student characteristics are lost. To resolve this issue a separate model will be estimated in addition to the previous one, using rich data on the students' own, family and household characteristics, which are available in the current dataset. Results can then be compared for both models using either pupil fixed effects or pupil characteristics. Hence two models are estimated, one controlling for pupil fixed effects and the other for pupil characteristics to see which teachers are doing better with the same pupils (within the same school) and which teachers are doing better with pupils with the same observed characteristics. This is an advantage of the current study over other studies where they cannot estimate both a pupil fixed effects specification and a specification with pupil characteristics in the same study. For example, the Azam and Kingdon (2015) study lacks student characteristics in their administrative data.

A second issue to consider when estimating the effects of teachers on pupil outcomes is the potential source of endogeneity due to the possible correlation between teacher unobservables such as ability, motivation and effort captured in the error-term and teacher characteristics used in an achievement equation. For example, teacher education and motivation may be positively related to each other. The studies by Kingdon (2006) and Aslam and Kingdon (2011), and other panel data studies, also face this problem. These studies incorporated pupil fixed effects in their equations, and used teacher characteristics as separate variables in an achievement equation.

The teacher fixed effects approach used in this study does not include teacher characteristics as a separate vector in the first stage. Rather, it includes teacher fixed effects, which pick up all invariant characteristics of teachers across pupils, whether observed or unobserved, in order to get an idea of the relative influence of a particular teacher on pupils' performance. This is aiming to identify the overall teacher quality, measured as the total effect of a teacher in a general way (Aaronson et al., 2007; Rockoff, 2004; and Rivkin et al., 2005).

Another methodological advantage of the teacher fixed effects approach is that it does not suffer from the problem of non- random attrition of pupils and teachers in a given sample over time that exists in panel data studies. Non-random attrition over time creates a bias, upward or downward, in the estimates of teachers' effects depending on the characteristics of the teachers leaving the survey (Rivkin, 2005). The current methodology is exploiting the variation within teachers across different pupils at a single point in time (in Year 9). So I do not have the problem of non-random attrition due to the nature of analysis.

3.4.2 What explains teacher quality?

The additional feature of the analysis is that it regresses measured teacher quality on to different teacher and school characteristics in order to see whether observed teacher and school characteristics have any explanatory power on estimated teacher quality obtained from equation (2). Therefore, by the end of the analysis, it leads to an equivalent destination to that of the earlier studies, namely which particular teacher characteristics raise pupils' performance, but in an unbiased and consistent way (Aronson et al., 2007).

$$\hat{\alpha}_j = \lambda Z_j + u_j$$

In this equation $\hat{\alpha}_j$ is the estimated teacher fixed effects, measured from the first stage regression using equation (2); Z_j is the vector measuring teacher characteristics such as age, education, training experience, process variables and school variables (class size, private school); and u_j is the error term. The second stage relates the estimated teacher fixed effect to the teacher's own characteristics and process variables by regressing measured teacher fixed effects on to teacher characteristics and classroom practices, in order to further know which specific characteristics drive the differences in performance picked up by these teacher fixed effects. So, the second stage determines which characteristics make a teacher better than other teachers.

Another concern faced by different researchers is isolating the pure teacher effects from other inputs to the educational process, such as school variables and class-room variables. Slater et al. (2012) mentioned that all teachers in their sample remain in the same school over two periods. Thus, it is impossible to separate the pure teacher effect and the school effect. The present study faces a similar issue, since teachers do not work in more than one school and so the school effect cannot be separately identified from the teacher fixed effects. In order to attempt to isolate the teacher effect, I also control for the quality of school in the second stage.

3.4.3 Descriptions of variables

Some of the variables used in this study are very different to those used in the previous literature investigating teacher characteristics and pupil performance in developed countries. For example pupil age is very important in the context of Pakistan. According to The World Bank data, the minimum age to start primary education in Pakistan is 5 years. However, there is no enforcement of this, parents can send their children late to schools, particularly in rural areas where children start their schooling late as compared to their counterparts in urban areas. The pupil age variable in my survey has a huge variation, the maximum is 25 years and the minimum is 15 years. The median value of age is 20. A description of variables used in this study is given in Table 3.2. Descriptive statistics of variables are given in Table 3.3.

Family income is the sum of income of all individuals in the pupil's household. Most of the households live in a joint family system, where any individual such as elder siblings in addition to father and mother may also be working and have an income and be sharing towards family expenditure. Therefore, after asking about father's and mother's education, a separate question was designed in the family level questionnaires, to ask, who else is working at home and what their income was.

Food frequency is measured as the number of meals taken per day, with the breakfast variable used additionally to see if the pupil has other meals but usually skips breakfast, as eating breakfast habitually boosts performance at school (Adolphus et al., 2013). It is very obvious that not having the basic three meals per day and missing breakfast can have a negative effect on the health of young people which further leads to poor academic performance due to them being malnourished. Food frequency and breakfast variables are likely to be highly correlated with the income variables but in this dataset the correlation between family income and breakfast and family income and food frequency is 0.09. The correlation between food frequency and breakfast is 0.28.

The mode of transport variable is used to proxy the level of poverty. The schools are not located in pupils' postcode areas. Therefore, rich pupils can use a car, bus or motorcycle to reach schools while poor pupils use either cycle or they walk to schools. Both variables measuring food taking frequency and mode of transport are taken as proxies for a disadvantaged background. In a country like Pakistan where more than half of the population lives below the poverty line[20] and do not have enough basic meals (breakfast, lunch and dinner), these variables in this context are very important to capture the effect of poverty/disadvantage. Generally, in Pakistan breakfast, lunch and dinner are considered as basic meals. However, due to poverty many people do not have all of them.

The tuition variable is very important in the education system of developing countries, as in Pakistan in the evening most teachers provide children with extra tuition and earn extra income in addition to their school job.

Studying time at home is also important as many students have to help at home in household chores, especially girls, which can negatively affect their performance at school.

The health problems variable is not measuring minor ailment, instead the question is asked in pupil questionnaires "do you have any chronic health problems affecting your studies".

In the teacher dataset, the teacher training variable is given a value of 1, if a teacher has received a specialist training called B.Ed or M.Ed. B.Ed is the Bachelor of Education, and is a one year specialist course offered to those who wish to take up the teaching profession. The minimum qualification required for B.Ed is the Bachelor degree. Similarly, M.Ed is the Masters of Education, and is a one year course. The minimum requirement for doing the M. Ed is masters. In all government schools, it is mandatory for teachers to have specialist training in order to teach in primary or high schools, while it is not necessary for the teachers of private schools.

[20] Pakistan Economic Survey (2013-2014)

Table 3.2 Description of variables

Variables	Description of variables used
Pupil and home level variables	
Score Year 9 (Dependent variable)	Examination score measured in Year 9
Pupil age	Age of pupil measured in years
Female	Dummy variable 1= female 0 = male
Log of family income	Log of income of all persons' earnings
Own house	Dummy variable 1= parents own house 0= rented house
Physical activity per week	Dummy variable 1= yes 0 = no
Study time at home per day	Measured in hours studying at home
Health problems affecting schooling	Dummy variable 1= yes 0 = no
Mode of transport	Dummy variable 1= if pupil reaches school by car or bus, 0= cycling or walk
Fathers' education	Father education measured in years of schooling
Mothers' education	Mother education measured in years of education
Family size	Number of persons living in house
Birth order	Individual's standing among other siblings
Siblings	Number of brothers and sisters
Food Frequency	Number of meals per day
Breakfast	Dummy variable 1= yes 0 = no
Tuition (private coaching)	Dummy variable 1= yes 0 = no
Urban	Dummy variable 1= urban 0 = rural
Subject Variables	
Subjectdum1(Language Y 9)	Dummy for subject 1
Subjectdum2 (English Y 9)	Dummy for subject 2
Subjectdum3 (Mathematics Y 9)	Dummy for subject 3
Subjectdum4 (Religion Y 9)	Dummy for subject4 – omitted category
Teacher Variables	
Teacher age	Age of teacher measured in years
Teacher education	Education of teacher measured in years of schooling
Teacher training	Dummy variable 1= if teacher has a specialist training 0 = no
Teacher experience	Teaching experience measured in years
Teacher experience2	Square of teaching experience
Teacher workload	Number of classes a teacher is teaching per day
Authoritarian Behaviour	Dummy variable 1= teacher has authoritarian behaviour in class 0 = Democratic behaviour
Class duration	Number of minutes spent in teaching a class in a day
Class size	Number of students in class room
Teacher salary	Monthly salary of teacher
School Variables	
Private school Year 9	Dummy variable 1= private school 0= public school in Year 9

Table 3.3 Descriptive summary of variables

Variables	Observations	Mean	S.D.	Min.	Max.
Pupil and home level variables					
Score Year 9	1876	37.80	16.19	3	75
Pupil age	1876	20.22	1.25	15	25
Female	1876	0.46	0.50	0	1
Log of family income	1876	9.39	0.98	6.90	13.81
Own house	1876	0.78	0.41	0	1
Physical activity	1876	0.88	0.30	0	1
Study time at home	1876	2.32	1.31	0	10
Health problems	1876	0.28	0.45	0	1
Mode of transport	1876	0.41	0.49	0	1
Fathers' education	1876	10.27	5.16	0	20
Mothers' education	1876	6.80	5.81	0	18
Family size	1876	7.70	3.10	3	32
Birth order	1876	2.95	1.78	1	10
Siblings	1876	4.58	1.99	0	16
Food Frequency	1876	2.79	0.58	1	6
Breakfast	1876	0.73	0.44	0	1
Tuition	1876	0.60	0.49	0	1
Urban	1876	0.92	0.27	0	1
Subject Variables					
Subjectdum1(Language Y 9)	1876	0.25	0.43	0	1
Subjectdum2 (English Y 9)	1876	0.25	0.43	0	1
Subjectdum3 (Mathematics Y 9)	1876	0.25	0.43	0	1
Subjectdum4 (Religion Y 9)	1876	0.25	0.43	0	1
Teacher Variables					
Teacher age	56	36.80	9.55	22	59
Teacher education	56	15.76	0.73	14	18
Teacher training	56	0.64	0.48	0	1
Teacher experience	56	11.61	8.40	0	33
Teacher workload	56	4.35	1.52	2	7
Authoritarian Behaviour	56	0.32	0.47	1	0
Class duration	56	39.73	1.13	35	40
Class size	56	47.32	31.35	12	110
Teacher salary	56	10071	4499.92	4500	15000
School Variables					
Private school	56	0.57	0.49	0	1

Figure 3.1 and 3.2 shows the average score across the 4 subjects for each pupil in private and public schools.

Figure 3.1 Average score of pupils in private schools using Kernel density

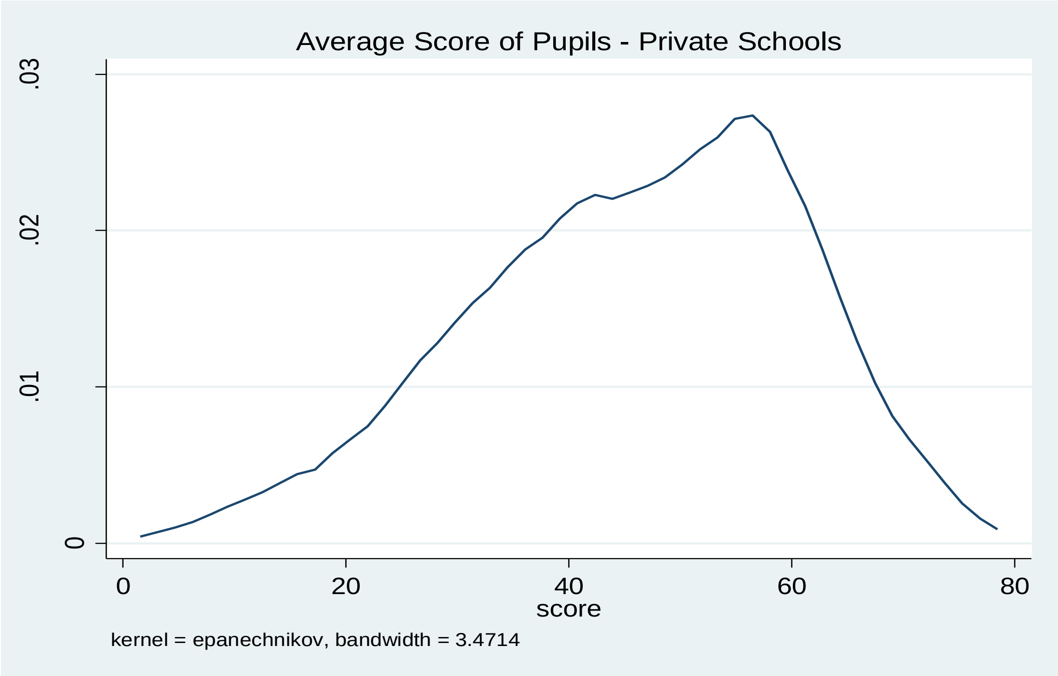
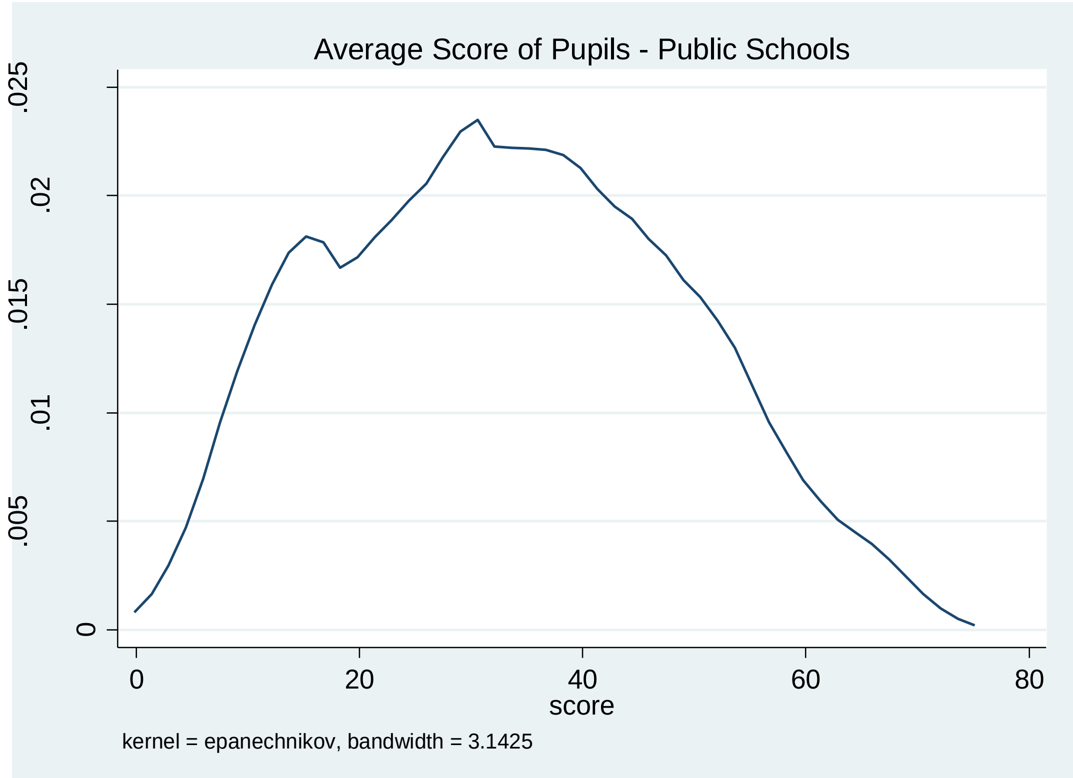


Figure 3.2 Average score of pupils in public schools using Kernel density



3.5 Results

This section considers what is the size of the teacher effects (3.5.1 below) - i.e. what is the difference in scores between having a good teacher and a bad teacher? How does this difference in scores compare to the overall distribution of scores. One way to quantify this would be to express the difference in scores between having a good teacher and a bad teacher in terms of the number of standard deviations of the overall test score distribution. I also used inter-quartile range (IQR) to describe teacher fixed effects.

3.5.1 Estimation of teacher effects (Baseline Model, first stage results)

First, I estimate the model which includes teacher dummies and subject dummies and the rich vector of individual and household characteristics as explanatory variables. The dependent variable is score in Year 9, with a standard deviation of 16.13, a maximum value of 75 and a minimum of 3. The score distribution has median 38 and mean 37.37, so it appears to be normally distributed and not truncated, exhibits substantial variance, all of which further suggests that score in Year 9 is the appropriate measure of pupil performance of the sampled population. The distribution of scores is given in column 1 of Table 3.4 and Figure 3.3 shows the distribution.

Figure 3.3: Distribution of Score Year 9 (Dependent variable)

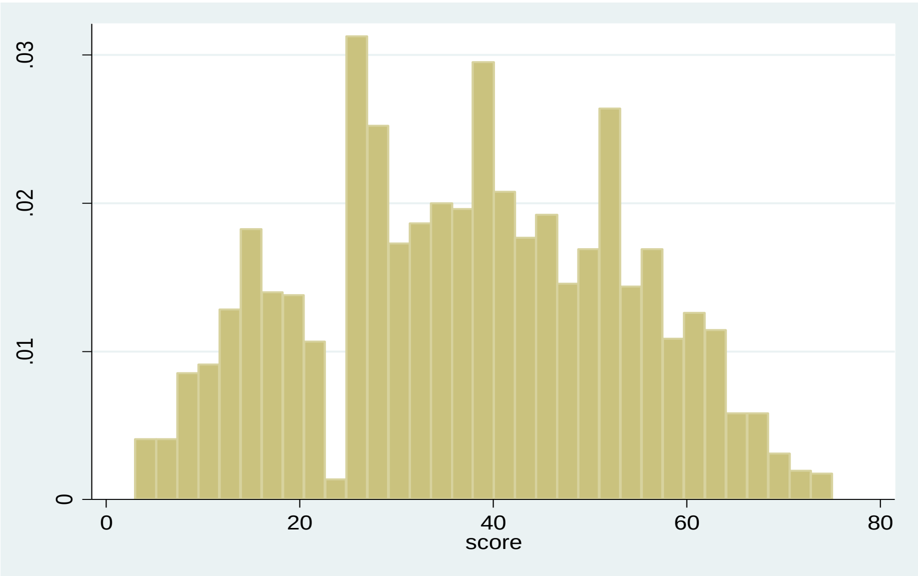


Table 3.4: Distributions of score Year 9 and teacher fixed effects

	Score Year 9	Teacher fixed effects1 Model-1	Teacher fixed effects 2 Model-2
1^{th} Percentile	6	-25.21	-43.44
5^{th} Percentile	11	-22.73	-37.24
10^{th} Percentile	15	-19.47	-32.08
25^{th} Percentile	25	-12.45	-20.05
50^{th} Percentile	38	-4.35	-7.31
75^{th} Percentile	50	2.49	2.54
90^{th} Percentile	59	8.32	12.79
95^{th} Percentile	63	11.53	18
99^{th} Percentile	70	21.45	37.31
99-1 gap	64	46.66	80.75
90-10 gap	44	27.77	44.78
75-50 gap	12	6.84	20.1
75-25 gap	25	14.94	22.59
50-25 gap	13	8.1	12.74
Standard deviation	16.13	10.67	17.88
Mean	37.36	-5.21	-8.77
R-square	-	0.57	0.74
P- vales for F-test	-	0.000	0.000
Teacher fixed effects	-	Yes	Yes
Subject fixed effects	-	Yes	Yes
Pupil characteristics	-	Yes	No
Pupil fixed effects	-	No	Yes
Observations	-	1876[21]	2361[22]

[21] The number of observations is different in teacher fixed effects model 1 and 2, because of missing values on some of the pupil characteristics.

[22] I did the pupil fixed effects on the same sample as the pupil characteristics model, i.e. 1876 observations, and found broadly the same results. I am not comparing the results between the two methods, but rather presenting a different way of deriving the teacher fixed effects. Given the number of fixed effects being estimated in the pupil fixed effects regression, I need as many observations as possible hence why I keep the sample at the largest possible size.

Table 3.5 below describes the first stage results of equation (2) using pupil variables, teacher fixed effects and subject fixed effects (to control for systematic variation in marks across subjects).

Table 3.5: Pupil- level regression with teacher and subject fixed effects

Variables	Coef.	S.E.
Pupils' age	-0.244	3.529
Pupils' age square	-0.027	0.087
Female	6.963	1.189***
Log of family income	1.090	0.399***
Own house	-0.689	0.634
Physical activity	0.497	0.875
Study time at home	-0.392	0.216*
Health status	-0.680	0.592
Mode of transport	-0.957	0.673
Fathers' education	0.009	0.068
Mothers' education	0.121	0.068*
Family size	-0.103	0.101
Birth order	0.126	0.180
Siblings	-0.207	0.188
Food Frequency	0.902	0.480*
Breakfast	1.799	0.622***
Tuition	-0.142	0.568
Urban	0.649	1.018
Subjectdum1(Language Y9)	14.059	4.900***
Subjectdum2 (English Y9)	7.868	4.105*
Subjectdum3 (Mathematics Y9)	5.893	4.900
Constant	37.364	35.756
Teacher fixed effects	Yes	
Pupil fixed effects	No	
Subject fixed effects	Yes	
Observations	1876	
R^2	0.57	

* p<0.10, ** p<0.05; *** p<0.01,
Standard Errors in parentheses

The important part of the analysis is the teacher fixed effects (Table 3.4). The importance of teacher quality can be measured by the variation in teacher fixed effects (Rockoff, 2004). Here each teacher is measured relative to the omitted teacher, who therefore has a relative fixed effect of zero, and the omitted teacher was deliberately chosen to be in the middle of the distribution.

There is significant variation in teachers' effects; some teacher effects are positive and highly significant while others are negative and significant. The estimated teacher fixed effects have a 10.66 standard deviation which is broad. Column 2 in Table 3.4 describes the teacher fixed effects in Model 1. Azam and Kingdon (2015) find a 0.511 standard deviation of estimated teacher fixed effects using a value added model to estimate teacher fixed effects and administrative data from one of the districts of India.

An exceptionally good teacher having a maximum score [23] (at the 99th percentile) is related to an increase of 1.32[24] standard deviations in pupils' test score, while an exceptionally bad teacher (defined as being at the 1st percentile) has a detrimental impact on scores of 1.55 standard deviations both relative to the reference teacher.

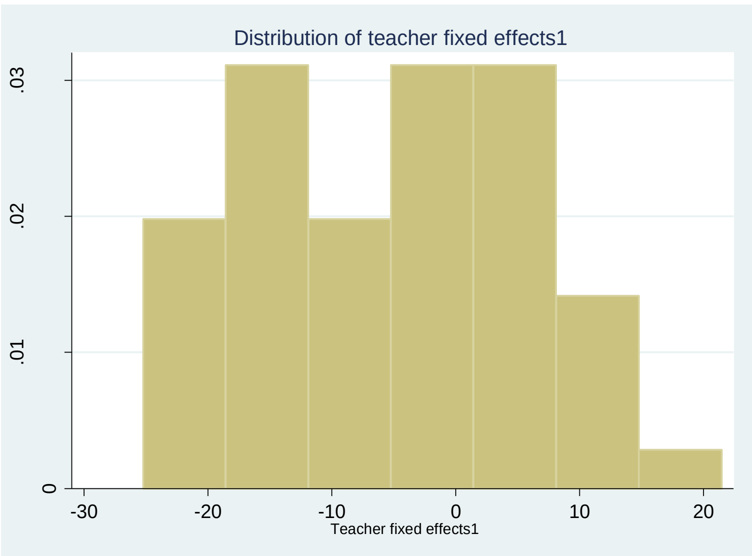
The most important thing to focus on when discussing these results is the gaps (75-25 gap) as it is the difference between the quartiles that is important. The actual size at each percentile will depend on which teacher was chosen as the omitted category i.e. the actual values at the quartiles depend on who the omitted teacher was when estimating the fixed effects, since they are all measured relative to that omitted teacher. So changing the omitted teacher would change the values at the quartiles. But the difference between the quartiles would not change. In effect, by measuring both relative to the omitted teacher then calculating the difference in 'performance' is the difference between the teacher at the 75th and the teacher at the 25th percentile.

[23] Different studies used different definitions for good and bad teachers, for instance, Slater et al. (2012) defined a good teacher as being at the 75th percentile and a bad teacher at 25th percentile, while other studies, Azam and Kingdon (2015) and Clotfelter et al (2010) defined a very good teacher as being at the 90th percentile and a very bad teacher as being at the 10th percentile. The present study uses the definition of good and bad teachers placing them at the 75th and 25th percentiles respectively.

[24] This number is obtained by dividing the teacher fixed effect number at the 99th percentile by the standard deviation of the dependent variable (pupils' score) as reported in column 2 of Table 4.

A good teacher defined as being at the 75th percentile is related to an increase in score by 0.15 of a standard deviation, while a bad teacher decreases the score by 0.77 of a standard deviation. Therefore, a pupil having been taught by a good teacher (75th percentile teacher) scores 0.92 of a standard deviation more than a pupil who is taught by a bad teacher (25th percentile teacher) leaving a significant effect on pupil performance. Clotfelter et al. (2010) find a difference of 0.23 standard deviations in the predicted student achievement between the good and bad teacher, whereas the education literature finds a moderate difference of 0.20 standard deviations (Azam and Kingdon, 2015). Also comparing the difference of 0.92 standard deviations in score, to other variables included in the model, teacher quality has the biggest impact in terms of standard deviation on score. Taking into account the whole distribution of teacher scores, it reveals a range from -25.21 to 21.45 with a -5.21 mean value. The standard deviation is 10.67, which is about 0.661 of the standard deviation of the pupil scores, depicting a large dispersion in teacher score from the mean value -5.21.

Figure 3.4: Distributions of teacher fixed effects in Model 1



Briefly considering the other coefficients in the education production function reported in Table 3.5, studying Language in Year 9 is related to an increase of 0.86[25] standard deviations in score, relative to Religion (the omitted category). Similarly, English as a subject is related to an increase in score up to 0.48 standard deviations, relative to Religion. Mathematics is not different to Religion scores in Year 9. These systematic differences show the importance of controlling for subject taught.

Most of the other coefficients on individual and household variables hold expected signs. The coefficients on pupil age and pupil age squared are insignificant. The coefficient on female is positive and significant at the 1% level of significance, showing that being female is related to an increase of 0.42 standard deviations in exam scores, relative to male. The above finding is similar to Slater et al. (2012) for England that females score higher than males. Also it is broadly accepted that females perform better than males. The coefficient on log of family income is also positive and highly significant. It shows that a 1 standard deviation increase in family income leads to an increase of 0.07 standard deviations in exam scores. There is an established link between income and child educational outcomes, as shown in the research undertaken in chapter 2. The studies of Blanden (2004) and Galindo-Rueda and Vignoles (2005) find this for UK. Higher income allows parents to assist their children through different facilities such as books, private coaching and other supplementary materials helping children in their education.

The coefficients on the variables indicating own house and physical activity are insignificant. The study time at home coefficient is negative and significant at the 10% level of significance; it means an additional hour of studying at home is associated with a 0.02 standard deviation decrease in score. This could arguably be due to those students who are already struggling in studies; they have to study more hours at home.

Health status and mode of transport have insignificant coefficients. Fathers' education is not related to score of Year 9, while the coefficient on mothers' education is significant at the 10% significance level meaning that an additional year increase in mothers' education has an impact of .0007 standard deviations increase in score.

[25] These standard deviations are obtained by dividing the respective coefficient values, β by the standard deviation of the dependent variable, score.

Alderman et al. (2001) find a strong positive link between maternal education and child achievement, however their study was based on primary school education.

Family size, birth order and number of siblings have insignificant coefficients, though family size and number of siblings have negative signs which mean that more persons at home and a higher number of siblings reduce the exam score. Food frequency measured as number of meals per day turns out to have a significant coefficient at the 10% level, showing that an additional intake of a meal leads to an increase of 0.05 standard deviations in exam scores. The coefficient on having breakfast is significant at 1% level of significance; this means having breakfast is related to an increase of 0.11 standard deviations in exam scores, relative to those who miss breakfast.

Tuition (extra coaching) is negatively related to exam scores and urban is positively related to score, though both have insignificant coefficients. The R^2 shows that 57% of the variation in the dependent variable is explained by these variables, which is relatively high.

3.5.2 Pupil fixed effects model (first stage results)

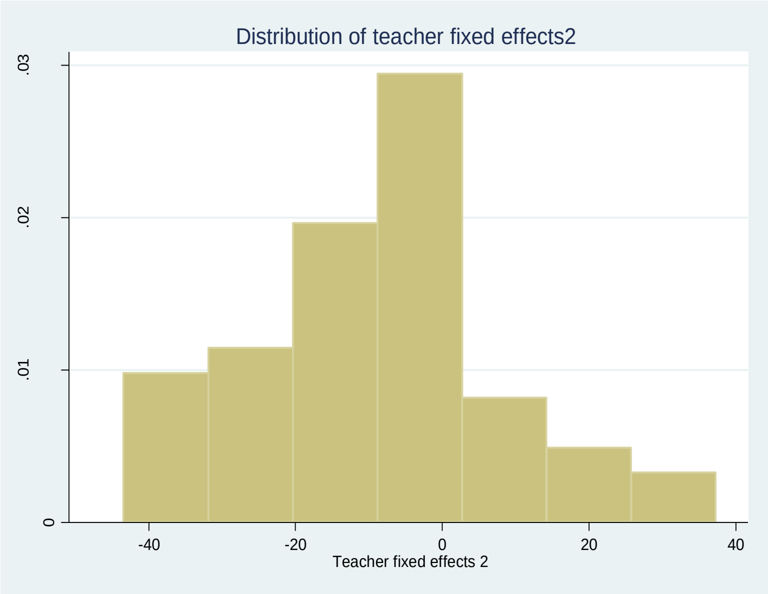
The above model may not be able to capture the full heterogeneity in students and family backgrounds. To capture those, I introduce pupil fixed effects into the model. I estimate the same model as earlier in this section, but instead of pupil characteristics, I use pupil fixed effects with teacher fixed effects and subject fixed effects in the first stage[26]. The intuition behind including pupil fixed effects rather than pupil characteristics, is to see which teachers are doing better with the same pupils (and so within the same school) and not just which teachers are doing better with pupils with the same observed characteristics. It is also important to control for pupil fixed effects, given that I do not have prior ability controls, so this is only way to fully control for the type of students each teacher has.

I did not use the school fixed effects, since each teacher only works within a single school, hence there is no variation for teachers across schools. Looking at the first stage results it does not make a huge difference compared to the previous specification, and

[26] The first stage results from the pupil fixed effects model are not presented in the chapter as there are no coefficients values in the model due to using pupil fixed effects, teacher fixed effects and subject fixed effects in the specification.

in fact some of the teacher fixed effects coefficients are larger in absolute size. The distribution of teacher fixed effects in model 2, shown in Figure 3.5, is more dispersed than the teacher fixed effects in model 1.

Figure 3.5: Distributions of teacher fixed effects in Model 2



The standard deviation of teacher fixed effects in model 1 is 10.67 and teacher fixed effects in model 2 is 17.88. The inter-quartile range (IQR) is 22.9 while it was 14.94 in the case of teacher fixed effects in model 1. This finding is contrary to previous studies, Slater et al (2012), find that the estimation error is smaller in the pupil fixed effects model relative to the pupil characteristic model, suggesting that the pupil fixed effects model is more precise at estimating teacher fixed effects. Azam and Kingdon (2015) find a 0.379 standard deviation of teacher fixed effects in model 2 (pupil fixed effects) and a 0.513 standard deviation of teacher fixed effects in model 1 (pupil characteristics), however, both studies use a control for prior ability, and it may be this fact that makes the results of this study different to theirs.

Using these new results, a pupil taught by a good teacher (being at the 75th percentile) has an increase in exam scores of 0.15 standard deviations, while a bad teacher (being at the 25th percentile) decreases the score of a pupil by 1.2 standard deviations. Therefore, a pupil taught by a good teacher (75th percentile) scores 1.35 standard

deviations more than a pupil who is taught by a bad teacher (25th percentile). The R^2 shows 74 % of the variation in the dependent variable is explained by these variables.
The results with the pupil fixed effects are theoretically more relevant, given that they control for the individual pupils' characteristics (at least their fixed characteristics that are constant across teachers). Rockoff (2004) highlights that the most credible way to estimate teacher effects is to regress test score on teacher dummy variables and controlling for variation in student characteristics and other classroom specific variables.

3.5.3 What explains teacher quality?

From the first stage results, it has been clear that teacher quality is an important determinant of pupil performance; however, it is important to know what explains variation in teacher quality. Is it the teacher observable variables or process related variables or school related variables that explain variation in teacher fixed effects? To examine this issue further, I regress estimated teacher fixed effects, against teacher and school variables.

Results are given in Table 3.6. The dependent variable is teacher fixed effects from model 1; using pupil characteristics, subject fixed effects and teacher fixed effects in the first stage regression. The amount of variation in the estimated teacher effect explained by these variables is low- R^2 is 0.20. Only 2 out of 11 variables have a significant coefficient, age having a non-linear relationship with teacher quality. As a teacher becomes older, beyond a certain point, it improves the teacher quality. It could be due to the fact that as a teacher becomes older he is likely equipped with more experience and teaching techniques, hence it may improve teacher quality. Model (4) uses a variable, teacher salary as an additional regressor, which is negatively related to estimated teacher quality, but has an insignificant coefficient. These results have an important implication that it is not observable characteristics that make a good teacher.

Table 3.6: Explaining teacher quality (teacher fixed effects in Model 1)

	Teacher Quality Model (3)	**Teacher Quality Model (4)**	**Teacher Quality Model (5)**
Teacher age	-3.57** (1.79)	-3.32** (1.69)	-3.39** (1.62)
Teacher age^2	0.043** (0.02)	0.04** (0.02)	0.04** (0.19)
Male teacher	-0.38 (3.52)	-1.37 (3.36)	0.59 (3.24)
Teacher education	-1.61 (2.30)	0.40 (2.32)	1.43(2.28)
Teacher training	-1.05 (4.56)	-2.29 (4.33)	-1.38 (4.17)
Teacher experience	1.50 (1.11)	1.76 (1.05)	1.80* (1.00)
Teacher experience^2	-0.04 (0.03)	-0.05 (0.03)	-0.04 (0.02)
Teacher workload	0.97 (1.18)	0.68 (1.12)	0.61 (1.07)
Authoritarian behaviour	0.10 (3.46)	-0.50 (3.27)	-0.64 (3.14)
Class duration	-0.45 (1.41)	-1.14 (1.36)	-1.09 (1.30)
Class size	-0.07 (0.09)	0.09 (0.11)	0.20 (0.12)
Teacher Salary (Rs)		-0.002 (0.00)	0.00006 (0.00)
Private school			27.59** (13.03)
Constant	97.66	97.01	37.59
Observations	53	53	53
R-squared	0.20	0.30	0.38
P- vales for F-test	0.51	0.18	0.07

Dependent variable is teacher fixed effects1 derived from table 5 using teacher fixed effects, subject fixed effects and pupil characteristics.
* *p*<0.10, ** *p*<0.05; *** *p*<0.01
Standard Errors in parentheses

As the first stage regression results do not control for school effects, so the estimates of teacher effects were not purged of the school effects, so it is important to control for school effects at the second stage. Strictly speaking, the teacher fixed effect estimated in the first stage is a combination of teacher and school effects, which is why it is important to control for school effects in the second stage. I estimate a third version of teacher quality controlling for school quality.

The highly significant private school coefficient (coded as 1= private school, 0= public school) in the second stage results shows that the teachers with the best results are those teaching in private schools. So the private school teacher does better than public school teachers with pupils with the same characteristics. This is partly a consequence of not controlling for prior attainment, so that private school teachers may get the best results because they get the best students, not because they are the best teachers. It is clear by looking at the average scores of students by type of school, in Figure 3.1 and 3.2, that the private school pupils on average have higher scores (the mean score of private school pupils is 45.95 and 33.35 for public school pupils) than those from public schools. This could be due to the fact that the private schools have higher fees and they set strict admission criteria and have a limited number of spaces. These schools set their own entry tests and interviews to assess the average calibre of pupils before offering admission to them. Generally, they make sure to admit only those who can ensure higher results.

Although the model uses a good set of pupil control variables, it would seem that they are not doing enough to equalize pupil characteristics, between school types, and the best pupils are still going to private schools (holding constant the control variables mentioned above). To investigate this issue further, I estimate again these three second stage specifications, against teacher fixed effects, controlling for pupil fixed effects rather than pupil characteristics in the first stage. Results are given in Table 3.7.

Table 3.7: Explaining teacher quality (teacher fixed effects in Model 2)

	Teacher Quality Model (6)	**Teacher Quality Model (7)**	**Teacher Quality Model (8)**
Teacher age	-8.78*** (2.62)	-8.35*** (2.49)	-8.54*** (2.47)
Teacher age^2	0.11*** (0.03)	0.10*** (0.03)	0.11*** (0.02)
Male teacher	3.96 (5.39)	2.44 (5.13)	3.35 (5.11)
Teacher education	-2.54 (3.26)	0.51 (3.33)	1.25 (3.33)
Teacher training	-11.42* (6.52)	-13.54** (6.22)	-12.17* (6.23)
Teacher experience	1.53 (1.59)	1.98 (1.51)	1.92 (1.49)
Teacher experience^2	-0.05 (0.04)	-0.06 (0.04)	-0.05 (0.04)
Teacher workload	1.01 (1.79)	0.65 (1.69)	0.50 (1.67)
Authoritarian behaviour	-0.45 (5.16)	-1.20 (4.88)	-1.54 (4.83)
Class duration	-1.03 (2.04)	-2.04 (1.97)	-1.99 (1.95)
Class size	0.02 (0.13)	0.26 (0.16)	0.38** (0.18)
Teacher Salary (Rs)		-0.003** (0.00)	-0.001 (0.00)
Private school			25.89 (18.54)
Constant	235.47	230.23	181.56
Observations	53	53	53
R-squared	0.40	0.48	0.51
P- vales for F-test	0.01	0.00	0.00

Dependent variable is teacher fixed effects2 derived from table 5 using teacher fixed effects, subject fixed effects and pupil fixed effects.
* *p*<0.10, ** *p*<0.05; *** *p*<0.01
Standard Errors in parentheses

Overall results remain the same, but now the private school coefficient is insignificant (Model 8), suggesting that the pupil fixed effects are more successfully picking up the differences in pupil abilities than the observable pupil characteristics. Model 7 has a negative significant coefficient on salary, which implies that those good performing teachers are rewarded less. This is due to the fact that in the current data the mean salary of private school teachers is 6375 Rs. while the mean salary of public school teachers' is more than double that of private schools teachers' salary i.e. 15000 Rs., and private schools are performing better than public schools, therefore, the coefficient on teacher salary is negative.

Teachers' salaries in private schools are a small fraction of the teacher salaries in government schools. Private school employers pay to the teacher what is called the market-clearing wage (i.e. the wage that individuals would get if they entered into the open market given their level of qualifications). In addition, it is based on the local unemployment rate, which is high among graduates. There are, therefore, a large number of educated people available, who are unemployed and are presumably willing to take jobs at low pay. Private schools are able to take advantage of the low market wage that prevails and offer low teacher salaries.

On the other hand, teachers' salaries in public sector are bureaucratically set and governed by Sixth Pay commission or following the Basic Pay Scale. Those salaries have quite a lot of rinse in them, meaning that they are non-productivity related salaries; they are same for all teachers across the board irrespective of whether a teacher is productive or not, performing well or not, regularly comes to school or is a chronically absentee. So arguably private school teachers have some leeway to reward teachers based on efforts rather than having a blanket salary structure based on seniority and qualifications only (Kingdon, 2017).

Teacher training is negatively related to teacher quality but at the 10% level. This is likely to be because of reverse causality, the lower quality teachers are put in for more training (Aslam and Kingdon, 2011).

A comparison of results using pupil variables and pupil fixed effects suggests that the model controlling for pupil fixed effects has not made that much difference to the results - presumably because there were not many selection effects of pupils to teachers anyway, so that controlling for unobserved pupil characteristics does not make much difference. Although I only surveyed one teacher per subject in the present study there was more than one teacher teaching each subject in year 9 within each school, so there could have been an issue related to the school choosing which pupil is taught by which teacher within schools. There is therefore a reason for estimating the pupil fixed effects specification, to control for any unobserved characteristics of pupils. The pupil fixed effects method is a robustness check as changing from the pupil characteristics methodology to pupil fixed effects, the results remain similar i.e. the choice of specification does not make a substantial difference to the results. There still could be selection effects of pupils into schools however, so it is not completely random which teacher teaches which pupil. This can be seen in the changing significance of the private school coefficient between specifications.

3.6 Conclusions

The present study uses the survey data of 611 pupils studying in Year 9 (aged 14-15 years), from private and public schools of Pakistan in 2008-9, in one of the districts of Punjab, Pakistan. The survey data is linked with administrative data of the student exam scores, to address the issue of teacher quality using a teacher fixed effects approach. In particular, it evaluates the teachers in terms of the exam results of their pupils.

This chapter investigates two questions: Do teachers matter and what characteristics make teachers good or bad relative to other teachers. Teacher fixed effects and then pupil fixed effects techniques to address heterogeneity in terms of pupils and home backgrounds are used for the estimation of teacher quality.

The results have shown that teachers matter a great deal within schools. A good teacher defined as being at the 75th percentile is related to an increase in score by 0.15 standard deviations, while a bad teacher decreases the score by 0.77 standard deviations. Therefore, pupils having been taught by a good teacher (75th percentile teacher) score 0.92 standard deviations more than the pupils who are taught by a bad teacher (25th percentile teacher) leaving a significant effect on pupil performance. Therefore, improving overall teacher quality in schools does seem a right direction for public policy. If teachers are randomly assigned across schools then school assignment is not very important. This may not be true as it would mean it does not matter which school you get in, if every school has an equal chance of having a bad teacher. So in reality it seems likely that teachers try to cluster according to quality. This question cannot be answered definitively as the present data set does not allow us to separate the pure teacher effect from the school effect within a school.

The findings of the present study confirm the previous findings of teacher quality for developed countries as found in many studies in the US that there exists considerable variation in teacher effectiveness, thus confirming the findings of developed countries in a developing country setting in Pakistan. Azam and Kingdon (2015) also confirm this for India. In their study, teacher effects capture the impact of spending two years with a teacher, while the current study captures the teacher effects of spending one year with the teacher.

The above finding of teacher quality suggests that teacher quality is also important for academic performance at high school level, and that family background and ability are not the only important factors. Within schools, teacher quality differs with the same pupils. This is a systematic effect averaged over all pupils, rather than anecdotal for just one pupil. Similarly it is not due to the general level of marks in the teacher's subject, since the equations control for the systematic variation in subjects marks through subject fixed effects. The same pupil can score a significantly different score in different subjects in a particular grade given different teacher quality. This debate has clear implications for improving teacher quality. Students can benefit by improving average teacher quality at schools.

Rivkin et al. (2005) find that teacher quality is related to socioeconomic gaps in outcomes. Such a finding is also useful here: teacher assignment could play an important role in alleviating unequal outcomes existing in gender and social groups. Since the teacher quality has an effect for the entire class therefore it has a strong effect as compared to any student based incentives.

The other part of the analysis in this chapter aimed to identify whether observed characteristics and classroom methods and practices explain the differences in teacher quality. Similarly, as found in many studies for the US and the UK and India, I find that teacher characteristics explain very little of teacher quality. So, observable teacher characteristics do not explain teacher quality well. Though the current dataset is small, many other authors with larger datasets confirm this (see particularly Kane et al. (2008)). Therefore, in a setting where teacher observed characteristics are not related to teacher quality, policies that reward teachers on the basis of performance are more effective than those based on hiring and retaining teachers with particular credentials.

This suggests that it is hard to identify good teachers ex ante, but administrative data can be used to identify them ex post. Despite the conclusion that measuring teacher fixed effects is a productive path, the task remains incomplete due to lack of adequate data. As Slater et al. (2012) suggest, a greater role for performance management and personnel policies in schools is implied.

Also, policies related to teacher recruitment and progression must be rethought and redesigned. It could be done probably by looking at pupil progress analysis in the probationary period, sharper increases in pay on the basis of performance, mentoring, and more stringent hiring procedures (Kane and Staiger, 2002).

Finally looking at the above results and findings from previous literature on teacher quality, it is very clear that further research is needed in order to identify high quality teachers, and then how to recruit and retain them once they are identified. Following the common practice, almost everywhere, compensating teachers solely on the basis of education, experience and training is less likely to yield an increase in teacher quality. So it is crucial to identify alternative sources of information on teacher quality in order to design policies to increase student achievement.

Some caveats apply to the conclusions of the present study. The analysed sample consists of private and public schools from one particular district of Punjab; hence it cannot be claimed for other parts of the country as there might be geographical dimensions which are not captured by the current study due to data limitations. Also the sample size is small, future research with larger datasets may uncover important differences of teacher performance and personality.

Nevertheless, the present study managed to find some important findings about teacher quality in a setting of a developing country with much reduced selection issues. Many other studies face the issue that pure teacher effects cannot be separated from school effects, and the present study also faces a similar issue since it is very rare for a teacher to work in two schools at once. But the current dataset does have some observable characteristics of schools that can control for school differences, which is rare in other leading datasets. In addition, prior academic achievement such as score in grade 8 is not available in the current dataset, as used by other studies to deal with non-random sorting of teachers to pupils. However, a helpful feature of the data allows estimating a pupil fixed effects model along with the teacher fixed effects approach used in this study to address the issue that teachers are not randomly assigned to students within schools.

Appendix A to Chapter 3: Matching of students to teachers

This appendix describes the linking of pupils to the teachers. There are 3 data files; the pupil data file has all information on all pupils' own and family level characteristics, plus a unique identifier for each pupil. It was in wide format containing 611 observations on pupils. Also each pupil has their subject score for 4 subjects in Year 9, in separate columns for each subject. The dataset was converted from wide format to long format - i.e. to multiple observations for each individual with one observation for each subject they do, giving 611* 4= 2444 observations. The second data file is teacher data, which has teacherno and schoolno as identifiers and has all information from the teacher questionnaires. The third file is school head data, which has schoolno as an identifier and all information about school related variables. In the teacher dataset file, and by cross tabulating teacher subject with schoolno, there is only one teacher observed for each subject within each school. So given that I know which school each of the pupils went to, then for each subject exam result I simply assign the teacher who teaches that subject in that school (i.e., for example in school 5, I know which teacher teaches each subject in school 5, so for each pupil in school 5, I assign their English score to the observed English teacher in Year 9, their maths score to the observed maths teacher in Year 9, etc.) Hence I created a teacherno variable in the pupil data file, and then merged teacher data into the pupil data using the teacherno variable then school data into the master file using the schoolno variable.

Appendix B to Chapter 3: Data collection process

Before the inauguration of the actual survey, this study was approved by the relevant university department board of studies meeting. Prior to undertaking the actual survey, a pilot study was carried out over two schools, a private and a public school, to check the feasibility of the actual survey. As a result of the pilot, the questionnaires were altered accordingly. For example, it was found that the pupil questionnaire was very lengthy and a few questions the pupils could not really answer and so were re-written, while a couple of questions had language ambiguities that were made clear in the updated version of the pupil questionnaire.

A random sampling technique was used to select the schools sample from the entire school population within the district under consideration. Random sampling was insured through the opt for lottery method, as all the schools names were written on pieces of papers and put into a box, in total 15 schools were drawn from the box, those include 6 public schools and 9 private schools. Prior to the launch of present survey, all schools were informed before the beginning of survey and acquired their consent to let the author enter the school with the assurance of confidentiality and to collect information from pupils, teachers and schoolheads. Among the 9 private schools, one school refused to collect the data, and therefore, that private school was dropped. The remaining 14 schools that accepted the idea were our sample.

Once the sample of private and public schools were chosen using random sampling, 611 Year 9 pupils were selected within schools. Further, to gather pupils' information, a briefing was given to all Year 9 pupils in different class rooms, a day before collecting the actual information. In the briefing, it was explained to them that this is a research project related to their education and they were requested to fill out the prescribed self-completion questionnaires. The contents of the questionnaires and the whole process were explained to them under the assurance of privacy and confidentiality. The time span to fill out the pupils' questionnaires was 60 minutes and all were done in the presence of the author in all schools and in all classrooms.

Moving on to the ethical practices, all Year 9 pupils within class rooms were given an opportunity to withdraw or not to participate if they did not wish to do so, therefore, the selected sample is based on their choice to participate in present survey. They were advised to ask their parents if they wished their children to contribute to the survey. This was also communicated to them in the briefing session that once they and their parents were happy to join the survey, they would be further requested to collect the information of their father, mother and siblings. However, it was observed that young pupils were excited about participation in the survey and mostly pupils participated except those who were absent on that particular date. A few did not participate but we do not have their information as it was optional for them to take part in the survey. The pupils' participation in the current survey was observed as an independently random sample, since there was no basis of their selection in the survey except that participation was not mandatory as we were aiming for the fact that pupils' participation in survey would be a random process.

Once pupils were chosen within school and class rooms, then we collected respective parents' information using self-completion questionnaires given to young pupils which were returned to the school authorities in a sealed envelope, so that school heads, and subject teachers, collected all relevant information. It was suggested to pupils that where the parents were illiterate, they could fill themselves that information on their behalf, or if the mother was illiterate then the father could fill out the relevant mother information or seek help from another literate adult or elder sibling at home in filling out the parents' questionnaire. All parents of sample pupils responded to the parents' questionnaires and we have no missing parents' information.

Similarly, self-completion teacher and school head questionnaires were given to subject and head teachers and the response rate was 100% with no missing subject teacher and head teacher information.

From school records, information that contained unique board registration numbers was obtained. To make sure that it was an authentic source, it was not asked of the pupils directly so there could be no chance of error, as it was long number composed of serial numbers, and letters. Subsequently, survey data were matched to the administrative data of pupils' exam scores of the higher education board using the unique pupil identifiers. Information on their exams scores were therefore observed in Year 9, Year 10, Year 11 to Year 12, over a four year period, to obtain longitudinal data of pupils exam scores.

During a follow-up survey, the distance variable was calculated by driving from each individual's home to the nearest post-compulsory educational institutions. Further details can be found in the data sections of Chapter 3 and 4.

To summarise, this is very comprehensive and rich dataset having information on all related aspects of young pupils' lives: education, families, subject teachers, school heads and exam sores. It is representative of rural, urban, private public, male, female and the mix of socioeconomic backgrounds. The setup of the survey was purpose-built with careful planning, meticulousness and almost no chance of measurement error

Chapter 4:

Selection Effects and Post-Compulsory Education in Pakistan

4.1 Introduction

Participation in post-compulsory education can be simply defined as continuing into further education beyond the age when it is no longer compulsory. There could be many benefits in doing so but most likely it could be related to higher future earnings. Substantial evidence is available that participation in higher education leads to higher future wages (see, amongst others, Oreopoulos, 2006; Leigh and Ryan, 2008; Neal and Johnson 1996; Walker and Zhu, 2003) and not going on welfare (Herrnstein and Murray, 2010). Therefore, it is important to increase participation in post-compulsory education and also it is an issue in education policy[27] how to increase participation in post-compulsory education. The above issue can be discussed in two perspectives, first, is how to increase the participation in post-compulsory education and the second is what are the factors that lead to the higher academic achievement of pupils once they are at the post-compulsory education level.

Regarding the participation in post-compulsory education, a considerable literature exists studying the factors that affect the decision to participate in post-compulsory education (for example, Ashford et al., 1993; Clark, 2011; Gray et al., 1993; Lenton , 2005; McIntosh, 2001; Payne, 1998 and Rice, 1999). Most studies have found that prior attainment and family socioeconomic status (SES) are the most important factors determining the participation in post-compulsory education. The intergenerational mobility literature has a direct relevance to this topic – suggesting that parents' education and income are positively related to children's participation in post-compulsory education (Chevalier, 2004). However, over and above the effect of these factors, geography and neighbourhood also determine the participation in college and

[27] Its importance can be seen in education policy, as in the UK, it is always a priority of government to increase school leaving age, most recently school leaving age has been increased to age 18 in 2015. However, it could be full-time education or spending 20 or more hours per week in training or volunteering.

university education or institutional choice (Gibbons and Vignoles, 2012; Gorard and Smith, 2006; Faggian and McCann, 2009; Long, 2004; Alam and Winters, 2009).

One important constraint while discussing participation in post-compulsory education is proximity to an institution providing post-compulsory education, which is seldom explored in research. Dickerson and McIntosh (2013) are the first[28] to study the impact of distance to an institution on the decision to participate in post-compulsory education in the UK. They find an overall small negative effect- higher distance reduces the participation in post-compulsory education. Therefore, one of the aims of the present study is to investigate the specific aspect of geography (distance to nearest institution providing post-compulsory education), along with other key drivers such as prior attainment and family SES, that influence participation in post-compulsory education, which hitherto has been given less attention relative to other determinants.

There are different reasons involved in studying this particular aspect of participation in post-compulsory education. One is the different cost associated with higher distances to the nearest Further Education institution. In financial terms it is the travel and reallocation costs involved, in terms of temporal, commuting time, and in psychological terms the inconvenience and unpleasantness of lengthy commuting every day. Human capital theory suggests that education is an investment decision taken by pupils anticipating that the present value of future benefits outweighs the present costs (Becker, 1994). Therefore due to costs involved in greater distances to nearest institutions, it has an obvious impact on the likelihood to participate in post-compulsory education.

Distance could have a differential effect on the participation in post-compulsory education in urban and rural settings. There are two things to be considered here. One, distances are likely to be less in urban areas relative to rural areas, and two, urban areas are equipped with better connectivity. So, living in an urban area served with better transport facilities, being distant for several miles from an institution might have a different impact to those living in rural areas for whom it may be hard to commute without private transport. It is important to know the link between distance and likelihood of participation as it is an issue for all pupils who do not live within easy

[28] Dickerson and McIntosh (2013) are the first to study the impact of distance on immediate post-compulsory education choices at age 16.

walking distance of their nearest education institution. So, type of area can make a difference in the decision to participate in post-compulsory education and availability of local schools/institutions delivering the desired program of study. Similarly, distance effects might vary by gender as it depends on whether parents allow their children to commute, especially in developing countries, where girls have less mobility and face more restrictions than boys regarding schooling (Lloyd et al., 2009).

Another reason for studying the impact of characteristics such as urban/rural on the likelihood to participate in post-compulsory education is that they are more amenable to change as compared to other long term determinants like prior attainment and home background. For instance, providing improved transport and subsidizing the travel cost both could have a clear effect in lowering the cost of education, which in turn could improve the likelihood to participate in post-compulsory education.

The other motivation of the current study is that it is jointly investigating the determinants of participation and performance in post-compulsory education. Previously no other study has investigated these two perspectives together: i.e. participation and performance at post-compulsory education level.

Pakistan is a developing country facing different challenges in its education system, high school dropouts being one of them. According to a local NGO[29], 35000 young people dropout from high school every year. Individuals having a high school diploma earn relatively more compared to those with a lower level of education (Frenette, 2006). Therefore, it is important to look at why young people drop out from education or for those participating in higher education, what are the main factors behind their participation. Pakistan is a good case for investigating the determinants of educational outcomes for the following reasons. First, despite the vast and growing literature examining educational outcomes in developing countries, it contributes to the literature on educational outcomes in Pakistan that is still in its infancy. Second, Pakistan is an interesting place to study the impact of accessibility and availability of higher education institutions on participation in post-compulsory education as Pakistan's educational system gives flexibility to everyone to choose any school, private or public according to their affordability. So there is vast variation in terms of enrolment and access across

[29] There are no official statistics available on school dropouts in Pakistan; this is the report of an NGO "Alif Ailaan" finding that Pakistan has the highest school dropout rate in the world.

schools in such an education system (Asahi, 2014). Also schools are not allocated according to catchment area as is the case in the UK. Therefore, there could be an issue of accessibility and availability of education institutes. Generally, poor people are most likely to send their children to public schools where the education is free and compulsory up to Grade 10 (age 15-16) and rich people are more likely to send their children to private schools which charge high fees.

The present study uses survey data on 611 pupils from one of the districts of Pakistan, studying in Year 12 at age 17-18 years, to explore the determinants of performance and participation in post-compulsory education, controlling for previous attainment in Year 9 and Year 10, the second last and last years of compulsory grades respectively. A Heckman model is used when analysing performance in order to control for potential sample selection bias resulting due to unobserved characteristics being associated with the decision to leave (or continue in) education after achieving compulsory education. Proximity to nearest post-compulsory institution measured as distance and living in an urban location are used as instruments for selection into post-compulsory schooling, since both measure the availability and accessibility of post-compulsory institutions. The results show a lack of significant selection effects, which suggests that participation and performance are not jointly determined. Participation is determined by the availability of post-compulsory institutions, while performance in Year 12 is driven by the ability which is measured by previous attainment. The results further highlight that distance reduces participation most for those living in rural areas.

To summarize the present study contributes to the literature in the following ways. (1) It is the first study (to the best of my knowledge) in the economics literature jointly investigating the determinants of participation and performance in post-compulsory education (Year 12) in developed and developing countries which further helps to understand whether participation and performance are jointly determined or they are independent processes. (2) It is the second study in the literature, and the first in developing countries, investigating the effects of proximity to nearest institution providing post-compulsory education. (3) It is also the first study in Pakistan looking at the determinants of performance in higher grades, Year 12 taken at age 17-18, conditional on participation. (4) The present study takes into account the differences of regional (urban/rural) disaggregation in the analysis of participation in Pakistan.

This study is different from the Dickerson and McIntosh (2013) study as they measure distance "as a crow files" which could be substantially different from real distance while this study measures the distance in a unique way by driving from pupils' homes to nearest education institution using data on homes and schools location. The latter arguably is the exact distance, previously all studies on distance have not measured distance using this method. It is similar to their study as it also looks at immediate post-compulsory education choices at age 16 but they do not consider performance in post-compulsory education.

The remainder of the chapter is structured as follows: section 4.2 reviews the relevant literature on performance and participation in post-compulsory education in developed and developing countries, section 4.3 describes the data and discusses the key variables used in the analysis, section 4.4 presents the methodology employed, description and summary statistics of variables used, section 4.5 discusses the results using OLS, Probit and Heckman selection methods, and section 4.6 discusses the policy implications and concludes.

4.2 Literature Review

This section reviews the literature, including developed and developing countries on determinants of academic achievements/performance in addition to participation in education.

4.2.1 Participation in Post-Compulsory Education

The participation in post-compulsory education is a question often explored in literature. Firstly, the level of education participation usually discussed is the post-compulsory (Dickerson and McIntosh, 2013) and university level (Gibbons and Vignoles, 2012; Spiess and Wrohlich, 2010; Frenette, 2006). Secondly, regarding factors, previous attainment, family SES, labour market conditions, and finally, the distance and rural-urban dichotomy are the ones most commonly discussed.

A study by McIntosh (2001) seeks to explain changes in the proportion of young people at the age of 16, 17 and 18 in four European countries (Germany, Netherlands, Sweden and England) who decide to participate in post compulsory education. The analysis further proceeds using the Engle and Granger two step method. It is a time series study rather than a cross section covering the period from 1960 to 1994. The findings reveal that prior academic attainment before the end of compulsory schooling is the key variable in explaining the continuation into post compulsory education. It implies that the better performance during compulsory schooling gives more confidence to female pupils in their own ability to proceed into further education than males. The impact of local labour market conditions measured as the level of youth unemployment has a small effect on the decision to participate into post compulsory education.

Similarly, Bradley and Lenton (2005) analyse the magnitude, timings and determinants of dropouts from post compulsory education in Britain using data from Youth Cohort Surveys (YCS) over the period 1985-1994. The results show that the risk of dropout depends on the young people's ethnicity, prior attainment, family socioeconomic status (SES) and the state of the labour market. Students are more likely to participate in education if the local labour market is weak as education becomes more attractive due to a high local unemployment rate. By the same token, Clark (2011) assesses the relationship between the local labour market and enrolment in post compulsory education in England. The key finding of the study is that local unemployment has a major impact on enrolment in post compulsory education in England. Clark finds a different result to earlier papers because he has a panel of regions and can estimate the unemployment effect off variation within regions over time.

Turning to papers looking at Pakistan, Lloyd et al. (2009) assess different factors responsible for primary and middle school dropout in Punjab, Pakistan over six years from 1997-2004. The dataset tracks longitudinal changes in both school and family environment. Results suggest girls dropout more than boys, and both school and family factors are significant. Attending a government primary school and the family experiencing unwanted birth (used as an instrument for family being affected by a household shock) will increase the likelihood of dropout, while living in a better-off neighbourhood, mother having any education and availability of post primary schooling reduce the chances of dropouts. For boys it is also true that living in a well-off neighbourhood will reduce the probability of dropping out.

An equally significant aspect of participation could be the geographical one discussed in the following literature. In terms of the post-compulsory participation education decision, Dickerson and McIntosh (2013) are the first to investigate the effect of distance to nearest education institution on a pupil's immediate post compulsory participation decision in England. Their study highlights the small overall effect of the distance variable; an additional kilometre in distance will lead to a 1.5 percentage points decrease in the probability of participation in post compulsory academic study. The marginal effects of the study show the strong correlation between post compulsory participation, prior attainment and family background. Also they found that distance affects women more than men. The distance has more significant impact on young people who are on the margin of participation in post compulsory education (according to their prior attainment and family background).

To be able to understand the effect of distance on the participation in university education, Frenette (2006) studies the link between distance from home to university on participation in university education using Canadian Household Survey data and matching that with a university postal codes database. The main finding of the study is that distance to nearest school negatively affects the probability of attending university. The pupils living within the commuting distance are more likely to attend university than the pupils who live far away, out of commuting distance. Further, distance to the school is found to matter more for those from lower-income families. In a more recent study, Frenette (2009) study shows that the creation of new universities increases the enrolment rate in local youth, which in turn suggests that distance does indeed matter.

Also, Spiess and Wrohlich (2010) analyse the impact of distance between home to the nearest university on the decision to participate in higher education using a German panel dataset. Estimates using a discrete choice model show that along with parental education and gender, distance indeed matters. Further they found that distance has a stronger negative effect for those who belong to lower income families, however, they do no control for the academic background of pupils. They measure three types of distance variable in their dataset: (1) distance to the nearest university, (2) distance to the nearest university of applied sciences, (3) the minimum of the two distances. They find that those who live further than 12 km are at a disadvantage in accessing university compared to those who live closer to the university, 6 km. Further, results show that the impact of distance is mainly driven by transaction costs rather than neighbourhood/spill over effects. They differentiate between transaction costs and neighbourhood effects using an interaction term between two variables "living in a university town" and the "student population density".

Additionally, Gibbons and Vignoles (2012) study the effect of distance on participation and choice in higher education (university) in England. Results find little effect of geographical distance on the participation decision, but a strong effect on the choice of institutions in England.

In the economics of education literature, proximity to relevant education institution is considered to be a reliable instrument, for example for education in a wage equation. The choice of instrument for schooling is generally an area of debate. It is often hard to find a valid exclusion restriction given data constraints; however, potential benefits justify the effort. The ideal experiment would be random assignment into education but that is rare. Dickerson and McIntosh (2013) suggest that in the field of returns to education, distance from an education institution is a valid instrument for education or qualifications acquired, while Card (1995), Dee (2004) and Rouse (1995) suggest proximity to college as a valid instrument for education attainment/schooling. The basic intuition for using distance as an instrument is that it substantially reduces the costs of attending a relevant academic institution, particularly for those from an unprivileged background. Also in lieu of distance, some other studies use a rural/urban dummy in analyzing participation in university or post-secondary education, such as Kane and Spizman (1994) and Christofides et al., (2001) for the US and Canada respectively. However these studies do not draw direct inferences on distance, the motivation behind them to prove to the reader that rural pupils live further from university than urban pupils.

There is a long lasting debate in the field of sociology of education discussing whether regional differences matter for education outcomes (Sixt, 2007) particularly considering segregation among rural and urban areas. For this the most commonly used variables are: population density or a rural/urban dummy. For example, Christofides et al., (2001) study the impact of rural and urban differences on the enrolment into higher education. The issue related to these studies is that they implicitly assume that youth living in urban areas are closer to education institutions than rural youth and hence more likely to participate in education.

Despite the fact that existing evidence suggests that distance clearly matters for the participation in education and education attainment, at the same time, the numerical effects of distance vary substantially among different studies as mentioned above.

4.2.2 Determinants of Performance

There is literature available on general determinants of performance, taking various determinants at a time. Kasirye (2009) investigates the impact of SES and school variables on the learning achievements of grade 6 (age 12) students using the standard OLS and school fixed effects education production function in Uganda. Due to high dropout in schools, the data has non–random allocation of children into schools. The base results of their study after controlling for selection do not change. They find that among the most important determinants of child educational outcomes is teacher training then teacher ability: the number of teachers with two years of compulsory teacher training has a strong impact. However, other teacher and head teacher characteristics have an insignificant influence on learning outcomes. Moreover, a child having its own place to sit in the classroom positively increases school performance while parental education has an impact on child educational outcomes only for boys, but not for girls as fathers' education is significant for males.

Another study by Garcia (2014) in one of the most developed and industrialized regions in Russia highlights the importance of gender, nationality, peers, and health status factors in explaining the educational outcomes of young people who are in Year 9 at age 15 of their study. Girls' performance in education is better than boys.

Engin-Demir (2009) estimates the effects of student, family and school characteristics on the academic achievement of 719 sixth, seventh and eighth grade school students between the ages of 12-14 in urban Turkey. The findings of the study indicate that among all sets of variables, student characteristics (for student characteristics they use hours devoted to study per week, female, and student well-being measured as a combined effect of two variables, 1) student perception about their teacher treatment in class ranked as bad, good or very good, 2) number of friends students have) are the most important variables explaining the large variation in student academic achievement. The family (fathers' education, family income) and school characteristics (student-teacher ratio, number of students in a class, and teacher education and in-service teacher training) are also significantly related to academic achievement but their effects are small. The variation in academic achievements is explained as 15 percent by student characteristics, 4.3 percent by school characteristics and 5.4 percent by family characteristics.

Aslam (2009) using a dataset from the Lahore district of Pakistan analyses the determinants of pupil performance across public and private schools and across subjects in grade 8 using an OLS educational production function. She finds that a large number of pupil, family and school related factors are important in explaining variation in pupil learning outcomes, however, it is not apparent which factors are most important: home background, school or teacher related.

Zhao and Glewwe et al. (2010) look at basic determinants of school attainment using household survey data from one of the less developed provinces of China, Gansu. Using number of years of schooling as an outcome variable, a censored ordered logit estimate demonstrates that mothers' education and attitudes towards children's education are strongly significant. The children whose mothers have 6 years of primary schooling are 1.4 times more likely to go to school as compared to children whose mothers have no education. Teacher experience at lower secondary level has a positive impact on school attainment. School resources especially having a science lab increase school attainment by 1.8 times.

In addition to the above literature, a body of the whole intergenerational mobility literature particularly Chevalier et al. (2013) and Black et al. (2005) supports that parents' education has a positive and significant impact on children's education.

To summarise, it can be seen from both strands of literature, that distance could be a potential deterrent in determining the participation into further education, further, there is no previous study that examines the performance in post compulsory education in terms of exams scores in Year 12 controlling for selection into that level of education, hence, the present work adds to the literature.

Also, distance is very rarely used as an instrument due to the complex data requirement, the current study adds to the empirical field of education economics where an instrument is required as a valid exclusion restriction to identify the exogenous variation to control for the selection into education. There is a criticism that distance and urban/rural location should not be used as interchangeably (Frenette, 2006), thus, the current analysis uses these two variables as separate controls in the selection equation while studying the impact of geographical constraints on the participation decision.

4.3 Data

This chapter uses survey data collected by the author on 611 pupils studying in Year 9 at age 14-15 in 2008-09 using a random sampling technique on private and public schools from one of the districts of Punjab, Pakistan. The first wave of data is collected from pupils, their families, subject teachers and school heads using pupil, family, teacher and school head questionnaires respectively in Year 9. Later the data on pupils are matched with administrative examination board data using unique pupil identifiers to get their test scores information. Further details can be found in the data section of the previous chapter. Overall, it is a comprehensive survey data set having information on the characteristics of individual pupils, their families, teachers, schools, exams scores and some geographic variables. Second, third and fourth waves collected the data again on exam scores of sampled pupils studying in Year 10 aged 15-16 in 2009-10, Year 11 aged 16-17 in 2010-11, and Year 12 aged 17-18 in 2011-12. The current study uses the pupils' score in Year 9, Year 10, and Year 12. Year 9 and Year 10 are the last two years of compulsory schooling as in Pakistan compulsory schooling is from Year 1 to Year 10. Year 11 and Year 12 are post-compulsory schooling years.

A follow-up survey was conducted in 2013-14 in order to create the distance variable. The variable distance to nearest post-compulsory education institution in this survey, measured in kilometres, was calculated using information about their actual home address which was recorded for all the sampled students at the time of the first survey. Also information is available on all education institutions providing post-compulsory education in that district in the dataset; therefore, the distance variable is created for all the sampled people by driving from their homes to the respective nearest feasible education institution, not the schools which they actually attended. It could be possible that the nearest education institution differed from the actual school attended. So the idea is to capture the accessibility for an individual to the nearest post-compulsory education institution. It is worth noting the effort in collecting the dataset which makes analysis accurate leaving less chance of measurement error.

The variable, Year 12 aggregated marks, used in this chapter benefits from certain properties. The outcome variable (marks achieved) is the end-of-course exam that includes the theoretical as well as practical syllabus that is designed and relevant to high school students. A more convincing aspect of the outcome variable is that the scores used in the study (Year 9 score in chapter 3 and Year 12 in chapter 4) are not low stake assessments, otherwise it might worry the researcher that students did not try to perform well in low-stakes tests (Kortetz, 2008), in this way low stakes exams may not be a true measure of students' true learning gains. The dataset also contains the prior academic performance score, which is vital in explaining educational outcomes, particularly investigating post-compulsory education performance.

Table 4.1 below gives a description of the sample of students having missing observations in the dataset. Initially, 611 pupils were randomly drawn from the student population at the selected schools. After that academic records of all sampled pupils have been tracked in Year 9, 10, 11 and 12. By the end of Year 12, 47% pupils are no longer participating in education.

The issue in the current dataset is that there appears to be two drop-out points, some are not completing Year 10 (or not doing the exams at least) while others are not participating in Year 11/12. Looking at Table 4.1, everyone except for one person who studies in Year 11 also studies in Year 12. So when I define the participation variable as whether they are observed in Year 12 or not, I am considering all those who are in post-compulsory education or not (there is just that one person who did participate after the end of compulsory education, in Year 11, but did not continue to Year 12, who is dropped from the dataset). So the group of non-participators in post-compulsory education include a group who drop out after Year 9 and a group who drop out after Year 10. The issue regarding the drop-out in Year 10 is addressed in the robustness section at the end of the chapter.

Table 4.1: Percentage of Schools Dropouts Pupils in Data

	Respondents	Percentage of original sample	Education dropouts
Initial Sample	611	-	-
Academic progress recorded in Year 9	591	96%	4%
Academic progress recorded in Year 10	441	72%	28%
Academic progress recorded in Year 11	326	53%	47%
Academic progress recorded in Year 12	325	53%	47%

Notes: The above table tracks the records of academic performance of individuals over a period of time from Year 9 to Year 12 collected through a survey and later recorded through the matched administrative dataset from BISE, Bahawalpur using a unique pupil identifier. The last column shows the percentage of pupils who dropout from education within the observed sample.

The education dropout in the dataset is not random sample attrition due to the facts explained below.

As far as education dropouts/missing pupils in the dataset as illustrated in Table 4.1 are concerned, they could be missing for reasons other than dropout: for example they could be missing because they have moved to another district. Using their unique pupil identifier (Board Registration Number) they can however be searched for in two other adjacent districts, Bahawalnagar[30] and Rahimyarkhan[31], along with the Bahawalpur district, as the BISE Bahawalpur, is responsible for registration of students and conducting exams in all these three districts of Punjab. Therefore, after searching for all missing pupils in the administrative data in all three districts, it could be ruled out that they are missing due to migration to another district. However it could be possible that they move to another district further away, where they no longer can be tracked. However, this is unlikely as the Bahawalpur district is an agricultural district not an industrial or Cantonment (under the control of military) area, so migration[32] is less likely to happen in the Bahawalpur district.

[30] Bahawalnagar, one of the districts in the Punjab province of Pakistan, consists of five sub-districts.

[31] Rahimyarkhan is one of the cities and districts in the Punjab province in Pakistan. Administratively, it is divided into four sub-districts.

[32] Migration means transfer of a student or a candidate from one institution/board to another institution/board for seeking of higher education. This type of migration is allowed by the board on the request of an applicant under a convincing reason with the consensus of both the heads of the institutions until 31st December of the academic year. This type of migration is allowed by the board if the candidate fulfils all the legal formalities specified by the board in this regard i.e. attestation of migration application from the last attended college, prescribed fees etc.

Second, they could be missing due to not responding to the survey, as there is sample attrition in general. But this is not the case as all information in the adopted questionnaires is collected at the beginning of the survey, and the unique pupil identifier as well so that in later sweeps their exams results are tracked through administrative data.

Third, sometimes results could be pending for some pupils due to: (1) Unfair means cases, such as cheating or communicating to another candidate or not following the exam's rules/instructions, misbehaving with the exams staff etc. Disciplinary action may follow in this regard against that candidate. (2) Results declared later due to incomplete documents or fees, this could happen if a candidate fails to provide required documents and fees. Results could be withheld until he/she clears the documents or fees. But all those numbers kept pending are given in the Board Gazette separate under the heading of Results Later. I checked this carefully and did not find a single case like this.

Fourth, it is more likely that they dropout from education, as comparing the characteristics of pupils who dropout to those who participated in Year 12, 85% pupils are from public school and their average prior attainment (Year 9 score) is 205 out of 525, which is lower than those who remain in the sample (298). 90% of dropout pupils are from rural areas, also they have a lower level of fathers' education, mothers' education and fathers' income and hence have a greater level of disadvantage.

Also it is possible that after failing[33] in exams due to lower scores they left that stream of education (that particular academic year) and reappear in the Supplementary[34] exam. In that case they would be given a different board registration number, which I do not know. However, I also searched for them in the administrative online data of the Supplementary Exam in all three districts, through inputting their names and fathers'

[33] Generally, in order to pass a particular subject/paper a candidate must obtain 33% of the total marks of each subject, otherwise, it would be considered as "Fail" and to pass the overall grade/year, he must score 33% of the total marks of all subjects of that year. In case of failure he has to appear again in the exam.

[34] The supplementary exam is the additional exam arranged for those who are not able to pass the annual exam at the first attempt or they do not appear in the exam due to other reasons such as illness, absence etc. Also those who pass the grade but want to again appear in the exam to improve their scores to get admission into a required institution to meet their eligibility or merit criteria, can take it as a second opportunity. Supplementary exams are scheduled after the specific duration of annual exams, so that students can get ready in order to appear in exams. Generally the exams rules stay the same as regular exams.

names in the search box and I did not find any of them there. As in the online administrative data available on the website of BISE, exam records can be searched either using pupil name and fathers' name or board registration number. Therefore, all the above arguments make the claim stronger that these are education dropouts that are observed, and all possible measures have been taken to trace these missing pupils had they still been in education. Therefore, the missing observations in Year 12 are almost certainly because they are no longer in education and not simply due to non response to a survey. It is therefore people making a choice whether they participate in post compulsory education or not which is why I need to control for selection to control for sample selection bias. To take up those techniques, the dataset has sufficient valid exclusion restrictions e.g. proximity to post-compulsory educational institution.

4.4 Methodology

The classic example of correcting a model for selection bias (using the Heckman sample selectivity correction approach) is the estimation of the wage equation where a wage equation is estimated for only those who participated in the labour market and data on wages are missing for those who did not enter into the market who are unemployed. Selection bias refers to a problem when an outcome equation is estimated for a restricted or sub sample for which data are not a random draw from a population; instead it has been observed for a particular restricted group or a sub sample of a population (Heckman, 1979). If this is the case then OLS estimation will give biased/misleading estimates if the same variables that determine selection into the sample also affect the outcome variable of interest. It is obvious that one cannot estimate the determinants of raw performance of pupils in a particular grade, unless sample selection has been taken into account. This is due to the fact that pupils who dropped out from education are more likely to be those with low ability, less motivated, from low quality schools and from unprivileged home backgrounds (Alivernini and Lucidi, 2011; Chowdry et al., 2013).

In order to overcome the omitted variable bias prior academic achievement Year 9 score is used as a proxy of ability. There could be an issue here, as Year 10 is the last year of compulsory education then it should be year 10 test scores that are used as measures of past ability rather than Year 9. The main analysis is done on the Year 9 sample, because of the larger sample size. Additional analysis in the robustness checks section is therefore taken using Year 10 score as a measure of past ability rather than Year 9.

Since not every pupil is observed in Year 12, and some people have already dropped out of education before entering into post-compulsory education, which is represented as participating in Year 12, this requires a correction for selection bias, where the selection is on whether the students are still in education in Year 12 or not. The Heckman Selection Model is used to account for this selectivity bias. Note that 28% of pupils dropped out in Year 10 and 47% in Year 11 as previously explained in Table 4.1. It is therefore important to determine whether there are two separate decisions that need to be modelled, i.e., decision 1 - taking the Year 10 exams or not, and decision 2 - participating in post-compulsory schooling in Years 11 and 12 or not. The analysis therefore checks the validity of treating the two decisions as a single decision to drop out, by separately considering the second decision in isolation, modelling the selection into post-compulsory education conditional of having completed the Year 10 exams.

The procedure has two steps i.e. it estimates two equations: The first is the participation/selection equation using a binary dependent variable which takes the value of 1 if the pupil is present in the estimation sample for Year 12, 0 otherwise. The second stage is to estimate the performance equation, accounting for the probability of selection into Year 12.

Participation Model:

Thus, the following model takes participation in Year 12/ post-compulsory education Pi as a binary choice variable.

Where

P_{it} = 1 if pupil participated in Year 12 examination

P_{it} = 0 if pupil did not participate in Year 12 examination

And the model of participation in post-compulsory education is

$P_{it} = f\,(Y_{it-k}, I_{it-k}, F_{it-k}, S_{it-k}, S_{it}, \text{distance}_{it-k}, \text{ur}_{it-k})$

Y_{it-k} is the prior attainment of pupil i measured as the total score in Year 9.

I_{it-k} is the measure of pupil attributes of pupil i in Year 9 such as age, health problems, gender and birth order.

F_{it-k} captures the family characteristics of pupil i measured in Year 9 such as fathers' education, mothers' education, fathers' income, parents' own house.

S_{it-k} captures the type of school attended (private or public school) by pupil in Year 9.

S_{it} is the type of institution attended in Year 12

The participation model has two additional variables: $distance_{it-k}$, distance to nearest post-compulsory education institution and $ur_{it-k,}$ a dummy for urban location in which a pupil lives.

The variable distance to nearest post-compulsory institution, measured in kilometres, is the "actual" measured distance from pupils' homes to the nearest post-compulsory education institution. Thus it is a continuous variable having a mean of 8 km. It is picked as an instrument as while this should affect the likelihood of attending post-compulsory grade i.e. Year 12, it should have no impact on Year 12 performance of pupils. To account for non linearity, log of distance is used in this study (Newbold and Brown, 2015).

Distance is a good instrument but is very rarely used as an instrument due to data availability while discussing post-compulsory education. However, there could be an issue related to this variable that it might be the case that the distance is higher but the public transport is better. To account for this, I also used another instrument, urban.

The variable named as urban is measured as a dummy variable having the value 1 if the pupil is living in an urban area and 0 if he/she is living in a village area (rural location), again this is a convincing instrument as rural areas have no access or limited access to post-compulsory schools and fewer transport facilities relative to urban areas. The urban location also has no direct effect on pupils' performance in Year 12 but it has a direct influence on participation into post-compulsory education. However, there could be differential resources between the two locations which impact on performance, such as kind of institution. Therefore, I used a separate control for kind of institution attended in Year 12 and Year 9 in the model. However, this is only one potential difference in resources, and there still could be other remaining differences between the two types of area.

Previously mentioned in the literature review, another potential criticism of using only the urban dummy in studies exploring variation in education participation due to regional differences assumes that the urban areas are nearer to school, college, and university. It may not be the case, due to some education institutions being easily accessed by nearby pupils while some medium size urban areas have no education institutions at all. Therefore, it can be argued that the urban dummy and distance cannot be used interchangeably. In response to the above criticism, the current study controls for both distance and the urban dummy separately.

Therefore, the present dataset has several variables which might be considered as convincing identifying variables. The main point in choosing an instrument is that the instrument determining the participation should not affect the outcome variable of performance in post-compulsory education. The correlation between instruments and participation and performance is described below in Table 4.2. It is clear that participation is strongly correlated with distance and the urban dummy, having a correlation coefficient of 0.86 and 0.91 respectively. Relative to participation, performance measured as Year 12 score is weakly correlated with these instruments. Clearly these statistics show that these are good instruments fulfilling the basic criteria of instruments.

Table 4.2: Correlation between instruments, participation and performance

Instruments	Participation (as 1, 0) First stage	Performance (Total marks Year 12) Second stage
Distance	-0.801(P-value, 0.000)	-0.197(P-value, 0.000)
Urban dummy	0.916(P-value, 0.000)	0.054(P-value, 0.324)

Notes: This table gives the Pearson correlation between instruments (Log of distance and urban dummy) and participation in to post-compulsory education and performance in Year 12, measured as total marks in Year 12.

However, it might be the case that using time spent on commuting might directly reduce the time available for school work, reducing achievement. Also, the different neighbourhood conditions in urban areas might influence achievement directly. Firm evidence on the issue requires some kind of a test, but that ideal test does not exist. Often researchers have to rely on an exclusion restriction that is empirically significant in the first stage probit selection equation and insignificant in the equation of interest. Therefore these instruments are tried in the second stage equation to see empirically if they affect the outcome variable, achievement. The details of these tests are given in the robustness checks section. Both instruments are found to have insignificant coefficients in the achievements model, further justifying our choice of exclusion restrictions and therefore, they seem to be valid exclusion restrictions in the current context.

Subsequently, the idea to use the above two variables as instruments is that both are measuring the availability of post-compulsory institutions. Both of these variables capture the access to a post-compulsory education institution. It is important to incorporate these two variables additionally into the participation model since these enable us to look at the incidence of participation in post-compulsory education across rural-urban geography. More importantly these are the exclusion restrictions for identifying the participation equation. However, further relevance of instruments is tested to check if they are genuine valid exclusion restrictions. They could have an effect on performance (presented in robustness section).

Additionally, the present study used the interaction of log of distance to nearest post-compulsory education institution and urban in the participation equation, to see whether distance matters more for those in a rural setting than in an urban setting which has a significant relevance to disadvantaged people. Individuals who live in rural areas typically have less access to public transport, higher distance to travel and a greater level of disadvantage. As a consequence, the participation of individuals living in rural areas may be more affected by the distance to the nearest school. In the current dataset, the average distance to nearest school for pupils from urban areas is 2km, while it is 16.6 km for pupils living in rural areas, which shows a huge difference regarding location. Further the interaction of log of distance and male and mode of transport are also used to see any particular differences.

Participation Equation:

The first stage equation is estimating whether or not a pupil is in education in Year 12 using a probit model since participation is a binary variable taking values 0 and 1.

$$P_{it} = \gamma_1 Z_{it} + \varepsilon_{it} \qquad (1)$$

Where Z_{it} denotes the vector of all those variables influencing the participation into post- compulsory education and γ_1 is the vector of parameters and ε_{it} is the disturbance term, normally distributed with mean zero and variance σ_{ε}^{2}.

Performance Model:

Generally economists have used a production function to approach the issues of school quality and examination of education performance. An educational production function takes into account the relationship between educational inputs and educational output also called an 'input-output' approach.

Instead of a contemporaneous specification, which assumes observed achievements are determined by only current inputs, an education production function with a modified traditional value-added model was used (Hanushek and Rivkin, 2012). It is modified as it does not use all past inputs, instead some current inputs are also used such as kind of institution attended in Year 12. It is called a value-added model as it takes into account the effect of prior attainment in an earlier period and it is important to control for this as it gives an indication of achievement that an individual brings to the classroom i.e. the effect of previous teachers and school, and individual ability (Todd and Wolpin, 2003).

$$Y_{it} = f\,(Y_{it-k}, I_{it-k}, F_{it-k}, S_{it-k}, S_{it}, u_{it})$$

Where

Y_{it} is the achievement of student i measured by the total score obtained by a pupil in Year 12.

u_{it} is the error-term, assumed to be distributed normally with mean zero and constant variance.

Therefore, the second stage equation can be written from the above performance model as:

Performance Equation:

$$Y_{it} = \beta X_{it} + u_{it} \qquad (2)$$

Where

Y_{it} is the academic performance measured by Year 12 total score of the i^{th} pupil. X_{it} is the vector of the i^{th} pupil's individual, family and school characteristics, containing all variables mentioned in the performance model except distance and the urban dummy, β is the vector of parameters and u_{it} is the disturbance term, normally distributed with mean 0 and variance $\sigma_{u.}^{2}$ Following the Heckman (1979) technique, the performance equation can be corrected for selectivity bias by introducing a term λ_i, called the inverse Mills ratio from the first stage regression, as an additional regressor in equation (2). Equation (3) is modelling Year 12 performance, once the Heckman lambda term is introduced.

So that

$$Y_{it} = \beta X_{it} + \delta\lambda_{it} + u_{it} \qquad (3)$$

Where

$$\lambda_i = \frac{\emptyset(Hi)}{\varphi(Hi)}$$

$\emptyset(.)$ is the standard normal probability density function and $\varphi(.)$ is the normal cumulative distribution function. The coefficient on the lambda term is the measure of sample selection bias. If this is statistically different from zero, the null hypothesis, "no sample selection bias" is rejected.

Often, in the Heckman model the main interest lies in the second stage equation, and the selection equation is just to control for bias, but in this case, the selection equation which is participation into post-compulsory education (who is still in education in Year 12) is interesting in itself. In other words, it is possible to argue that the Heckman model has served a dual purpose, one, it controlled for selection bias, so that performance conditional on participation can be studied, and the second, it shows why people dropped out from education (why they did not participate into Year 12).

Though using Heckman methodology, the parameters of interest in the model can be identified via non-linearity (when an exclusion restriction is not utilised) of the probit function, it is preferred to incorporate a valid exclusion restriction as it leads to less multicollinearity not only among predictors but also between error terms, due to less correlation between the X vector and the inverse Mills ratio (Bushway et al., 2007). The crucial step for the Heckman model to work in the presence of an exclusion restriction (to 'identify' the model) is, there must be at least one variable in the participation/selection equation, that must not be in the performance equation. This variable is called the instrument/exclusion restriction. Hence it facilitates the identification.

A lack of suitable exclusion restriction leads to biased parameter estimates due to potential endogeneity and OLS is not appropriate to use. The instrument and exclusion restriction are very similar to each other (Angrist and Krueger, 2001). Hence in the existing case, it must have at least one variable that will directly affect participation (whether a pupil is still in education in Year 12 or not), but will have no impact on their performance in post-compulsory education (once they are there).

The description of variables used in this study is given in Table 4.3.

The Year 12 score is the sum of scores in all 7 subjects offered. Among them, four subjects are compulsory (Language, English, Religion and Pakistan Studies) and 3 are optional depending upon the field of study opted. In total five fields of study exist in Year 12. These are the Premedical group, Pre-engineering group, General science group, Humanities group, and Commerce group. Maximum total marks in Year 12 are 1150. The total score for each subject is 200 except Islamic Studies and Pakistan Studies, these two subjects have 75 each.

Prior Attainment (Year 9 score) is the overall/aggregate score in 7 subjects being offered in Year 9 in schools. In total 4 subjects are compulsory (Language, English, Religion and Mathematics) while 3 subjects are elective, different in two fields of study. In Year 9 only two fields of study are offered: Science and Arts. The maximum total score in Year 9 is 625.

The type of institution in Year 12 variable captures the kind of institution attended in Year 12. There are three categories for this variable: appearing in the exam after attending private school, public school and as an independent candidate. Candidates attending a private or public institution can send their exams admissions as a “regular” candidate as they are on the roll of their respective attended institution, in this case, that institution sends their admission on their behalf; while those who do not attend any formal institution due to non-availability of higher education institution within travel distance or due to health problems appear as independent candidates. Generally, rural youth and females or disabled prefer to appear as an independent candidate. They do self-study and preparation for exams, and send their exams admissions independently through their respective registered board.

There are certain subjects/disciplines which cannot be taken as a private candidate, particularly science subjects such as physics, chemistry, and biology. As these subjects have practical exams along with theoretical exams, attending regular classes is the requirement for them, so these cannot be taken by the private candidates. In the current dataset, 45.88% pupils attended a private institute, 43.38% a public institute and 10.77% of pupils appeared as independent candidates.

In Year 9 there are only two categories of type of institute attended: private and public school. This is due to the reason that the time when the survey was initiated in Year 9, all the sampled pupils were taken directly from schools. So there is no category of the independent candidate in Year 9. I use the separate control for type of institution attended in Year 9 and Year 12 in the participation and performance equation respectively. In Year 9, 69.98% of pupils are from public schools.

Table 4.3: Description of Variables

Variables	Description
Score in Year 12 (dependent variable)	Total marks obtained by pupil in Year 12
Pupil age	Age of pupil measured in years
Male	Dummy variable equals 1 if male and 0 for female
Own House	Dummy variable equals 1 if parents own a house
Number of Siblings	Number of brothers and sisters pupil has
Fathers' Education	Father's education measured in years
Mothers' Education	Mother's education measured in years
Birth Order	Pupil's standing among other siblings
Health Status	Dummy variable equals 1 if pupil has a chronic disease
Type of Institution in Year 12 (Reference category: Private Institution)	1 if pupil attended private institution, 2 for public Institution and 3 if appeared in exam as an independent candidate not through any institution
Number of Rooms	Number of rooms available at home
Private Coaching (tuition)	Dummy variable equals 1 if pupil takes private coaching other than school
Mode of Transport	Dummy variable equals 1 if pupil uses car or bus to reach school and 0, if he/she walks.
Study Time at Home per day	Dummy variable equals 1 if pupil studies at home more than 2 hours per day
Private School Year 9	Dummy variable equals 1 if pupil attended private school in Year 9 and 0 for public school.
Prior Attainment Year 9	Academic total score of Year 9
Father Income per month	Income of father measured in Rupees
Log of Distance	Distance measured in km, from home to nearest secondary institution, not the school which they actually attended
Urban	Dummy variable (0/1) equals 1 if pupil is living in urban area and 0 for rural area

Notes: Number of rooms is checked by adjusting for household size but results remain the same.

Table 4.4: Summary Statistics of Variables

Variables	Obs.	Mean	S.D.	Min.	Max.
Score in Year 12 (out of total marks for Year 12, 1150)	292	630.88	167.62	232	1020
Pupil age (Years)	553	15.20	1.26	10	20
Pupil age missing	553	0.11	0.31	0	1
Male	553	0.49	0.50	0	1
Own house	553	0.78	0.41	0	1
Number of siblings	553	4.69	2.04	0	16
Fathers’ education (Years)	553	10.01	5.13	0	20
Mothers’ education (Years)	553	6.44	5.77	0	18
Fathers’ income per month (Rs.)	553	18839.25	46595.21	0	1000000
Birth order	553	3.06	1.85	1	11
Health problems	553	0.28	0.45	0	1
Public institution Y12	292	0.44	0.50	0	1
Private institution Y12	292	0.45	0.50	0	1
Independent institution Y12	292	0.11	0.31	0	1
Number of rooms	553	4.40	2.53	1	18
Private Coaching (Tuition)	553	0.59	0.49	0	1
Mode of transport	553	0.57	0.49	0	1
Study hours per day	553	0.75	0.43	0	1
Private school Year 9	553	0.30	0.46	0	1
Prior attainment Year 9 (Total score in Year 9 is 625)	553	253.20	87.78	57	449
Distance (kms)	553	8.43	8.15	0.5	40
Urban	553	0.55	0.50	0	1

Notes: 1: The number of observations here are different to those in table 1.The difference is due to usable sample. 2: Pupils’ age missing variable is to allow for the low number of observations on this variable. 3: To get a relevance the mean income would be equal to 140£ per month approximately.

4.5 Results and Discussions

This section is divided into two sub-sections. Section 4.5.1 presents the results from the participation model and Section 4.5.2 discusses the raw and selection corrected results of performance analysis.

4.5.1 Determinants of Participation in Post-Compulsory Education, Probit Model

Table 5 represents the probit analysis of equation (2) showing determinants of participation in post-compulsory education, the dependent variable is measured as a binary variable, having a value of 1, if a pupil participated in a post-compulsory grade and a value of 0, if he/she dropped out i.e. did not participate in post-compulsory education. The second column shows the marginal effects and standard errors are in parentheses.

The coefficient on log of distance shows that the further pupils live from a post-compulsory institution, the less likely they are to undertake post-compulsory education. So a 1% increase in distance reduces the likelihood of participation by 0.07 percentage points (or equivalently, we could say a 10% rise in distance reduces the likelihood of participation by 0.7 percentage points.

The coefficient on the urban variable shows that pupils living in an urban locality are more likely to participate in post-compulsory education, living in an urban area increases the probability of participation in post-compulsory education by 30 percentage points relative to those living in a rural area. This is in line with the general hypothesis that urban areas have more transport available and more facilities. Both coefficients are highly significant and used as instruments at the first stage in the participation equation.

Mother’s education is negatively related to the likelihood of participation in post-compulsory education, a one year increase in mother’s education, decreases the probability of participation in the post-compulsory grade, by 0.6 percentage points. This is surprising and is in contrast to the hypothesis that mothers’ education has a positive impact on children’s educational outcomes (Chevalier, 2004). This could be due to the fact that controls for the factors that the mother's education effect works through (for example children of well-educated mothers being highly able and/or going to good schools), are included and so could capture some of the positive influence of mother's education.

Private school attended in Year 9, which is measured as a dummy variable, equals 1 for private school and 0 for public school, has a negative coefficient, implying that attending private school in Year 9, decreases the likelihood of participation in post-compulsory education, by 10.31 percentage points holding other things constant. This is also a surprising result and it could also be due to controlling for prior ability as it could take away the private school effect.

As previously described, the distance variable was calculated by driving from each individual home to the nearest post-compulsory educational institutions (details are described on page 111). The average distance for public schools is 10.87Km and for private schools is 3.3Km

Looking at the private school effect, shows the counter intuitive result for private schools in the participation model, that those in private school in Year 9 are less likely to participate post-compulsory. This could be because those in private school are in communities without a nearby public school, and so have fewer opportunities to participate post-compulsory. The descriptive statistics of distance by type of school are consistent with this (average distance for public schools is 10.87Km and for private schools is 3.3Km).

Additionally, an interaction between distance and private school in the participation model was tried and the coefficient is positive and significant (results are not reported). This implies that the negative impact of distance is less for those going to private school. This is also corroborated by the fact that the average distance of private schools is less than public schools. A final possibility is that the expected private school advantage is due to high attainment students in such schools, so that a positive private school effect on participation might emerge if we no longer for control for prior attainment. However after dropping the prior attainment variable, the private school result remains the same (results available from author on request).

The mode of transport variable has a positive effect on the probability of participation in post-compulsory education, using bus/car as a mode of transport increases the likelihood of participation in post-compulsory education by 6.2 percentage points. As the mode of transport is used as a proxy for accessibility and also it could be an indication of poverty, so it implies that accessibility and a lower level of poverty have a positive effect on participation in post-compulsory education and vice versa.

Studying at home more than two hours decreases the probability of participating in post-compulsory education by 6.9 percentage points. It could be due to those struggling in their studies, need more time to catch-up.

Prior ability has no significant effect on participation in post-compulsory education. This result is different from previous findings for developed countries (McIntosh, 2001).

The overall results suggest that participation in post-compulsory education is mainly determined by the distance to the nearest educational institution and living in an urban area. This is interesting; both variables measure the availability of post-compulsory education institutions for young people and have clear implications for youth education in remote areas. If more institutions are made available to pupils living far from post-compulsory education institutes, participation in education can be increased. These results are consistent with the results from previous studies (Dickerson and McIntosh, 2013; Frenette, 2006; Spiess and Wrohlich, 2010).

Though results from the present analysis suggest that lower distance increases participation, however, the reason driving the distance effect is not clear. It could be either financial cost or commuting cost due to frequent long travelling. So if higher distance reduces the incentive to participate in further education due to financial cost then provision of subsidized travel to those participating in further education would be an optimal policy.

However, if distance affects the decision to participate because of commuting cost, then the effective policy would be to improve the quality of transport services through increasing the frequency of travel services and reducing the travel time, perhaps giving the priority to public transport on congested roads. Admittedly, it is much easier to target financial travel subsidies than general travel services.

Table 4.5: Probit Model: First stage, Participation in Post-Compulsory Education

Variables	Marginal Effects	Standard Errors
	Model1	
Log of distance	-0.071***	(0.029)
Urban dummy	0.303***	(0.070)
Age	0.002	(0.009)
Age missing	0.021	(0.154)
Male	-0.011	(0.026)
Own house	-0.003	(0.027)
Number of siblings	-0.009	(0.008)
Fathers’ education	0.004	(0.003)
Mothers’ education	-0.006*	(0.003)
Fathers’ income	-2.26	(1.230)
Birth order	0.014*	(0.008)
Health status	-0.036	(0.023)
Private school Year 9	-0.103***	(0.044)
Number of rooms at home	-0.002	(0.004)
Mode of Transport	0.061**	(0.027)
Private coaching	-0.001	(0.021)
Study time home	-0.069**	(0.032)
Prior attainment Year 9	-0.000	(0.000)
Observations	553	

Table 5 presents the Marginal Effects (M E) on the probability of participation into post-compulsory education. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Model 1 is the first stage using the log of distance and urban dummy as instruments.
The main analysis was done on the year 9 sample, because of the larger sample size compared to Year 10 sample.

Model 1 used log of distance and urban dummy variables determining the participation in post compulsory education in the selection equation, as in Table 4.5.

Table 4.6 gives the first stage results for interaction effects between distance to nearest compulsory institution and the urban dummy, on the probability of participation in post-compulsory education. An interaction between distance and urban is tried in the first stage selection equation, to see if distance has more impact in rural areas where transport is not as good.

The coefficient on the urban dummy is insignificant which suggests that if pupils live at zero distance from a school (literally next door to a school) then it does not make any difference to their participation whether they live in an urban area or rural area- the school is right there for them.

Table 4.6: First stage results: Probit model, participation in post-compulsory education, Interactions Effects (Distance*urban)

Covariates	M.E	(S.E)
Log of distance	-0.229*	(0.134)
Interaction, Log of distance X urban dummy	0.174	(0.135)
Urban dummy	-0.101	(0.318)
Prior attainment Year 9	-0.00	(0.000)
Number of observations	553	

Marginal effects (M.E) on the probability of participation in post-compulsory education.
Standard errors (SE) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

The interaction term has a positive coefficient, it suggests distance has a greater impact in rural areas, though it is statistically insignificant due to the high standard errors.

It can be argued that the natural way to include mode of transport would be to interact it with log of distance: having a car, for instance, will have a bigger effect if you live further away. This has been tried and results remain the same (Table 4.7) and the interaction term also has a significant coefficient (though when we include the separate term mode of transport along with its interaction with distance, results remain the same except the interaction term becomes insignificant due to the higher standard error). The interaction term has positive coefficient meaning that using car/motor cycle, bus as a mode of transport (more accessibility) will lead to a higher effect on participation rates the larger the distance needed to be travelled relative to using these modes: cycle/walk as transport.

Table 4.7: First stage results: Probit model for participation in post-compulsory education, Interactions Effects (Distance*Mode of transport)

Covariates	M.E	(S.E)
Log of distance	-0.079***	(0.033)
Urban dummy	0.304***	(0.071)
Interaction, Log of distance X Mode of transport	0.015	(0.024)
Mode of transport	0.044	(0.038)
Prior attainment Year 9	-0.000	(0.000)
Number of observations	553	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Additionally the interaction between distance and male is tried to see if distance reduces participation more for girls relative to boys. The interaction term has a negative coefficient which suggests that distance has a more negative impact for boys than girls. This is surprising and contrary to our hypothesis since in Pakistan, girls face more restrictions regarding their mobility and education particularly in higher classes due to cultural and religious issues and due to gender inequality widespread in most Asian countries (Ali et al., 2011). However, the male interaction has an insignificant coefficient. Results are presented in Table 4.8.

It could be expected that the interaction terms would attract significant coefficients, but the interaction terms are quite difficult to find a significant coefficient for, particularly in a small sample, because by definition they are correlated with other explanatory variables which in turn raises standard errors. Therefore, none of the interaction terms have significant coefficients due to small sample size.

In all three models using interaction terms it should be noted that the coefficient on total marks Year 9 which capture the effect of prior attainment is highly significant in the performance equation and insignificant in the participation equation.

Table 4.8: First stage results: Probit model for participation in post-compulsory education, Interactions Effects (Distance*Male)

Covariates	**M.E**	**(S.E)**
Log of distance	-0.068***	(0.029)
Urban dummy	0.297***	(0.070)
Interaction, Log of distance X Male	-0.013	(0.025)
Male	0.040	(0.039)
Prior attainment Year 9	-0.000	(0.000)
Number of observations	553	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.5.2 Determinants of Performance in Post-Compulsory Education

4.5.2.1 Determinants of Performance using OLS

The first column in Table 4.9 displays the coefficients of determinants of performance in post-compulsory education (Year 12) produced through OLS regression of equation (1). The dependent variable in this model is total score in Year 12, having a SD of 167 and a mean of 630.

Among the different determinants of performance in the post-compulsory grade, the statistically significant coefficients are on the variables private institution in Year 12, taking an exam as an independent candidate in Year 12, study time at home and prior attainment in Year 9.

The results imply that attending a public school/institution relative to a private institution in the post-compulsory grade decreases performance on the exam by 0.24 of a standard deviation of marks. Similarly, appearing in the exam as an independent candidate not through an institution is associated with a decrease of 0.12 of a standard deviation of marks in the post-compulsory grade. These findings are consistent with previous literature: Aslam (2009); Kingdon (1996), stating the hypothesis regarding the relative efficiency of private and public schools report that private schools are better than public schools. This could be mainly due to poor resources at public school, higher proportion of pupils at schools relative to low number of teachers etc.

Study at home more than 2 hours is associated with positive performance in post-compulsory education; studying time at home has a positive impact on performance by 0.07 of a SD of marks in the post-compulsory grade. The variable most strongly associated with performance in post-compulsory education is prior attainment in Year 9. It implies that a 1 SD increase in prior attainment in Year 9 is associated with an increase of 0.79[35] of a SD of marks in the post-compulsory grade. It is also in accord with previous literature such as Dickerson and McIntosh (2013) that highlights the positive effect of past educational performance on current educational outcomes. All other variables in this model have statistically insignificant coefficients. The R^2 has a value of 0.67, and overall, the F test is significant. However, these are raw coefficients on all variables in the model, as selection bias is not taken in to account.

[35] These standard deviations are obtained by multiplying the respective coefficient values, beta, to its standard deviation and then dividing the product by the standard deviation of the dependent variable.

4.5.3.2 Refined Determinants of Performance in Post-Compulsory Education (controlling for selection into post- compulsory education using Heckman Selection Model)

Table 4.9 describes the determinants of performance (Outcome equation), three models are estimated, the only difference is, model 2 gives the OLS, raw coefficients of determinants of performance in post-compulsory education, while, the later models 3 and 4 account for sample selection.

The log of distance and urban dummy are used as instruments in the first stage in model 3. The log of distance and interaction between the log of distance and urban dummy are used as instruments in the first stage in model 4.

It is possible that results reported in model 2 of table 4.9 are biased, due to the fact that every student is not observed in Year 12 and 47% of pupils do not participate in post-compulsory education, hence leaving the sample non-random. The pupils observed in post-compulsory education may be more able, motivated and with better human capital relative to those who dropped out. There is a test below for the table of $\rho=0$, which is not rejected. Results show no selection effects as the test of rho (rho is the correlation between the errors in the two equations) in the Heckman model is insignificant in model 3. The value of χ^2 is 1.91, with P-value 0.167, hence, the null hypothesis ($\rho=0$, that is there is no sample selection bias) cannot be rejected. Similarly, the value of χ^2 is 1.84 in model 4, with P-value 0.174, again suggesting that the null hypothesis of no sample selection, i.e., $\rho=0$, is not rejected.

Table 4.9: Determinants of Performance in Post-compulsory Education: Second stage results, Heckman correction model

Variables	Model 2 (OLS) Coefficients S.E	Model 3 (Heckman) Coefficients S.E	Model 4 (Heckman) Coefficients S.E
Age	-2.797	-3.893	-3.785
	(5.507)	(5.427)	(5.429)
Age missing	-35.540	-50.963	-49.292
	(85.154)	(83.876)	(83.922)
Male	-3.818	-7.389	-7.686
	(14.535)	(14.335)	(15.743)
Own house	-8.274	-10.835	-11.212
	(16.077)	(15.718)	(15.743)
Number of siblings	4.588	4.009	4.011
	(4.627)	(4.546)	(4.551)
Fathers' education	-1.625	-1.640	-1.646
	(1.699)	(1.657)	(1.660)
Mothers' education	1.543	1.457	1.402
	(1.518)	(1.478)	(1.482)
Fathers' Income	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Birth order	-3.373	-2.451	-2.409
	(4.672)	(4.351)	(4.357)
Health status	-13.984	-14.679	-15.002
	(13.642)	(13.357)	(13.385)
Public Institute Year 12	-79.223***	-78.294***	-78.114***
	(14.558)	(14.381)	(14.393)
Independent candidate Year 12	-67.353***	-64.904***	-64.862***
	(23.390)	(22.864)	(22.865)
Number of rooms at home	3.6555	2.918	2.884
	(32.870)	(2.829)	(2.835)
Mode of transport	-16.141	-15.388	-15.356
	(13.471)	(13.117)	(13.133)
Private coaching	-20.247	-17.298	-17.164
	(12.553)	(12.261)	(12.282)
Study time home	5.972*	22.070*	21.651
	(114.349)	(14.108)	(14.148)
Prior attainment Year 9	1.520***	1.510***	1.513***
	(0.104)	(0.103)	(0.103)
Constant	256.552**	279.284***	277.861***
	(92.642)	(91.788)	(91.911)
Observations	295	553	553
Rho	-	0.386	0.416
	-	0.246	0.257
Lambda	-	37.790	40.828
	-	24.542	25.807

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1,
Model 2- OLS
Model 3 - Heckman selection corrected model (Second stage) - χ^{2} = 1.91, P- value = 0.167, Probit: First stage uses the log of distance and urban dummy as instruments.
Model 4 - Heckman selection corrected model (Second stage) using interaction term between log of distance and urban - χ^2 = 1.84, P-value = 0.174, Probit: First stage uses the log of distance, urban dummy and its interaction

Also by looking at all the results in model 2, 3 and 4 suggests that broadly results do not change which is an indication of no selection bias. This suggests that performance once in post-compulsory education is mainly caused by ability (the coefficients on prior attainment in Year 9 are also useful here as they are significant in the performance equation). Tables 4.10, 4.11 and 4.12 show the second stage results of using interaction terms (distance*urban, distance*mode of transport and distance*male in first stage respectively). The inclusion of the interaction terms in the first stage does not change any second stage results.

Table 4.10: Second stage results: Determinants of Performance in post-compulsory education: Dependent variable Year 12 score

Covariates	**Coefficients**	**(S.E)**
Prior attainment Year 9	1.512***	(0.103)
Rho	0.416	(0.257)
Lambda	40.828	(25.807)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 1.84, P- value = 0.174
*** p<0.01, ** p<0.05, * p<0.1

Table 4.11: Second stage results: Determinants of Performance in post-compulsory education: Dependent variable Year 12 score

Covariates	**Coefficients**	**(S.E)**
Prior attainment Year 9	1.512***	(0.103)
Rho	0.419	(0.248)
Lambda	41.100	(24.877)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 2.12, P- value = 0.14
*** p<0.01, ** p<0.05, * p<0.1

Table 4.12: Second stage results: Determinants of performance in post-compulsory education: Dependent variable Year 12 score

Covariates	**Coefficients**	**(S.E)**
Prior attainment Year 9	1.510***	(0.103)
Rho	0.397	(0.245)
Lambda	38.941	(24.465)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 1.99, P- value = 0.15
*** p<0.01, ** p<0.05, * p<0.1

It should also be noted that in all the above models using interactions terms, the lambda reported in the last rows of table 4.10, 4.11 and 4.12, which is the coefficient on the Inverse Mills Ratio, is not significant which shows the lack of selectivity. Rho is the correlation coefficient between the error terms in the two equations (participation and performance) and is also of interest - again it is not significant, suggesting they are two unrelated processes.

4.6 Robustness checks

In this section several robustness checks are applied to see if the results are affected radically.

4.6.1 Change in sample used to Year 10

This section checks whether results remain robust to some change in the specification, for example the sample used, using those who participated in Year 10 rather than Year 9 as the basis of the sample. The issue is having a sample of those who participated until the end of Year 10. One potential criticism of the main analysis of this chapter is that there could be a further selection issue as there is one more year of compulsory schooling involved, which is Year 10, after that being considered (Year 9) and from which some individuals drop out. The analysis is therefore redone focussing on the participation choices of only those who made it to the end of Year 10 and sat the final exams in compulsory schooling.

The analysis is redone for Year 10 completers only, to compare the results to those obtained when using the full sample from Year 9. If the results are qualitatively similar, this would suggest that the process influencing the continuation decision at the end of Year 10 is the same as that influencing the continuation decision earlier in Year 9 - i.e. there is only one decision/process rather than two processes.

So the only thing that differs is that the group of non-participators in post-compulsory education include a group who drop out after Year 9 and a group who drop out after Year 10. It is the fact that the sample is restricted to those who complete year 10 that is relevant. The usable sample when using Year 10 marks is 417. The results for first stage equation are given in table 4.13. Looking at the distance coefficient in the first stage equation in Table 4.13, it is significant at the 10% level in the participation equation and also is not significantly different from the distance coefficient when using the Year 9 sample in Table 4.5. The urban dummy coefficient also has a similar result when using the Year 9 sample. The second stage results from the Heckman model are presented in Table 4.14 and they also show broadly the same pattern as when using the Year 9 sample.

Table 4.13: Participation equation, Probit model: first stage using Year 10 sample

Variables	Marginal Effects	Standard Errors
Log of distance	-0.046*	(0.027)
Urban dummy	0.279***	(0.000)
Age	0.002	(0.009)
Age missing	0.029	(0.146)
Male	-0.015	(0.026)
Own house	-0.035	(0.030)
Number of siblings	-0.014*	(0.008)
Fathers' education	0.003	(0.003)
Mothers' education	-0.004	(0.003)
Fathers' income	6.13	(1.18)
Birth order	0.018**	(0.008)
Health status	-0.025	(0.024)
Private school Year 9	-0.079**	(0.040)
Number of rooms at home	-0.001	(0.005)
Mode of Transport	0.060**	(0.027)
Private coaching	-0.027	(0.022)
Study time home	-0.062**	(0.030)
Prior attainment Year 9	-0.000	(0.000)
Observations	417	

1. Table 4.13 presents the marginal effects (M E) on the probability of participation into post-compulsory education. 2. Standard errors in parentheses. 3. *** p<0.01, ** p<0.05, * p<0.1

Table 4.14: Determinants of Performance in Post-compulsory Education: Second stage results using Year 10 sample

Variables	Model 5 (OLS) Coefficients S. E.	Model 6 (Heckman) Coefficients S.E.	Model 7 (Heckman) Coefficients S.E.
Age	-1.204	-3.159	-2.978
	(5.838)	(5.740)	(5.7444)
Age missing	-8.619	-38.516	-36.090
	(90.126)	(88.59)	(88.666)
Male	-9.426	-16.212	-16.699
	(17.181)	(15.767)	(15.787)
Own house	4.925	2.061	1.009
	(17.181)	(16.670)	(16.742)
Number of siblings	8.654*	7.104	6.916
	(5.042)	(4.953)	(4.966)
Fathers' education	-0.092	-0.462	-0.469
	(1.882)	(1.830)	(1.834)
Mothers' education	-0.204	0.098	-0.181
	(1.710)	(1.660)	(1.663)
Fathers' Income	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Birth order	-2.267	-0.042	0.143
	(4.897)	(4.839)	(4.848)
Health status	-15.006	-12.360	-12.769
	(14.336)	(13.972)	(13.992)
Public Institute Year 12	-69.071***	-69.281***	-68.802***
	(15.738)	(15.370)	(15.347)
Independent candidate Year 12	-67.402***	-70.863***	-70.811***
	(26.654)	(25.717)	(25.707)
Number of rooms at home	2.144	1.8161	1.807
	(3.163)	(3.081)	(3.085)
Mode of transport	-4.694	2.787	-2.896
	(14.666)	(14.222)	(14.245)
Private coaching	-32.238	-30.981	-31.134
	(13.415)	(13.029)	(13.054)
Study time home	12.043*	11.188*	10.695
	(15.400)	(14.222)	(15.016)
Prior attainment Year 9	0.784***	0.775***	0.778***
	(0.057)	(0.056)	(0.056)
Constant	116.505	154.836	152.652
	(103.271)	(102.272)	(102.390)
Observations	266	417	417
Rho	-	0.387	0.513
	-	0.230	0.229
Lambda	-	43.668	51.293
		23.411	23.604

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1,
Model 5- OLS
Model 6 - Heckman selection corrected model (Second stage) - $\chi^{2} =$ 2.38, P- value = 0.123, Probit: First stage uses the log of distance and urban dummy as instruments.
Model 7 - Heckman selection corrected model (Second stage) using interaction term between log of distance and urban - χ^2 = 2.70, P-value = 0.1001, Probit: First stage uses the log of distance, urban dummy and its interaction.

After finding the same pattern of results as when looking at Year 9, it can be argued that the two decisions (complete compulsory schooling and proceed into post-compulsory education) do not need to be modelled separately, but rather there is really one decision and basically to treat the "complete Year 10" decision and "enter Year 11" decision as a single decision whether to continue or not - whether to participate in post-compulsory education or not, with those choosing not to do the Year10 exams simply dropping out a little earlier because they may know that they are not going to continue anyway.

4.6.2 Using different dependent variable

The results are also robust to change in the dependent variable, using a common subject score in Year 12, English Year 12 instead of total marks in Year 12 and controlling for a respective subject score in Year 9 (English Year 9) and Year 10 (English Year 10) separately. The results are presented in Table 4.15 and 4.16. These tables show that the results remain qualitatively the same when using subject score in Year 12 compared to total Year 12 score.

Table 4.15: First stage results[36]: Probit model for participation in post-compulsory education using Year 9 sample

Covariates	**M.E**	**(S.E)**
Log of distance	-0.837***	(0.351)
Urban dummy	3.562***	(0.723)
Prior attainment Year 9 (English Year 9 score)	0.011	(0.012)
Number of observations	552	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 9 sample: Dependent variable Year 12 English subject score

Covariates	Coefficients	(S.E)
Prior attainment Year 9 (English Year 9 score)	0.995***	(0.078)
Rho	0.268	(0.195)
Lambda	3.538	(2.607)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 1.54, P- value = 0.21
*** p<0.01, ** p<0.05, * p<0.1

[36] In the interest of brevity, only coefficients of interest are shown in most of the results tables and both first stage and second stage are put together

Table 4.16: First stage results: Probit model for participation in post-compulsory education using Year 10 sample

Covariates	**M.E**	**(S.E)**
Log of distance	-0.719**	(0.392)
Urban dummy	3.683***	(0.806)
Prior attainment Year 10 (English Year10 score)	-0.004	(0.0134)
Number of observations	416	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 10 sample: Dependent variable Year 12 English subject score

Covariates	Coefficients	(S.E)
Prior attainment Year 10 (English Year10 score)	0.885***	(0.067)
Rho	0.336	(0.219)
Lambda	4.335	(2.875)

Heckman selection corrected model (ρ = 0) - χ^2 = 1.70, P- value = 0.19
*** p<0.01, ** p<0.05, * p<0.1

4.6.3 Distance to the nearest school of the same kind

It can be argued that it is not the distance to the nearest school that matters, but the distance to the nearest school of the type the family would consider (for example arguing that if a family would only consider going to a private school, then the fact that a public school is only 1 mile away would be irrelevant, it would be how close the nearest private school is that would be relevant).

To address the potential criticism of the main results that the distance variable was not reflecting a real choice for some people, who wanted to send their children to a different type of school, a new distance variable is created and used in performing this further analysis. So the idea in calculating the new distance variable for young people is whether the school for which distance is measured is the same type as what they want to attend in the post-compulsory phase.

The results using the new distance variable in Table 4.17 show that when we restrict the sample to those for whom the distance variable does reflect their choice of type of school, the qualitative nature of the results is unaffected.

Table 4.17: First stage results: Probit model for participation in post-compulsory education using Year 9 sample

Covariates	**M.E**	**(S.E)**
Distance to nearest same type of school	-0.862***	(0.352)
Urban dummy	3.351***	(0.728)
Prior attainment Year 9	0.000	(0.002)
Number of observations	460	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 9 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 9	1.604***	(0.121)
Rho	0.350	(0.259)
Lambda	33.041	(24.843)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 1.55, P- value = 0.21
*** p<0.01, ** p<0.05, * p<0.1

4.6.4 Check for valid exclusion restrictions

It can be argued that distance and the urban dummy could also affect performance (e.g. if living far away makes you miss classes), and so they are not valid exclusion restrictions.

To confirm, these instruments are tried in the performance equation, then tested empirically, to see if they attract significant coefficients in the performance equation. The findings show that they attracted insignificant coefficients in the performance equation. So the results, given in Table 4.18 and 4.19 empirically support them being genuine exclusion restrictions in the particular context of this study. The results are robust for the Year 10 sample as well, which are given in Table 4.20 and 4.21.

One other potential concern would be about where people live being endogenous - i.e. they have specifically chosen their house location for the access it gives to schools- as people choosing where to live could potentially affect the instruments: distance and urban, for reasons that also might affect participation. For example, if those who value education more choose to deliberately live near schools. Distance, in this case, would not be a random exogenous variable but would have been endogenously determined.

Evidence suggests that this does happen in the UK (Allen et al., 2010). For example, in Britain, most schools are allocated on the basis of the catchment area: it is mainly due to the fact that places are usually offered first to children, who live nearest to a school (Gibbons, 2012).

It is less relevant in Pakistan though as such decision-making is not usual in Pakistan. Generally, people in Pakistan do not choose their house and move location for this reason as it is hard for them to afford due to the additional cost involved in relocation. Also, there is no official policy to assign school places on the basis of catchment area as there are no defined catchment areas or postcodes. Parents can choose any school depending upon mainly their affordability and then preferences. So this criticism is probably not very important in the existing case.

Table 4.18: First stage results: Probit model for participation in post-compulsory education using Year 9 sample

Covariates	**M.E**	**(S.E)**
Log of distance	-2.461***	(0.229)
Prior attainment Year 9	0.001	(0.001)
Number of observations	553	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 9 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 9	1.506***	(0.103)
Urban dummy	19.192	(81.126)
Rho	0.239	(0.347)
Lambda	23.333	(34.25)

Heckman selection corrected model ($\rho = 0$) - $\chi^2 =$ 0.55, P- value = 0.48
*** p<0.01, ** p<0.05, * p<0.1

Table 4.19: First stage results: Probit model for participation in post-compulsory education using Year 9 sample

Covariates	**M.E**	**(S.E)**
Urban dummy	5.00***	(0.494)
Prior attainment Year 9	-0.000	(0.002)
Number of observations	553	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 9 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 9	1.503***	(0.103)
Log of distance	-8.244	(14.250)
Rho	0.394	(0.239)
Lambda	38.661	(23.998)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 2.01, P- value = 0.15
*** p<0.01, ** p<0.05, * p<0.1

Table 4.20: First stage results: Probit model for participation in post-compulsory education using Year 10 sample

Covariates	**M.E**	**(S.E)**
Log of distance	-2.403***	(0.264)
Prior attainment Year 10	0.000	(0.001)
Number of observations	417	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses
. *** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 10 sample: Dependent variable total score in Year12

Covariates	Coefficients	(S.E)
Prior attainment Year10	0.778***	(0.057)
Urban dummy	63.081	(70.677)
Rho	0.485	(0.292)
Lambda	48.798	(30.499)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 1.38, P- value = 0.24
*** p<0.01, ** p<0.05, * p<0.1

Table 4.21: First stage results: Probit model for participation in post-compulsory education using Year 10 sample

Covariates	**M.E**	**(S.E)**
Urban dummy	4.987***	(0.576)
Prior attainment Year 10	-0.000	(0.001)
Number of observations	417	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 10 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 10	0.775***	(0.057)
Log of distance	0.913	(14.903)
Rho	0.370	(0.247)
Lambda	38.828	(24.998)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 21.60, P- value = 0.20 *** p<0.01, ** p<0.05, * p<0.1

4.6.5 Different combinations of the instruments

The results are also robust to different combinations of the instruments. The following uses either an urban dummy or log of distance separately as a sole restriction and compares the results to those given earlier using both variables together as exclusion restrictions. The results are robust across Year 9 (Tables 4.22 and 4.23) and Year 10 sample (Tables 4.24 and 4.25). The results remain qualitatively similar to those presented earlier.

Table 4.22: First stage results: Probit model for participation in post-compulsory education using Year 9 sample

Covariates	**M.E**	**(S.E)**
Log of distance	-2.456***	(0.227)
Prior attainment Year 9	0.001	(0.001)
Number of observations	553	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 9 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 9	1.506***	(0.102)
Rho	0.181	(0.243)
Lambda	17.651	(23.841)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 0.56, P- value = 0.45
*** p<0.01, ** p<0.05, * p<0.1

Table 4.23: First stage results: Probit model for participation in post-compulsory education using Year 9 sample

Covariates	**M.E**	**(S.E)**
Urban dummy	4.968***	(0.489)
Prior attainment Year 9	0.001	(0.002)
Number of observations	553	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 9 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 9	1.530***	(0.101)
Rho	0.358	(0.234)
Lambda	35.104	(23.8344)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 01.85, P- value = 0.17
*** p<0.01, ** p<0.05, * p<0.1

Table 4.24: First stage results: Probit model for participation in post-compulsory education using Year 10 sample

Covariates	**M.E**	**(S.E)**
Log of distance	-2.364***	(0.255)
Prior attainment Year 10	0.000	(0.001)
Number of observations	417	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 10 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 10	0.777***	(0.056)
Rho	0.275	(0.298)
Lambda	27.395	(29.918)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 0.75, P- value = 0.38
*** p<0.01, ** p<0.05, * p<0.1

Table 4.25: First stage results: Probit model for participation in post-compulsory education using Year 10 sample

Covariates	**M.E**	**(S.E)**
Urban dummy	4.963***	(0.568)
Prior attainment Year 10	0.000	(0.001)
Number of observations	417	

Marginal effects (ME) on the probability of participation into post-compulsory education.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Second stage results: Determinants of Performance in post-compulsory education using Year 10 sample: Dependent variable total score in Year 12

Covariates	Coefficients	(S.E)
Prior attainment Year 10	0.786***	(0.056)
Rho	0.358	(0.234)
Lambda	35.104	(23.8344)

Heckman selection corrected model (ρ = 0) - $\chi^2 =$ 2.07, P- value = 0.15
*** p<0.01, ** p<0.05, * p<0.1

In all results tables for the Year 9 and 10 samples estimated so far, the lambda, which is the coefficient on the Inverse Mills Ratio, is not significant showing the lack of selectivity. Rho is the correlation coefficient between the error terms in the two equations (participation and performance) and is also not significant in all results tables, suggesting they are two unrelated processes. Also, there is a test below all of the results tables of ρ=0, which is not rejected. These strongly support that participation and performance are two independent processes and there is no selectivity bias.

On top of this, significant coefficients on the log of distance and urban dummy in every participation model suggest that participation is mainly driven by accessibility. Given that we consider all of these factors in a multivariate setting (participation and performance), it could be due to the fact that at age 14-15 (participation model) distance from school is more relevant relative to those at age 17-18 (performance model).

Also the prior attainment has an insignificant coefficient in all estimated participation models while it has a significant coefficient in all performance models after controlling for selection into post-compulsory education, this strongly supports that after taking into account the selection into post-compulsory education performance depends on ability.

4.7 Conclusions

Participation and performance in education are discussed in the literature through different perspectives. A number of studies have explored participation in higher education in developed countries with the main findings being that previous attainment and family background are the variables most associated with participation (Lenton, 2005; McIntosh 2001). However, a very small number of studies have discussed the geography of education with respect to participation in post-compulsory education (Dickerson & McIntosh, 2013). An enormous literature is available discussing different determinants of education outcomes in developed and developing countries (Engin-Demir, 2009; Aslam and Siddiqui, 2003). However, participation and performance in post-compulsory education have not been considered jointly before.

Therefore, to fill the gap existing in the literature on the field of education, this chapter jointly explored the determinants of participation and performance in post-compulsory education controlling for the selection effects using unique survey data on pupils studying in Year 12 at age 17-18 from one of the districts of Punjab, Pakistan in 2011-12. The present survey data is also linked with administrative data on student exams scores, together with the teachers, schools and family levels survey data.

The two main features of this study are 1) Investigate the performance controlling for the selection into post-compulsory education. 2) Participation and performance are studied using a control for prior ability, which is rare.

Additionally, this chapter focuses on the impact of accessibility and availability of institutions on participation in education. Accessibility is defined as the distance from pupils' homes to the nearest post-compulsory education institutions and availability is defined as the location where pupils live – urban or rural area. The distance (km) variable is created in the follow-up survey conducted in 2013-14 using actual travel distance by driving from pupils' homes to the nearest post-compulsory education institutions using information on the location of their homes collected in 2008-09 at the time the survey was initiated.

Further, an interaction of accessibility and availability i.e. interaction between the log of distance and urban is used to see whether distance affects participation differently across urban – rural locations.

The Heckman selection model is used to study participation and performance jointly using distance and urban as exclusion restrictions. The value of the coefficient of rho, which tells about the bias associated due to any non-random selection of the post-compulsory participating sample, is insignificant suggesting that participation and performance are two completely different processes. The key variables strongly associated with participation are distance and urban while performance is mainly associated with ability and type of institution attended i.e. the coefficients on prior ability in Year 9 is highly significant in the performance equation and insignificant in the participation equation.

Thus those participating in post-compulsory education are not necessarily the most able, but those with the best access to post-compulsory education. It suggests that participation in further and higher education can be increased by giving more access to further education institutions within travel distance to young pupils, either by increasing the number of further education institutions or upgrading the existing secondary schools up to the level of further education institutions. Up-gradation of existing institutions would likely be less costly than establishing new institutions. As distance is a deterrent to participate in further education, then, distance learning can be considered as a policy response. Due to the rapid increase in technology, it could be possible by giving remote access to libraries, developing websites via posting teaching materials, use of podcasts and video of lectures and classes, online submission of assignments, all these can make distance learning an engaging possibility.

With respect to distance, the current analysis is not able to identify the cause of reduced participation due to higher distance. If distance reduces participation due to longer travel time that involves longer commuting and mental boredom, then, reducing the travel time by increasing the amount of public transport is an effective policy. If distance reduces participation due to financial issues, then giving a subsidy, grant or loan to the young pupils would be an appropriate option for example, the Canada Student Loan Program (CSLP) provides loans to poor pupils (Frenette, 2006).

Further, the results find that participation is more an urban phenomenon than rural, with distance reducing participation more for those living in rural areas. Provision of more further education institutions in rural areas and targeting rural pupils for further education subsidy more than urban pupils as they already have the greater level of disadvantage would be an optimal policy. The distance learning policy discussed in the above scenario may not be appropriate in rural areas as there is currently only limited availability of technology. The most rural areas in developing countries do not have access to the internet or very low strength signals, without which distance learning might not be possible.

Given the unique nature of the analysis describing the participation and performance in detail, it is obvious that the sample size is small. However, the current analysis expects reduced bias due to controlling for prior ability and selection in post-compulsory education institutions. Also the distance and urban dummy variables seem to be strong exclusion restrictions under this particular study, however, one should be cautious before generalising to other parts of Pakistan, or indeed to other countries. For-example, Bahawalpur district is an agricultural state in South Punjab, while other metropolitan areas may exhibit different patterns of participation or performance.

Finally, regarding the econometric point of view, the paper increases the confidence of the reader and researcher proving distance as a valid exclusion restriction where they require an instrument for educational outcomes such as qualification acquired or graduation in estimated wage equations and has relevance to any empirical field of study.

Chapter 5:

Conclusions

This thesis has analysed three related, yet independent empirical chapters, each exploring important topics in the area of economics of education. Throughout the three chapters that form the core of this thesis, I have looked at the determinants of pupils' educational outcomes in high stakes exams in order to contribute to the existing literature and the gaps that are present within it. At present the evidence on intergenerational mobility of education and teacher quality is less clear cut and arguably undeveloped particularly within the teacher quality field in developing countries. Also, performance in post-compulsory education after taking selection into post-compulsory education into account has never been discussed before. Clearly, there is scope to build upon existing research and explore these issues together. To summarise, Chapter 2 looked at the effect of fathers' education and mothers' education on the education of their children. Chapter 3 explored two questions: "what is the impact of being taught by high quality and low quality teachers on pupils' performance" and "what characteristics of teachers make them high performing and low-performing teachers". Chapter 4 investigated the performance in post-compulsory education after controlling for the selection into post-compulsory education.

5.1 Thesis Summary

The first empirical chapter, chapter 2, examined the impact of parents' education on children's education where children's education is measured in the main analysis as their GCSE exams score (GCSE total points score) which is the sum of total performance across all GCSEs taken at age 16. The data used were from England, using the first wave of the Longitudinal Study of Young People in England (LSYPE) matched to the National Pupil Database (NPD) to perform the analysis.

The main concern in estimating the intergenerational mobility of education is to isolate the effects of parents' education due to parents' genes and parents' parenting style. These are "nature" and "nurture" effects respectively. IV methodology is used, and the raising of the school leaving age (RoSLA) of 1972 in England is used as an instrument to identify the exogenous variation in parents' education that is orthogonal to parents' genes. If we still find the effect on children's education then it could only be due to the effect of upbringing.

The chapter contributed to the existing literature on intergenerational mobility in several ways. Firstly, it used a more comprehensive measure of pupil performance in England at age 16. Most previous studies are limited in using achieving five or more GCSEs at grade C or above as a measure of pupil outcomes that does not include the information that we have in the NPD. Secondly, we estimated the effect of fathers' education and mothers' education separately using separate specifications, which helps to evaluate the effect of each parent's education specifically. In addition the strength of the current analysis is that it estimates a model controlling for the education of both parents which is less common in existing literature.

The initial results from OLS models show that there is a positive effect of fathers' education and mothers' education on pupils' outcomes. The results from the IV model confirm the findings from the OLS model, while additionally, results suggest that when controlling for both parents' education, then mothers' education is positively related to children's education while effects of fathers' education are not significant. Interestingly, these findings are robust when splitting the sample across daughters and sons.

Remaining in the field of education and similarly focusing on the determinants of performance, Chapter 3 answers two main questions about teacher quality. Firstly, does teacher quality matter? Secondly, what determines teacher quality? Using a survey dataset linked to the administrative dataset of the exam board in one of the districts of Punjab, Pakistan.

The survey data has detailed information on pupils' individual, family, peers, schools and teachers characteristics. The outcome variable is measured as the exam score of Year 9 pupils at age 14-15 in 2008-2009 using a unique pupil identifier through exam board data, namely the Board of Intermediate and Secondary Education (BISE).

Using a standard education production function, a teacher fixed effects approach is employed that allows us to pick up the effect of all teacher invariant characteristics across pupils. The key point in the analysis is that I observe different teachers teaching the same pupils in the same schools. Therefore, pupil and school characteristics are in effect being held constant, and any variation in pupil performance across subjects for the same pupil is attributed to the effects of different teachers. The first stage estimates of teacher fixed effects tell about how each teacher does with the pupils (within a school) relative to other teachers with the same pupils.

The analysis in this chapter also exploits the within pupils variation i.e., the pupil fixed effects approach addressed the issue of non-random assignments of teacher to students. If more able pupils are assigned to more able teachers then it would be a threat for the validity of the analysis as the resulting estimates would be biased upward or downward. The pupil fixed effects approach also serves as the proxy for prior ability which is not available in the dataset. Further, in the second stage, the estimated teacher effect is regressed on teacher variables in order to know which specific characteristics drive the differences in performance picked up by these teacher fixed effects.

First stage results from the teacher fixed effects approach show that there is a considerable variation in teachers' effects. Some teacher effects are positive and highly significant while others are negative and significant making them high performing and low performing teachers. A good teacher defined as placed at the 75th percentile is related to an increase in pupil test scores by 0.15 of a standard deviation, while a bad teacher decreases the score by 0.77 of a standard deviation. Therefore, a pupil having been taught by a good teacher placed at the 75th percentile teacher scores 0.92 of a standard deviation more than a pupil who is taught by a bad teacher placed at the 25th percentile teacher leaving a significant effect on pupil performance.

The results from the pupil fixed effects approach show that, a pupil taught by a good teacher placed at the 75th percentile has an increase in exam scores of 0.15 standard deviations, while a bad teacher placed at the 25th percentile decreases the score of a pupil by 1.2 standard deviations. Therefore, a pupil taught by a good teacher placed at the 75th percentile scores 1.35 standard deviations more than a pupil who is taught by a bad teacher placed at the 25th percentile. The results from pupil fixed effects are more pertinent given that they control for unobserved pupils' characteristics. The second stage results show that observable teacher characteristics explain little of this variation in teacher quality.

The final empirical chapter, chapter 4, examines the effect of distance to nearest post-compulsory education institutes on the immediate participation in post-compulsory education and studies the socio-economic factors affecting achievement in post-compulsory education (Year 12) conditional on that participation. There is a limited literature that investigates the effect of distance on the immediate participation in post compulsory education. A study by Dickerson and McIntosh (2013) is the only study to examine the impact of distance on immediate post-compulsory education choices at age 16 but they do not study performance in post-compulsory education.

The data used in this chapter is the same as in chapter 3. The first wave collects the data on the 611 pupils and their family, school and teachers. Second, third and fourth waves collected the data again on exams scores on sampled pupils studying in Year 10 aged 15-16 in 2009-10, Year 11 aged 16-17 in 2010-11, and Year 12 aged 17-18 in 2011-12. The current study uses the pupils' score in Year 9, Year 10, and Year 12. Year 9 and Year 10 are the last two years of compulsory schooling as in Pakistan compulsory schooling is from Year 1 to Year 10. Year 11 and Year 12 are post-compulsory schooling years. The distance variable is created during the follow-up survey later.

A value-added achievement production function is used to estimate the determinants of performance controlling for selection into post-compulsory education. The Heckman selection model is used to account for the selection bias resulting from a choice whether pupils continue or not after completing the certain level of education.

The analysis presented several findings. Firstly, distance has a negative impact on the likelihood of participation in post-compulsory education. Secondly, living in urban area has a positive effect on the probability of participation in post-compulsory education. Thirdly, after taking into account the selection into post-compulsory education, prior ability has a positive effect on the performance. Finally, a lack of significant selection effects shows that participation and performance are not dependent on each other. Several checks are carried out to gauge the validity of instruments used and distance and urban dummy are found as genuine exclusion restrictions.

5.2 Policy Implication and Avenues for Future Research

As education is important for later life outcomes (e.g. employment, earnings, health), so it is very important to understand what the determinants of education outcomes are, in order to design programs that are most effective at improving such outcomes.

The findings of chapter 2 indicated that after taking into account the endogeneity of parents' education, there is a positive association between parents' education and children' education when using the raising of the school leaving age in 1972 as an instrument for parents' education. This finding implies that intergenerational spillovers may provide a compelling argument for policies designed to encourage young people to remain in schools for longer and reinforce the policies to increases the participation in education. In this regard the recent reform of education in the UK in 2015 that increases the *de facto* education/training participation age until 18 seems a promising step in the appropriate direction.

Since the RoSLA effects identify a substantial nurture component to intergenerational mobility, it is possible to design policies that improve pupil performance for those with parents who have low education by providing support to improve parenting style to match that which high education parents are able to provide for their children. If intergenerational mobility had no nurture effects then such policies would be ineffective.

Future research that could be done on the true transmission mechanism if there is any, for example the mothers affected as a result of school leaving age policies are less likely to smoke or drink during pregnancy. Obviously, this new research requires a rich dataset.

Additionally, future work to disentangle the effect of reform for men and women from marriage market effects may be important as it helps to account for the asymmetry across fathers and mothers. So far it is not clear as reform may have an effect on the marriage options of individuals that in turn may affect them differently since the estimates on fathers' education seem noisy. This needs to be explored further.

The results from the first part of the analysis of chapter 3 corroborate the previous findings of teacher quality in developed countries, US and UK (Hanushek and Rivikin, 2006; Aaronson et al., 2007; Slater et al., 2012) and for a developing country, India (Azam and Kingdon, 2015) that there is a substantial variation in teacher effects even after controlling for subject fixed effects. Teacher quality therefore is very important for pupils' performance in high stakes exams in a developing countries context. This means that teachers matter and within schools teacher quality varies a lot. There could be the most effective teachers having a positive effect on their pupils' performance and there could be less effective teachers having a negative effect on pupils' performance. This suggest that pupils could reap benefits from polices that improve the overall teaching quality at schools. It is also an indication to policy makers that it is important to improve average quality of teaching. Given that teacher quality plays an important role in explaining pupils' performance, teacher assignment could be used as a means for reducing unequal educational outcomes across gender and social groups. Improving teacher quality seems less costly compared to student based incentives as improving teacher quality would be applicable for the entire class.

The findings from the second part of chapter 3 show that, as found in US studies, teacher observable characteristics do not explain variation in teacher quality in our dataset. This implies that it is hard to identify good teachers *ex ante*, but administrative data can be used to identify them *ex post*. This brings room for performance management and personnel policies at schools within public policy. Furthermore, teacher progression policies may be radically rethought if *ex ante* discrimination is hard. In addition, polices to hire teachers on the basis of resume characteristics should be reviewed and improved.

As far as policy implications of chapter 4 are concerned, the analysis suggested that distance to the nearest education institution and living in a rural areas limit the participation in post-compulsory education. However, the current analysis is unable to identify the mechanism of these effects being negative. There could be two options to take in response: if distance has a negative effect due to the financial cost associated with it, then providing a subsidy on travelling would be relevant. If distance is negatively affecting participation due to long frequent commuting, then improving the road infrastructure and transport would be an effective policy so that the travel time could be reduced. However, improving overall transport would be more costly.

In addition, together building more schools and improving transportation network can have substantial positive effects both for the rural youth and those living far from schools as it facilitates the accessibility.

The research carried out in chapter 4 serves an econometric advantage that distance to nearest education institution seems a valid instrument for participation in post-compulsory education as found by Dickerson and McIntosh (2013), with an assumption that household location is judged as exogenous. This is relevant to any empirical field where an instrument is required for education.

Finally, each chapter of this thesis individually contributed to the literature on a wider perspective, factors affecting pupils' performance in high-stakes exams with a particular focus on intergenerational mobility of education in the UK and teacher quality and determinants of participation and performance in post-compulsory education in Pakistan. At present, given the lack of consensus in the field of intergenerational mobility, teacher quality and virtually non-existent research in the field of economics of education in developing countries particularly in Pakistan, it develops scope to build upon the existing research. In its entirety, this thesis sheds light on the wider determinants of pupils' performance in high stakes exams including in the subareas of parents, teachers and socio-economic determinants and call for further research to investigate how policy makers may intervene to increase pupils' performance. The research presented in all three chapters therefore provides a step further towards closing the gaps in the previous literature within the field of economics of education. We hope that analysis within this thesis provides important suggestions for future researchers, education economists and stimulates research in this area.

Questionnaire

(STUDENT)

PART – I

Child Characteristics

Name : ___________________________ Father's name: ___________

1. Date of Birth: ____________________
2. Sex: ___Male / Female
3. Group: _________ Science / Arts
4. Class Roll No. ___________
5. Board Examination Roll No. ________
6. Height: ___________
7. Weight: __________________________
8. Age: _____________
9. Do you participate in extra curricular activities? Yes / No
10. If yes, what type of activities?
    a. Debates
    b. Drama
    c. Music
    d. Painting
    e. -----------
11. In which position you come in class.
    a. 1st
    b. 2nd
    c. 3rd
    d. -----------
12. What is your hobby?
    a. Stamp collection
    b. Coin collection
    c. Gardening
    d. ______________________
13. Do you perform any physical activity? Yes / No
14. If yes, what type of physical activity.
    a. Games
    b. Walk
    c. Exercise
    d. -----------

15. How much time you spend in a week on such activity. __________Hours

16. What is your wish to become a.

    a. Doctor

    b. Engineer

    c. Pilot

    d. Teacher

    e. ------------

    f. None

17. Do you take tuition? Yes / No.

18. If yes, for how much time? ___________Hours

19. Do you get daily homework from school? Yes / No

20. If yes, how much time you spend on homework daily ________ Hours.

21. Who helps you in your study?

    a. Mother

    b. Father

    c. Brother

    d. Sister

    e. -----------

22. Are you suffering from any chronic disease? Yes / No

23. If yes what type of disease.

    a. Cough

    b. Cold

    c. Week eyesight

    d. Asthma

    e. -----------

24. Does your disease affect your work? Yes / No

25. How do you go to school?

    a. Bicycle

    b. Motor Cycle

    c. Car

    d. school bus

    e. walk

26. How long does it take for you to reach to school from home? ___________Hours

27. Do you get pocket money? Yes/no

28. If yes, how much Rs_______ daily.

PART – 2

Family Characteristics

1. Father's Education:____
2. Father's Profession ____
3. Father's Income: Monthly (Rs.) _________________________________
4. Mother's Education:____
5. Mother's Profession ____
6. Mother's Income: Monthly (Rs.) _________________________________
7. Locality where you live.____
8. Residence:
    a. Rental
    b. Own
    c. Sharing
9. How many rooms in your home
10. Total Family Size____________
11. a. Number of sisters ________
    b. Number of Brothers ______
12. Information about children from first child to youngest.

S.No	Age	Sex	Class / Education	School / Occupation	Monthly income

13. What is your number among children?_______________

14. Who else is the earning member in the family not accounted above?____________________

15. Total monthly of all members as mentioned in no 14.

16. How many time you take eat in a day.

    a. One Time
    b. Two Times
    c. Three Times
    d. Four Times
    e. ___________

17. Do you take breakfast daily? Yes/no

18. If no, how many times in a week?_________

19 Do you take eatable in school break daily? Yes/no

20. Weekly nutritional chart. (food items consumed).

Food items	**Yes / No**	**Days in a Week**	**Quantity**
Milk			
Yogurt			
Bread			
Meat			
Vegetable			
Fruit			
Juice			
Egg			
Rice			
Pulses			

Questionnaire

SCHOOL HEAD

Name: ______________________________ School Name:________________

Age: ________________________________ Gender: ____________________

Qualification: ________________________ Training: ____________________

Total Teaching: _______________________ Administration Experience: _____

Experience in present school: ____________ Present Position in School: ______

Tuition Fee per Month (9^{th} Class): _________ Other Expenditure (9^{th} Class): ___

Total enrolment of school: ______________ 9^{th} Class enrolment: ____________

How many teachers are in school? Male ____ Female ______

No of teachers teaching 9^{th} class: ____________

Qualification of 9^{th} Class Teachers / Training _________________________________

Monthly salary of teachers: _______________

School efforts to improve skills of teachers.__________________________________

____

Do you think that school infrastructure is enough according to the needs of students.

Yes / No

Facilities:

Sr. No.	Facilities	Yes / No	If Yes How Many
01	Canteen		
02	Library		No of Books
03	Computer Lab		No of Computers
04	Sports Complex		Area of Sports Complex
05	Play ground		Area of play ground
06	Class Rooms		No of class rooms

Note: Please tick the (√) the box that comes closest to your views.

Responses: SA=Strongly Agree, A= Agree, UD= Un decided, DA= Disagree, SDA= Strongly Disagree.

Sr.No	Statement	SDA	DA	UD	A	SA
	Honesty					
1	I am honest with my profession.					
2	I advise my teacher to be dutiful.					
3	I develop nobleness in students personality					
	Competent					
4	I am competent as school head.					
5	I make efforts to enhance competency of my staff.					
	Forward looking					
6	I always do futuristic planning.					
7	I look forward to shape new school environment.					
	Inspiration					
8	I inspire my staff to work hard.					
9	I inspire students to achieve their goals.					
	Intelligent					
10	I know my responsibilities.					
11	I do my duties skillfully.					
	Fair Minded					
12	I take decision impartially.					
13	I am fair-Minded toward student's promotion.					
14	I am fair-Minded toward teacher's promotion.					
	Broad-Minded					
15	I accept criticism with open mind.					
16	I promote broad mindedness in the staff.					
	Courageous					
17	I take bold steps for development of school.					
18	I face difficulties with courage.					
	Straight-Forward					
19	I am straightforward in my dealing.					
	Imaginative					
20	I am not imaginative for new ideas.					

21	**Planning** **I** divide my time appropriately.					
22	I always prepare the plan before doing my work.					
23	**Organizing** I organize things properly.					
24	**Staffing** I appoint the staff according to requirement.					
25	**Directing** I give direction to the teachers to increase their efficiency.					
26	I give direction to the students to develop their confidence.					
27	**Co-ordination** I always co-ordinate with teachers.					
28	I always co-ordinate with students.					
29	**Innovating** I introduce innovative skills of the teachers.					
30	I introduce innovative skills of the students.					
31	**Controlling** I control all academic activities nicely.					
32	**Budgeting** I formulate school budget properly.					
33	**Decision-Making** I need to be authoritative as school head.					
34	I have decision making power to solve the school problems					
35	**Human Relation** I always consult my teachers on important school matters.					
36	I solve school problems friendly.					
37	**Reporting** I highlight school performance.					
38	I covey to teachers all new information.					
	Communication					

39	I communicate my ideas to the students.					
40	I communicate modern pedagogy to the teachers.					
41	**Leading** I develop leadership qualities among teachers.					
42	I develop leadership qualities among students.					
43	**Problem solving** I solve teacher's problems properly.					
44	I solve student's problems properly.					
45	**Motivation** I motivate teachers to develop leadership qualities.					
46	I motivate students to play active role in school activities.					
47	**Vision** I have vision about the future needs.					
48	**Voice** I talk my teachers politely.					
49	**Credibility** I have credibility among teachers.					
50	**Commitment** I am committed to my job.					
51	**Devotion** I am devoted to my profession.					
52	**Assumption** I don't believe on assumptions only.					
53	**Management of school Plant** I take special care of school plants.					
54	**Co-curricular activities** I arrange co-curricular activities for personality development.					
55	**Management of Human Resources** I assign duties to the teachers according to their expertise.					
56	**Management of Financial Recourses** I spend financial resources with care. .					

Questionnaire

(School Teacher)

Please check or list the relevant response.

1. Name School Teacher: ______________________________

   Designation: ______________________________

   School Name: ______________________________

   Age ______________________________

   Sex: _____________ Subject: ________________

   a. Male

   b. Female

2. Qualification: a. B.A/B.Sc Teaching Training

   c. M.A/M.Sc a. B.Ed

   d. M.Phil/Ph.D b. M.Ed

   e. Other c. PTC

   d. Others

3. Teaching Experience: a. Less than 5 years

   b. 6 – 10 years

   c. 11 – 15 years

   d. 16 – 25 years

   e. __________________

4. Duration of Class: a. 40 min

   b. 50 min

   c. 1 hour

   d. _____________________

5. Daily work Load: a. 2 Classes / day

   b. 3 Classes / day

   c. _______________________

6. Class Size (Number of Students) ____________________ Students

7. Which teaching ethod you usually use in classroom?

   a. Lecture method

   b. Discussion method

   c. Panel discussion

   d. Demonstration method

e. Activity method

f. Any other method

8. Do you prepare your lesson plan?
   a. Fortnightly
   b. Weekly
   c. Daily
   d. No formal planning

9. Do you teach the students by keeping their individual mental caliber in your mind?
   a. Yes
   b. No
   c. Some times

10. Do you give assignments in terms?
   a. Problems
   b. Memorizing facts
   c. Doing Text Book exercises
   d. Writing articles on certain topics.

11. How many times do you organize panel discussion during a month.
   a. Once
   b. Twice
   c. Thrice
   d. More than that

12. Check the types of tests you give to your pupils?
   a. Oral
   b. Essay type
   c. Objective Type
   d. Objective plus Easy type

13. Are you satisfied with the results of class students?
   a. Yes
   b. No
   c. To some extent

14. In your opinion which one is the best way of keeping class discipline?
   a. Authoritarian
   b. Democratic

15. What type of behavior do you adopt while teaching?
   a. Democratic

b. Sympathetic

c. Aggressive

16. Have you attended any refresher courses?

a. Yes

b. No

17. Are you satisfied with the teaching course that you are teaching during a job?

a. Yes

b. No

c. To some extent

18. Do you give the students a free hand to express?

a. Yes

b. No

c. Some times

19. Are all the basic facilities of classroom sufficient for the current strength of students?

a. Yes

b. No

c. Some extend

20. Which areas of service rules are giving you satisfaction for your work?

a. Salary package

b. Leaves granted

c. Job security

21. In orde4r to perform your duties smoothly school administration supports you

a. Yes

b. No

c. Some Times

22. To achieve a good academic career what is important for students?

a. Teachers

b. Parents

c. School

d. School Heads

e. Others

23. Why do the children normally miss school.

a. Illness

b. No Transport

c. House Work

d. Others

24. What suggestion are needed to improve quality of education?

i. ______________________________

ii. ______________________________

iii. ______________________________

Bibliography

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of labor Economics*, *25*(1), 95-135.

Adolphus, K., Lawton, C. L., & Dye, L. (2013). The effects of breakfast on behavior and academic performance in children and adolescents. *Frontiers in human neuroscience*, *7*.

Ailaan, A. (2014). Million *broken promises: The crisis of Pakistan's out-of-school children*.
Retrieved 5 June 2015 from https://d3n8a8pro7vhmx.cloudfront.net/alifailaan/pages/475/attachments/original/1415177621/25_Million_Broken_Promises_Report.pdf?1415177621

Alderman, H., Orazem, P. F., & Paterno, E. M. (2001). School quality, school cost, and the public/private school choices of low-income households in Pakistan. *Journal of human Resources*, 304-326.

Ali, T. S., Krantz, G., Gul, R., Asad, N., Johansson, E., & Mogren, I. (2011). Gender roles and their influence on life prospects for women in urban Karachi, Pakistan: a qualitative study. *Global Health Action*, *4*(1), 7448.

Alivernini, F., & Lucidi, F. (2011). Relationship Between Social Context, Self-Efficacy, Motivation, Academic Achievement, and Intention to Drop Out of High School: A Longitudinal Study. *The Journal of Educational Research, 104*(4), 241-252.

Allen, R., Burgess, S., & Key, T. (2010). *Choosing secondary school by moving house: school quality and the formation of neighbourhoods*: The Centre for Market and Public Organisation working paper no. 10/238.

Allensworth, E. M., & Easton, J. Q. (2007). What Matters for Staying On-Track and Graduating in Chicago Public Highs Schools: A Close Look at Course Grades, Failures, and Attendance in the Freshman Year. Research Report. *Consortium on Chicago School Research*.

Alm, J., & Winters, J. V. (2009). Distance and intrastate college student migration. *Economics of Education Review*, *28*(6), 728-738.

Anderberg, D., & Zhu, Y. (2014). What a difference a term makes: the effect of educational attainment on marital outcomes in the UK. *Journal of Population Economics*, *27*(2), 387-419.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *The Journal of Economic Perspectives*, *15*(4), 69-85.

Antonovics, K. L., & Goldberger, A. S. (2005). Does increasing women's schooling raise the schooling of the next generation? Comment. *American Economic Review*, 1738-1744.

Asahi, K. (2014). *The Impact of Better School Accessibility on Student outcomes*: Discussion Paper No. 0156 , Spatial Economics Research Centre, London School of Economics.

Ashford, S., Gray, J. and Tranmer, M. (1993*). The introduction of GCSE exams and changes in post–16 participation.* Employment Department Research Series, Youth Cohort Report no. 23.

Aslam, M. (2009). The relative effectiveness of government and private schools in Pakistan: are girls worse off? *Education Economics, 17*(3), 329-354.

Aslam, M., & Kingdon, G. (2011). What can teachers do to raise pupil achievement? *Economics of Education Review, 30*(3), 559-574.

Aslam, M., & Siddiqui, R. (2003). The Determinants of Student Achievement in Government and Private Schools in Pakistan. *The Pakistan Development Review*, 841-876.

Association, N. E. (2012). Research spotlight on academic ability grouping. *Retrieved November, 17*, 2012.

Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., et al. (2009). Evaluating the impact of performance-related pay for teachers in England. *Labour Economics, 16*(3), 251-261.

Azam, M., & Kingdon, G. G. (2015). Assessing teacher quality in India. *Journal of Development Economics, 117*, 74-83.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2007). Enhanced routines for instrumental variables/GMM estimation and testing. *Stata Journal, 7*(4), 465-506.

Becker, G. S. (1964) Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, New York: Columbia University Press

Becker, G. S. (1994). Human capital revisited *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)* (pp. 15-28): The University of Chicago Press.

Becker, G. S., & Tomes, N. (1994). Human capital and the rise and fall of families: The University of Chicago Press.

Behrman, J. R., Foster, A. D., Rosenweig, M. R., & Vashishtha, P. (1999). Women's schooling, home teaching, and economic growth. *Journal of Political Economy, 107*(4), 682-714.

Behrman, J. R., Khan, S., Ross, D., & Sabot, R. (1997). School quality and cognitive achievement production: A case study for rural Pakistan. *Economics of Education Review, 16*(2), 127-142.

Behrman, J. R., & Rosenzweig, M. R. (1999). “Ability” biases in schooling returns and twins: a test and new estimates. *Economics of Education Review, 18*(2), 159-167.

Behrman, J. R., & Rosenzweig, M. R. (2002). Does increasing women's schooling raise the schooling of the next generation? *The American Economic Review, 92*(1), 323-334.

Behrman, J. R., & Rosenzweig, M. R. & Zhang, J (2005). Does increasing women's schooling raise the schooling of the next generation? Reply. *The American Economic Review, 95*(5), 1745-1751.

Behrman, J. R., Ross, D., & Sabot, R. (2008). Improving quality versus increasing the quantity of schooling: Estimates of rates of return from rural Pakistan. *Journal of Development Economics, 85*(1–2), 94-104.

Bingley, P., Christensen, K., & Jensen, V. M. (2009). Parental Schooling and Child Development: Learning from Twin parents. *The Danish National Centre for Social Research Working Paper no: 07–2009*.

Bjorklund, A., Jantti, M., & Solon, G. (2007). *Nature and nurture in the intergenerational transmission of socioeconomic status: Evidence from Swedish children and their biological and rearing parents*: National Bureau of Economic Research.

Bjorklund, A., Lindahl, M., & Plug, E. (2004). Intergenerational effects in Sweden: What can we learn from adoption data? *IZA Discussion Paper No. 1194*.

Björklund, A., Lindahl, M., & Plug, E. (2006). The origins of intergenerational associations: Lessons from Swedish adoption data. *The Quarterly Journal of Economics, 121*(3), 999-1028.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2003). *Why the apple doesn't fall far: Understanding intergenerational transmission of human capital*: National Bureau of Economic Research.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. *The American Economic Review, 95*(1), 437-449.

Blanden, J., & Gregg, P. (2004). Family income and educational attainment: a review of approaches and evidence for Britain. *Oxford Review of Economic Policy, 20*(2), 245-263.

Blanden, J., Buscha, F., Sturgis, P., & Urwin, P. (2012). Measuring the earnings returns to lifelong learning in the UK. *Economics of Education Review, 31*(4), 501-514.

Bohlmark, A., & Lindahl, M. (2007). The impact of school choice on pupil achievement, segregation and costs: Swedish evidence. *IZA Discussion paper no: 2786*

Bound, J., & Solon, G. (1998). *Double trouble: on the value of twins-based estimation of the return to schooling*: National Bureau of Economic Research.

Bradley, S., & Lenton, P. (2007). Dropping out of post-compulsory education in the UK: an analysis of determinants and outcomes. *Journal of Population Economics, 20*(2), 299-328.

Brown, S., Mcintosh, S., & Taylor, K. (2011). Following in Your Parents’ Footsteps? Empirical Analysis of Matched Parent–Offspring Test Scores*. *Oxford Bulletin of Economics and Statistics, 73*(1), 40-58.

Bushway, S., Johnson, B. D., & Slocum, L. A. (2007). Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *Journal of Quantitative Criminology, 23*(2), 151-178.

Card, D. (1995). *Using geographical variation in college proximity to estimate the returns to schooling.* In L. N. Christofides, E. Kenneth Grant, & R. Swidinsky (Eds.), Aspects of labour market behavior: Essays in honour of John Vanderkamp, Toronto.

Carneiro, P. & Heckman, J. (2003). Human capital policy. *NBER Working Paper No. 9495*

Carneiro, P., Meghir, C., & Parey, M. (2007). Maternal Education, Home Environments and the Development of Children and Adolescents. *IZA Discussion Papers no. 3072*.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *The American Economic Review, 104*(9), 2593-2632.

Chevalier, A. (2004). Parental education and child's education: A natural experiment. *IZA Discussion Papers no:1153*.

Chevalier, A. (2004). Parental education and child's education: A natural experiment. *IZA Discussion Papers* no: 1153).

Chevalier, A., Harmon, C., O' Sullivan, V., & Walker, I. (2013). The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics, 2*(1), 1-22.

Chevalier, A., Harmon, C., O'Sullivan, V., & Walker, I. (2013). The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics, 2*(1), 8.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(2), 431-457.

Christofides, L. N., & Hoy, M. (2001). Family income and postsecondary education in Canada. *Canadian Journal of Higher Education, 31*(1), 177-208.

Clark, D. (2011). Do Recessions Keep Students in School? The Impact of Youth Unemployment on Enrolment in Post-compulsory Education in England. *Economica*, *78*(311), 523-545.

Clark, D. (2011). Do Recessions Keep Students in School? The Impact of Youth Unemployment on Enrolment in Post-compulsory Education in England. *Economica*, *78*(311), 523-545.

Clotfelter, C., Ladd, H., & L. Vigdor, J. (2007a). How and Why Do Teacher Credentials Matter for Student Achievement? *National Center for the Analysis of Longitudinal Data in Education Research (CALDER)working paper no: 02*.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of human Resources, 41*(4), 778-820.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007b). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of human Resources, 45*(3), 655-681.

Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). What makes great teaching. *Review of the underpinning research. Durham University: UK.*

Dearden, L., Machin, S., & Reed, H. (1997). Intergenerational mobility in Britain. *The Economic Journal*, 47-66.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics, 88*(9), 1697-1720.

Devereux, P. J., & Hart, R. A. (2010). Forced to be Rich? Returns to Compulsory Schooling in Britain*. *The Economic Journal*, *120*(549), 1345-1364.

Dickerson, A., & McIntosh, S. (2013). The impact of distance to nearest education institution on the post-compulsory education participation decision. *Urban Studies, 50*(4), 742-758.

Dickson, M., Gregg, P., & Robinson, H. (2016). Early, late or never? When does parental education impact child outcomes? *The Economic Journal, 126*(596).

Engin-Demir, C. (2009). Factors influencing the academic achievement of the Turkish urban poor. *International Journal of Educational Development, 29*(1), 17-29.

Ermisch, J., & Francesconi, J. (2004). Intergenerational mobility in Britain: new evidence from the British Household Panel Survey. *Generational Income Mobility in North America and Europe*, 147-189.

Faggian, A., & McCann, P. (2009). Universities, agglomerations and graduate human capital mobility. *Tijdschrift voor economische en sociale geografie, 100*(2), 210-223.

Field, F. (2010). *The Foundation Years: preventing poor children becoming poor adults, The report of the Independent Review on Poverty and Life Chances*: The Stationery Office.

Frenette, M. (2006). Too far to go on? Distance to school and university participation. *Education Economics, 14*(1), 31-58.

Galindo-Rueda, F., & Vignoles, A. (2005). The declining relative importance of ability in predicting educational attainment. *Journal of human Resources, 40*(2), 335-353.

García, E. D. T. (2014). Determinants of educational outcomes: Analysis of the Republic of Tatarstan. *Communist and Post-Communist Studies, 47*(1), 39-47.

Gibbons, S. (2012). Big ideas: valuing schooling through house prices. *Centrepiece, 17*(2), 2-5.

Gibbons, S., Machin, S., & Silva, O. (2008). Choice, competition, and pupil achievement. *Journal of the European Economic Association*, *6*(4), 912-947.

Gibbons, S., & Vignoles, A. (2009). Access, Choice and Participation in Higher Education. CEE DP 101. *Centre for the Economics of Education (NJ1)*.

Gibbons, S., & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional science and urban economics, 42*(1), 98-113.

Glewwe, P., Ilias, N., & Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics, 2*(3), 205-227.

Glewwe, P., & Kremer, M. (2006). Chapter 16 Schools, Teachers, and Education Outcomes in Developing Countries. In E. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. Volume 2, pp. 945-1017): Elsevier.

Gorard, S., & Smith, E. (2006). Beyond the 'learning society': what have we learnt from widening participation research? *International Journal of Lifelong Education, 25*(6), 575-594.

Gray, J., Jesson, D. and Tranmer, M. (1993) *Boosting post-16 participation in full-time education: a study of some key factors.* Youth Cohort Report No. 20, Employment Department Research Series, London.

Hanushek, E. A. (2002). Publicly provided education. *Handbook of public economics, 4*, 2045-2141.

Hanushek, E. A. (2006). Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries*. *The Economic Journal, 116*(510), C63-C76.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*(5), 527-544.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). The Market for Teacher Quality. NBER Working Paper No. 11154. *National Bureau of Economic Research*.

Hanushek, E. A., & Rivkin, S. G. (2006). Teacher quality. *Handbook of the Economics of Education, 2*, 1051-1078.

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annu. Rev. Econ., 4*(1), 131-157.

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annu. Rev. Econ., 4*(1), 131-157.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica, 47*(1), 153-161.

Herrnstein, R. J., & Murray, C. (2010). *Bell Curve: Intelligence and Class Structure in American Life*: Free Press.

Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature, 49*(3), 615-651.

Hoxby, C. M. (1996). How teachers' unions affect education production. *The Quarterly Journal of Economics*, 671-718.

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica, 62*(2), 467-475.

Kane, J., & Spizman, L. M. (1994). Race, Financial Aid Awards and college Attendance. *American Journal of Economics and Sociology, 53*(1), 85-96.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27*(6), 615-631.

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *The Journal of Economic Perspectives, 16*(4), 91-114.

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation*: National Bureau of Economic Research.

Kasirye, I. (2009). Determinants of learning achievement in Uganda. *Economic Policy Research Centre, Uganda*.

Kingdon, G. (1996). The Quality and Efficiency of Private and Public Education: A Case-Study of Urban India. *Oxford Bulletin of Economics and Statistics, 58*(1), 57-82.

Kingdon, G. (1996). The quality and efficiency of private and public education: a case-study of urban India. *Oxford Bulletin of Economics and Statistics, 58*(1), 57-82.

Kingdon, G., & Teal, F. (2010). Teacher unions, teacher pay and student performance in India: A pupil fixed effects approach. *Journal of Development Economics, 91*(2), 278-288.

Kingdon, G. G. (2006). Teacher characteristics and student performance in India: A pupil fixed effects approach,Gobal Poverty Research Group,Working Paper Series no:059.

Kingdon, G. G., & Teal, F. (2007). Does performance related pay for teachers improve student performance? Some evidence from India. *Economics of Education Review, 26*(4), 473-486.

Kingdon, G. G., *The Private Schooling Phenomenon in India*: A Review. IZA Discussion Paper No. 10612. Available at SSRN: https://ssrn.com/abstract=2940602

Koretz, Daniel. 2008. *Measuring Up: What Educational Testing Really Tells Us*. Cambridge, MA: Harvard University Press. pp. 7-14.

Krueger, A., & Ashenfelter, O. (1992). *Estimates of the economic return to schooling from a new sample of twins*: National Bureau of Economic Research.

Ladd, H., Clotfelter, C., & Vigdor, J. (2007). How and why do teacher credentials matter for student achievement. *NBER Working Paper, 142786*.

Lavy, V. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of political Economy, 110*(6), 1286-1317.

Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Economics of Education Review, 29*(3), 480-488.

Leigh, A., & Ryan, C. (2008). Estimating returns to education using different natural experiment techniques. *Economics of Education Review, 27*(2), 149-160.

Lenton, P. (2005). The school-to-work transition in England and Wales. *Journal of Economic Studies, 32*(2), 88-113.

Lloyd, C. B., Mete, C., & Grant, M. J. (2009). The implications of changing educational and family circumstances for children's grade progression in rural Pakistan: 1997–2004. *Economics of Education Review, 28*(1), 152-160.

Løken, K. V. (2010). Family income and children's education: Using the Norwegian oil boom as a natural experiment. *Labour Economics, 17*(1), 118-129.

Long, B. T. (2004). How have college decisions changed over time? An application of the conditional logistic choice model. *Journal of econometrics, 121*(1), 271-296.

Machin, S., & McNally, S. (2008). The literacy hour. *Journal of Public Economics, 92*(5), 1441-1462.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth*. *The Quarterly Journal of Economics, 107*(2), 407-437.

Mayer, S. E. (1997). *What money can't buy: Family income and children's life chances*: Harvard University Press.

McIntosh, S. (2001). The demand for post-compulsory education in four European countries. *Education Economics, 9*(1), 69-90.

Meghir, C. (2005). Educational reform, ability, and family background. *The American Economic Review, 95*(1), 414-424.

Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. *Journal of Development Economics, 99*(2), 486-496.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy, 66*(4), 281-302.

Mincer, J. A. (1974). Schooling and earnings *Schooling, experience, and earnings* (pp. 41-63): National Bureau of Economic Research: New York.

Mulligan, C. B. (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy, 107*(S6), 184-224.

Muralidharan, K., & Sundararaman, V. (2011). Teacher opinions on performance pay: Evidence from India. *Economics of Education Review, 30*(3), 394-403.

Neal, D., & Johnson, W. (1996). The Role of Premarket Factors in Black-White Wage Differences. *Journal of Political Economy, 104*(5), 869-895.

Newbold, K. B., & Brown, W. M. (2015). The Urban–Rural Gap in University Attendance: Determinants of University Participation Among Canadian Youth. *Journal of Regional Science, 55*(4), 585-608.

Nguyen, A., & Getinet, H. (2003). Intergenerational mobility in educational and occupational status: evidence from the US.

OECD. (2007). Qualifications and lifelong learning. In Policy brief.

Oreopoulos, P. (2006). The compelling effects of compulsory schooling: evidence from Canada. *Canadian Journal of Economics/Revue canadienne d'économique, 39*(1), 22-52.

Oreopoulos, P., Page, M. E., & Stevens, A. H. (2003). *Does human capital transfer from parent to child? The intergenerational effects of compulsory schooling*: National Bureau of Economic Research.

Oreopoulos, P., Page, M. E., & Stevens, A. H. (2006). The intergenerational effects of compulsory schooling. *Journal of Labor Economics, 24*(4), 729-760.

ONS. (2011). Earnings by qualifications, 2011

Pakistan Economic Survey 2013-2014. (2014). *Ministry of Finance*. Pakistan.

Payne, J. (1998) Routes at sixteen: trends and choices in the nineties. Research Report No. 55, Department for Education and Employment, London.

Plug, E. (2004). Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *The American Economic Review, 94*(1), 358-368.

Plug, E., & Vijverberg, W. (2003). Schooling, family background, and adoption: Is it nature or is it nurture? *Journal of Political Economy, 111*(3), 611-641.

Pronzato, C. (2012). An examination of paternal and maternal intergenerational transmission of schooling. *Journal of Population Economics, 25*(2), 591-608.

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education economics, 12*(2), 111-134.

Rice, P. (1999). The impact of local labour markets on investment in further education: Evidence from the England and Wales youth cohort studies. *Journal of Population Economics, 12*(2), 287-312.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247-252.

Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education, 4*(4), 537-571.

Rouse, C. E. (1995). Democratization or diversion? The effect of community colleges on educational attainment. *Journal of Business & Economic Statistics, 13*(2), 217-224.

Sacerdote, B. (2000). *The nature and nurture of economic outcomes*: National Bureau of Economic Research.

Sacerdote, B. (2007). How large are the effects from changes in family environment? A study of Korean American adoptees. *The Quarterly Journal of Economics, 122*(1), 119-157.

Sawada, Y., & Lokshin, M. (2009). Obstacles to school progression in rural Pakistan: An analysis of gender and sibling rivalry using field survey data. *Journal of Development Economics, 88*(2), 335-347.

Schütz, G., Ursprung, H. W., & Wößmann, L. (2008). Education policy and equality of opportunity. *Kyklos, 61*(2), 279-308.

Silles, M. A. (2011). The intergenerational effects of parental schooling on the cognitive and non-cognitive development of children. *Economics of Education Review, 30*(2), 258-268.

Slater, H., Davies, N. M., & Burgess, S. M. (2012). Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England. *Oxford Bulletin of Economics and Statistics, 74*(5), 629-645.

Solon, G. (1999). Intergenerational mobility in the labor market. *Handbook of labor economics, 3*, 1761-1800.

Spiess, C. K., & Wrohlich, K. (2010). Does distance determine who attends a university in Germany? *Economics of Education Review, 29*(3), 470-479.

Stinebrickner, T. R., & Stinebrickner, R. (2007). *The causal effect of studying on academic performance*: National Bureau of Economic Research.

Stock, J. H., & Yogo, M. (2002). Testing for weak instruments in linear IV regression. *National Bureau of Economic Research Technical Working Paper Series no:284*.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In identification and inference for econometric models: essays in honor of Thomas Rothenberg, ed. Andrews DW, Stock JH, 80–108: Cambridge University Press.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal, 113*(485), F3-F33.

Trillingsgaard, T., & Sommer, D. (2016). Associations between older maternal age, use of sanctions, and children's socio-emotional development through 7, 11, and 15 years. *European Journal of Developmental Psychology*, 1-15.

U. N. D. P (2003). *Thailand Human Development Report*: United Nations Development Programme.

Vignoles, A., & Meschi, E. (2010). *The determinants of non-cognitive and cognitive schooling outcomes.* Centre for the Economics of Education

Walker, I., & Zhu, Y. (2003). Education, earnings and productivity: recent UK evidence. *Labour Market Trends, 111*(3), 145-152.

Zhao, M., & Glewwe, P. (2010). What determines basic school attainment in developing countries? Evidence from rural China. *Economics of Education Review*, *29*(3), 451-460.