# IMPROVING SCHOOL ATTENDANCE BY RAISING SCHOOL QUALITY

by

Muhammad Jehangir Khan

*A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University*

*Economics Section of Cardiff Business School, Cardiff University*

December 2012

# DECLARATION

This work has not previously been accepted in substance for any degree and

is not concurrently submitted in candidature for any degree.

Signed …… (candidate)

Date ……

**STATEMENT 1**

This thesis is being submitted in partial fulfillment of the requirements for the degree of ……(insert MCh, Md, **MPhil,** PhD etc, as appropriate)

Signed …… (candidate)

Date ……

**STATEMENT 2**

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by footnotes giving explicit references.

Signed …… (candidate)

Date ……

**STATEMENT 3**

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate)

Date ……

## Acknowledgement

I would like to thank and express heartfelt gratitude to my supervisor Professor James Foreman-Peck for his continuous support throughout my PhD. He has been not only a mentor but a dear friend. His inspirational, patient and supportive role helped me along this fulfilling and challenging long road.

I thank the Pakistan Institute of Development Economics (PIDE) for funding my study. Without this support I would have been unable to study at Cardiff University.

I also thank my parents for their inspirational support throughout my life. My siblings, Haris and Huzaifa, have been the best of friends along this long journey. And last, but not the least, my wife, who shares my passion for study. Despite her illness over the last one and half year, she has always prioritised my study.

I am most grateful to all these supporters.

Muhammad Jehangir Khan

# IMPROVING SCHOOL ATTENDANCE BY RAISING SCHOOL QUALITY

Abstract

Using a two generational model of investment in education, this thesis shows that school quality enhancement is very highly valued by the average rural Pakistani family, and by those below the official poverty line. Corollaries are that quality enhancement will be an effective policy for boosting school attendance and that subsistence poverty is not a major reason for keeping children out of school.

The empirical strategy is to demonstrate (chapter 4) that school quality variables influence perceived child cognitive achievement. Parental perceptions are critical rather than objective performance (if they diverge) because it is parents that decide whether their children should attend school. These results provide support for the finding in the following chapter (5) that some school quality variables also significantly influence school attendance. Note that the empirical focus of this study is on the more important attendance rather than the commonly discussed school enrolment.

The data source, the 2004-5 Pakistan Rural Household Survey, PRHS-II, is the only one currently available for Pakistan that makes available a multitude of school quality measures. Identification in the empirical models of school attendance and cognitive achievement is achieved first by establishing and utilising a distinctive feature of the Pakistani environment; the arbitrary and random allocation of resources to state schools economy and society. The quality variable library is confirmed exogenous with the only instrument in the data set suitable on grounds of relevance (correlation with library), community population. If school resources were adjusted to school need, school attendance would not influence the staff-student ratio. But instrumenting this ratio changes its sign in the predicted fashion. For cognitive achievement, cmp (conditional (recursive) mixed process estimator) (Roodman 2009) is implemented to endogenise staff-student ratio in ordered probit models. Since we have either controlled for endogeneity or established the exogeneity of the quality measures in the attendance and cognitive achievement equations, we can be confident that the parameter estimates correctly capture the impact of school quality variables. The possibility that other variables, not instrumented in the school attendance or cognitive achievement equations, are endogenous does not bias these estimates. Nor does including a range of extra community characteristics in the school attendance and cognitive achievement equations affect the school quality parameter values estimated.

School quality measures must be measured at the community/village level to capture the options for school non-attenders. This reduces the precision with which the standard errors can be estimated. Normally it is appropriate to cluster standard errors at the village level, and these results are presented.

Judging by parental assessment of child performance, in rural Pakistan private schools achieve better outcomes than state schools. Despite comparatively low private school fees, the average rural household with three children is unable to afford to send their children to private schools. Simply expanding private schooling provision then is not a solution to the currently poor education available to these households. In any event such expansion is not without its problems because there are typically arbitrary political or regulatory barriers to establishing private schools. Therefore this thesis suggest that improving (perceived) state school quality may be the most effective strategy for improving human capital in rural Pakistan.

## Contents

# List of Tables

# List of Figures

# Chapter 1 Introduction

Human capital accumulation is a critical element in economic development (Mincer 1958; Becker 1975, 1994; Mankiw, Romer and Weil 1992; Abbas and Foreman-Peck 2008a). This is why in the majority of poor, developing countries there is also a low level of educational attainment and high child labour participation. Child work competes with schooling (Weiner 1991; Dasgupta 1995; Basu and Van 1998; Baland and Robinson 2000; Ravallion and Wodon 2001; Rosati and Rossi 2003; Emerson and Knabb 2006; Edmonds 2007; Hou 2010). In this way family circumstances have a significant impact on the accumulation of human capital.

Economic models of the family as a decision making unit (Becker 1965; Becker and Lewis 1973; Becker and Tomes 1976; Schultz 1997; Gao and Wahid 2009[1]; Angrist et al. 2010[2]), not only show them choosing child quantity and quality along with other goods in consumption, but also determining the allocation of family members' time and other resources in the production activities of the family. Families here maximize their collective utility subject to the production function for each argument of the utility function (including child quality and quantity) and a time budget constraint[3]. The family as a decision maker bears a similar relation to household production as does the firm to market production in traditional economic theory. The welfare of each member of the family is normally a part of a unified family welfare function. According to Schultz (1974, 1997), this unified welfare function[4] involves "overheads," and nonmarket (shadow) prices which play an important role in households' producer and consumer activities, including the bearing and rearing of children.

The recognition that the family is an economic unit implies that its consumption behaviour represents collective or joint households' decision[5]. Each member's allocation of time

[1] Gao and Wahid (2009) focus only on child quality.
[2] Angrist et al. (2010) is an empirical test, using a quasi experimental setup, of the quantity-quality trade-off, which theoretically evaluated by Becker and Lewis (1973) and Becker and Tomes (1976).
[3] Time constraint is a sum of market time, time spent at home to produce each argument of the utility function and leisure time. I have given a simple picture here. For detail see Becker and Tomes (1976) and Schultz (1997)
[4] Shadow prices are the opportunity cost of the market goods and household member's time inputs used to produce a unit of the commodity (e.g. Quality and quantity). These are part of the full income (which is treated as exogenous) but not of production constraint. These shadow prices are fixed by market set prices and wages and do not depend on the bundle of commodities consumed by the household. Otherwise, these shadow prices will depend on parent preferences and returns to scale and cease to be exogenous (Schultz 1997). In the setup of Becker and Tomes (1976) these shadow prices are the part of the full income and their specified utility function is maximized subject to full income but in Schultz (1997) the specified utility function is maximized subject to production function for each commodity and a time budget constraint. Here I have only reported the setup used by Schultz.
[5] Here we are referring to unitary model of household's decision, where parents make decisions on the part of their children.

between market and nonmarket activities is best understood in the context of household members' interdependence of needs and activities, as well as their potential and characteristics. This behaviour of the family differentiates their roles, along with investment in its member's physical and human capital. The household's production process is characterized by complementarities, substitution and comparative advantages in household member's endowment of differential skills and earning capabilities. The level of these endowments is typically only fixed in the short run but not in the long run (in a more complete perspective).

Even when these endowments are genetically determined, marital selection would make these endogenous to the family (Mincer and Polachek 1974). In line with Mincer and Polachek (1974) endowments can be augmented by investing in the human capital of household's members including children. It is evident that by the time children enter the first grade of formal schooling, there are significant differences in their mathematical and verbal competences. Different factors are responsible including; (1) human capital accumulated before the child reach the age of six years. The pre-school stock of human capital reflects the influence of varying inputs of time and other resources of parents, siblings and the child, (2) inherent ability. Accumulation of pre-school human capital is as important as acquiring it through formal schooling or on the job training (Leibowitz, 1974).

Childhood in developed economies is a time for school learning and play rather than participation in earning activities, but the high level of child labour in poor countries suggest that there it is not. Recently renewed efforts have been made by the international community to achieve 'Education for All (EFA)[6]' by the year 2015. The situation spurs further theoretical and empirical investigation to isolate different factors central to parental decisions to invest in child schooling. Most of these efforts have been purely theoretical (Basu and Van 1998; Ranjan 1999; Baland and Robinson 2000; Edmonds 2007). Ranjan (1999), Baland and Robinson (2000) and Edmonds (2007) focussed on credit market imperfections on the assumption that households are unable to borrow and lend optimally over the period in which their children ideally would attend school. Alternatively, Basu and Van (1998), Ray (2000a, 2000b) and Bhalotra (2007) assigned responsibility to household level subsistence poverty.

[6] Education for All (EFA) is a global commitment under the auspices of United Nation to provide basic quality education to children, youth and adults. The World Education Forum (Dakar, 2000), comprised of 164 governments, identified six goals and made a commitment to achieve Education for All (EFA) by 2015.

Theory provides competing models based on alternative assumptions. The preference between them may be purely on the basis of subjective judgment of plausibility. But only the confrontation of these competing models with data quantifies and therefore compares them on a more objective basis. The present thesis contributes five chapters to the literature on school quality, school outcomes, school attendance and cognitive abilities. It explains the extent of child schooling in Pakistan in the most plausible fashion. It combines all the three elements of sound economic analysis; economic theory, statistical methods and good data. More specifically, this thesis builds a model of household's child schooling decision. It derives implications for school quality and attendance in the presence and absence of household's subsistence considerations, thereby deriving testable predictions of the role of poverty in the parental schooling decision.

The testing and estimation of the model is based on the Pakistan Rural Household Survey-II of 2004-2005 (PRHS-II). This data set is unique as it has a wide range of indicators of school quality for the schools attended by children in the surveyed households. Such evidence is most often absent in a developing country context. Pakistan is a developing country with a lower level of child schooling and higher child work participation than among neighbouring south Asian countries. Pakistan also continues to experience high population growth and the Total Fertility Rate (TFR) is higher than in neighbouring countries.

The analysis proceeds as follows; chapter two discusses the supply of human capital and reports evidence on the impact of school quality on child school performance in terms of higher earnings in the labour market, cognitive achievement (test scores on standardised tests), and school attendance, as well as the impact of the accumulated human capital on economic growth. The proposition that school quality is an important determinant of educational outcomes appeals to common sense but is a controversial matter in the literature both from developed and developing countries (Case and Deaton 1999). Much micro- and macro- level evidence is presented concerning the quantification of the productivity effects of higher school quality and hence the impact of accumulated human capital on earnings, test scores and economic growth. The reported evidence concludes with high productivity effects of education at the micro-level but no obvious consensus at the macro-level.

Chapter three provided a detailed description of PRHS-II survey. It discusses school quality indicators often absent in a developing country survey. School quality indicators in rural Pakistan are very poor and in some villages the average student-teacher ratio across schools is

as high as 79 pupils. The data also shows pro-male bias in terms of level of education, literacy, current school enrolment and work outside the home on non-family farms in rural Pakistan. Family wellbeing (consumption expenditure) as well as child education is higher in Punjab than in Sindh. The reported estimates on adult literacy, child education, households' consumption, credit rationing and school quality are consistent with other national level surveys (RHPS-2012) and government statistics (GoP 2011), as well as with education estimates from the World Bank.

Chapter four investigates the impact of school quality on child cognitive abilities in the PRHS-II sample. Parental demand for child schooling depends on how much it contributes to their children's skills, which affect earnings in the future. Similarly this study utilizes new measures of child cognitive abilities available in the data set and provides a new framework/methodology to the literature for evaluating cognitive achievement. Particularly it endogenises student-teacher ratio in the context of ordered probit model while implementing cmp.

Chapter five builds a model of the household child schooling decision. Specifically it derives implications for school quality in the presence/absence of household's subsistence considerations. The model predicts that if a household chooses to send its children to school because of an improvement in school quality, it cannot have been below subsistence before the quality increased. In particular it evaluates the income equivalent gain from raising school quality in rural Pakistan.

Chapter six contributes to the debate about whether school quality is better improved by the state or private provision of education in rural Pakistan. This study controls for private school status in the child cognitive production functions, before going on to examine the phenomenon of private school expansion in rural Pakistan. It investigates other important factors; school quality differences between the two educational services providers and the per child cost to the household of each service provider. Chapter six also develops school cost consequences for average household fertility in the data set and its bearing on school attendance in private schools in rural Pakistan.

Finally chapter 7 concludes with policy recommendations based on the evidence presented.

# Chapter 2 Human Capital and Schooling

In this chapter we discuss the concept of human capital which underpins the importance assigned to schooling as an investment in both individual and collective future prosperity. The objective is to show how the idea of school quality modifies, or at least shows in a new light, many of the standard findings of researchers in the area of human capital. This then provides the background and the motivation for the more detailed examination of school quality in later chapters.

Section 2.1 introduces the concept, while section 2.2 outlines perhaps the oldest and most widely researched area of human capital, the Mincer equation, explaining how these results on rates of return can be affected by variation in school quality. The second strand of human capital research discussed in section 2.3 is the macroeconomics of human capital. There we suggest that the international results showing no or a negative effect of human capital on output, may be deceptive because they fail to take into account the heterogeneous quality of generally state-provided education in the sampled countries. As an illustration of the source of such quality variation, and a context for subsequent analysis, section 2.4 describes the supply and organization of education in Pakistan.

## 2.1 The Concept of Human Capital

The OECD defines human capital as 'the stock of competences, knowledge and other attributes embodied in individuals to perform economic activity (OECD, 1998, p.9). In a broader perspective it is described as a 'combination of individuals' own innate talents and abilities and the skills and learning they acquire through education and training (health is also frequently counted among these)' (OECD 2007)

Human capital theorists largely focus on advantages that are developed through education and experience or training, because 'innate ability' by definition is fixed. Here it is important to differentiate acquired ability or advantage from innate ability or talent. Acquired advantage is a characteristic that develops after birth but is not congenital. Analogous is the competitive advantage a country or firm gains through accumulated knowledge, in contrast to that from natural resources. Innate talent or ability must have been in one from birth, existing naturally rather than acquired; potentials that are present in some children but not others. Individuals apparently differ to some extent in their ability to understand complex ideas, to effectively adapt to the environment, to learn from various experiences such as playing a musical

instrument or brick laying, to engage in various forms of reasoning, and to overcome obstacles by taking thought, in the same way that their personalities differ.

Human capital is accumulated by investing in five different kinds of learning environments in the economy that vary over the life cycle (Ederer et al. 2007);

1) Informal education by parents including cultural adaptations;
2) Formal school education;
3) Formal university and higher education
4) Informal learning on the job
5) Formal and informal adult education: Skills which adults acquire outside of their daily work environment, which are nevertheless either directly or indirectly job-related such as management training

Other classifications can be found which divide the above stages into three broad categories of learning;

1) Formal education which takes place in institutions called schools and usually involves young people who have not yet begun their working lives.
2) Non-formal[7] education as organized programmes of learning taking place out of school. These are shorter in nature and narrowly focused, more concerned with applied knowledge and may teach occupational skills or literacy or citizenship.
3) Informal education takes place outside any institutional framework or organized program. People learn many important things at home, in the community or on the job.

Human capital focuses on individual agents - whereas social capital focuses on relationships between them and the networks they form[8]. The accumulation of human capital depends upon values and behaviour patterns which govern the acquisition, deployment and effectiveness of skills. Human capital is measured in terms of earning power, or indirectly by supposed contributors to that earning power, especially by the duration of schooling or the level of qualifications. The direct outcomes of human capital are income and productivity and the indirect outcomes are health and civic activity.

[7] According to Coombs and Ahmed (1974), the distinction between formal and non-formal education is mainly administrative. Formal education is linked with schools and other training institutions; whereas, non-formal linked with community groups and other organizations. Informal consist of what is not covered by the aforementioned two categories e.g. interaction with friends, family and colleagues. Fordham (1993) states that the definitions of these education categories/classification do not imply hard and fast categories. Particularly, there is some overlap/confusion between non-formal and formal. According to infed (2012) the distinction between formal and non-formal education is a matter of some debate.

[8] Social capital is defined 'in terms of networks, norms and trust, and the way these allow agents and institutions to be more effective in achieving common objectives' (Schuller 2001, p.4).

The inadequacy of measures of human capital such as schooling is often acknowledged (Behrman and Birdsall 1983; Behrman et al. 1997). But its easy measurement and the availability of large data sets ensures level of qualification continues to dominate as a standard measure. However, there are many attributes and skills that are developed during childhood outside school, as noted above. Therefore human capital is not easy to measure fully and the standard measures used like duration of schooling and levels of qualification are far from capturing the extent of human capital (OECD, 1998).

Despite this caveat to its measurement, there are strong theoretical and empirical reasons for linking schooling to higher returns in the form of higher wages and productivity for the individual, which fuel economic growth (dating back to Mincer (1958) and Becker (1975, 1994)). Human capital is represented as similar to tangible or physical means of production and one can invest in human capital in the form of education, training and medical care. Hence the output one gets or produces partly depends on the rate of return on the human capital stock one owns.

The next section focuses on the productivity-enhancing role of human capital investment and the inherent problems until recently in its quantification, discussed above.

### 2.2 Microeconomics of Productivity Enhancement by Human Capital: Mincer

Following Harmon et al. (2003), the human capital model underlying the Mincer specification is that the individual invests in education optimally. That is, up to the point where the present value of the extra earnings from a little more schooling should equal the cost, primarily the wage that would have been earned over the extra schooling period. Workers compare the present value of lifetime earnings associated with different levels of education. They continue education as long as the marginal benefits of schooling outweigh the marginal costs. The net present value of an extra year of schooling compared with no schooling is given by:

$$NPV = \sum_{i=1}^{L} \frac{E_0 - E_N - C}{(1+r)^{i-1}} + \sum_{i=L+1}^{R} \frac{E_W - E_N}{(1+r)^{i-1}}$$

where r is the internal rate of return to investment in education. ($E_0 - E_N$) is the forgone earning while educating for another extra year. ($E_W - E_N$) are the benefits accrued to the individual after an extra year of schooling. C is the direct cost of schooling. R is the number

of potential years in the labour force. When the log of earnings is linear in years of schooling, the coefficient on schooling is the internal rate of return.

$$\ln(E_S) = \alpha + \beta S + u \qquad (1)$$

Where $Ln(E_s)$ is the log wage of individual i, (S) is individual i's, years of schooling. The coefficient $\beta$ is the Mincerian private return to schooling. This coefficient indicates that an additional year of schooling raises the wage by $\beta$ per cent on average.

Certain conditions must holds for OLS estimates of the coefficient of schooling $\beta$ to be unbiased. First, individual wages must equal their marginal product. Second, variables correlated with schooling and also affecting wages should not be excluded from the analysis (Jones 2001); for instance ability, family background, school quality. Thirdly, in the derivation of equation (1) it is assumed that for the decision of an extra year of schooling the time period 'T' of earning should be large enough to offset the additional costs (assumed to be forgone earnings only). If the time period in the market workforce is small for some reason as in the case of female labour force participants than males, as in Pakistan, then $\beta$ would be biased (Abbas and Foreman-Peck 2008b)[9].

The conditions discussed above represent possible departures from the standard Mincer model assumptions and certainly influence the wage-schooling associations. Individuals with different ability and family background, differ in the returns they secure from a given course of education (determinants of ability are genetic factors and the amount and quality of time input to children) and in the funds available to finance the courses. Those with more ability and funds will secure more education but the result for equation (1) will be different for each individual. The coefficient $\beta$ would be biased upward in the absence of these variables. More able receive rents to ability. Those born poor will not be able to borrow due to imperfection of the human capital market, and will receive less education than those born rich. The long run supply of skilled workers would slope more steeply upward to the extent that these factors are important. Therefore we would have a possible departure from the assumptions of equation (1).

[9] Abbas and Foreman-Peck (2008b) shows that returns to female education outside of agriculture and to successive levels of education are remarkably higher in Pakistan. The apparent higher returns, in each case, are interpreted primarily by two basic departures from the Mincer model; firstly, due to parental unwillingness to invest in female education because earning opportunities are limited/ not available for female workers; and secondly, poor quality of primary education in Pakistan.

When school quality is taken into account, it is difficult to properly measure schooling and years of schooling completed is a crude or an inappropriate measure. The family as a concerned decision-taker pays great attention to the quality of school when deciding on schooling. Those with more years tend to have higher quality of years. Therefore ignoring school quality leads to an upward bias in measurement of the average rate of return to a year of education.

Different variables are often used in the literature for capturing school quality; for instance class size, average expenditure per pupil, student teacher ratio. Similarly there are numerous other factors affecting wage earning correlation as we move away from its inbuilt competitive structure; for instance the impact of pecuniary (consumption aspects like status, job security) and non pecuniary aspects (working conditions), post school investment such as on the job training, dispersion in rates of return such as restriction of competition due to occupational licensing and finally signalling.

The largest cost to the individual is the forgone earnings in the next best alternative of the time spent in education. But the amount of forgone earning depends on the available job prospects in both developed and least developed countries. On the one hand the children in least developed countries enter the job market at a very early age, hence increasing the opportunity cost of education. On the other hand mass unemployment lowers the opportunity cost of education in these nations. Similarly in developed nations compulsory schooling regulation and minimum work age laws reduce the opportunity cost of education.

Critics (Solmon l973; Wachtel 1975) view schooling as an inadequate measure of the amount of education obtained. They term school quality as a 'missing ingredient' in the human capital model. Psacharopolous (1985) noted that the dimensions of school quality in less developed countries were typically missing in quantitative measures of education. There are a few studies on educational quality, however, which have shown the social and economic importance of school quality. In a country with lower gross national product, school quality has a higher impact on student learning and factors of socioeconomic status appear to have a smaller effect on learning (Heyneman and Loxley 1983).

In developing countries a much lower social rate of return to primary years of schooling expansion is revealed once quality is taken into account. In countries where school quality is very low, then increasing quality can have a higher rate of return than increasing quantity (Behrman and Birdsall 1983). Card and Kruger (1992) used data on workers' earnings from the US 1980 census to calculate the rate of return to schooling for cohorts of workers born in

a particular state between 1920 and 1929. After analyzing this data they concluded that children born in states that offer better schools had a substantially higher private rate of return to schooling. Card and Kruger findings motivated a great deal of research to determine the robustness of the correlations.

Any simple formulation of school earning relationship may not represent the actual contribution of schooling. Therefore different studies in the past have been undertaken to tackle this problem in a number of different ways. Test scores have been added to the earning equations to control for ability bias (Boissiere et al. 1985; Glewwe 1996). Twins studies were used to tackle the problem of unobserved family characteristics and differences in genetics (Butcher and Case 1994; Ashenfelter and Krueger 1994; Bonjour et al. 2003). Test scores were used as measures of human capital rather than years of schooling (Boissiere et al. 1985; Glewwe, 1996; Jolliffe 1998; Alderman et al. 2009) and natural experiments (Angrist and Krueger 1991; Harmon and Walker 1995; Kane and Rouse 1995) to randomly calculate return to extra years of schooling. The literature surveyed here definitely shows that these factors (school quality, ability, family background) affect schooling decisions and consequent return to education.

The schooling model discussed above is based on the idea that education increases productivity and this increase in productivity increases wages. But a second line of thought is that more educated workers receive higher wages not necessarily because education has taught workers any valuable skills but because the level of educational attainment signals a workers' productivity to employers. The signalling model hypothesizes that education signals the worker's innate ability without increasing workers productivity. This theory reflects the idea that educational attainment is correlated with other unobserved characteristics (innate ability) of the worker that existed before he (or she or their family) made schooling decisions. If the signalling relationship holds, then the coefficient on schooling in the schooling model does not report the true impact of education on productivity. Consequently the social return to education will be driven below the private return.

However, in the signalling theory education still reflects productivity, even if it does not cause all of it (Jones 2001). Therefore it is extremely difficult to empirically establish whether education plays a productivity enhancing role, a signalling role or a combination of both. It is very important to resolve this paradox as the two views have different policy implications. Competitive employers use screening to the extent to allocate the right people to the right jobs and only to help them find the more productive employees. Therefore sorting could be shown to potentially benefit society in terms of raising profits and national income

but some people may be worse off and it also widens income dispersion. However, compensation must be extracted from the gainers of the screening to the losers so that no one is worse off even though income dispersion has increased. Therefore if signalling holds, the coefficient on standard Mincer earning equation (1) would be overstated, even if it represents part of it.

One way of distinguishing between signalling and human capital is by comparing Mincer equations between the self-employed and the employed (Abbas and Foreman-Peck 2008b). The schooling of the self-employed signals to no-one, whereas their education may well enhance their productivity and earnings. If, as Abbas and Foreman-Peck found for Pakistan, the returns to schooling are similar for both groups it is reasonable to assume that education is productivity enhancing for the individual, not merely signalling.

The Mincer earnings function could either overstate the social benefits of education due to a correlation between education and unobservable characteristics and of signalling effects or understate the role of education because it is unable to take account of positive externalities and enhancement in productivity due to better match between workers and jobs. However, despite these different possible shortcomings, there is a consensus in favour of significant return to human capital, which breaks down with macroeconomic studies.

### 2.3 Macroeconomics and Human Capital

Romer (1986) and Lucas's (1988) "endogenous growth" broadened the concept of capital to include human capital. Endogenous growth models offset diminishing-returns to capital by focusing on "ideas gaps" and learning externalities to explain why some countries are richer than others. In their models technology and human capital are both "endogenous" to the system. The main thrust of the idea is that there is increasing rather than decreasing returns to all forms of investment taken together.

Neo-classical growth models modified to include human capital are best summarized by Mankiw, Romer and Weil (MRW) (1992). Their standard approach is to treat human capital or the average years of schooling of the labour force as an ordinary input in the production function. MRW (1992) propose an augmented neoclassical production function (Cobb Douglas) in the form;

$$Y = K^{\alpha} H^{\gamma} (AL)^{1-\alpha-\gamma} \qquad (2)$$

Here Y is the output, K represents capital, L labour, and A level of technology. L and A assumed to grow at the rate n and g exogenously.

They used a measure of human capital accumulation (the percentage of the working-age population in secondary school) as an additional explanatory variable in their cross-country regressions. They showed that the stock of human capital was in fact dependent upon both financial savings ($s_K$) and on human capital investment ($s_H$). The steady state values of human capital (h*) and physical capital (k*) are;

$$\tilde{k}^* = \left( \frac{s_K^{1-\gamma} s_H^{\gamma}}{n + g + \delta} \right)^{\frac{1}{1-\alpha-\gamma}} \qquad \tilde{h}^* = \left( \frac{s_K^{\alpha} s_H^{1-\alpha}}{n + g + \delta} \right)^{\frac{1}{1-\alpha-\gamma}}$$

The inclusion of human-capital accumulation lowered the estimated effects of saving and population growth to the values predicted by the augmented Solow model. They contend that the augmented Solow model provides an almost complete explanation of why some countries are rich and other countries are poor – although the MRW model is dependent upon exogenous technical progress, as is the original Solow model.

A reaction to this type of literature has been the work by Benhabib and Spiegel (1994) and Spiegel (1994), or Pritchett (2001), and more recently Bils and Klenow (2000), according to which the role of human capital in economic growth has been vastly overstated. Benhabib and Spiegel (1994) and Spiegel (1994), used a standard growth accounting approach and incorporated initial per capita income and estimates of years of schooling from Kyriacou (1991), and found a negative coefficient on growth of years of schooling. Similarly Lau et al. (1991) estimated the effect of education by level of schooling (primary vs secondary) for five regions and found that primary education has an estimated negative effect in Africa and Middle East North Africa, insignificant effect in south Asia and Latin America, and positive and significant effect only in East Asia.

Pritchett (2001) used two cross national data sets (Barro and Lee 1993 and Nehru and Dhareshwar 1993) and creates an estimate of growth rate of per worker educational capital. Barro and lee (1993) estimate the educational attainment of the population aged 25 and above using census or labour force data and create a full panel of five yearly observations over the period 1960 to 1985 and filling in the missing data using enrolment rates for a large number of countries. This data set has now been extended to 2000, with further refinement to the data set (Barro and Lee 2000). Nehru and Dhareshwar (1993) cumulate enrolment rates into to annual estimates of the stock of schooling of the labour force age population, using perpetual inventory method and created annual observations for 1960-1987. Using the estimates of schooling from these data sets, Pritchett (2001) creates a measure of human capital, using the

microeconomic specification of earnings developed by Mincer. He used the macroeconomic model worked out by MRW (1992) and concluded that cross-national data shows no association between human capital accumulation and per capita growth rate.

Of particular interest for developing countries in general and Pakistan in particular are models where international technology diffuses to the economy of interest at a pace that depends upon the availability of human capital to take advantage of it (Rosenzweig 1995; Abbas and Foreman-Peck 2008a).

Macroeconomic evidence linking education and macroeconomic growth appears inconclusive. Temple (2000) doubts these results of no effect of education, for convincing reasons of measurement error, outliers and incorrect specification. The emerging macroeconomic analysis, very recently, has focused on the direct test of productivity effects of education as in Jones (2001)[10]. But this area still has problems and weaknesses with its statistical techniques and modelling (Temple 2000).

A likely reason for ambiguous results with cross-national comparisons is the supply of education. This is largely under the control of governments in most countries with varying degrees of effectiveness. The assumption that one year of schooling is as effective in one country as in another is unlikely to be valid[11]. The quality of education both within countries and between them almost certainly varies markedly – and justifies the focus of this thesis. Education systems differ; to understand their individual contributions to human capital and economic growth we need to be aware of the institutional detail. For this reason we now describe the position in Pakistan.

### 2.4 Overview of Schooling in Pakistan

Pakistani educational performance is poor relative to neighbouring countries measured by Gross Enrolment Ratio (GER) and Net Enrolment Ratio (NER). In 2004-05 the country's adult literacy rate stood at 49.9 percent compared to Sri Lanka (90.7 percent), India (61 percent), Iran (82.4 percent) and Indonesia (90.4 percent) (GoP 2009). The Human Development Index for Pakistan was 0.55, which is lower than that of some countries in the

[10] Jones (2001) utilizes an econometric model to simultaneously estimate a wage (an earning) function and a production function for the workers and for the firm where they are working. The approach allows for the direct comparision of the relative productivity and relative wage of workers with different level of education in Ghanaian manufacturing. The result shows that education is positively correlated with productivity and firms pay according to workers productivity.

[11] Although some panel studies have been attempted they are typically of necessity 'short' and so dominated by the cross-national variation.

region but slightly better than that for Bangladesh and Nepal. Pakistan's performance is weaker on the Global Competitiveness Index (GCI) for health and education components than major competitors like India, China, Bangladesh, Malaysia and Sri Lanka. In terms of the quality of educational outcomes, Pakistani students are well below the international scaled mean of 495 in the TIMSS (Trends in International Mathematics and Science Study). This means Pakistani student performance is poorer than their counterparts in other countries in the region, except Iran (GoP 2009).

The current estimates show moderate improvement in adult literacy ratio (57.7 percent in 2010-2011) relative to the estimate of 49.9 percent in 2004-05. Low educational performance is always coupled with regional (rural and urban as well as provincial) and gender disparities in education provision. Adult literacy rate for males and females is 69.5 percent and 45.2 percent respectively. Literacy remains higher in urban areas (73.2 percent), while the rate for rural areas is 49.2 percent (GoP 2011). The major reasons offered for the poor performance on these indicators have included the Government's education policy and the lack of confidence in public sector education establishments to provide quality education. Cheating in exams is a widespread practice, ghost schools and absentee teachers are common phenomena (GoP 2009)

Education provision is largely financed by the public sector[12], which spends 2 percent of the GDP on education. The low proportion of government expenditure on education shows the low priority Pakistan places on education. It spends relatively less on education than other countries in the region. Education spending in India amounts to 3.3 percent of the GDP, Iran (5.2 percent), Bangladesh (2.6 percent) and Nepal (3.3 percent). (GoP 2009).

In Pakistan the formal system of education consists of: primary level (5 years of education), middle level (3 years of education), high level (2 years of education; also called Matric/GCSE) and higher secondary level schooling (2 years of FA/FSc. Level/A' Level). The Ministry of Education at the federal/national level formulates education policies and implements major action plans. However, under the Devolution Plan 2001, the system has been decentralized and the delivery and management responsibility for education has shifted from provinces to the district administration. In the new setup, most school education affairs are dealt with by the executive district officers (EDOs). To ensure smooth functioning of the system, their major duties include: policy implementation, supervision and monitoring of

[12] Private sector is allowed to undertake education provision in Pakistan. However, detail on private sector provision has been delayed until chapter number 6.

schools. The executive education officer is supported by district education officers (DEOs), deputy district education officers (Dy.DEOs) and other staff at lower levels. However, policy formulation, and the budgetary requirements of the district governments for education delivery and teacher training are still largely the prerogative of the central as well as provincial governments[13].

The major portion of the education budget is spent on teacher's salaries (71.45 percent under the recurrent head). Only a meagre amount is spent on quality improvements (28.55 percent is spent under the development head), such as curriculum development, teacher training, monitoring and supervision, and physical infrastructure (GoP 2011). Most schools are sparsely equipped and library facilities, computer resources, sports and recreation facilities are poor. The paucity of these facilities can be judged from the fact that a large number of schools are without buildings (9.8 percent), without boundary walls (32 percent), without drinking water facilities (35 percent), without toilet facilities (35.5 percent) and around 60 percent without electricity (GoP 2011). Estimates of school quality indicators, based on data utilized in this thesis (Pakistan Rural Household Survey (PRHS-II) 2004-05), also portrays the same story that: overall 59 percent of schools are without playgrounds, 78.63 percent are without libraries, without toilet facility (38 percent), without boundary wall (38 percent), without potable drinking water (35 percent) and 49 percent of schools are not electrified.

For political and fiscal decentralization, the Local Government Ordinance (2001)[14] assigned powers, responsibilities and service delivery functions to three levels of local development services; at district, tehsil[15] and union council level. Under the Local Government Ordinance (LGO), citizen community boards (CCBs) were to be established to institutionalize bottom up demands for developmental projects. The provision for the establishment of CCBs in the LGO supposedly is the realization by civil society of the importance of local area development and service delivery but its implementation has been disappointing. The CCB is empowered to establish and support School Management Committees (SMCs) to enable education service providers to respond to civil society (parents) concerns and views, and to encourage greater community participation. School Management Committees (SMCs) have been formed on the logic that parents and local community have better information about the

[13] In Pakistan, every year, allocations are made for education spending in federal budget, as well as, in provincial budgets. Provinces manage education matters of primary, middle and secondary level of education. However, higher secondary education matters are still dealt by the federal education ministry.
[14] Also called Devolution Plan 2001
[15] Tehsil represents the administrative subdivision of a district. Each tehsil represents the second lowest tier of local government in Pakistan. Administratively it is further subdivided into a number of Union councils.

needs and quality of local schools than central and/or sub-national governments. SMCs are entrusted with responsibilities for managing how a school appears. However, the practical implementation and performance of the arrangements are unsatisfactory (HRCP and CEF 2005).

### 2.5 Conclusion

By any reckoning the Pakistani educational system has a low average quality. This must have implications for the rates of return achieved for schooling there and also for international cross-sections of the contribution of human capital to macroeconomic performance. In failing to take into account this quality, the second group of studies will be misleading. In the first group, we can anticipate for Pakistan and similar countries a high rate of return to schooling for the few who are able to climb the educational ladder. Poor school quality will prevent most individuals progressing in the manner laid down in the simple versions of the Mincer model. The observed high returns[16] in Pakistan are due to a scarcity of supply rather than the strength of demand for human capital. Attempted reforms of the educational system show no signs of rectifying this problem.

[16] See Abbas and Foreman-Peck (2008b) ,also explained earlier.

# Chapter 3 Data

## 3.1. Introduction

The empirical analysis of this thesis is based on the Pakistani Rural Household Survey, second wave, of 2004-5, (PRHS-II). This data source offers a unique combination of child level data and a wide range of school quality indicators, essential for the present study. In the present chapter we explain the relationship of PRHS-II to other Pakistani data sources, before going on to describe the key variables for this study both at the Pakistani level and in an international context.

### Pakistani Data Sources

Three major data bases are available for Pakistan, namely population censuses, registration statistics and household survey data. The population censuses have not been undertaken regularly and the registration of births and deaths have been inadequate historically (Nayab and Arif 2012). The irregularity in conducting censuses has necessitated the undertaking of nationally representative household surveys periodically with the financial support of different donor agencies such as the World Bank.

The (Pakistani) Federal Bureau of Statistics (FBS) has periodically conducted household surveys such as the Household Income Expenditure Survey (HIES), the Labour Force Survey (LFS) and the Pakistan Demographic Survey (PDS). Different socio-economic aspects have been covered in each survey because each usually has different policy objectives. However, these surveys have filled what would otherwise be a Pakistani data vacuum.

All these surveys are cross-sectional so that they cannot be used to investigate the dynamism of different economic process. This places a premium upon longitudinal or panel data sets where the same households are visited and interviewed at different times periods.

For this reason the International Food Policy Research Institute (IFPRI) conducted a panel survey of 800 households in rural Pakistan covering five years from 1986 to 1991. The sample covered four rural districts with no representation from Balochistan province. Subsequently the Pakistan Institute of Development Economics conducted the Pakistan Socio-Economic Survey (PSES) comprised of two years panel (1998-99 and 2001) in rural and urban Pakistan. Both these panels were discontinued after the above mentioned rounds.

The Pakistan Institute of Development Economics (PIDE) revisited the IFPRI panel households in 2001, after a gap of 10 years, with the financial assistance of the World Bank.

The sample was increased to 16 districts from all four provinces. The first two rounds of the panel survey named the 'Pakistan Rural Household Survey-PRHS' were undertaken in 2001 and 2004 in rural areas. For the third round of the survey, which was carried out in 2010, an urban sample was added and the survey was renamed the 'Pakistan Panel Household Survey-PPHS.

The PRHS-I (2001) was administered in all four provinces of Pakistan; whereas PRHS-II (2004) was restricted to the two large provinces, Sindh and Punjab due to security concerns. The left-over households in PRHS-II from Khyber PakhtunKhwa (KP) and Balochistan were visited in the PPHS-2010 after a lapse of approximately ten years. The urban sample in PPHS 2010 was selected from the same 16 districts included in PRHS-I (2001).

### 3.1.1 Selection of Districts and Primary Sampling Units (PSUs)

In PRHS-I and II, a village was considered as a Primary Sampling Unit (PSU). There were 43 PSUs (villages) in the IFPRI sample. In PRHS-I and II, 98 more PSUs were selected randomly from the 12 new districts (Nayab and Arif 2012).

The sample size of the three rounds of PRHS panel survey is provided in Table 3.1. Table 3.1 also shows the split households in 2004 and 2010 of the survey rounds. A split household is defined as a new household where at least one member of the previous panel household has established a new household permanently. The movement of the member from the panel household is due to the marriage of a female member or to a son or brother deciding to live separately from the panel household.

In PRHS, households split only within a sampled village (PSU) were interviewed. Those households who split or moved out of the sampled villages were not followed due to high cost involved in this process (Arif and Farooq 2012). Arif and Farooq (2012) state that, firstly, all the PRHS sampled households are not strictly panel households. Only rural households from Sindh and Punjab are panel households as they are covered in all three rounds. Secondly all the sampled households covered in 2001 were re-interviewed in 2010. So for the analysis of all rural areas in Pakistan panel data are available for 2001 and 2010. Thirdly for the analysis of 2004 and 2010 panel data are restricted to Sindh and Punjab only. Lastly, the split households are not panel households and matching with the original panel households is not straightforward.

Table 3.1: Households Covered over the Three Rounds of PRHS

| | PRHS 2001 | PRHS 2004 | | | PPHS 2010 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Panel house-holds | Split house-holds | Total | Panel house-holds | Split house-holds | Total Rural house-holds | Urban house-holds | Total Sample |
| Pakistan | 2721 | 1614 | 293 | 1907 | 2198 | 602 | 2800 | 1342 | 4142 |
| Punjab | 1071 | 933 | 146 | 1079 | 893 | 328 | 1221 | 657 | 1878 |
| Sindh | 808 | 681 | 147 | 828 | 663 | 189 | 852 | 359 | 1211 |
| KP | 447 | - | - | - | 377 | 58 | 435 | 166 | 601 |
| Balochistan | 395 | - | - | - | 265 | 27 | 292 | 160 | 452 |

Source: Adopted form Arif and Farooq (2012)

The PRHS-I (2001) sample size was 2721 households. The same households were followed in PRHS-II (2004), only in Punjab and Sindh as mentioned earlier, and comprised of 1614 households. The PRHS-II also interviewed 293 split households; leading to a sample of 1907 households (Table 3.1). A total of 2198 panel households were re-interviewed in the PPHS-2010 in all four provinces. Accounting for the split households (602 households), the total rural sample increased to 2800 households. The urban sample (1342 households) was selected randomly from the 16 panel districts, leading to a total of 4142 households in PPHS-2010.

Most recently in 2012, the Innovative Development Strategies (IDS), the Pakistan Strategy Support Programme (PSSP), the International Food Policy Research Institute (IFPRI), have jointly revisited the above mentioned (IFPRI) sample while expanding the sample to 19 districts spread over three provinces of rural Pakistan: Sindh, Punjab, and Khyber Pakhtunkhwa (KPK). The survey is named as 'Pakistan Rural Household Panel Survey 2012 (round I). The sample size is 2,124 households. However, 34 households were dropped due to refusal and a total 2,090 households were surveyed (560 in Sindh, 1,340 in Punjab, and 224 in KPK). A brief comparison is undertaken on major indicators between 'Pakistan Rural Household Panel Survey-RHPS 2012 (round I)' and the data utilized in the current thesis 'PRHS-II (2004)'.

### 3.1.2 PRHS-II Descriptive Statistics and Comparisons

The PRHS-II survey collected information on agricultural-related activities, credit, employment, and several demographic events as well as households' consumption

information for the month preceding the survey [17][18]. We restricted our attention to only PRHS-II (second round of the panel survey) as the PRHS-I (2001)[19] data has not been properly process and at the time of writing is not in usable condition for the major school quality indicators, as well as for other household indicators. Most importantly the quality attributes of schools attended by children of the surveyed households in the villages were not available in PRHS-I sample. The PPHS-2010 has been brought to usable condition only recently in 2012 and has not yet been released to general public use for economic analysis.

In Pakistan schooling formally begins at the age of 5 and is compulsory until the age of 14[20]. PRHS-II does not include hours of work information. On the other hand, more appropriately for the focus of the present study PRHS-II does have indicators about the schools that the children in the survey actually attended.

Some of the PRHS-II households had no children less than 15 years of age. Therefore the final sample is reduced to 1434[21] households (with 3950 children). A distinction has also been drawn between landless and non-landless households. The descriptive statistics on land holdings and household consumption are presented for those households having school age children (either currently at school or not) to make possible comparisons between the selected households. Hence the data in current thesis has been restricted to school age children and later on merged with their personal characteristics (child- level characteristics e.g. age gender, education, literacy and numeracy) and their respective household characteristics e.g. consumption expenditure, landholdings and types of tenure, parental education, literacy and numeracy, for the subsequent econometric analysis undertaken in the later chapters. However, a crude comparison is undertaken with other national level statistics or surveys wherever possible. Comparison with other countries (at the same level of development as Pakistan) on

[17] As mentioned earlier the data was gathered jointly by Pakistan Institute of Development Economics and World Bank . The first round of PRHS was launched in September 2001 to January 2002 covering two crop seasons of Kharif 2000 and Rabi 2000/01. The more reliable second round PRHS-II, used here, was launched in August 2004 to October 2004 covering the same crop seasons of Kharif 2004 and Rabi 2004/05.

[18] Of these households 933 were from the Punjab province and 681 from Sindh.

[19] We requested PIDE for both PRHS-I and PRHS-II but only PRHS-II was provided due to data issues (cleaning, processing) with PRHS-I dataset.

[20] The Pakistani constitution prohibits employment of children below the age of 14 years in any factory or mine or any other hazardous employment. The Employment of Children Act (ECA), 1991 prohibits employment of children in 13 specific sectors. But the prohibition against employing children in hazardous labour, and the regulations governing the working conditions of children under 14 do not apply to family run establishments.

[21] 668 households in Sindh and 766 households in Punjab.

the selected indicators is possible only once comparable data are restricted to the specific age children (age 5-15) and arranged by relevant indicators.

Household surveys are nowadays available for different countries but different policy indicators are not always present in these surveys, such as school based information on school quality indicators connected to the children of the sampled households in the surveys. In this chapter, we have undertaken a brief comparison of selected Pakistani education indicators of interest from the World Bank, with those from Bangladesh, Ethiopia, India, Nepal, Peru and Iran. Literacy and child education are discussed in the following section 3.2, village level school quality indicators are presented in section 3.3, household consumption expenditure and landholdings are considered in section 3.4, credit rationing is discussed in section 3.5, a brief international comparison on selected development indicators is undertaken in section 3.6, and section 3.7 concludes.

## 3.2 Literacy and Education

### 3.2.1 Literacy

Literacy is the ability to read, write and carry out simple addition and subtraction. The PRHS-II survey asks two main questions to examine individual literacy: whether the individual is able to read a simple letter or the newspaper with understanding and is able to do simple addition and subtraction. The answers to these questions have been reported in three categories:1) yes easily, 2) yes with difficulty, and 3) not at all.

We define four levels of literacy in this section to make our estimates comparable with 'Pakistan Rural Household Panel Survey 2012 (round I)': 1) a person who can easily read and perform simple addition and subtraction is called 'literate', 2) a person who can read and do simple addition and subtraction with difficulty is called 'literate with difficulty', 3) a person that can perform only one task (that can do simple calculation but cannot read at all) is defined as 'partially literate', 4) an individual who cannot perform any of these activities is defined as 'illiterate'.

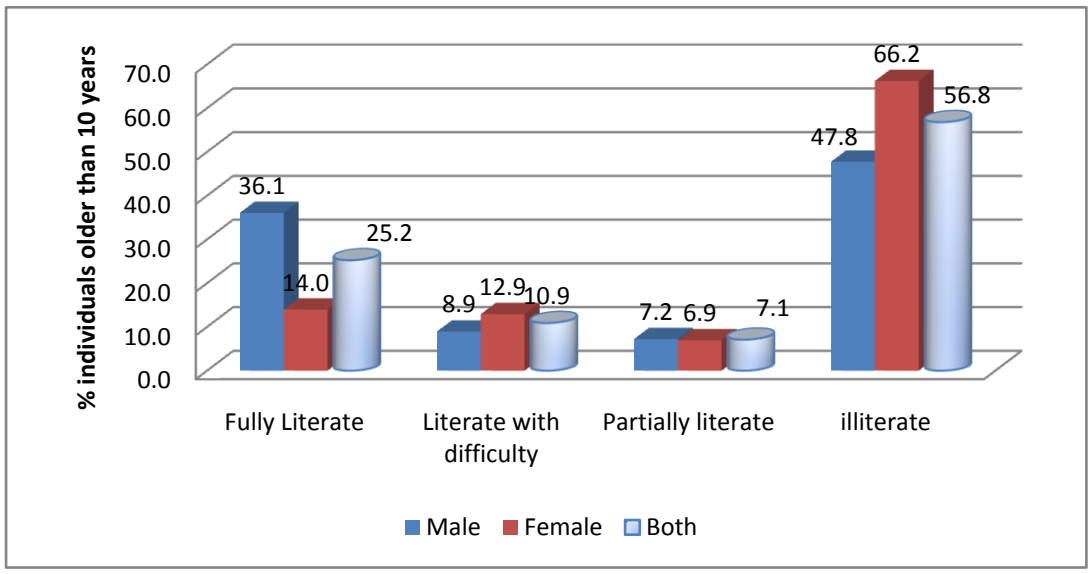
The literacy information is reported for each household member age 11 and above. It is important to highlight that the estimates of literacy level and literacy rate from 'Pakistan Rural Household Panel Survey 2012 (round I)' are for the whole sample. By contrast the estimates from PRHS-II presented here are of children aged 11 to 15 and parents of children age 5-15. This is because the data have been restricted to school age children for later econometric analysis, as well as to draw comparison between those households who send

their children to school and those who do not. However, the estimates presented below shows similar outlines to those presented in 'Pakistan Rural Household Panel Survey-RHPS 2012 (round I)': males are more literate than female, and the literacy rate is higher in Punjab than in Sindh.

The distribution of literacy levels for the specified individuals is presented in Figure 3.1. The figure shows that 25 percent of individuals are able to read and calculate easily. The proportion of females (14 percent) is lower than males (36 percent) in this group. Nearly 57 percent of individuals are completely illiterate. They are not able to read and calculate. The proportion of illiterate females is higher (66 percent) than males (48 percent). A similar observation is reported by RHPS 2012 as presented in Figure 3.2.

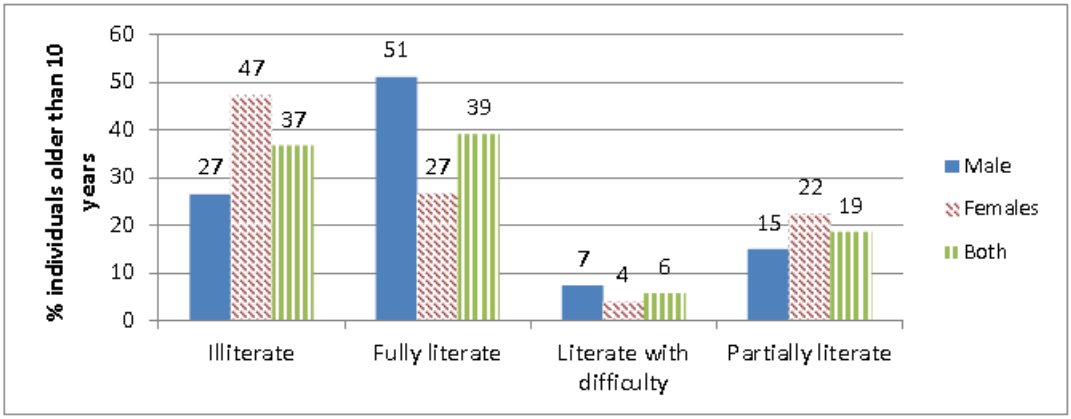
Nearly 10 percent are literate with difficulty and another 7 percent are partially literate.

Figure 3.1: Distribution of Individuals by Literacy Level across Gender (PRHS-II)




Source: Calculated from PRHS-II

Figure 3.2: Distribution of Individuals by Literacy Level across Gender (RHPS-2012)



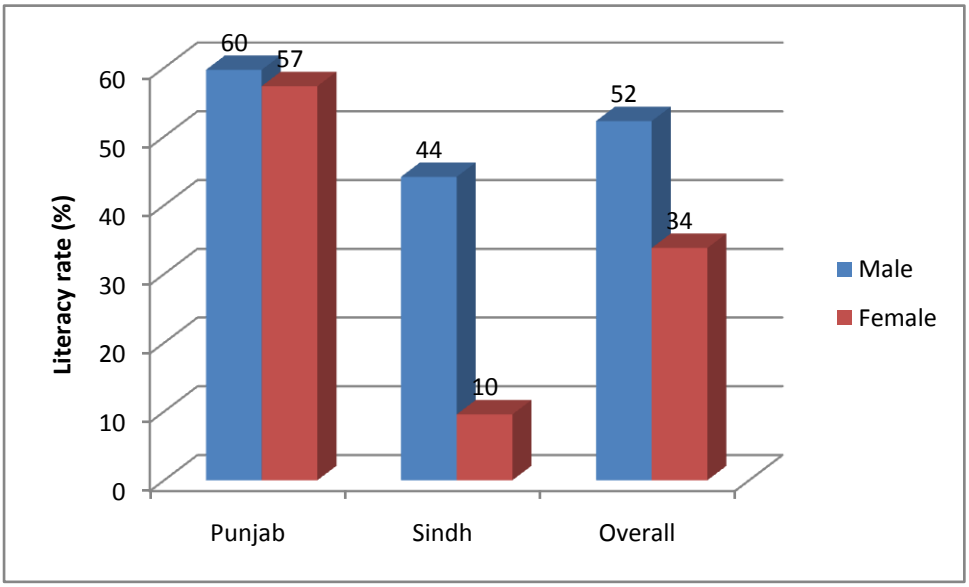

Source: Adopted from RHPS-2012

### 3.2.2 Literacy Rate

The literacy rate is defined as the number of persons older than 10 years who are able to read and perform simple addition and subtraction with or without difficulty as a proportion of all persons older than 10 years. The overall literacy rate is 43 percent in Figure 3.3 (males 52 percent and females 34 percent).

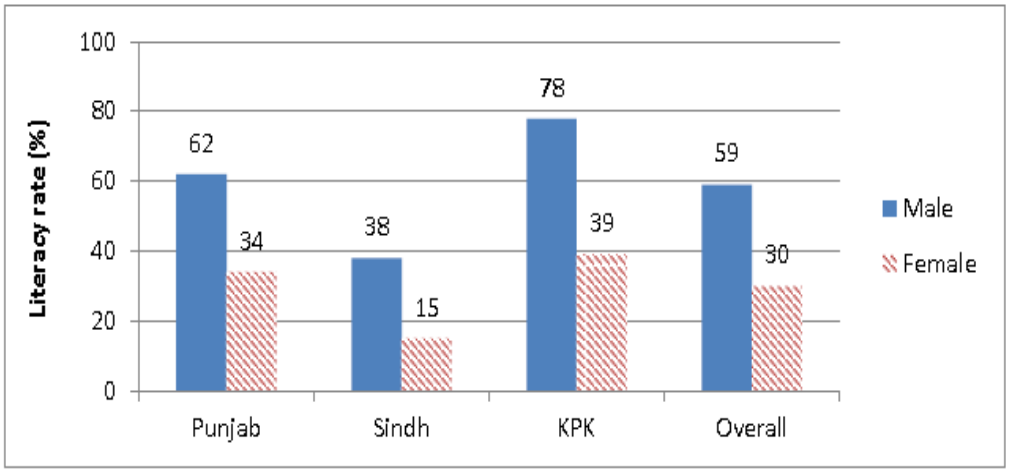
The provincial distribution shows that literacy is lower in Sindh (44 percent for males and an extraordinary 10 percent for females). A similar observation is reported by 'Pakistan Rural Household Panel Survey-RHPS 2012 (round I)' as presented in Figure 3.4.

Figure 3.3: Literacy Rate by Province and Gender (PRHS-II)




Source: Calculated from PRHS-II

Figure 3.4: Literacy Rate by Province and Gender (RHPS-2012)




Source: Adopted from RHPS-2012

### 3.2.3 Education Level

The PRHS-II survey allows the reporting of completed years of schooling for children and their parents in Table 3.2. The table shows that 59 percent of individuals never attended school, 26 percent of individuals have less than primary level of education, 3 percent have less than middle level of education, and 1.4 percent individuals have higher than matric (secondary) qualifications. The provincial disaggregation shows that there is a higher proportion of individuals with no qualification in Sindh than in Punjab. The proportion of individuals with less than primary level of education is also higher in Sindh. The proportion with less than primary education is higher in PRHS-II as these are reported for population age 5 years and old, whereas the RHPS-2012 proportions (Table 3.3) are reported for age 11 years and above.

Table 3.2: Level of Education by Province (PRHS-II)

| | Sindh | Punjab | Both |
|---|---|---|---|
| Never attended school | 61.1 | 57.7 | 59.3 |
| Less than primary(<5 years) | 28.2 | 25.3 | 26.7 |
| Completed primary(=5 years) | 5.5 | 6.3 | 5.9 |
| Uncompleted middle (>6 but < 8) | 1.7 | 4.1 | 2.9 |
| Completed middle (=8) | 0.7 | 2.1 | 1.4 |
| Uncompleted matric (>9 but < 10) | 0.3 | 1.1 | 0.7 |
| Completed matric(=10) | 1.0 | 2.1 | 1.6 |
| Higher than matric | 1.4 | 1.3 | 1.4 |
| All | 100.0 | 100.0 | 100.0 |

Source: Calculated from PRHS-II

Table 3.3: Level of Education by Province (RHPS-2012)

| | Punjab | Sindh | KPK | Overall |
|---|---|---|---|---|
| Never attended school | 54.8 | 65.7 | 50.3 | 56.7 |
| Less than primary (<5 year) | 7.9 | 7.6 | 5.4 | 7.6 |
| Completed primary (=5 years) | 11.2 | 7.7 | 8.8 | 10.2 |
| Uncompleted middle(>6 but <8) | 4.1 | 1.0 | 5.4 | 3.6 |
| Completed middle (=8) | 8.0 | 2.6 | 5.0 | 6.6 |
| Uncompleted matric (>9 but <10) | 1.7 | 0.3 | 3.8 | 1.6 |
| Completed matric(=10) | 7.4 | 8.4 | 14.2 | 8.2 |
| Higher than matric | 4.9 | 6.8 | 7.1 | 5.5 |
| All | 100 | 100 | 100 | 100 |

Source: Adopted from RHPS-2012

The provincial disaggregation across gender (Table 3.4) indicates that majority of females have no qualifications or never attended school. Among females who have been to school, the majority of them have less than primary level of education and about 3.4 percent have completed primary level of education. These figures show the same pattern reported in 'Pakistan Rural Household Panel Survey 2012 (round I)' (Table 3.3 and Table 3.5).

Table 3.4: Level of Education by Province and Gender (PRHS-II)

| | Sindh | | Punjab | | Both | |
|---|---|---|---|---|---|---|
| | male | female | male | female | male | female |
| Never attended school | 48.5 | 73.9 | 48.1 | 68.3 | 48.3 | 71.1 |
| Less than primary(<5 years) | 33.4 | 22.9 | 28.9 | 21.4 | 31.1 | 22.2 |
| Completed primary(=5 years) | 8.6 | 2.4 | 8.0 | 4.4 | 8.3 | 3.4 |
| Uncompleted middle (>6 but < 8) | 3.1 | 0.3 | 5.1 | 2.9 | 4.2 | 1.6 |
| Completed middle (=8) | 1.3 | 0.2 | 3.0 | 1.1 | 2.2 | 0.6 |
| Uncompleted matric (>9 but < 10) | 0.7 | 0.0 | 1.8 | 0.5 | 1.2 | 0.2 |
| Completed matric(=10) | 1.8 | 0.2 | 3.3 | 0.7 | 2.6 | 0.5 |
| Higher than matric | 2.6 | 0.2 | 1.8 | 0.8 | 2.2 | 0.5 |
| All | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

Source: Calculated from PRHS-II

Table 3.5: Level of Education by Province and Gender (RHPS-2012)

| | Punjab | | Sindh | | KPK | | Overall | |
|---|---|---|---|---|---|---|---|---|
| | Male | Females | Male | Females | Male | Females | Male | Females |
| Never attended school | 40.1 | 70.1 | 53 | 78.2 | 26.2 | 73.6 | 41.5 | 72.1 |
| Less than primary | 10.3 | 5.3 | 11 | 4.3 | 8 | 2.8 | 10.2 | 4.9 |
| Completed primary | 13 | 9.2 | 9.8 | 5.6 | 6.2 | 11.4 | 11.7 | 8.6 |
| Uncompleted middle | 6.9 | 1.3 | 1.9 | 0 | 9.4 | 1.5 | 6.1 | 1 |
| Completed middle | 11.3 | 4.6 | 2.6 | 2.6 | 8.5 | 1.6 | 9.2 | 3.9 |
| Uncompleted matric | 2.5 | 0.9 | 0.6 | 0 | 7.6 | 0.1 | 2.6 | 0.6 |
| Completed matric | 10.3 | 4.4 | 10.8 | 6 | 23 | 5.7 | 11.6 | 4.9 |
| Higher than matric | 5.6 | 4.3 | 10.3 | 3.3 | 11.1 | 3.3 | 7 | 4 |
| All | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Source: Adopted from RHPS-2012

### 3.2.4 Current Schooling of Children

Among 3950 children between the age range 5-15 years surveyed in PRHS-II, 49.4 percent are in the age group 5-9 years (primary school age), 25.9 percent are of middle school going age (10-12 years), and 24.7 percent children are in the age group 13-14 years( high school age). About 49.4 percent of children are currently in school (57.1 percent boys and 40.8 percent girls). Among the total current enrolled children, 50 percent are in the age group 5-9 years, 57 percent are in the age group 10-12 years old, and 40 percent are in the age group 13-15 years.

The provincial level disaggregation shows that attendance rate is higher in Punjab than in Sindh. Similarly, girls are more deprived in terms of enrolment in Sindh than in Punjab (Table 3.6). The same observation is also reported by RHPS-2012 (Table 3.7).

Table 3.6: Attendance Rate among Children (age 5-15) by Gender and Province (PRHS-II)

| | Sindh | Punjab | Full sample |
|---|---|---|---|
| **Boys** | **52.0** | **61.9** | **57.1** |
| 5-9 years | 51.1 | 60.1 | 55.7 |
| 10-12 years | 66.4 | 70.6 | 68.8 |
| 13-15 years | 42.1 | 56.2 | 49.2 |
| **Girls** | **33.3** | **48.9** | **40.8** |
| 5-9 years | 37.4 | 52.0 | 44.2 |
| 10-12 years | 38.3 | 50.2 | 44.4 |
| 13-15 years | 18.1 | 40.8 | 29.0 |
| **Both** | **42.7** | **56.0** | **49.4** |
| 5-9 years | 44.4 | 56.5 | 50.3 |
| 10-12 years | 51.3 | 60.8 | 56.4 |
| 13-15 years | 31.3 | 49.6 | 40.3 |

Source: Calculated from PRHS-II

Table 3.7: Attendance Rate among Children (age 5-14) by Gender and Province (RHPS-2012)

| Age group | Punjab | Sindh | KPK | Full sample |
|---|---|---|---|---|
| **Boys** | 69.1 | 34.6 | 84.9 | 60.5 |
| 5-9 years | 65.2 | 26.8 | 82.1 | 55.1 |
| 10-12 years | 78.1 | 49.3 | 91.5 | 71.4 |
| 13-14 years | 63.7 | 35.5 | 84.1 | 58.3 |
| **Girls** | 54.6 | 19.5 | 67.9 | 45.5 |
| 5-9 years | 56.9 | 21.0 | 63.0 | 45.9 |
| 10-12 years | 61.4 | 19.0 | 77.6 | 50.0 |
| 13-14 years | 39.2 | 12.7 | 64.7 | 37.2 |
| **Both** | 62.1 | 27.3 | 77.2 | 53.3 |
| 5-9 years | 61.2 | 23.9 | 73.8 | 50.7 |
| 10-12 years | 70.5 | 34.1 | 84.4 | 61.2 |
| 13-14 years | 50.5 | 27.0 | 75.6 | 48.1 |

Note: Total number of children in RHPS-2012 is 3574 between the age ranges 5-14 years.
Source: Adopted from RHPS-2012

### 3.2.5 Child Schooling and Work

While the survey asks about child employment participation, the questions only address work away from home, such as on someone else's farm. There is no enquiry about child employment on a households' own farm. This is likely to be an important oversight for the study of child labour (Bhalotra and Heady 2003). The omission implies that child work in agricultural and domestic activities is under-reported in the survey; Table 3.8 shows that a large proportion of both sexes (nearly half of females and more than one third of males) do not attend school and do not work away from home. Female children are fully represented in the sample[22] and are more than twice as likely as males to be working away from home. Table 3.9 reports the distribution of children who work away from home by income

[22] When compared with gender proportions of Pakistan's population given in Pakistan Economic Survey (2006-7, Table 12.2), the estimates based on current data set are exactly equivalent to that of the economic survey 2006-7.

percentile. Only 30-32 percent of these children belong to the bottom quintile while the rest belong to the top quintiles.

Table 3.8: Kharif Season Children's Allocation to Schooling, Work and Both, by Gender, Pakistan 2004/05

| | | Male | Female | Total |
|---|---|---|---|---|
| Schooling Only | Count | 1,165 | 745 | 1,910 |
| | Percent | 55.85 | 39.97 | 48.35 |
| Working Away from Home Only | Count | 66 | 161 | 227 |
| | Percent | 3.16 | 8.64 | 5.75 |
| Possibly or Probably Working at Home | Count | 828 | 943 | 1,771 |
| | Percent | 39.69 | 50.59 | 44.84 |
| Both Schooling and Work | Count | 27 | 15 | 42 |
| | Percent | 1.29 | 0.8 | 1.06 |
| Total | Count | 2,086 | 1,864 | 3,950 |
| | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table 3.9: Children Working away from Home by Household Income net of Child Earnings (Percentile)

| Household Income/Expenditure (percentile) | children working away from home | |
|---|---|---|
| Total Household Income net of Child Earnings | Number | Percent |
| ≤10 % (Rs. 31920) | 39 | 14.50 |
| 10% to 25% (Rs. 31921-Rs.46804) | 49 | 18.22 |
| 25% to 50% (Rs. 46805-Rs. 74695) | 67 | 24.91 |
| 50% to 75% (Rs. 74696-Rs.109901 | 75 | 27.88 |
| >75% (Rs.109902) | 39 | 14.50 |
| Total | 269 | 100 |
| Household Income net of Child Earnings (Equivalised)[23] | | |
| ≤10 % (Rs. 5899) | 38 | 14.13 |
| 10% to 25% (Rs. 5900-Rs.8707) | 41 | 15.24 |
| 25% to 50% (Rs. 8708-Rs. 12946) | 87 | 32.34 |
| 50% to 75% (Rs. 12947-Rs.19716) | 64 | 23.79 |
| >75% (Rs.19717) | 39 | 14.50 |
| Total | 269 | 100 |

Source: Calculated from PRHS-II

### 3.3 Community Level School Quality Indicators

Sample statistics for local school quality over the 94 villages are reported in Table 3.10. The 'number of schools in the community' (average 7) measures ease of access to schooling and variety of provision. School quality variables show a wide range of variation between communities and a high average deprivation (Table 3.10). 78 percent of schools are without libraries, 59 percent are without playgrounds, 38 percent without toilet facilities, 39 percent without a boundary wall, 35 percent without potable drinking water and 49 percent of schools are not electrified. The mean student-teacher ratio is about 34 and the maximum is 79 in rural communities/villages.

These estimates are consistent with those of GoP (2011) which reports that most schools are sparsely equipped and library facilities, computer resources, sports and recreation facilities are poor.

[23] Equivalence scales correct for the fact that the needs of the each household grow with each additional member but not proportionally. Needs for space, electricity and so on are not twice as high for a household with two members as for a single individual. With equivalence scales each household member is assigned a value in proportion to their needs. The OECD equivalence scale or **t**he Oxford scale assigns a value of 1 to the first member of the household and of 0.7, 0.5 to each additional adult and child, respectively, in the household. Specifically the formula is ; $H = 1 + \alpha(n_A - 1) + \beta n_C$ , where $\alpha = 0.7$ and $\beta = 0.5$ and, $n_A$ and $n_c$ are the number of adults and children in the household.

Table 3.10: School Quality by Community PRHS-II Sample 2004-5

| Village School Quality | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| Percentage of schools with a library | 21.63 | 21.55 | 0 | 85.71 |
| Percentage of schools with a playground | 40.51 | 29.84 | 0 | 100 |
| Student-teacher ratio | 33.73 | 13.03 | 13.86 | 79.07 |
| Number of schools in a community (village) | 7.047 | 2.951 | 1 | 14 |
| Percentage of schools with a toilet facility | 61.50 | 28.31 | 0 | 100 |
| Percentage of schools with a toilet facility for girls | 34.48 | 26.74 | 0 | 100 |
| Percentage of schools with a furniture for students | 73.33 | 27.20 | 0 | 100 |
| Percentage of schools with a blackboard/chalk | 90.16 | 19.03 | 0 | 100 |
| Percentage of schools with a potable drinking water facility for children | 65.46 | 32.20 | 0 | 100 |
| Percentage of schools with a boundary wall | 61.45 | 30.17 | 0 | 100 |
| Percentage of schools with textbooks for students | 74.33 | 26.89 | 0 | 100 |
| Percentage of electrified schools | 51.05 | 31.23 | 0 | 100 |

Source: Calculated from PRHS-II
Std. Dev: stands for standard deviation

## 3.4 Household Consumption and Landholdings

### 3.4.1 Household Expenditure (Consumption)

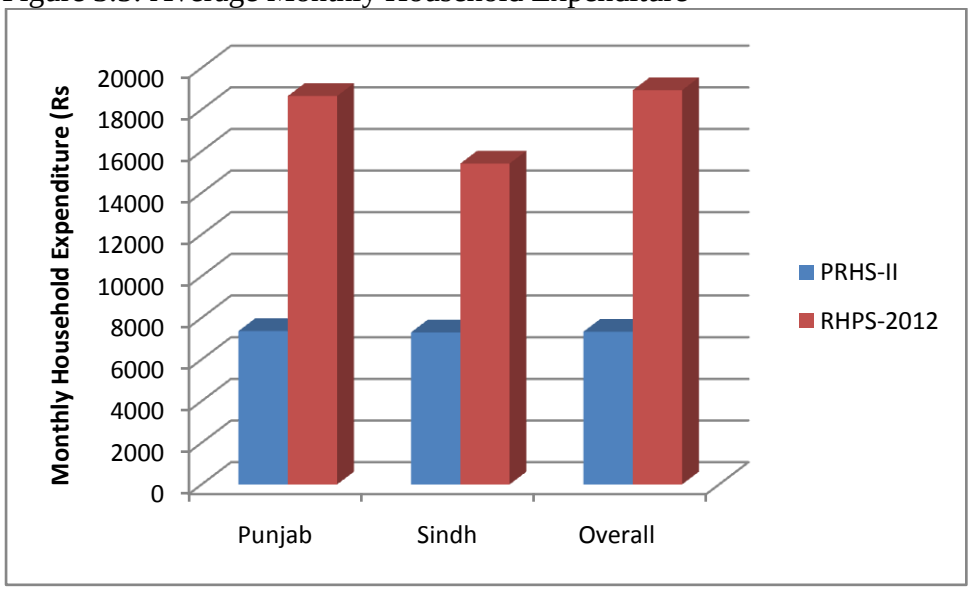
Household expenditure and income shows the economic wellbeing of a household in a society. The measurement and conceptual advantages of using expenditure rather than income is widely acknowledged (Bhalotra 2003). Consumption expenditure is a better measure of economic wellbeing as households smooth their consumption over time. The information on consumption expenditure is used to estimate poverty. However, Bhalotra (2007) shows that household consumption expenditure net of child earnings is an improved indicator from the viewpoint of household subsistence. The PRHS-II survey reports a detailed module on household consumption expenditure that has been calculated in the manner as specified in Bhalotra (2007) and this variable is named 'household income net of child earnings' in this thesis.

PRHS-II shows that monthly average household expenditure/income net of child earnings is Rs.7337 over both provinces (1434 households); 7302 in Sindh (668 households), and 7368 (766 households) in Punjab. The estimates show that households in rural Punjab[24] spend

[24] The poverty profile of rural Sindh and Punjab constructed on the basis of the current data set shows that rural Sindh is poorer than rural Punjab. 38 percent of the families are below the official poverty line in rural Sindh and 31 percent in rural Punjab. Official poverty line was Rs.878.64 per adult equivalent per month in 2004-05 in Pakistan (GoP, 2007-2008).

relatively more than households in Sindh; an observation also noted by RHPS 2012. Comparing with RHPS-2012, our estimates are lower. However, these are consistent with the findings of RHPS-2012 (Figure 3.5) as the prices of consumer goods (durable and non-durable goods) have more than doubled over the intervening period of about eight years in Pakistan as the CPI (Consumer Price Index) increase has averaged more than 10 percent a year from 2004 to 2012, reaching a peak of 21 percent in 2008 (GoP 2012, Table 7.3B). Also the estimates in RHPS-2012 are for the whole sample while our estimates are for the selected range of households who have children in schools. We have also deducted child earnings.

Figure 3.5: Average Monthly Household Expenditure




Source: Calculated from PRHS-II

### 3.4.2 Land Holdings

In the PRHS-II only agricultural households were asked about the land tenancy arrangements. Agricultural Households are defined as farm households that cultivate land, such as owners who cultivate their own land, sharecroppers, who cultivate land on a predetermined contract which define the division of cost of cultivation and output in the last year, and tenants who rent-in land on fixed rent. About 61 percent (879 households) of the total households are agricultural households. So data for renting indicates how much of the landowner's property was rented out and/or share-cropped, not the amount of renting in the sample as a whole. The closest approximation to an agricultural labourer in the data set is identified by the answer to the question 'Do you work on someone else farm?' Hence the relation between type of parental land tenure and occupation and schooling or child work is difficult to establish.

57.11 percent (819) of the households are landless. The remaining 42.89 percent of households (615) are landowners. The landholding of the smallest landowner is 0.06 acres while the maximum landholding being 235 acres of land. The mean plot size is just less than 10 acres. About 65 percent of landless households owned farm animals, whereas 20 percent owned non-farm animals. Landless households in rural Pakistan usually keep farm animals to get milk products for household consumption. They also keep non-farm animals to diversify their income generation activities e.g. to rent out non-farm animals in the local market or for agricultural operations. Almost 87 percent of landowning households also own farm animals, while about one third own non-farm animals.

Mean annual household expenditure (household income net of child earnings (equivalised income)) for landless families is Rs.15315 (Pakistan Rupees) with standard deviation of 10008 (Table 3.11). At first sight what is surprising about Table 3.11 is that landowning households do not spend very much more than the landless. At least in this rural survey land is not the principle that divides the rich and the poor, although there are very large and rich landowners in Pakistan.

The household expenditure variable has been made consistent over all households in the data set by using the OECD (1982) Equivalence Scale. In the present sample there is a maximum of 42 members in one household. It is therefore important to control for household size as well as composition.

Table 3.11: Land Ownership Status and Household Expenditure (Normalized for Numbers Equivalent)

| Figures in Pakistani Rupees | Landless Households | Land Owning Households |
|---|---|---|
| Observations | 819 | 615 |
| Mean | 15315 | 18645 |
| Standard deviation | 10008 | 15349 |
| Minimum | 1997 | 1978 |
| Maximum | 81721 | 156364 |

Source: Calculated from PRHS-II

### 3.5 Credit Rationing

Households require credit to avoid unexpected income fluctuations, to smooth consumption over time or to undertake investments. Without adequate credit consumption smoothing is restricted and this may lead to children being put to work in order to maintain consumption in the face of negative shocks. Jacoby (1994) reported from Peru that credit constraints negatively affect children's primary schooling attainment. A study by Fuwa et al (2009) from rural India also found that credit market failure leads to enormous reallocations of a child's time in different activities such as schooling, work, leisure and households chores. The negative impact of credit constraint amounts to a 60 percent decrease in average schooling time of the rural children. Studies have also viewed child labour decisions as a part of risk management strategy (Jacoby and Skoufias 1997) to minimize the impact of shocks such as bad harvest due to calamities or job loss, when access to credit is not available to smooth their consumption.

A number of studies (Ersado 2005) have found credit access to be very difficult to measure from LSMS (Living Standard Measurement Surveys), which only reports whether a given household has a loan or bank account or not. It is not a good measure of access to credit because the households which did not have a loan account might have access to credit but have no borrowing needs (Ersado 2005). Some recent studies have recognized these potential issues and relied on some sort of exogenous income shocks to infer the effects of credit constraints on household behaviour. Beegle et al (2006) used self reported crop shocks to analyse the effect of credit constraint on child labour. But Fuwa et al (2009) noted the flaws with this approach in that it mixes substitution and wealth effects. Productivity shocks change households' shadow prices consisting of substitution and wealth effects. A negative shock to farm productivity reduces demand for farm production inputs whereas the wealth effect increases the labour supply to farming, if outside employment opportunities are limited. According to Fuwa et al. (2009), the approach adopted by Beegle et al. (2006) may not identify the effect of a credit constraint (which is the marginal utility value of current wealth relative to future) on child labour but the joint effect of price changes and consequent wealth changes. To address these issues the data set's (Pakistan Rural Household Survey) special module on credit access draws a distinction between credit constrained and unconstrained households in a similar spirit to that of Fuwa et al. (2009). The downside is that the module is only available for agricultural households. Nevertheless, access to the commercial bank

(distance) at the community level is also available as a proxy for access to the formal credit market for the whole sample.

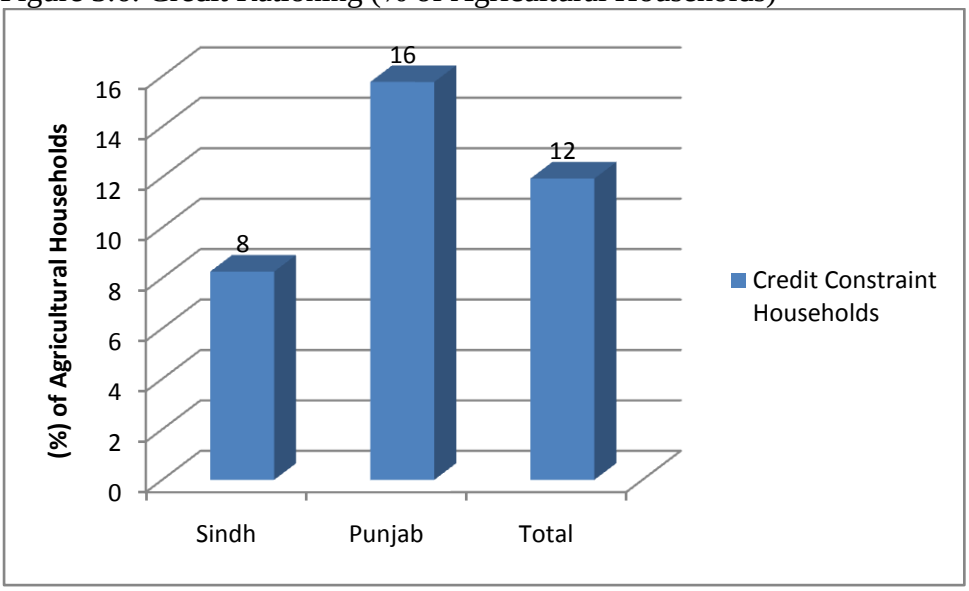
For the identification of liquidity constrained households, heads were asked about household's experience with lenders over the last Kharif 2004 season prior to the survey. Whether the household had attempted or not to get a loan over the last Kharif 2004 season was used to identify credit constrained households. For those households which tried to borrow money, it was determined whether or not they were successful in getting the requested amount under the specified conditions. If the response was 'yes', the household was identified as unconstrained. Those households which were unable to borrow as much as requested or whose requests were rejected were classified as credit constrained households.

Those households that did not attempt to get a loan were further interviewed about the reasons for not trying to get one. The choice set was: (1) Inadequate collateral, (2) Past default with this lender, (3) Bad credit history, (4) Lender too far away, (5) Lender has worse terms than other available sources, (6) Lender's procedures are too cumbersome, (7) Need to pay bribes, (8) Do not borrow at all, (9) Other.

We used Fuwa et al.'s (2009) broad definition of credit constraint. Respondents who choose (1) to (7) were classified as credit constrained households. The remaining households who did not try to borrow were regarded as unconstrained. We used the broader definition of credit constrained because 92.5 percent of those households who did not attempt to get a loan indicated that they did not borrow at all (choice number 8 above), and there remained a limited number of households who could possibly be credit constrained.

The proportion of households that can be viewed as credit constrained is 12 percent over both provinces. The provincial disaggregation shows that credit constraint households are 16 percent in Punjab as opposed to 8 percent in Sindh (Figure 3.6). A similar observation is also reported in RHPS-2012 that a larger proportion of households in Punjab (28 percent) attempted to get a loan over the last year.

Figure 3.6: Credit Rationing (% of Agricultural Households)




Source: Calculated from PRHS-II

### 3.6 International Comparison

We now move to compare selected development indicators on education from the World Bank for Pakistan and other countries at the same level of development (Table 3.12). Pakistan spends less on education (2 percent of GDP) than Bangladesh, India, Nepal, Peru and Iran. The highest spending has been reported for Iran (5 percent) followed by India (3.30 percent). Pakistan is the third highest country in terms of very young population (68 percent of working age population). The highest age dependency ratio is reported for Ethiopia (86 percent), followed by Nepal (70 percent). Fertility rate (births per women) is higher in Ethiopia (5 children) followed by Pakistan (4 children). Pakistan has less poverty than other countries, except Peru and Iran in 2005. Nepal has the highest rural population (86 percent) followed by Ethiopia (85 percent), Bangladesh (75 percent), India (71 percent) and Pakistan (65 percent).

Primary school enrolment rate is lowest in Ethiopia, followed by Pakistan. In Pakistan 63 percent of primary school age children were enrolled in schools in 2004 (73 percent boys and 53 percent girls). The highest enrolment rate has been reported for Nepal (100 percent for both boys and girls) (Table 3.13). Secondary school enrolment rate in 2004 was much the lowest in Pakistan (30 percent) and females had lower enrolment (26 percent) than males (35 percent). The World Bank figure for Pakistan's adult literacy rate is about 49 percent in 2005 (64 percent for males and 34 percent for females) and is consistent with our estimates from PRHS-II. The lowest adult literacy rate is in Ethiopia (29 percent (42 percent males and 18

percent females)). Pakistan has the second highest student-teacher ratio (38 pupils) in 2004 for primary schools. The highest student-teacher ratio has been reported for Bangladesh (40 pupils). The secondary school student-teacher ratio in Pakistan is much poorer (42 pupils) and is higher than in Bangladesh, India and Peru.

Across this international range of human capital indicators Pakistan is closest to the bottom on average, despite the country's relatively high GDP per capita. It is very apparent that for Pakistan education is an unusually low budgetary priority in an international context.

Table 3.12: World Bank Development Indicators for Pakistan (Comparison with Developing Countries)

| Indicator Name | Year | Pakistan | Bangladesh | Ethiopia | India | Nepal | Peru | Iran, Islamic Rep. |
|---|---|---|---|---|---|---|---|---|
| Public spending on education, total (% of GDP) | 2004 | 1.95 | 2.25 | - | 3.30 | 3.17 | 2.83 | 4.87 |
| Age dependency ratio, young (% of working-age population) | 2004 | 67.67 | 57.26 | 86.32 | 52.82 | 70.32 | 52.06 | 40.59 |
| Fertility rate, total (births per woman) | 2004 | 3.87 | 2.70 | 5.29 | 2.88 | 3.48 | 2.73 | 1.87 |
| GDP per capita (constant 2000 US$) | 2004 | 562.28 | 417.95 | 136.62 | 536.60 | 234.78 | 2248.41 | 1844.52 |
| Household final consumption expenditure per capita (constant 2000 US$) | 2004 | 403.28 | 309.32 | 108.42 | 332.92 | - | 1587.61 | 934.29 |
| Poverty headcount ratio at $1.25 a day (PPP) (% of population) | 2005 | 22.59 | 50.47 | 38.96 | 41.64 | - | 8.55 | 1.45 |
| Poverty headcount ratio at $2 a day (PPP) (% of population) | 2005 | 60.31 | 80.32 | 77.63 | 75.62 | - | 20.48 | 8.03 |
| Poverty headcount ratio at national poverty line (% of population) | 2005 | 23.90 | 40.00 | - | 37.20 | - | 55.60 | - |
| Rural population (% of total population) | 2004 | 65.79 | 74.77 | 84.49 | 71.08 | 85.22 | 25.36 | 33.15 |

Source: World Bank (World Development Indicators (WDI)).
(-): Shows data not available for countries on selected indicators.

Table 3.13: World Bank Education Indicators for Pakistan (Comparison with Developing Countries)

| Indicator Name | Year | Pakistan | Bangladesh | Ethiopia | India | Nepal | Peru | Iran, Islamic Rep. |
|---|---|---|---|---|---|---|---|---|
| Net enrolment rate, primary (% of primary school age children) | 2004 | 63.32 | - | 50.79 | 94.98 | - | 99.60 | 96.62 |
| Net enrolment rate, primary, female (% of primary school age children) | 2004 | 52.76 | - | 46.58 | 93.46 | - | 99.87 | 95.11 |
| Net enrolment rate, primary, male (% of primary school age children) | 2004 | 73.43 | - | 54.97 | 96.37 | - | 99.34 | 98.07 |
| Net enrolment rate, secondary (% of secondary school age children) | 2004 | 30.44 | 44.00 | - | - | - | 68.82 | - |
| Net enrolment rate, secondary, female (% of secondary school age children) | 2004 | 26.26 | 44.91 | - | - | - | 68.51 | - |
| Net enrolment rate, secondary, male (% of secondary school age children) | 2004 | 34.45 | 43.14 | - | - | - | 69.12 | - |
| Literacy rate, adult female (% of females ages 15 and above) | 2005 | 35.37 | - | 17.98 | - | - | 82.45 | 76.80 |
| Literacy rate, adult female (% of females ages 15 and above) | 2006 | 39.61 | - |  | 50.82 | - | 83.45 | 77.24 |
| Literacy rate, adult male (% of males ages 15 and above) | 2005 | 64.06 | - | 41.94 | - | - | 93.72 | 88.01 |
| Literacy rate, adult male (% of males ages 15 and above) | 2006 | 67.65 | - |  | 75.19 | - | 94.23 | 87.27 |
| Literacy rate, adult total (% of people ages 15 and above) | 2005 | 49.87 | - | 29.82 | - | - | 87.91 | 82.44 |
| Literacy rate, adult total (% of people ages 15 and above) | 2006 | 54.15 | - | - | 62.75 | - | 88.70 | 82.33 |
| Student-teacher ratio, primary | 2004 | 37.50 | - | - | 40.20 | 35.83 | 23.30 | 22.60 |
| Student-teacher ratio, primary | 2005 | 38.34 | 47.04 | - | - | 39.71 | 23.03 | 21.74 |
| Student-teacher ratio, secondary | 2004 | 41.86 | 27.37 | - | 32.70 |  | 15.21 | - |

Source: World Bank (World Development Indicators (WDI)).
(-): Shows data not available for countries on selected indicators.

## 3.7 Conclusion

This chapter provided a detailed description of PRHS-II survey and estimates of adult literacy in general and child education in particular. It investigated rural households' consumption and credit rationing. Estimates of different indicators of school quality in rural Pakistan have also been presented.

The school quality indicators show that school quality in rural Pakistan is very poor and prospects for a good education are dismal. In some villages the average student-teacher ratio across schools is as high as 79 pupils. School quality variables show a wide range of variation between communities as well. Females compared to males are deprived in terms of level of education, literacy and current school enrolment. More female children compared to male children work outside the home on non-family farms in rural Pakistan. The small proportions of children working outside the home come from households in lower quintile of household expenditure (equivalised). A very small proportion of households could be regarded as credit rationed in rural Pakistan. The Punjab stands ahead of Sindh in child education as well as family wellbeing (consumption expenditure).

Education indicators from the World Bank tell the same story that child enrolment in Pakistan is much worse than other developing countries. Pakistan spends relatively less on education consequently the student-teacher ratio is higher than in other developing countries. The estimates of school quality and child and household characteristics presented in the chapter are consistent with other national level surveys (RHPS-2012) and government statistics (GoP 2011), as well as with education estimates from the World Bank. The next chapters undertake a rigorous investigation and quantification of child schooling and cognitive achievement in rural Pakistan with reference to school quality.

**Glossary**

A brief definitions of concepts used are as follows;

**Agricultural Households** are defined as farm households that cultivate land, such as owner who cultivate their own land, sharecroppers, who cultivate land on a predetermined contract which define the division of cost of cultivation and output in the last year, and tenants who rent-in land on fixed rent.

**Enrolment rate** is the number of students of a specific age group enrolled, irrespective of the level of education, as a percentage of the number of population in that age group.

**Net Enrolment rate** is the number of students of a specific age group enrolled in a given level of education as a percentage of the corresponding age group population.

**Literacy rate** is defined as the number of persons older than 10 years who are able to read and perform simple addition and subtraction with or without difficulty as a proportion of all persons older than 10 years.

**Household** is a domestic unit of all those persons who live together and eat from a common meal. It may be consist of one member or more and may not be related to each other.

**Household expenditure** refers to the expenses incurred by individual members on service and goods.

**Equivalence scales** correct for the fact that the needs of the each household grow with each additional member but not proportionally. Needs for space, electricity and so on are not twice as high for a household with two members as for a single individual. With equivalence scales each household member is assigned a value in proportion to their needs. The OECD equivalence scale or **t**he Oxford scale assigns a value of 1 to the first member of the household and of 0.7, 0.5 to each additional adult and child, respectively, in the household. Specifically the formula is; $H = 1 + \alpha(n_A - 1) + \beta n_c$, where $\alpha = 0.7$ and $\beta = 0.5$ and , $n_A$ and $n_c$ are the number of adults and children in the household.

**Large farmers** usually cultivate more than 25 acre of land.

**Farm animals** include the cow, buffalo, goat and sheep, whereas **non-farm animals** include the horse, donkey, camel and bullock.

**Province** is an administrative unit in Pakistan. There are four provinces in Pakistan: Punjab, Sindh, Khyber PakhtunKhwa, and Balochistan. There are also two federally administered territories: Federally Administered Tribal Areas, and Islamabad Capital Territory.

**Village** refer to a small community or group of houses in the rural areas of Pakistan, usually smaller than a town and embrace common way of life.

**Student teacher ratio** is calculated, here, as the total number of enrolled students in all schools in a village divided by the total number of teachers in these specific schools.

**Percentage of schools with a Library** is calculated as the number of schools in a village with a library divided by total number of schools the specific village and multiplied by 100. The same procedure has been undertaken for the calculation of other school inputs e.g. playground, electrified schools etc.

# Chapter 4 School Quality and Cognitive Achievement in Rural Pakistan

## 4.1 Introduction

The objective of this chapter is to evaluate the impact of school quality on children's cognitive abilities. The demand for schooling as an investment depends upon by how much parents expect formal education to enhance skills and therefore future income. Parental assessment of school quality is thus a vital element in their choice of 'child quality', and a critical influence on whether parents opt to send their children to work or to school.

In this chapter we assume that parental judgements of their children's cognitive achievements is a indicator of parent's views on school quality. If those children that have attended school are judged[25] by their parents to have higher cognitive achievements than those who have not, then schools are perceived as adding value. The main research interest is in establishing a link between objective measures of school quality and children's cognitive achievements. If such a link is found we can move with more confidence to assessing the valuation placed by parents on school quality-according to these objective measures-and how such valuations affect willingness to send their children to school. Children's cognitive abilities are indeed found to be affected by our school quality measures, but perhaps surprisingly, not by teacher experience or qualifications.

Identification in the cognitive achievement and school attendance model are established first by utilising a distinctive feature of the Pakistani environment; the arbitrary and random allocation of resources to state schools economy and society. The quality variable library is confirmed exogenous with the only instrument in the data set suitable on grounds of relevance (correlation with library), community population. If school resources were adjusted to school need, school attendance would not influence the staff-student ratio. But instrumenting this ratio changes its sign in the predicted fashion. For cognitive achievement, cmp is implemented to endogenise staff-student ratio in ordered probit models. Since we have either controlled for endogeneity or established the exogeneity of the quality measures in the attendance and cognitive achievement equations, we can be confident that the parameter estimates correctly capture the impact of school quality variables. The possibility that other variables, not instrumented in the school attendance or cognitive achievement equations, are

[25] Parents most often ask their children to read newspaper for them as well as to carryout simple calculations pertaining to daily life activities (e.g. to buy a commodity in the local market or to count plants in an agricultural plot) in rural Pakistani villages or communities.

endogenous does not bias these estimates. Nor does including a range of extra community characteristics in the school attendance and cognitive achievement equations affect the school quality parameter values estimated.

This chapter is organized as follows. Section 4.2 describes child cognitive achievement in reading and maths ability for the age range 11-15 in rural Pakistan. The methodology for assessing the impact of school quality on cognitive achievement is presented in section 4.3. In section 4.4 results and discussion from the main econometric findings are set out, while section 4.5 concludes.

## 4.2 Cognitive Achievement

In the Pakistan Rural Household Survey (PRHS-II) 2004-05[26] parental evaluation of child cognitive ability in reading and mathematics ability is only available for persons aged 11 years or more. Our concern is with people who can be classified as children, and who are expected to be at school, so the present study is confined to children of age 11 to 15. This is also consistent with the PRHS definition of child. The measures of child cognitive achievement in the survey are also available for children who are not currently at school. This property rules out some sample selection concerns in measuring school effects. Questions about the impact of sample selection are often raised for studies that only concentrate on school-going children or whose sample is based on schools instead of households and that have not counted children who might not be at school because of low school quality.

Parental evaluation of child cognitive ability in reading and mathematics is reported as an ordinal three category variable in the PRHS survey. Parental response to the question whether the child is able to read a simple letter or newspaper with understanding is reported as; 'not at all', 'yes with difficulty' and 'yes easily'. Similar parental responses to the question whether the child is able to do simple addition and subtraction are collected. These responses are ranked from lowest (1) to highest (3).

These measures of child cognitive ability in reading and mathematics are not based on test scores on standardized tests from an independent assessor but, instead are the subjective evaluation of the parents concerned. There is no doubt that cognitive ability measures based on test scores are more accurate objective measures, but they were not available. However, in the present study we are interested in what influences parental behaviour in sending or not

[26] A detailed discussion the data set is undertaken in chapter 3.

sending their children to school. For such behaviour it is parental perceptions that matter, not objective reality – if they differ. Cognitive measures based on parental evaluation could be questioned on the grounds that in a society with low levels of literacy, parents might not be able to assess whether their children are learning or not; also they might not be able to assess whether they are getting educational services for which they are entitled. Andrabi et al (2002) reported, from Pakistan, that parents are able to infer quality variation between schools and fees respond in predictable ways to school inputs. It might be argued that if the parent cannot read then they may describe their child reading ability as good when in fact it is poor– and a literate person would describe the same child as being poor at reading. We include parental literacy in reading and numeracy as a regressor to test whether it makes any difference to the coefficients of interest. The results reported in Table 4.8 show that our main inference about school quality remains the same and parental literacy behaves the same as parental education (for comparison see Table 4.10).

The descriptive statistics in Table 4.1 show that more than half of the children (53.49 percent) were not in school.

Table 4.1: Child School Participation Status Age: 11-15

| Currently in School? | Freq. | Percent | Cum. |
|---|---|---|---|
| | | | |
| No | 859 | 53.49 | 53.49 |
| Yes | 747 | 46.51 | 100 |
| | | | |
| Total | 1,606 | 100 | |

Source: Calculated from PRHS-II

Table (4.2) shows child reading ability by current school attendance status. Current school attendance is associated with higher reading ability (56.53 percent). The reading ability figure for children who are not currently at school is 10.71 percent. Current school absence is associated with no reading ability (83.0 percent). The absence of reading ability figure for children who are currently at school is 25.03 percent. Not surprisingly a child's reading ability is also influenced by age. Older children who are currently at school are more likely to be able to read (Table A1 given in appendix A). However, their ability to read easily does not show any obvious relationship with age. This pattern is due to the fact that there are children not currently in school but who have ever been to school[27]. Therefore their cognitive ability

[27] There are three groups of children: currently in school, currently not in school but have ever been to school, and have not been to school at all since school age (child school age is 5).

appears in the same table for children who are not currently at school/have not ever been to school.

Table 4.2: Reading Ability by Current School Attendance (age: 11-15)

| Is able to read a simple letter or the newspaper with understanding | | Currently in School? | | |
|---|---|---|---|---|
| | | No | Yes | Total |
| Not at all | Count | 713 | 187 | 900 |
| | Percent | 83.0 | 25.03 | 56.04 |
| Yes, with difficulty | Count | 54 | 137 | 191 |
| | Percent | 6.29 | 18.34 | 11.89 |
| Yes, easily | Count | 92 | 423 | 515 |
| | Percent | 10.71 | 56.63 | 32.07 |
| Total | Count | 859 | 747 | 1,606 |
| | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table (4.3) presents children's reading ability by gender. Male children are more likely to have a higher reading ability (39.58 percent) than female children (23.54 percent). This difference is due to higher proportion of female not currently at school than male counterpart. Children's reading ability is also represented by age and gender (appendix A, Table A2). Reading ability of children increases with their age. However, the pattern is much more obvious for boys than girls.

Table 4.3: Reading Ability by Gender

| Is able to read a simple letter or the newspaper with understanding | | Gender | | |
|---|---|---|---|---|
| | | Male | Female | Total |
| Not at all | Count | 391 | 509 | 900 |
| | Percent | 45.78 | 67.69 | 56.04 |
| Yes, with difficulty | Count | 125 | 66 | 191 |
| | Percent | 14.64 | 8.78 | 11.89 |
| Yes, easily | Count | 338 | 177 | 515 |
| | Percent | 39.58 | 23.54 | 32.07 |
| Total | Count | 854 | 752 | 1,606 |
| | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table (4.4) gives children's mathematics ability by current school attendance status. Current school attendance is associated with higher maths ability (59.57 percent). The figure for children who are not currently at school is 18.86 percent. This is much lower than children currently attending school. Similarly current school absence is associated with inability to undertake simple addition and subtraction (71.01 percent). The proportion of children who are currently at school/attending school but are not able to do simple addition and subtraction is 24.63 percent. Child mathematics ability is also cross-tabulated by age and current school attendance status (Table A3 given in appendix A). The pattern is not so obvious; the inability to calculate of children who are currently at school decreases with age. However their ability to calculate easily does not show any obvious relationship with age. Inability to calculate of children, who are currently not at school, increases as they become older whereas, the ability to calculate easily also increases with age. This pattern may stem from children who have currently left school.

Table 4.4: Maths Ability by Current School Attendance

| Is able to do simple addition and subtraction? | | Currently in School? | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| Not at all | Count | 610 | 184 | 794 |
| | Percent | 71.01 | 24.63 | 49.44 |
| Yes, with difficulty | Count | 87 | 118 | 205 |
| | Percent | 10.13 | 15.8 | 12.76 |
| Yes, easily | Count | 162 | 445 | 607 |
| | Percent | 18.86 | 59.57 | 37.8 |
| Total | Count | 859 | 747 | 1,606 |
| | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table (4.5) presents children's mathematics ability by gender. Being male is associated with higher maths ability (45.67 percent) than being female (28.86 percent). Child maths ability is also tabulated by age and gender (appendix A Table A4). Boys' ability to do simple addition and subtraction is increasing in age. However, the pattern is not obvious in the case of girls.

Table 4.5: Maths Ability by Gender

| Is able to do simple addition and subtraction? | | Gender | | Total |
|---|---|---|---|---|
| | | Male | Female | |
| Not at all | Count | 339 | 455 | 794 |
| | Percent | 39.7 | 60.51 | 49.44 |
| Yes, with difficulty | Count | 125 | 80 | 205 |
| | Percent | 14.64 | 10.64 | 12.76 |
| Yes, easily | Count | 390 | 217 | 607 |
| | Percent | 45.67 | 28.86 | 37.8 |
| Total | Count | 854 | 752 | 1,606 |
| | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

As noted in the earlier tables on child literacy and numeracy and current school attendance, children who are not currently at school did have fairly positive cognitive ability in reading and mathematics. This pattern is due to those children who have at some time been to school but have now left. Tables 4.6 and 4.7 investigate this pattern. A child's reading and mathematics ability is represented for those children who have some time been to school/or not, the time he/she was of school age. Children who have been to school have higher maths/reading ability. However children who have never been to school since school age are not able to read a simple letter or the newspaper with understanding (100 percent) but a small proportion of children were able to undertake simple addition and subtraction.

The descriptive analysis presented above points to a number of potential correlates of child cognitive ability in reading and mathematics. Children's cognitive ability is higher in current school attendance as well as increasing with child age. The descriptive statistics also show that cognitive ability of male children is higher than female children, especially maths ability. Similarly children who are not currently at school but have ever been to/have not ever been to school have reported positive cognitive outcomes. Ignoring them in the econometric analysis would certainly cause sample selection consequences; a common characteristic of those studies using school-based surveys (Arif and Saqib 2003; Das et al. 2006; Aslam 2009).

Table 4.6: Reading Ability and Ever been to School

| Is able to read a simple letter or the newspaper with understanding | | Ever been to School or not? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Not at all | Count | 308 | 592 | 900 |
| | Percent | 30.37 | 100 | 56.04 |
| Yes, with difficulty | Count | 191 | 0 | 191 |
| | Percent | 18.84 | 0 | 11.89 |
| Yes, easily | Count | 515 | 0 | 515 |
| | Percent | 50.79 | 0 | 32.07 |
| Total | Count | 1,014 | 592 | 1,606 |
| | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table 4.7: Mathematics Ability and Ever been to School

| Is able to do simple addition and subtraction? | | Ever been to School or not? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Not at all | Count | 289 | 505 | 794 |
| | Percent | 28.5 | 85.3 | 49.44 |
| Yes, with difficulty | Count | 166 | 39 | 205 |
| | Percent | 16.37 | 6.59 | 12.76 |
| Yes, easily | Count | 559 | 48 | 607 |
| | Percent | 55.13 | 8.11 | 37.8 |
| Total | Count | 1,014 | 592 | 1,606 |
| | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

## 4.3 Modelling Methodology

### 4.3.1 Conceptual Framework and Literature Review

If parents choose to ensure child school participation when school quality is higher, then school quality must translate into child cognitive achievement. However, to quantify the link between school quality and cognitive achievement, controls for other important child and household level characteristics must be included in the model along with school quality.

Our investigation focuses on estimating production functions for cognitive achievement (Angrist and Lavy 1999; Alderman et al. 2001; Arif and Saqib 2003; Das et al. 2006; Urquiola 2006; Aslam 2009) reported as parental evaluation of child ability in reading and

numeracy. Major direct determinants of the cognitive achievement are level/years of schooling, school quality, parental input into the schooling process and a child's potential to learn.

With regard to school quality-related variables[28] we assess the contribution of student-teacher ratios and proportion of schools with a library, while controlling for level/years of schooling, family background/characteristics. These school quality variables are community-level averages/expected values. Community-level measurement of school quality is necessary because we cannot measure at the child level the student-teacher ratio or other school quality measures for a child who does not attend school. We also include as an explanatory variable for cognitive achievement the number of schools in a local community.

School lack of availability or accessibility in the community entails potentially significant opportunity-costs to the parents in terms of child time going to and from school to home. Alderman et al. (2001), Ersado (2005) and Filmer (2007) focussed on number of schools in an area to proxy school accessibility/availability. However, even if schools are available to households in the local area they may not have the desired quality attributes. Then number of schools alone would not be able to proxy school accessibility/availability. In such circumstances, the nearest school would have much more disutility to parents than one further away. Consequently we regard number of schools simply as a control to ensure our two quality coefficients are identified [29]. These capacity variables reflect the educational infrastructure available to all households in the community and could be an important determinant in the explanation of child cognitive achievement. However, to distinguish school capacity/infrastructure effect from school crowding effect, i.e. a lot of students in relation to teachers, we also control for number of pupils per teacher. For instance, a few schools might be available in a particular community having the desired school quality characteristics but might be over-crowded due to high parental evaluation and therefore child participation.

The characteristics of the specific school a child attends are determined jointly with the schooling decision, hence are endogenous to schooling (Deaton 1988; Alderman et al. 2001).

[28] I tested for an exhaustive list of school quality attributes like toilet facility for girls/boys, school boundary wall, electrified schools, furniture for students, textbooks for students, teacher qualifications, teacher experience but student teacher ratio playground and Library stood robust/significant in the coming regressions.

[29] Further detail on the exogeneity/endogeneity of our school quality correlates are delayed until chapter 5 because we undertook Instrumental variable analysis to control for the potential endogenous variables in the ordinal analysis but other diagnostic tools for the evaluation of endogeneity or testing for over-identifying restrictions have not been available in ordered specifications (Instrumental Variable Ordered Probit specification) until recently.

But community-level school quality variables/expected values of school quality inputs, are not endogenous to household-level child schooling decisions (Alderman et al. 2001; Ersado, 2005).

Alderman et al. (1996) and Behrman et al. (2008) assumed the availability and quality of local schools as exogenous for their estimates. Because the availability and quality of local schools in rural Pakistan are determined by district and higher level decisions they are not in direct response to village characteristics and household demands. If school services are provided without consideration for local level demands for school services, the supply of teachers will be exogenous to child cognitive ability/ school participation. However, this will not be so for the student-teacher ratio, unless and until the supply of teachers is adjusted spontaneously with demand for child schooling. Student-teacher ratios will be positively correlated with enrolments, and enrolments in turn positively correlated with socioeconomic status of parents and school performance (Angrist and Lavy 1999; Case and Deaton 1999; Hoxby 2000; Urquiola 2006; Das et al. 2011). We tested for these claims in our sample and found that the student-teacher ratio is endogenous to child cognitive ability and child school participation. The school quality measure, i.e. the proportion of schools with a library is exogenous (further detail presented in chapter 5, please also consult footnote 29).

Parental inputs into the schooling process, time and other resources devoted to helping children in education, are a matter of household choice which is not directly observed in the data (Behrman et al.1997). To control for such inputs, parental level of education is included in the estimating equations along with household expenditure. Expenditure (household income net of child earnings (equivalent)[30]) is usually preferable to income as it is less volatile and reflects permanent income (Behrman et al.1997; Bhalotra 2003). In addition, there is no control available in the data for individual's potential to learn/pre-school ability that is independent of the current level of schooling. The possibility is that the coefficient on schooling would be positively biased for this reason.

### 4.3.2 Empirical Methodology

The framework adopted for modelling the outcome variable on child cognitive achievement is an ordered probit model, suitable when the dependant variable is categorical. Multiple factors contribute to children's cognitive attainment in different degrees/proportions, hence

[30] Detailed discussion on equivalent scales and different measures of household income/expenditure is presented in chapter 5. Similarly, household expenditure or household income has been used interchangeably throughout the thesis. However, it is household expenditure that is determined not household income.

multivariable modelling is an informative and an especially useful tool for understanding cognitive ability. The ordered probit model is extremely useful because, unlike OLS regression, it allows for unequal differences between ordinal categories in the dependant variable (McKelvey and Zavoina 1975; Greene 2003). For example it does not assume that the difference between 'not able to read' and 'able to read with difficulty' is the same as the difference between 'able to read with difficulty' and 'able to read easily'. Ordered probit model especially takes into account the qualitative difference between levels of child cognitive ability.

The ordered probit model is more appropriate than unordered multinomial models because unordered multinomial models do not account for the ordinal nature of the categorical variable (in this case, child cognitive achievement). Another variant for ordered analysis is the ordered logit model. However, practically ordered probit and ordered logit model give very similar results (Hill et al. 2011). The error term in ordered probit model is normally distributed while in the case of ordered logit model it is logistically distributed. The diagrammatic representation of ordered probit model is given in figure 4.1.

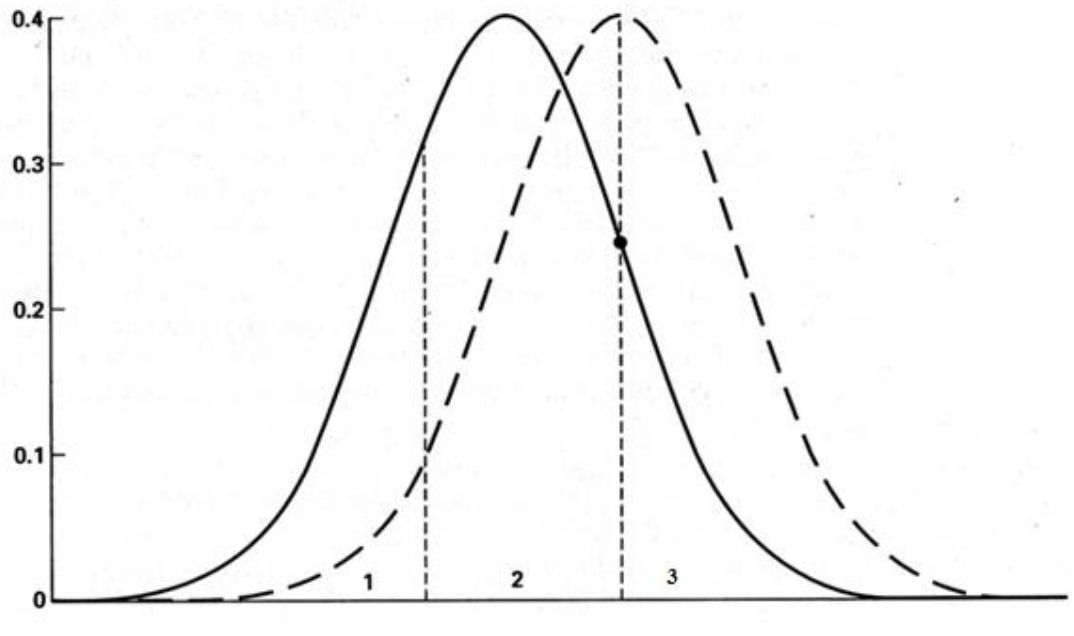



Figure 4.1: Ordered Probit Model; Changes in Explanatory Variables Shift Distribution of Reading/Maths Performance over the Three Categories

Attributes of children, family background/characteristics and school characteristics are chosen as explanatory variables. The dependant variable is child cognitive ability in reading and mathematics. Child reading or mathematics ability is classified in three categories with the most able children coded as 3 (yes easily); the next category as yes with difficulty (coded as 2); and the last category as not at all (coded as 1).

Let $y_i^*$ be the unobserved latent variable representing child cognitive ability (in reading and mathematics) in household '$i$' coded 1,2 and 3; '$b$' is the vector of estimated parameters and '$e_i$' represents the unobserved disturbance term having mean zero and unit variance (assumed to be normally distributed) and $X_i^{'}$ is the vector of explanatory variables. The ordered probit model takes the form;

$$y_i^* = X_i^{'}b + e_i \qquad (4.1)$$

The true dependant variable, in the present case parental opinion, is unobservable. The observed outcome $y_i$ is related to the index $y_i^*$ by means of (m-1) threshold values $\tau_1 < \tau_2$, where '$\tau_1 < \tau_2$' are the two thresholds between which observed categorical responses are estimated. If the opinion about child cognitive ability in reading and mathematics is in the lowest category, then $y_i^* \le \tau_1$ and consequently the alternative 'not at all' is chosen. Similarly when $\tau_1 < y_i^* \le \tau_2$, then alternative 'yes, with difficulty' is chosen and if the sentiment is in the highest category, then $y_i^* > \tau_2$ is chosen. It is represented as;

$$y = \begin{cases} 1 \quad (not, atall) \quad \cdots \quad \cdots if \cdots y_i^* \le \tau_1 \\ 2\,(yes, withdiffic\,ulty) \cdots\cdots if \cdots \tau_1 < y_i^* \le \tau_2 \\ 3 \quad (yes, easily) \quad \cdots \quad \cdots if \cdots y_i^* > \tau_2 \end{cases}$$

Once a particular probability distribution is assumed for $y_i^*$ or $e_i$, then one is able to calculate probability ($P$) of each outcome, where '$\Phi(.)$' is the cumulative distribution of the ordered probit model, as;

$$P(y=1) = P(y_i^* \le \tau_1) = P(X_i^{'}b + e_i \le \tau_1)$$

$$P(y=1) = P(e_i \le \tau_1 - X_i^{'}b)$$

$$P(y=1) = \Phi(\tau_1 - X_i^{'}b)$$

The probability that y=2 is given as;

$$P(y=2) = P(\tau_1 < y_i^* \le \tau_2) = P(\tau_1 < X_i'b + e_i \le \tau_2)$$

$$P(y=2) = P(\tau_1 - X_i'b < e_i \le \tau_2 - X_i'b)$$

$$P(y=2) = \Phi(\tau_2 - X_i'b) - \Phi(\tau_1 - X_i'b)$$

Similarly the probability that y=3 is given as;

$$P(y=3) = P(y_i^* > \tau_2) = P(X_i'b + e_i > \tau_2)$$

$$P(y=3) = P(e_i > \tau_2 - X_i'b)$$

$$P(y=3) = 1 - \Phi(\tau_2 - X_i'b)$$

The general expression for probability of three categories is given as;

$$P(y=n) = \Phi(\tau_n - X_i'b) - \Phi(\tau_{n-1} - X_i'b) \quad \text{n=1, 2, 3}$$

Where

$Xb = (b_1\,Female\ (yes) + b_2\,Age\ (Years) + b_3\,Household\ Expenditure + b_4\,Mother\ Level\ of\ Education + b_5\,Father\ Level\ of\ Education + b_6\,Child\ Level\ of\ Education + b_7\,Currently\ in\ School\ (Yes) + b_8\,Student\ teacher\ ratio + b_9\,Number\ of\ Schools\ in\ a\ village$

$+ b_{10}\,Proportion\ of\ schools\ with\ a\ library)$

The estimation is undertaken by maximum likelihood (McKelvey and Zavoina 1975; Greene 2003; Hill et al. 2011). In case of a random sample of three individuals with the first one not able to read a simple letter or the newspaper with understanding ($y_1$=1), the second one able to read a simple letter or the newspaper with understanding but with difficulty ($y_2$=2), and the third able to easily read a simple letter or the newspaper with understanding ($y_3$=3), then the likelihood function is given as;

$$L(b, \tau_1, \tau_2) = P(y_1 = 1) \times P(y_2 = 2) \times P(y_3 = 3)$$

These probabilities depend on the unknown thresholds $\tau_1$ and $\tau_2$, as well as the parameters '$b$' of the index function. The threshold values show the range of the normal distribution that is associated with the specific values of the observed response variable. The parameters '$b$' capture the effect of changes in the explanatory variables on the underlying scale. The marginal effect of changes in the independent variables is given by;

$$\frac{\partial P(y_i = n)}{\partial X_i} = (\Phi(\tau_{n-1} - X_i^{'}b) - \Phi(\tau_n - X_i^{'}b))b \qquad \text{n=1, 2, 3}$$

The sign of the parameter '$b$' always indicates the direction of the marginal effect of the highest category and is the opposite of the marginal effect of the lowest category. However marginal effects are particularly important where the effects of explanatory variables on the middle categories are ambiguous when only parameter estimates are available (as the figure 4.1 should demonstrate) (Hill et al. 2011).

The starting point for econometric analysis, is to assume that '$e_i$' satisfies the classical conditions given below;

1. $E(e_i \setminus X_i^{'}) = 0$ (regressors are exogenous)
2. $E(e_i^{\,2} \setminus X_i^{'}) = \sigma^2$ (conditional homoskedasticity)
3. $E(e_i e_j \setminus X_i^{'} X_j^{'}) = 0 \;\; i \neq j$, (conditionally uncorrelated observations)

Assumption 1 is important to get unbiased (consistent) estimates of '$b$'. Assumption 2 and 3 is essential to determine the form of the variance-covariance matrix (VCE) of the estimator '$b$'. Assumptions 1-3 results in '$b$' being asymptotically normally distributed with the default VCE. Under assumption 1-3, the estimator '$b$' is fully efficient. However, different assumptions about the data generating process, such as hetroskedasticity, results in different estimates of the VCE (Cameron and Trivedi 2009). When assumption 1 and 3 holds, but not 2, we get heteroskedastic uncorrelated errors. Then it is essential to estimate heteroskedasticity-robust estimator of the VCE of the regression estimator. This estimator was introduced by (White 1980) and also called robust/sandwich estimator. According to Cameron and Trivedi (2009), binary outcome model are unusual in a sense that there is no advantage in using the robust VCE of the MLE if the data is independent over 'i' and the density function is correctly specified. The ML default standard errors are obtained by imposing the restriction that the variance of a binary variable is $p(1-p)$. If the density function is correctly specified, then the robust VCE estimator is not required. If the default VCE form of the ML and robust option gives substantially different variances of the same estimator, then we may infer a miss-specified functional form of the density.

However, dependence between observations may arise due to cluster sampling and it is important to use the appropriate procedure that the cluster-robust approach which generalizes the White robust/sandwich estimator of the VCE. When errors for different observation are

correlated, assumption 3 does not hold. In such case both default and robust estimates of the VCE are invalid. In cross-sectional data it usually arises when errors are clustered. Clustered errors are errors that are uncorrelated across clusters or groups but are correlated within a cluster. For example, clustering arises when sampling is of independent units (villages in our case) but errors for individuals (children in our case) within the unit are correlated. Our school quality correlates are village/community level averages of schools in a village and village level school quality regressors do not vary within a village. A small error correlation for children in each village would lead to a great downward bias in the default standard errors, as well as in heteroskedasticity-robust standard errors. Hence, given the importance of cluster standard errors we estimated cluster standard errors of the regression estimators over the villages.

As noted above, assumption 1 is important to get unbiased (consistent) estimates of '$b$' that is the model error is uncorrelated with the regressors i.e., $E(e_i \setminus X_i^{'}) = 0$. If this assumption fails the ordered probit estimator is inconsistent, and henceforth the marginal effects no longer gives causal interpretation. This is a fundamental problem as marginal effects are potentially informative of economic policy (Cameron and Trivedi 2009). In such circumstances it is important to use the instrumental variable approach which provides consistent estimators under the assumption that valid instruments exists. The instruments (Z) are variables that are correlated with (X) and must satisfy $E(e_i \setminus Z) = 0$, that the instruments are uncorrelated with the regression error term.

We suspect that the student-teacher ratio is endogenous in our model (4.1). Let suppose a latent-variable model with dependant variable $y_1^*$ (cognitive achievement) in the structural equation (4.3). The endogenous variable (regressor) is represented by $y_2$ (student-teacher ratio) in equation (4.4) below. Both these endogenous variables are modelled in exogenous variables $Z_i^{'} = \left[X_{1i}^{'} X_{2i}^{'}\right]$. $X_1$ serves as instruments for itself and $X_2$ is the instrument for $y_2$, and the instrument must satisfy the moment condition that

$$E(e_i \setminus Z_i^{'}) = 0 \qquad (4.2)$$

in the equations;

$$y_{1i}^{*} = by_{2i} + X_{1i}^{'}\eta + e_i \qquad (4.3)$$

$$y_{2i} = X_{1i}^{'}\pi_1 + X_{2i}^{'}\pi_2 + v_i \qquad (4.4)$$

Where $i = 1,.....N$; $X_1$ is a vector ($K_1 \times 1$) of exogenous variables; and $X_2$ is a $L_1 \times 1$ vector of additional instrumental variables (IV) that affect $y_2$ but excluded from (4.3) as these IVs do not affect $y_1$ directly. To ensure identification it is required that $L_1 \geq 1$. The dependant variable $y_1^*$ is latent and not directly observed. However, the ordered outcome $y_1$ is observed as presented above.

Equation (4.3) (child cognitive achievement) is of main interest. The second equation (4.4) is a first-stage equation or reduced-form equation for the student-teacher ratio and serves as a source of identification. It also shows the strength of the instruments and goodness of fit of the reduced form equation. The reduced form equation explains the variation in the endogenous variable (student-teacher ratio) in terms of exogenous variables including IV $X_2$ (average number of teachers per school with secondary qualification and average teacher experience per school in the community) that are excluded from the structural equation (cognitive achievement). The excluded instruments are essential for identifying the parameters of the structural equation. The estimation procedure depends on the specification of the structural and reduced form equations and can be simultaneous or sequential. In the case of the IV probit model (presented in chapter 5), the structural model approach completely specifies the distributions of $y_1^*$ (cognitive achievement) and $y_2$ (student-teacher ratio) in (4.3) and (4.4), respectively. It is assumed that $e_i$ and $v_i$ are jointly normally distributed. The assumptions imply that $e_i \setminus v_i = \rho v_i + u_i$, where, $E(u_i \setminus v_i) = 0$.

A test of the null hypothesis of the exogeneity of $y_2$ is to test that $H_0 : \rho = 0$, as then $e_i$ and $v_i$ are independent (Wooldridge 2002; Cameron and Trivedi 2009). Consistent estimation requires both the normality and homoskedasticity of $e_i$ and $v_i$. The default procedure for IV probit model delivers ML estimates and estimation is simultaneous. An alternative procedure is sequential estimation of (4.3) and (4.4) with the two-step sequential estimator. The difference from the ML estimator is only computational as both approaches make the same distributional assumption (Cameron and Trivedi 2009). But computation of marginal effects is not allowed after the two step procedure and only available for ML estimator. The test of over-identifying restrictions is only available after the two-step estimation framework and has been estimated in the next chapter.

As mentioned the instruments must satisfy condition (4.2). This condition cannot be tested in the just identified case, but a test is possible in the over-identified case. When there is more

than one instrumental variable, we can test whether some of them are correlated with the structural error. For example if we have two instrumental variables for a single endogenous variable (here student-teacher ratio), we have one over-identifying restriction. However, even in the over-identified case instrument validity depends on economic theory, persuasive argument or norms established in the previous related empirical works (Greene 2003; Cameron and Trivedi 2009; Wooldridge 2009). Similarly the instruments must be relevant in that after controlling for the remaining exogenous variables $X_1$, the instruments $X_2$ must explain significant variation in $y_2$. Generally the stronger the association between the instruments (average number of teachers per school with secondary qualification and average teacher experience per school in the community) and endogenous regressor (student-teacher ratio) the better will be the identification of the model.

Estimation of the IV ordered probit model (cognitive achievement function) is here undertaken with the cmp setup. cmp is a seemingly unrelated (SUR) estimation programme (Roodman, 2009). cmp consistently estimates parameters in the limited information case when only the final stage is fully specified. Equations for earlier stages include instruments to address endogeneity and generally omit influential variables. This chapter uses cmp to control for the endogeneity of the student-teacher ratio in child reading and mathematics achievement functions while using two instruments; average number of teachers per school with secondary qualification and average teacher experience per school in the community. Currently there is no test available in the cmp setup to test for over-identified restrictions. However, we have tested for them in the next chapter (5) for the IV probit model for child school attendance.

### 4.4 Results and Discussion

Child cognitive achievement is estimated using Instrumental Variable (IV) ordered probit regression, to control for the endogeneity of important policy correlates of child cognitive reading and mathematics abilities[31]. The estimation of cognitive ability is undertaken separately for reading and mathematics over the whole range of child age (11-15).

Household expenditure is dropped as we tested for zero restriction ( $H_0; \beta_j = 0$, against $H_1; \beta_j \neq 0$ ), for household expenditure in the equation. And the results justified the exclusion of the variable from all reported equations in the subsequent analysis. A similar

[31] The implementation of Instrumental Variable ordered probit regression in Stata is possible with user written command cmp (Roodman, 2009), which has been followed in subsequent Instrumental Variable analysis.

result is reported by Case and Deaton (1999) for white children (ethnic group) in South Africa. Household resources have no effect on the educational attainment of white children but a significant positive effect on the educational achievement of black children. Case and Deaton argue that education is cheap enough and funding is not a constraint for whites but is a serious constraint for the poorer blacks. A similar result appears to be true in rural Pakistan, as nearly 90 percent of the children in the sample participate in government schools. Government schooling is virtually free in Pakistan with minimal school fees comparatively to private schooling. Hence private expenditure (household expenditure) on education is unlikely to be a constraint on households in rural Pakistan. The result strengthens our claim that lower school participation is not due to poverty but to lower school quality in Pakistan.

### 4.4.1 Whole Sample Results

#### 4.4.1.1 Reading Performance

Equation (1) Table 4.10 presents estimates on child reading ability while controlling for family background, number of schools in the community and correlates on school quality (student-teacher ratio and proportion of schools with a library). Instrumental variable estimates of the same model are reported as equation (2). Estimates of cluster standard errors are reported in square brackets in the same equations. Student-teacher ratio is instrumented for reasons given earlier. Our chosen instruments for student-teacher ratio (average number of teachers per school with secondary qualification and average teacher experience per school in the community) are valid as the relevant p-value $< 0.001$ (0.1 percent level of significance) in the first stage regression reported on Instrumental Variable equation in the same table. As expected, the significance of the IVs falls with the cluster standard errors. Nonetheless average teacher experience per school in the community is significant at 5 percent level of significance. Some researchers have used teacher qualification and teacher experience as school quality correlates in cognitive achievement functions and school participation equations, but there is no consensus that they are suitable measures of school quality (see Behrman et al. 1997 for a review). In the current data set we did not find coefficients on teacher qualification and teacher experience significantly different from zero in either the child cognitive function or the school participation equation (see Table 4.9). Further justification of why teacher qualification and teacher experience are valid instruments for student-teacher ratio in the context of Pakistan is deferred until chapter 5.

Unfortunately at the moment there is no test available for endogeneity or over-identifying restrictions in the case of instrumental variable ordered probit model other than looking at the

first stage regression. However, we do test over-identifying restrictions for the same instruments for student-teacher ratio in the instrumental variable child participation equations in chapter 5. The results reported there do support the validity of the chosen instruments.

Some striking generalization directly flows from the estimates reported of child cognitive ability in reading (Table 4.10). The simple ordered probit results show that the estimates of school quality variables are in accordance with the conceptual framework presented earlier, except for student-teacher ratio being endogenous to cognitive ability. Once the endogeniety of student-teacher ratio is controlled, the previous insignificant impact becomes significantly different from zero. However, the same relationship does not hold with cluster standard errors.

The statistical significance and direction of the impact on child cognitive ability in reading of other school quality variable 'proportion of school with a library' is similar in the simple ordered probit and IV ordered probit equations in the same table. Library proportion is highly significant (at 1 percent level of significance) with cluster robust estimation. The results indicate that library availability in village schools strongly affects children's ability to read easily. Library provision is more important for child reading ability than greater student-teacher interaction, that stems from lower student-teacher ratios or lower class sizes.

The measure of school accessibility or availability, i.e. number of schools in the community, significantly positively affects child reading ability. The results remain the same in both specifications, as well as with cluster robust standard error estimates ($p < .01$).

Controls for child level and household level characteristics behave consistently in both equation 1 and 2. Child cognitive ability in reading is lower for females but the estimated effect is insignificant. Child cognitive ability significantly increases with child age and parental level of education significantly improves child cognitive ability. However, the statistical significance and marginal effect of maternal level of education is higher than that of paternal level of education. Child level of education and child enrolment status increases the likelihood that the individual child has better cognitive ability. The estimated coefficients for these two variables are highly significant and positive in both equations; with or without cluster robust estimates of standard errors.

Specifically, the marginal effect estimates presented in Table 4.11 on outcome 3 (able to easily understand a simple letter or the newspaper with understanding) shows that females are less able to read easily by 2.5 percent equation (2) while keeping other correlates at the sample mean. However, we cannot reject the hypothesis of equal marginal effects between

the sexes. The marginal effect of maternal level of education is higher (8.1 percent) than paternal level of education (2.1 percent). A one year increase in child age significantly increases the ability to read easily by 3.5 percent while keeping other correlates at the sample mean. A one level increase in child education increases ability to read easily by 48 percent. Similarly the ability of children enrolled in school to read easily is higher (16) percent than children not at school currently. The mean value of library proportion is 0.205. An increase in library proportion up to 100 percent would raise the variable by 0.795. Increase in library proportion by 0.795 would result increase in child cognitive ability in reading by 24.1 percent (0.795×0.303=0.241).

#### 4.4.1.2 Mathematics Performance

The same analysis is undertaken for child cognitive ability in mathematics. Whole sample ordered probit and Instrumental variable (IV) ordered probit estimates are presented in equations 3 and 4 in Table (4.10). Marginal effects on child mathematics ability are given in Table 4.11 (equations 3 and 4).

Major findings from whole sample maths ability regression analysis are given as follows. The results show that the estimates of school quality variables are consistent in both equations (equations 3 and 4). Contrary to estimate of student-teacher ratio on child reading ability, student-teacher ratio in the case of maths ability is significant at 5 percent level of significance with cluster robust estimates of standard errors (equation 4). The parameter estimate of library proportion is insignificant, in the case of mathematics ability, with cluster robust estimates of standard errors. Bigger class size more drastically worsens child mathematics ability than in the case of reading. This might be because greater interaction or attention is needed on the part of the teacher to be fully abreast of each child's knowledge of mathematics. Evidently it is not possible in bigger classes. The results also indicate that the marginal effects of school quality are higher in the case of mathematics achievement than for reading (Table 4.11). A one standard deviation decrease in student-teacher ratio, approximately 13 pupils, increases child ability to calculate easily by 0.31 standard deviations. To add plausibility to our estimates we have compared our estimates with those reported on standard test scores in the literature. Urquiola (2006) found, for Bolivia, an improvement of up to 0.3 standard deviations for a one standard deviation reduction in class size (approximately 8 students). For Israel, Angrist and Lavy (1999), found effects ranges from 0.10-0.20 for a 10 student reduction in student-teacher ratio. Using a 10-student reduction, Urquiola (2006) report the effect sizes range between 0.19 and 0.35. Using the 10 student

reduction as a benchmark, for comparability, the effect found here is 0.24. The comparison shows that our results are broadly similar to that in Angrist and Lavy (1999) and Urquiola (2006).

The parameter estimate of the measure of school accessibility or availability; number of schools in the community, is insignificantly different than zero with cluster robust standard error estimate in child mathematics ability (equation 4).

Some of the parameter estimates of controls for child level and household level characteristics for maths are different from those for reading. Girls' cognitive ability in mathematics is significantly lower than boys by 4.7 percent (equation 4). Maternal level of education is significant at 5 percent level of significance. However, the estimated effect of paternal level of education is insignificant (equation 4).

### 4.4.2 Boys' and Girls' Cognitive Achievement Functions

The results on whole sample cognitive ability, reported earlier, show that there are significant gender differences in child cognitive achievement at least in mathematics. Therefore the same models are estimated separately over boys and girls to calculate the magnitude of likely differences on the vector of coefficients for each control accounted for in the regression model. Recent studies in Pakistan (Kingdon and Aslam 2008; Aslam 2009) have found significant pro-male bias in within household's allocation of educational expenditures in both the enrolment decision and the decision about how much to spend on the already enrolled children. Their findings reveal that boys are more likely than girls to be sent to quality schools. Girls' differential treatment adversely affects their cognitive ability in reading and mathematics.

#### 4.4.2.1 Boys' Ability in Reading and Mathematics

The estimates of ordered probit and Instrumental Variable ordered probit regression for boys' cognitive ability in reading and mathematics are provided in Table 4.12. Similarly, their respective marginal effects are presented in Table 4.13.

Major findings from boys' reading and maths ability regressions analysis are as follows. The estimates show that boys' ability to read and calculate significantly increases with school quality. However, the statistical significance and relevant magnitude of parameter estimates (marginal effects) of school quality correlates are more pronounced (higher) in the case of mathematics ability than in the reading ability (Tables 4.12 and 4.13). School quality

variables ‘student-teacher ratio’ and ‘proportion of schools with a library’ behave similarly to the case of whole sample regression estimates.

Other controls in the model show similar relationship across boys’ ability to read and calculate as reported above (whole sample analysis). The marginal effect of maternal level of educations is higher than paternal level of education; however, their relevant estimates are significant only for maternal level of education in the case of boys’ ability to numerate easily (equation 2 and 4, Tables 4.12 and 4.13). The apparent higher magnitude of the parameter estimates of maternal level of education and school quality correlates, in the case of maths ability than reading, might be interpreted as that more time input on the part of the mothers and greater specialized help on the part of the teachers is required for children to learn mathematics. The parameter estimate of the variable ‘number of schools in the community’ behaves the same as in the case of whole sample analysis; the estimated effect is significant for boys’ reading ability but insignificant for their mathematics ability.

#### 4.4.2.2 Girls’ Ability in Reading and Mathematics

The estimates of ordered probit and Instrumental Variable ordered probit regression on girls’ cognitive ability in reading and maths are provided in Table 4.14. The respective marginal estimates are provided in Table 4.15.

The main findings from girls’ reading and maths ability regression analysis are as follows. The parameter estimate of school quality measure ‘student-teacher ratio’ is insignificant in the case of girls’ reading ability (equation 2); with or without cluster robust estimates of the standard errors. However, the parameter estimate of ‘student-teacher ratio’ is negatively significantly different than zero in the case of girls’ mathematics ability (equation 4). The estimated effect of library proportion is significant and positive in the case of reading ability but insignificant in the case of girls’ mathematics ability- libraries may be important for providing reading material, but obviously less so for maths. These results are equivalent to those reported for whole sample and boys only sample estimates.

Results on other control in the models show that maternal level of education is significant only in the case of girls’ reading ability (equation 2). The estimated effect of paternal level of education is insignificant in girls’ reading as well as maths ability (equations 2 and 4). The parameter estimate of the variable ‘child age’ is significant in girls’ reading ability but insignificant in maths ability. Current school attendance does not have the same effect across girls’ reading and maths ability. Its estimated effect is significant in the case of reading ability

but insignificant in the case of maths ability (equations 2 and 4). Other controls in the model show similar representation across girls' ability to read and numerate. Their relevance (magnitude) and significance remain the same as reported in the case of boys and whole sample analysis.

## 4.5 Conclusion

The objective of this chapter was to evaluate the impact of school quality in child cognitive production functions. It implements Roodman's cmp to control for the endogeneity of student-teacher ratio in an IV ordered probit model. Parental evaluation of their child cognitive ability was quantified while using school quality indicators that are usually not available in any developing country surveys. The results show that school quality is of prime importance in child cognitive achievement in reading and mathematics ability while controlling for other child level and household level characteristics. We exploited measures of child cognitive achievement in reading and mathematics based on subjective evaluation of the parents concerned as we are mainly interested to investigate which factors influences parental beliefs, for these must affect the child school attendance decision.

Robertson and Symons (2003) reported for England that it is not school quality but parental quality (socioeconomic status and parental time) and peer group effect that have the strongest effects on child educational achievement. They found that conventional school quality inputs, such as class size, teacher experience and private school, perform badly (explain poorly) and were often perversely signed. The results we find for rural Pakistan are very different. Boys' and girls' cognitive ability as perceived by their parents in reading and mathematics significantly increase with higher school quality. In summary school quality significantly affects the likelihood of higher child cognitive achievement over the whole sample and in the boys' and girls' analysis.

Table 4.8: Ordered Probit Estimates of Parental Evaluation of Child Reading and Mathematics Ability with Parental Literacy (Reading and Numeracy)

| | (1)<br>Reading Ability | (2)<br>Mathematics Ability |
|---|---|---|
| main | | |
| Child Sex (female) | -0.0937 | -0.114 |
| | (0.0820) | (0.0756) |
| Child Age | 0.141*** | 0.115*** |
| | (0.0306) | (0.0271) |
| Mother's Level of Education | 0.159 | -0.115 |
| | (0.141) | (0.0962) |
| Father's Level of Education | 0.0849 | 0.0769 |
| | (0.0547) | (0.0520) |
| Child's level of Education | 1.920*** | 1.422*** |
| | (0.125) | (0.0819) |
| Currently in school (yes) | 0.652*** | 0.259** |
| | (0.0979) | (0.0920) |
| Number of schools in a Community or village | 0.0657*** | 0.0446*** |
| | (0.0143) | (0.0127) |
| Student Teacher Ratio | -0.00353 | -0.00740* |
| | (0.00320) | (0.00295) |
| Proportion of Schools with a Library | 1.265*** | 1.054*** |
| | (0.213) | (0.203) |
| Father Literacy | -0.00993 | |
| | (0.0630) | |
| Mother Literacy | 0.226+ | |
| | (0.134) | |
| Father Numeracy | | -0.0396 |
| | | (0.0584) |
| Mother Numeracy | | 0.802*** |
| | | (0.0547) |
| cut1 | | |
| _cons | 6.927*** | 5.369*** |
| | (0.522) | (0.436) |
| cut2 | | |
| _cons | 7.581*** | 5.966*** |
| | (0.531) | (0.441) |
| Pseudo R-Square | 0.442 | 0.371 |
| Wald chi2 (11) | 562.4 | 728.4 |
| p | 1.55e-113 | 4.34e-149 |
| N | 1606 | 1606 |

Standard errors in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

Table 4.9: Ordered Probit Estimates of Parental Evaluation of Child Reading and Mathematics Ability with Teacher Qualification and Teacher Experience

| | (1)<br>Reading Ability | (2)<br>Mathematics Ability |
|---|---|---|
| Child Sex (female) | -0.101 | -0.163* |
| | (0.0812) | (0.0705) |
| Child Age | 0.142*** | 0.0965*** |
| | (0.0306) | (0.0264) |
| Mother's Level of Education | 0.328** | 0.292** |
| | (0.113) | (0.101) |
| Father's Level of Education | 0.0790* | 0.0397 |
| | (0.0375) | (0.0325) |
| Child's level of Education | 1.922*** | 1.178*** |
| | (0.126) | (0.0762) |
| Currently in school (yes) | 0.652*** | 0.289** |
| | (0.0984) | (0.0891) |
| Number of schools in a Community or village | 0.0603*** | 0.0328** |
| | (0.0149) | (0.0124) |
| Student Teacher Ratio | -0.00406 | -0.0146*** |
| | (0.00318) | (0.00276) |
| Proportion of Schools with a Library | 1.219*** | 1.362*** |
| | (0.217) | (0.197) |
| Average Teacher Experience per School in the Community | -0.000150 | -0.0154 |
| | (0.0104) | (0.00869) |
| Average Number of teacher per school with secondary qualification | -0.312 | -0.349 |
| | (0.232) | (0.181) |
| cut1 | | |
| _cons | 6.749*** | 3.447*** |
| | (0.543) | (0.421) |
| cut2 | | |
| _cons | 7.403*** | 3.955*** |
| | (0.551) | (0.425) |
| Pseudo R-Square | 0.441 | 0.284 |
| Wald chi2 (11) | 549.0 | 594.7 |
| p | 1.12e-110 | 1.95e-120 |
| N | 1606 | 1606 |

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.10: Ordered Probit and Instrumental Variable (IV) Ordered Probit Estimates of Parental Evaluation of Child Reading and Mathematics Ability (Whole Sample (age 11-15))

| | (1) Reading Ability | (2) IV. Reading Ability | (3) Mathematics Ability | (4) IV. Mathematics Ability |
|---|---|---|---|---|
| Child Sex (female) | -0.101<br>(0.0817)[ 0.0986] | -0.0954<br>(0.0780)[ 0.0921] | -0.169<br>(0.0705) * [0.0833] * | -0.124<br>(0.0545) * [0.0760] |
| Child Age | 0.140<br>(0.0305) *** [ 0.0343] *** | 0.131<br>(0.0315) *** [ 0.0440] ** | 0.0977<br>(0.0263) *** [0.0310] ** | 0.0705<br>(0.0211) *** [0.0367] + |
| Mother's Level of Education | 0.322<br>(0.115) ** [ 0.115] ** | 0.303<br>(0.113) ** [ 0.121] * | 0.288<br>(0.102) ** [0.0972] ** | 0.208<br>(0.0885) * [0.105] * |
| Father's Level of Education | 0.0825<br>(0.0373) * [ 0.0425] + | 0.0785<br>(0.0352) * [ 0.0426] + | 0.0404<br>(0.0323)[0.0386] | 0.0317<br>(0.0251)[0.0317] |
| Child's level of Education | 1.921<br>(0.125) *** [ 0.123] *** | 1.814<br>(0.177) *** [ 0.355] *** | 1.165<br>(0.0749) *** [0.0813] *** | 0.865<br>(0.122) *** [0.356] * |
| Currently in school (yes) | 0.646<br>(0.0979) *** [0.121] *** | 0.602<br>(0.106) *** [ 0.173] *** | 0.285<br>(0.0888) ** [0.120] * | 0.198<br>(0.0719) ** [0.122] |
| Number of schools in a Community or village | 0.0664<br>(0.0143) *** [ 0.0196] *** | 0.0635<br>(0.0145) *** [ 0.0201] ** | 0.0338<br>(0.0117) ** [0.0240] | 0.0243<br>(0.00988) * [0.0220] |
| Student Teacher Ratio | -0.00342<br>(0.00319)[ 0.00474] | -0.0288<br>(0.0160) + [ 0.0403] | -0.0146<br>(0.00275) *** [0.00564] ** | -0.0628<br>(0.00684) *** [0.0271] * |
| Proportion of Schools with a Library | 1.302<br>(0.211) *** [ 0.226] *** | 1.136<br>(0.251) *** [ 0.395] ** | 1.454<br>(0.190) *** [0.308] *** | 0.866<br>(0.213) *** [0.573] |
| cut1<br>_cons | 6.873<br>(0.520) *** [ 0.537] *** | | 3.726<br>(0.399) *** [0.542] *** | |
| cut2<br>_cons | 7.526<br>(0.529) *** [ 0.534] *** | | 4.233<br>(0.403) *** [0.556] *** | |
| First Stage Regression | | Student Teacher Ratio | | Student Teacher Ratio |
| Average Number of teacher per school with secondary qualification | | -5.716<br>(1.608) *** [ 8.463] | | -4.242<br>(1.501) ** [7.677] |
| Average Teacher Experience per School in the Community | | 1.031<br>(0.114) *** [0.489] * | | 1.035<br>(0.115) *** [0.521] * |
| Constant | | 32.00<br>(0.513) *** [ 2.294] *** | | 31.66<br>(0.474) *** [1.988] *** |
| lnsig_2<br>_cons | | 2.527 ***<br>(0.0176) *** [ 0.122] *** | | 2.527<br>(0.0177) *** [0.123] *** |
| atanhrho_12<br>_cons | | 0.349<br>(0.236)[ 0.560] | | 0.833<br>(0.189) *** [0.603] |
| cut_1_1<br>_cons | | 5.600<br>(1.183) *** [ 2.657] * | | 0.953<br>(0.681)[2.165] |
| cut_1_2<br>_cons | | 6.217<br>(1.229) *** [ 2.768] * | | 1.327<br>(0.725) + [2.315] |
| Pseudo R-Square | 0.441[0.441] | | 0.283[0.283] | |
| Wald chi2 (9) | 546.0[653.0] | 647.3[544.2] | 601.3[314.3] | 1265.2[237.8] |
| p | 8.14e-112[8.51e-135] | 1.46e-133[1.99e-111] | 1.10e-123[2.46e-62] | 1.02e-266[3.71e-46] |
| N | 1606[93] | 1606[93] | 1606[93] | 1606[93] |

Standard errors in parentheses, cluster standard errors in square brackets (93 clusters in village or community)
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

Table 4.11: Ordered Probit and Instrumental Variable (IV) Ordered Probit Marginal Effects of Parental Evaluation of Child Reading and Mathematics Ability (Whole Sample (age 11-15))

| | (1) Reading Ability | (2) IV. Reading Ability | (3) Mathematics Ability | (4) IV. Mathematics Ability |
|---|---|---|---|---|
| Child Sex (female) | -0.0257 (0.0206)[ 0.0246] | -0.0254 (0.0206)[ 0.0243] | -0.0616 (0.0255)$^{*}$ [ 0.0296]$^{*}$ | -0.0471 (0.0204)$^{*}$ [ 0.0278]$^{+}$ |
| Child Age | 0.0357 (0.00782)$^{***}$ [ 0.00911]$^{***}$ | 0.0349 (0.00793)$^{***}$ [ 0.00970]$^{***}$ | 0.0357 (0.00963)$^{***}$ [ 0.0113]$^{**}$ | 0.0267 (0.00787)$^{***}$ [ 0.0134]$^{*}$ |
| Mother's Level of Education | 0.0822 (0.0297)$^{**}$ [ 0.0296]$^{**}$ | 0.0809 (0.0300)$^{**}$ [ 0.0297]$^{**}$ | 0.105 (0.0376)$^{**}$ [ 0.0357]$^{**}$ | 0.0787 (0.0334)$^{*}$ [ 0.0384]$^{*}$ |
| Father's Level of Education | 0.0210 (0.00951)$^{*}$ [ 0.0107]$^{*}$ | 0.0210 (0.00929)$^{*}$ [ 0.0107]$^{*}$ | 0.0148 (0.0118)[ 0.0139] | 0.0120 (0.00951)[ 0.0118] |
| Child's level of Education | 0.490 (0.0348)$^{***}$ [ 0.0416]$^{***}$ | 0.484 (0.0303)$^{***}$ [ 0.0508]$^{***}$ | 0.426 (0.0277)$^{***}$ [ 0.0291]$^{***}$ | 0.328 (0.0430)$^{***}$ [ 0.126]$^{**}$ |
| Currently in school (yes) | 0.168 (0.0259)$^{***}$ [ 0.0347]$^{***}$ | 0.163 (0.0250)$^{***}$ [ 0.0367]$^{***}$ | 0.104 (0.0324)$^{**}$ [ 0.0440]$^{*}$ | 0.0750 (0.0269)$^{**}$ [ 0.0447]$^{+}$ |
| Number of schools in a Community or village | 0.0169 (0.00362)$^{***}$ [ 0.00507]$^{***}$ | 0.0170 (0.00367)$^{***}$ [ 0.00503]$^{***}$ | 0.0124 (0.00426)$^{**}$ [ 0.00883] | 0.00922 (0.00371)$^{*}$ [ 0.00825] |
| Student Teacher Ratio | -0.000873 (0.000815)[ 0.00120] | -0.00769 (0.00476)[ 0.0117] | -0.00532 (0.00101)$^{***}$ [ 0.00206]$^{**}$ | -0.0238 (0.00287)$^{***}$ [ 0.0109]$^{*}$ |
| Proportion of Schools with a Library | 0.332 (0.0554)$^{***}$ [ 0.0556]$^{***}$ | 0.303 (0.0590)$^{***}$ [ 0.0760]$^{***}$ | 0.532 (0.0702)$^{***}$ [ 0.109]$^{***}$ | 0.328 (0.0779)$^{***}$ [ 0.208] |
| *N* | 1606 | 1606 | 1606 | 1606 |

Marginal effects; Standard errors in parentheses, cluster standard errors in square brackets
(d) for discrete change of dummy variable from 0 to 1
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$
Note: 93 clusters in village or community

Table 4.12: Ordered Probit and Instrumental Variable (IV) Ordered Probit Estimates of Parental Evaluation of Child Reading and Mathematics Ability (Boys only Sample (age 11-15))

| | (1) Reading Ability | (2) IV. Reading Ability | (3) Mathematics Ability | (4) IV. Mathematics Ability |
|---|---|---|---|---|
| Child Age | 0.144 (0.0397)*** [0.0400]*** | 0.125 (0.0376)*** [0.0455]** | 0.105 (0.0357)** [0.0357]** | 0.0797** (0.0293)** [0.0397]* |
| Mother's Level of Education | 0.269 (0.174)[0.182] | 0.238 (0.156)[0.168] | 0.431 (0.176)* [0.179]* | 0.330 (0.158)* [0.185]+ |
| Father's Level of Education | 0.111 (0.0572)+ [0.0597]+ | 0.0927 (0.0486)+ [0.0576] | 0.0413 (0.0475)[0.0545] | 0.0269 (0.0399)[0.0439] |
| Child's level of Education | 1.853 (0.149)*** [0.145]*** | 1.626 (0.215)*** [0.346]*** | 1.269 (0.104)*** [0.102]*** | 0.985 (0.153)*** [0.335]** |
| Currently in school (yes) | 0.627 (0.124)*** [0.144]*** | 0.550 (0.126)*** [0.176]** | 0.252 (0.115)* [0.125]* | 0.194 (0.0948)* [0.120] |
| Number of schools in a Community or village | 0.0642 (0.0182)*** [0.0210]** | 0.0571 (0.0171)*** [0.0225]* | 0.0389 (0.0159)* [0.0228]+ | 0.0293 (0.0137)* [0.0226] |
| Student Teacher Ratio | -0.00476 (0.00398)[0.00423] | -0.0415 (0.0135)** [0.0274] | -0.0160 (0.00354)*** [0.00561]** | -0.0615 (0.00866)*** [0.0264]* |
| Proportion of Schools with a Library | 1.311 (0.283)*** [0.312]*** | 0.977 (0.307)** [0.421]* | 1.570 (0.265)*** [0.331]*** | 0.967 (0.290)*** [0.544]+ |
| cut1 | | | | |
| _cons | 6.673 (0.666)*** [0.666]*** | | 4.123 (0.565)*** [0.611]*** | |
| cut2 | | | | |
| _cons | 7.382 (0.680)*** [0.670]*** | | 4.702 (0.572)*** [0.617]*** | |
| First Stage Regression | | Student Teacher Ratio | | Student Teacher Ratio |
| Average Number of teacher per school with secondary qualification | | -6.223 (2.198)** [8.834] | | -4.629* (2.159)* [8.615] |
| Average Teacher Experience per School in the Community | | 1.256 (0.163)*** [0.571]* | | 1.253 (0.166)*** [0.608]* |
| Constant | | 31.79 (0.680)*** [2.189]*** | | 31.44 (0.651)*** [2.028]*** |
| lnsig_2 | | | | |
| _cons | | 2.516 (0.0242)*** [0.128]*** | | 2.517 (0.0242)*** [0.129]*** |
| atanhrho_12 | | | | |
| _cons | | 0.535 (0.230)* [0.425] | | 0.763 (0.215)*** [0.525] |
| cut_1_1 | | | | |
| _cons | | 4.552 (1.261)*** [2.243]* | | 1.463 (0.879)+ [2.152] |
| cut_1_2 | | | | |
| _cons | | 5.175 (1.325)*** [2.363]* | | 1.913 (0.934)* [2.296] |
| Pseudo R-Square | 0.399[0.399] | | 0.284[0.284] | |
| Wald chi2 (8) | 295.7[311.2] | 442.3[286.0] | 301.7[272.3] | 640.8[220.0] |
| p | 3.41e-59[1.68e-62] | 1.63e-90[3.92e-57] | 1.75e-60[3.15e-54] | 3.91e-133[3.83e-43] |
| N | 854[93] | 854[93] | 854[93] | 854[93] |

Standard errors in parentheses, cluster standard errors in square brackets (93 clusters in village or community)
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

Table 4.13: Ordered Probit and Instrumental Variable (IV) Ordered Probit Marginal Effects of Parental Evaluation of Child Reading and Mathematics Ability (Boys only Sample (age 11-15))

| | (1) Reading Ability | (2) IV. Reading Ability | (3) Mathematics Ability | (4) IV. Mathematics Ability |
|---|---|---|---|---|
| Child Age | 0.0502 (0.0140)*** [ 0.0142]*** | 0.0448 (0.0132)*** [ 0.0156]** | 0.0416 (0.0142)** [ 0.0142]** | 0.0316 (0.0116)** [ 0.0158]* |
| Mother's Level of Education | 0.0938 (0.0612)[ 0.0637] | 0.0853 (0.0560)[ 0.0597] | 0.171 (0.0700)* [ 0.0710]* | 0.131 (0.0626)* [ 0.0734]+ |
| Father's Level of Education | 0.0386 (0.0199)+ [ 0.0206]+ | 0.0332 (0.0173)+ [ 0.0202]+ | 0.0163 (0.0188)[ 0.0215] | 0.0107 (0.0158)[ 0.0174] |
| Child's level of Education | 0.646 (0.0564)*** [ 0.0610]*** | 0.583 (0.0671)*** [ 0.112]*** | 0.503 (0.0417)*** [ 0.0408]*** | 0.391 (0.0605)*** [ 0.133]** |
| Currently in school (yes) | 0.212 (0.0398)*** [ 0.0469]*** | 0.192 (0.0397)*** [ 0.0545]*** | 0.0994 (0.0448)* [ 0.0491]* | 0.0768 (0.0373)* [ 0.0471] |
| Number of schools in a Community or village | 0.0224 (0.00631)*** [ 0.00747]** | 0.0205 (0.00596)*** [ 0.00790]** | 0.0154 (0.00630)* [ 0.00907]+ | 0.0116 (0.00543)* [ 0.00898] |
| Student Teacher Ratio | -0.00166 (0.00139)[ 0.00148] | -0.0149 (0.00517)** [ 0.0103] | -0.00632 (0.00141)*** [ 0.00223]** | -0.0244 (0.00346)*** [ 0.0104]* |
| Proportion of Schools with a Library | 0.457 (0.100)*** [ 0.108]*** | 0.350 (0.106)*** [ 0.143]* | 0.622 (0.105)*** [ 0.130]*** | 0.383 (0.115)*** [ 0.216]+ |
| *N* | 854 | 854 | 854 | 854 |

Marginal effects; Standard errors in parentheses, cluster standard errors in square brackets
(d) for discrete change of dummy variable from 0 to 1
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$
Note: 93 clusters in village or community

Table 4.14: Ordered Probit and Instrumental Variable (IV) Ordered Probit Estimates of Parental Evaluation of Child Reading and Mathematics Ability (Girls only Sample (age 11-15))

| | (1) Reading Ability | (2) IV. Reading Ability | (3) Mathematics Ability | (4) IV. Mathematics Ability |
|---|---|---|---|---|
| Child Age | 0.135 (0.0484) ** [0.0557] * | 0.131 (0.0498) ** [0.0588] * | 0.0915 (0.0389) * [0.0445] * | 0.0635 (0.0312) * [0.0463] |
| Mother's Level of Education | 0.354 (0.148) * [0.149] * | 0.341 (0.156) * [0.161] * | 0.211 (0.120) + [0.128] + | 0.145 (0.108)[0.115] |
| Father's Level of Education | 0.0520 (0.0504)[0.0543] | 0.0471 (0.0506)[0.0551] | 0.0414 (0.0432)[0.0482] | 0.0359 (0.0330)[0.0398] |
| Child's level of Education | 2.066 (0.229) *** [0.217] *** | 1.973 (0.325) *** [0.429] *** | 1.036 (0.115) *** [0.117] *** | 0.744 (0.193) *** [0.412] + |
| Currently in school (yes) | 0.649 (0.160) *** [0.174] *** | 0.635 (0.158) *** [0.195] ** | 0.354 (0.143) * [0.163] * | 0.226 (0.119) + [0.181] |
| Number of schools in a Community or village | 0.0719 (0.0236) ** [0.0340] * | 0.0677 (0.0257) ** [0.0387] + | 0.0278 (0.0174)[0.0338] | 0.0192 (0.0145)[0.0280] |
| Student Teacher Ratio | -0.00122 (0.00516)[0.00819] | 0.0213 (0.0316)[0.0485] | -0.0131 (0.00437) ** [0.00757] + | -0.0635 (0.0112) *** [0.0310] * |
| Proportion of Schools with a Library | 1.267 (0.325) *** [0.390] ** | 1.265 (0.324) *** [0.405] ** | 1.303*** (0.275) *** [0.447] ** | 0.772 (0.314) * [0.716] |
| cut1 _cons | 7.312 (0.865) *** [0.956] *** | | 3.557 (0.556) *** [ 0.721] *** | |
| cut2 _cons | 7.881 (0.874) *** [0.966] *** | | 3.984 (0.562) *** [0.734] *** | |
| First Stage Regression | | Student Teacher Ratio | | Student Teacher Ratio |
| Average Number of teacher per school with secondary qualification | | -5.742 (2.170) ** [6.997] | | -4.102 (2.083) * [6.901] |
| Average Teacher Experience per School in the Community | | 0.812 (0.163) *** [0.436] + | | 0.836 (0.160) *** [0.448] + |
| Constant | | 32.32 (0.762) *** [2.310] *** | | 31.88 (0.695) *** [2.033] *** |
| lnsig_2 _cons | | 2.536 (0.0258) *** [0.117] *** | | 2.537 (0.0258) *** [0.117] *** |
| atanhrho_12 _cons | | -0.304 (0.448)[0.685] | | 0.884 (0.336) ** [0.791] |
| cut_1_1 _cons | | 7.761 (0.803) *** [0.982] *** | | 0.693 (1.103)[2.577] |
| cut_1_2 _cons | | 8.305 (0.800) *** [ 0.975] *** | | 0.996 (1.171)[2.741] |
| Pseudo R-Square | 0.467[0.467] | | 0.251[0.251] | |
| Wald chi2 (8) | 239.1[248.2] | 225.8[227.5] | 278.8[147.8] | 571.6[147.8] |
| p | 3.55e-47[4.12e-49] | 2.32e-44[1.00e-44] | 1.36e-55[5.72e-28] | 3.01e-118[5.72e-28] |
| N | 752[93] | 752[93] | 752[93] | 752[93] |

Standard errors in parentheses, cluster standard errors in square brackets (93 clusters in village or community)
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

Table 4.15: Ordered Probit and Instrumental Variable (IV) Ordered Probit Marginal Effects of Parental Evaluation of Child Reading and Mathematics Ability (Girls only Sample (age 11-15)

| | (1) Reading Ability | (2) IV. Reading Ability | (3) Mathematics Ability | (4) IV. Mathematics Ability |
|---|---|---|---|---|
| Child Age | 0.0187** (0.00669) ** [ 0.00767] * | 0.0199* (0.00802) * [ 0.00928] * | 0.0281 (0.0119) * [ 0.0135] * | 0.0221 (0.0102) * [ 0.0141] |
| Mother's Level of Education | 0.0490* (0.0221) * [ 0.0222] * | 0.0519* (0.0251) * [ 0.0244] * | 0.0646 (0.0370) + [ 0.0397] | 0.0503 (0.0366)[ 0.0365] |
| Father's Level of Education | 0.00721 (0.00704)[ 0.00773] | 0.00715 (0.00761)[ 0.00806] | 0.0127 (0.0132)[ 0.0147] | 0.0125 (0.0113)[ 0.0130] |
| Child's level of Education | 0.286 (0.0347) *** [ 0.0368] *** | 0.300 (0.0499) *** [ 0.0693] *** | 0.318 (0.0356) *** [ 0.0326] *** | 0.259 (0.0534) *** [ 0.110] * |
| Currently in school (yes) | 0.104 (0.0323) ** [ 0.0366] ** | 0.111 (0.0325) *** [ 0.0416] ** | 0.112 (0.0468) * [ 0.0544] * | 0.0797 (0.0402) * [ 0.0577] |
| Number of schools in a Community or village | 0.00997 (0.00332) ** [ 0.00463] * | 0.0103 (0.00382) ** [ 0.00493] * | 0.00852 (0.00533)[ 0.0105] | 0.00669 (0.00492)[ 0.00942] |
| Student Teacher Ratio | -0.000169 (0.000716)[ 0.00113] | 0.00323 (0.00563)[ 0.00861] | -0.00402 (0.00133) ** [ 0.00228] + | -0.0221 (0.00525) *** [ 0.0135] + |
| Proportion of Schools with a Library | 0.176 (0.0502) *** [ 0.0571] ** | 0.192 (0.0638) ** [ 0.0779] * | 0.400 (0.0848) *** [ 0.134] ** | 0.268 (0.0969) ** [ 0.217] |
| *N* | 752 | 752 | 752 | 752 |

Marginal effects; Standard errors in parentheses, cluster standard errors in square brackets
(d) for discrete change of dummy variable from 0 to 1
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$
Note: 93 clusters in village or community

## Appendix A: **Tabulated Descriptive**

Table A1: Child Reading Ability by Age and Current School Attendance (age: 11-15)

Currently in School? and Is able to read a simple letter or the newspaper with understanding

---------- No ----------

| Age (Years) | Not at all | Yes, with difficulty | Yes, Easily | Total |
|---|---|---|---|---|
| 11 | 18.51 | 11.11 | 2.17 | 16.3 |
| 12 | 18.93 | 7.41 | 8.7 | 17.11 |
| 13 | 22.72 | 27.78 | 20.65 | 22.82 |
| 14 | 16.13 | 20.37 | 25 | 17.35 |
| 15 | 23.7 | 33.33 | 43.48 | 26.43 |
| Total | 100 | 100 | 100 | 100 |

Currently in School? and Is able to read a simple letter or the newspaper with understanding

---------- Yes ----------

| Age (Years) | Not at all | Yes, with difficulty | Yes, Easily | Total |
|---|---|---|---|---|
| 11 | 40.11 | 36.5 | 19.39 | 27.71 |
| 12 | 21.93 | 22.63 | 17.73 | 19.68 |
| 13 | 20.86 | 23.36 | 23.4 | 22.76 |
| 14 | 10.16 | 10.22 | 20.09 | 15.8 |
| 15 | 6.95 | 7.3 | 19.39 | 14.06 |
| Total | 100 | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table A2: Child Reading Ability by Age and Gender

Sex and Is able to read a simple letter or the newspaper with understanding

---------- Male ----------

| Age (Years) | Not at all | Yes, with difficulty | Yes, Easily | Total |
|---|---|---|---|---|
| 11 | 20.72 | 28 | 13.61 | 18.97 |
| 12 | 18.93 | 19.2 | 15.68 | 17.68 |
| 13 | 18.93 | 24 | 24.26 | 21.78 |
| 14 | 18.67 | 12 | 21.89 | 18.97 |
| 15 | 22.76 | 16.8 | 24.56 | 22.6 |
| Total | 100 | 100 | 100 | 100 |

Sex and Is able to read a simple letter or the newspaper with understanding

---------- Female ----------

| Age (Years) | Not at all | Yes, with difficulty | Yes, Easily | Total |
|---|---|---|---|---|
| 11 | 24.75 | 31.82 | 21.47 | 24.6 |
| 12 | 20.04 | 16.67 | 16.95 | 19.02 |
| 13 | 24.95 | 25.76 | 20.34 | 23.94 |
| 14 | 11.98 | 15.15 | 19.21 | 13.96 |
| 15 | 18.27 | 10.61 | 22.03 | 18.48 |
| Total | 100 | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table A3: Child Maths Ability by Age and Current School Attendance

| Currently in School? and Is able to do simple addition and subtraction? | | | | |
|---|---|---|---|---|
| ---------- No ---------- | | | | |
| Age (Years) | Not at all | Yes, with difficulty | Yes, easily | Total |
| 11 | 18.85 | 12.64 | 8.64 | 16.3 |
| 12 | 18.36 | 16.09 | 12.96 | 17.11 |
| 13 | 23.93 | 21.84 | 19.14 | 22.82 |
| 14 | 16.07 | 18.39 | 21.6 | 17.35 |
| 15 | 22.79 | 31.03 | 37.65 | 26.43 |
| Total | 100 | 100 | 100 | 100 |
| Currently in School? and Is able to do simple addition and subtraction? | | | | |
| ---------- Yes ---------- | | | | |
| Age (Years) | Not at all | Yes, with difficulty | Yes, easily | Total |
| 11 | 38.04 | 38.98 | 20.45 | 27.71 |
| 12 | 21.74 | 22.03 | 18.2 | 19.68 |
| 13 | 21.2 | 23.73 | 23.15 | 22.76 |
| 14 | 12.5 | 5.93 | 19.78 | 15.8 |
| 15 | 6.52 | 9.32 | 18.43 | 14.06 |
| Total | 100 | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table A4: Child Maths Ability by Age and Gender

| Sex and Is able to do simple addition and subtraction? | | | | |
|---|---|---|---|---|
| ---------- Male ---------- | | | | |
| Age (Years) | Not at all | Yes, with difficulty | Yes, easily | Total |
| 11 | 21.24 | 27.2 | 14.36 | 18.97 |
| 12 | 17.7 | 21.6 | 16.41 | 17.68 |
| 13 | 19.76 | 21.6 | 23.59 | 21.78 |
| 14 | 20.06 | 10.4 | 20.77 | 18.97 |
| 15 | 21.24 | 19.2 | 24.87 | 22.6 |
| Total | 100 | 100 | 100 | 100 |
| Sex and Is able to do simple addition and subtraction? | | | | |
| ---------- Female ---------- | | | | |
| Age (Years) | Not at all | Yes, with difficulty | Yes, easily | Total |
| 11 | 24.84 | 28.75 | 22.58 | 24.6 |
| 12 | 20.22 | 16.25 | 17.51 | 19.02 |
| 13 | 25.93 | 25 | 19.35 | 23.94 |
| 14 | 11.65 | 12.5 | 19.35 | 13.96 |
| 15 | 17.36 | 17.5 | 21.2 | 18.48 |
| Total | 100 | 100 | 100 | 100 |

Source: Calculated from PRHS-II

**Appendix B: cmp Code for Instrumental Variable (IV) Ordered Probit**

cmp setup

cmp (Reading Ability = Child Sex (female) Child Age Mother's Level of Education Father's Level of Education Child's level of Education Currently in school (yes) Number of schools in a Community or village Student Teacher Ratio Proportion of Schools with a Library) (Student Teacher Ratio = Average Number of teacher per school with secondary qualification Average Teacher Experience per School in the Community), ind( $cmp_oprobit $cmp_cont) tech(dfp) nolr vce(cluster vcode)

# Chapter 5 Child Education and School Quality in Rural Pakistan

## 5.1 Introduction

In the preceding chapter we have shown that, in effect, parents believe that our measures of school quality raise their children's cognitive achievement. Using these quality measures, the present chapter considers a policy of increasing school quality to raise school attendance, and reduce child labour, in the context of a formal model. A higher quality boosts family income in the next generation at the cost of the present generation that must pay for the schooling and forgo child earnings. When quality is sufficiently low the enhancement of future family income will be inadequate to compensate for the present loss, so children will not be sent to school. This chapter shows that, by simulating the model estimated on PRHS-II, school quality enhancement is very highly valued by the average rural Pakistani family, and by those below the official poverty line. A corollary is that quality enhancement will be an effective policy for boosting school attendance and that subsistence poverty is not a major reason for children failing to be sent to school.

Identification in the school attendance model is established first by utilising a distinctive feature of the Pakistani environment; the arbitrary and random allocation of resources to state schools economy and society. The quality variable library is confirmed exogenous with the only instrument in the data set suitable on grounds of relevance (correlation with library), community population. If school resources were adjusted to school need, school attendance would not influence the staff-student ratio. But instrumenting this ratio changes its sign in the predicted fashion. Since we have either controlled for endogeneity or established the exogeneity of the quality measures in the attendance equation, we can be confident that the parameter estimates correctly capture the impact of school quality variables. The possibility that other variables, not instrumented in the school attendance equation, are endogenous does not bias these estimates. Nor does including a range of extra community characteristics in the school attendance equation affect the school quality parameter values estimated.

The chapter is organised as follows; the next section, 5.2, describes the model and section 5.3 explains the empirical equations on which the tests and simulation are based. Section 5.4 presents the results while a concluding section 5.5 considers the implications of the findings.

## 5.2 The Model

The theoretical framework for estimating the impact of school quality on school attendance is that of an intergenerational family optimising over two periods. In the first period parents earn income. They also decide how to allocate their children's time between schooling (human capital accumulation), *(S)* (which boosts second period consumption) and child labour, $E_c$ (which raises first period consumption)[32]. For simplicity, schooling is the only form of intergenerational savings[33] in the model. In the second period, the first period children become adult income earners. Children can be expected to support their surviving parents in old age. These children enhance family consumption in period 2 and therefore raise family wellbeing. On the other hand because they are costly they might lower family consumption in period 1, in part depending on whether they go to school or work.

If parents choose child labour in the first period, the child earns '$w_c$' wage in the first period and unskilled adult wage '$w_u$' in the second period. If parents choose schooling in the first period, the child earns zero wages and a wage rate in the second period determined by the quality of human capital accumulated in the first period. This quality is measured by the increase in the wage rate, by a skill premium ψ, dependent on school quality. Parental wealth or income is assumed to be '$w_p$' at the time they decide on child work or school participation or leisure.

Children cost the parents even if they do not incur schooling expenses. This per child cost is $H_0$, while $H_1$ is per child schooling investment in human capital (including the price of, or expenses incurred on, school related inputs/essentials such as school fees, books, tuition, meals, transportation, or clothes).

Where n is the number of children in the first period household, $C_1$ is first period consumption and $C_2$ second period, household consumption in the two periods is defined as follows:

[32] In the formulation of children as income earning assets we follow a literature that includes Emerson and Knabb (2006), Baland and Robinson (2000), Basu and Van (1998), Dasgupta (1995) and Weiner (1991). Ignoring leisure apart from a constant for sleep is a realistic assumption for children in rural Pakistan.

[33] Saving/capital markets in the model can be allowed for other than child schooling investment. However their inclusion is avoided due to two reasons. First, positive saving in the first period would ensure positive schooling in the household without any further insights. Second, saving might be chosen simultaneously with child schooling. Choosing the efficient level of human capital investment would require the household to equate rate of return on human capital investment and market interest rate on positive assets. This would complicate our exposition without gaining any further insights.

$$C_1 = w_p + n.w_c(1-S) - n.H_0 - n.H_1 S$$

$$C_2 = n.w_u(1+\psi.S) \qquad (5.1)$$

Subsistence consumption is captured by a Stone-Geary utility function representing the household's preferences;

$$U = (C_1 - \alpha)^{\gamma_1}(C_2 - \beta)^{\gamma_2} \qquad (5.2)$$

where $\gamma_1 + \gamma_2 = 1$. $\gamma_1$ and $\gamma_2$ are the parameters of the intertemporal elasticity of substitution for consumption over the subsistence level, and also marginal budget shares. If $\gamma_1$ is greater than $\gamma_2$, the family assigns more importance to period 1 consumption than to a similar volume of consumption in period 2. Either they have a positive rate of time discount, or are less than perfectly altruistic towards the next generation, or perhaps both. $\alpha$, $\beta$ are the exogenous level of subsistence for $C_1$ and $C_2$. When $C_1 - \alpha$ or $C_2 - \beta < 0$ their contribution to U is undefined; the household below subsistence does not survive by definition. Subsistence levels vary with the number of adult persons in the household. In the first period child subsistence is covered by the child cost term $H_0$. In the second period where n is the number of children in period 1, $\beta = \beta_0 + n.\beta_1$.

The representative household then maximises utility from consumption first by choosing 'S' given n (see appendix A for calculus). The schooling optimum is[34];

$$S^* = -\frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u\psi} + \frac{\gamma_2(w_p + n(w_c - H_0) - \alpha)}{n.(w_c + H_1)} \qquad .....(5.3)$$

Improved school quality (greater $\psi$), and therefore higher second period earnings, increases schooling (assuming second period consumption without education is greater than the minimum subsistence consumption $(((\beta_0+\beta_1 n)/n.w_u) < 1)$ (5.4). The numerator of the above inequality is second period minimum consumption and the denominator is second period family income and consumption without education.

[34]Non-negative schooling requires that the second term of (5.3) must be at least as large as the second in absolute value. It should be noted that there are two interpretations of S –either the chances of school attendance or the allocation of time to school attendance.

The positive sign of 5.4 holds whether first period parental earnings alone are greater or less than subsistence. In the latter case, however, child earnings must be sufficient to keep the household above subsistence (to ensure non-negative schooling and survival).

$$\frac{\partial S^*}{\partial \psi} = \frac{\gamma_1}{\psi^2}\{1 - \frac{(\beta_0 + \beta_1 n)}{n.w_u}\} \qquad ...(5.4)$$

From (5.3) it is apparent that a higher children's allowance[35], by reducing $H_0$ costs for the household, increases schooling regardless of school quality. Hence when school quality is low this policy, along with allowances directly linked to school attendance, is not invariably conducive to greater human capital accumulation. Also inspection of (5.3) shows that greater impatience or reduced altruism reduce schooling and increase child labour ($\gamma_1$ increases relative to $\gamma_2$). The positive effect of the children's allowances on schooling increases with altruism or a lower time discount factor whereas the positive effect of school quality increases with $\gamma_1$.

Turning to the response of schooling to child wages (5.5), if the household is below subsistence with schooling costs, $nH_1 + nH_0 + \alpha > wp$, then the expression is positive; higher child wages increase schooling for these households. This is essentially the equivalent of Bhalotra's (2007) negative response of child hours worked to higher child wages. When parental wages are higher than subsistence and schooling costs, then increases in child wages decrease schooling.

$$\frac{\partial S^*}{\partial w_c} = \frac{\gamma_2(-w_p + nH_1 + nH_0 + \alpha)}{n.(w_c + H_1)^2} \qquad ...(5.5)$$

Higher parental earnings ($w_p$) increase schooling (5.6):

$$\frac{\partial S^*}{\partial w_p} = \frac{\gamma_2}{n.(w_c + H_1)} \qquad ....(5.6)$$

Schooling becomes less likely the more children there are in the household (5.7):

$$\frac{\partial S^*}{\partial n} = -\frac{\gamma_1 \beta_0}{w_u \psi . n^2} - \frac{\gamma_2 (w_p - \alpha)}{(w_c + H_1).n^2} \qquad ...(5.7)$$

[35] Children's allowance can be easily allowed in the first period budget constraint or in the denominator of the second ratio in 5.3.

If the household is above subsistence in the first period, the second term is negative. The first term of the partial derivative is also negative. This is consistent with Burney and Irfan's (1991 Table 5) empirical result for Pakistani rural children.

Assuming the representative household has optimised

$U = U(S, \psi, w_p, w_c, w_u, b)$,

the utility equivalent increase in parental earnings or income ($dw_p$) for a given improvement in school quality ($d\psi$) can be calculated:

$dw_p = -(\partial w_p/\partial S)\, dS - (\partial w_p/\partial \psi)\, d\psi$

If empirical estimates can be made that approximate these partial derivatives, then it will be possible to get an idea of the income equivalent gains from a given improvement in school quality.

### 5.3 Empirical Specification Variable Selection and Estimation

The model as derived above yields an estimating equation for child schooling as a function of school quality, household size and other variables.

The econometric specification of modelling households' inter-temporal consumption decision to send the child to school or work can be explicitly obtained from equation (5.3). We used binary probit model to evaluate the likelihood of child schooling or work decisions conditional on different community level school quality, child and household level characteristics.

Let $y_i^*$ be the unobserved latent variable representing the chances of child schooling in household '$i$'. The utility difference between child schooling or work decision depends upon the price of child human capital accumulation (school quality), and different child and household level characteristics. Thus, the utility difference of household '$i$' from child schooling could be written as a function of observed characteristics '$X_i^{'}$' and unobserved characteristics '$e_i$'.

$$y_i^* = X_i^{'} b + e_i \qquad (5.8)$$

Where '$b$' is a vector of parameters and '$e_i$' represents the unobserved disturbance term having zero mean and unit variance. The observed choices $y_i$ are based on comparison of

attitude towards child school participation $y_i^*$ relative to certain unknown threshold value which has been normalized to zero in the present model. However the zero threshold is of no consequence once intercept is accounted for in the vector ' $X_i^{'}$ '. Child schooling is undertaken ( $y=1$ ) if the utility difference exceeds some threshold level ( $y_i^* > 0$ ) and zero otherwise ( $y_i^* \leq 0$ ). Hence we have the probability of child schooling as:

$$\Pr(y_i = 1/X) = \Pr(y_i^* > 0/X) = \Pr\left[e_i > -(X_i^{'}b)/X\right]$$

$$\Pr(y_i = 1/X) = 1 - F\left[-X_i^{'}b\right] = F(X_i^{'}b)$$

Linear probability models are also widely used for the estimation of response probability in the literature. But the two important drawbacks with linear probability models are that the fitted probabilities can be less than zero or greater than one and the partial effect of any explanatory variable is constant. Similarly the variance of the error term is heteroskedastic in the case of linear probability model. These problems are usually overcome by using the most sophisticated binary response models; probit or logit model.

The magnitude of each ' $b$ ', in the model, is not much use as compared to the coefficients of the linear probability model. The probit model tries to estimate the effect of each explanatory variable on the probability of success but viewed as complicated due to the nonlinearity of ' $F(.)$ '. Therefore calculus is called for to calculate the partial effect of a continuous variable on the response probability. In case of continuous variable the partial derivative is calculated as:

$$\frac{\partial \Pr(y)}{\partial x_j} = \phi(X^{'}b)b_j \qquad (5.9)$$

There are two common ways to estimate the partial effect of a continuous variable on the response probability; Average Marginal Effects (AMEs) and Marginal Effects at sample Mean (MEMs). Both, AMEs and MEMs are estimated[36]. Greene (2003) shows that the estimates from both AMEs and MEMs would be the same in large samples but may not be in

[36] For computing marginal effects, one usually evaluates equation (5.9) at the sample mean of the data (Marginal Effect at Sample Means, MEMs). Alternatively the marginal effects may be evaluated at every observation and then these individual marginal effects are averaged across the sample (Average Marginal Effects, AMEs).

moderate and small samples. Current practice favours AMEs, when possible because econometric software packages were not supporting or able to calculate AME until recently.

But Bartus (2005) shows that MEMs and AMEs can be different, even in considerably large samples. MEMs might either underestimate or overestimate AMEs, depending largely on the sign of the second derivative of the density function. Therefore AMEs has also been estimated along with MEMs in the study. Nevertheless due to similar results only AMEs are discussed and reported in the proceeding sections.

Estimation is undertaken by using the method of maximum likelihood estimation. The likelihood function is given by;

$$f(y_i = 1/ X_i; b) = \left[F(X_i^{'}b\right]^{y}\left[1 - F(X_i^{'}b\right]^{1-y}$$

**Instrumental Variable (IV) Probit Model Specification**

As noted in chapter 4, assumption 1 is important to get unbiased (consistent) estimates of '$b$' that is the model error is uncorrelated with the regressors i.e., $E(e_i \setminus X_i^{'}) = 0$ . If this assumption fails the probit estimator is inconsistent, and the marginal effects no longer can be given causal interpretation. This is a fundamental problem as marginal effects are key to economic policy (Cameron and Trivedi 2009). In such circumstances it is important to use the instrumental variable approach which provides consistent estimator under the assumption that valid instruments are exists. The instruments (Z) are variables that are correlated with (X) and must satisfy; $E(e_i \setminus Z) = 0$, that the instruments are uncorrelated with the regression error term.

We suspect, on *a priori* grounds, that student-teacher ratio and fertility are endogenous in our model (5.8). We postulate a latent-variable model with dependant variable $y_1^*$ (school attendance) in the structural equation. The endogenous variables are $y_2$ (student-teacher ratio) and $y_3$ (fertility). These three endogenous variables are modelled in exogenous variables $Z_i^{'} = \left[X_{1i}^{'} \quad X_{2i}^{'} \quad X_{3i}^{'}\right]$. Where, $X_1$ serves as instruments for itself and $X_2$ is the instrument for $y_2$ (student-teacher ratio) and $X_3$ is the instrument for $y_3$ (fertility), and the instrument must satisfy the moment condition that;

$$E(e_i \setminus Z_i^{'}) = 0 \qquad (5.10)$$

All the three equations are specified as;

$$y_{1i}^{*} = b_1 y_{2i} + b_2 y_{3i} + X_{1i}^{'}\eta + e_i \qquad (5.11)$$

$$y_{2i} = X_{1i}^{'}\pi_1 + X_{2i}^{'}\pi_2 + X_{3i}^{'}\pi_3 + v_i \qquad (5.12)$$

$$y_{3i} = X_{1i}^{'}\theta_1 + X_{2i}^{'}\theta_2 + X_{3i}^{'}\theta_3 + \omega_i \qquad (5.13)$$

Where $i = 1,.....N$; $X_1$ is a vector ($K_1 \times 1$) of exogenous variables; $X_2$ is a $L_1 \times 1$ and $X_3$ is a $L_2 \times 1$ vector of additional instrumental variables (IV) that affect $y_2$ and $y_3$ but excluded from (5.11) as these IVs do not affect $y_1$ directly. To ensure identification it is required that $L_1 + L_2 \geq 2$. The dependent variable $y_1^{*}$ is latent and not directly observed. However, the binary outcome $y_1$ is observed as presented above.

Equation (5.11) is the structural equation of principal interest. Equation (5.12) and (5.13) are first-stage or reduced-form equations and serve as a source of identification. The first stage equations also show the strength of the instruments and goodness of fit of the reduced form equation. The reduced form equations explain the variation in the endogenous variable in terms of exogenous variables, including IV $X_2$ and $X_3$, that are excluded from the structural equation. The instruments are essential for identifying the parameters of the structural equation. The estimation procedure depends on the specification of the structural and reduced form equations and can be simultaneous (joint) or sequential.

In the present case, the IV probit model, the structural model approach completely specifies the distributions of $y_1^{*}$, $y_2$ and $y_3$ in (5.11), (5.12) and (5.13). It is assumed that; $e_i$ , $v_i$ and $\omega_i$ are jointly normally distributed. The assumptions imply that $e_i \setminus v_i, \omega_i = \rho_1 v_i + \rho_2 \omega_i + u_i$, where, $E(u_i \setminus v_i, \omega_i) = 0$. A test of the null hypothesis of the exogeneity of both $y_2$ and $y_3$ is to test that $H_0 : \rho_j = 0$, as then $e_i$, $v_i$ and $\omega_i$ are independent. Consistent estimation requires both the normality and homoskedasticity of $e_i$, $v_i$ and $\omega_i$. The default procedure for IV probit model delivers ML estimates and estimation is simultaneous (joint). An alternative procedure is sequential for the estimation of (5.11), (5.12) and (5.13) and is called the two-step sequential estimator. It is similar to the ML estimator; the difference is only computational as both assume the same distributional assumption (Cameron and Trivedi 2009). But computation of marginal effects is not allowed after the two step procedure and only available for ML estimator. The test of over-identifying restrictions is only available after the two-step estimation framework and has been estimated here.

As mentioned the instruments must satisfy condition (5.10). This condition cannot be tested in the just identified case, but a test is possible in the over-identified case. When there is more than one instrumental variable, we can test whether some of them are correlated with the structural error. For example if we have four instrumental variables for two endogenious variables, we have two over-identifying restriction. However, even in the over-identified case instrument validity depends on economic theory, persuasive argument or norms established in the previous related empirical work (Greene 2003; Cameron and Trivedi 2009; Wooldridge 2009). Also the instruments must be relevant in that after controlling for the remaining exogenous variables $X_1$, the instruments $X_2$ and $X_3$ must explain significant variation in $y_2$ and $y_3$, respectively. Generally the stronger the association between the instruments and endogenous regressors the better will be the identification of the model.

This chapter uses the ML estimation framework to estimate IV probit model to control for the endogeneity of student-teacher ratio and fertility in child school attendance equation while using four instruments; average number of teachers per school with secondary qualification and average teacher experience per school in the community for student-teacher ratio, and proportion of males of aged over 60 in the household and number of females aged 16-59 in the household for fertility. We have also accounted for cluster robust estimates of the standard errors along with default estimates of the standard errors for reasons reported in chapter 4. For the identification of other school quality variable 'proportion of schools with a library' we have exploited the distinctive characteristic of Pakistani 'random allocation of educational resources' in the communities in the same spirit as shown by Alderman et al. (1996) and Behrman et al. (2008). The random allocation of educational inputs to Pakistani state schools has been widely acknowledged e.g. a detailed discussion on the education budget of Pakistan has been presented by HRCP and CEF (2005) and discussed in chapter 2 (section 2.4).

**Variable Selection**

All school quality related variables are community- level averages/expected values of school quality inputs used by households in the community. This specification is chosen on the grounds first, that it is the only way the influence of quality on school non-attendance can be taken into account, and second that the averages do not reflect the characteristics of the specific school a child attends and hence for most quality variables are exogenous to the school attendance decision (Alderman et al. 2001). The characteristics of the specific school a

child attends are determined jointly with the schooling decision, hence are endogenous to schooling (Deaton 1988; Alderman et al. 2001; Ersado 2005)[37].

Nonetheless quality might be correlated with community level observables or unobservables; one community might be politically and economically more active or better endowed and school quality might be a response to these characteristics (Rosenzweig and Wolpin 1986; Behrman and Birdsall 1988; Pitt et al. 1993; Gershberg and Schuermann 2001). We therefore test for such a relation following Alderman et al. (1996) and Behrman et al. (2008). We assess whether, for sample villages, school availability and school quality indicators, depend on average village-level household income, measures of political influence, population and other village-level characteristics (Appendix B). Apart from population (negatively!), they do not.

While investigating the effect of schools availability (quantity) on child enrolment in 21 least developing countries, Filmer (2007, p.902) argues that "purposive placement of schools might bias the results towards a small number, but neither the results presented here nor existing studies suggest that this is the main reason". For instance if Pakistani school supply is directed at needs in the local communities, then parameter estimates on number of schools in the community in a school attendance equation will be lower than the true value. However, as (appendix B) shows, school availability does not depend on village level characteristics. Our focus is school quality, not school quantity, but our school quality variable, proportion of schools with a library does correlate with population in the local communities. We instrumented 'proportion of schools with a library' in a separate regression with only one valid instrument - population. The results show that school quality variable 'proportion of schools with a library' is exogenous to child schooling as expected (the Wald test for evaluating the null hypothesis that library is exogenous to child schooling is accepted (as p-value > 0.05) in equation 2 (Table 5.1).

[37] Specific household level or school level characteristics (which are particular to the school in which the child actually educated) most often suffer from endogeneity. For example, household child schooling expenses are only incurred for those children for whom no child labour decision are made. These variables are endogenous to child labour decisions. To avoid this problem household level school related variables are averaged either over relevant geographic units or by using community level school quality variables whenever possible (Ersado 2005).

Table 5.1: Exogeneity/Endogeneity Test of Library (Proportion) in the Child Schooling Equation: Probit and Instrumental Variable (IV) Probit Estimates.

| | Model 1 | |
|---|---|---|
| | (1) Schooling | (2) IV. Schooling |
| **Proportion of Schools with a Library** | 0.764*** | 2.177 |
| | (6.02) | (1.15) |
| Controls for child, household level characteristics are included in the estimation | | |
| athrho | | |
| _cons | | -0.248 |
| | | (-0.71) |
| lnsigma | | |
| _cons | | -1.780*** |
| | | (-158.18) |
| Pseudo R-Square | 0.154 | |
| Wald chi2 (18) | 662.6 | 749.9 |
| p | 4.96e-129 | 1.45e-147 |
| N | 3950 | 3950 |
| Wald Test of Exogeneity (/athrho = 0) | | |
| Chi2(1) | | 0.50 |
| P | | 0.4783 |

*z* statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Equation 2: Instrumented: Proportion of schools having Library.
Instruments: Variables List + Population in the Community

The student-teacher ratio is endogenous as discussed in chapter 4 (Angrist and Lavy 1999; Case and Deaton 1999; Hoxby 2000; Urquiola 2006; Das et al. 2011)[38]. Also, large households with many children may be more inclined to supply child labour and less willing to send children to school. Children from bigger households are more likely to work, as resources per individual are smaller in bigger households (Patrinos and Psacharopoulos 1997). But such a relation does not continue once household income per capita is controlled for (Bhalotra 2003). Alternatively a correlation between household size and child schooling or labour would exist if fertility is encouraged by child work prospects; hence desired household size is endogenous (Rosenzweig and Evenson 1977; Cain 1977; Singh and Schuh 1986; Eswaran 1998; Bardhan and Udry 1999). Notwithstanding, household size is usually treated as exogenous in empirical studies (Ray 2001; Rosati and Rossi 2003; Ersado, 2005; Bhalotra 2007; Hou 2010).

Again in developing countries households expand upwards (grandparents) and horizontally (uncles, aunts) and not just downwards (children). For this reason, household size is not a good indicator of fertility. A simple count of the number of children in the household –as we

[38] Among these studies, Hoxby, did not find any effect of class size on test scores.

use here–is a better indicator of fertility (Bhalotra 2003). We therefore use instrumental variables to test for the endogeneity of household size/ fertility along with student-teacher ratio in the IV probit child schooling equations to identify their impact on school attendance.

Equation (5.3) in the theoretical section contains some variables for which data is unavailable or does not vary in cross-sectional data, such as child wages ($w_c$). In addition there are other controls that should be included in empirical models intended to isolate the impact of school quality. These include determinants identified in child schooling and labour empirical work such as child sex (gender), child age, household head age and sex, parental educational level and ownership of productive assets, such as land and animals. The effect of age on child labour has been reported from many empirical studies (Alderman et al. 2001; Ray 2001; Rosati and Rossi 2003; Ersado, 2005; Hou 2010) for different countries as positive and quadratic. Child labour increases with age because child labour productivity is increasing in age but at a decreasing rate. Child labour also may become more socially acceptable or be viewed as less harmful as the child age advances (Bhalotra 2003).

A gender differential in child schooling and work participation has also been widely noted, in Pakistan (Bhalotra and Heady 2003; Hou 2010) as well as other developing countries (Ray 2000a, 2000b; Rosati and Rossi 2003; Ersado, 2005), on account of the different return to education or remittance propensities among boys and girls. If parents perceive boys as a source of old age support for them, then there would not only be a differential enrolment and employment rates among boys and girls but also the type of activity in which they participate would be different. Girls usually help their mothers with child care and housework. A gender dummy, as well as separate equations for each, is used to account for these possibilities. Bhalotra (2003) reported that the gender dummy from different Asian countries is significant, which indicates girls' greater work participation for reasons that are not captured by the other covariates in the model.

A measure for "domestic responsibilities" has also been incorporated, reflecting the number of very young children in the household (age less than five), which could be detrimental to the schooling, particularly of girls. Female headed households are represented by a dummy (sex of the household head) to account for any vulnerability[39] of the female headed

[39] The significance of female headship is culturally-specific. It depends upon the attitude towards widowhood, sex-biases in property rights, and women status among other things. It is argued that female headed households have weak economic prospects because of lower average education of women, limited labour market access (market segmentation), wage differentials (or discrimination) and sharper credit constraints. This has a negative impact on child schooling.

households and any possible heterogeneity in the preferences of women and men. Parental level of education of both mothers and fathers is included in child schooling equation. Educated mothers have a greater say in household decision making and may shift household resource allocation towards child schooling (Behrman et al. 1997; Galasso 2000; Bhalotra 2007). Educated mothers also may be more altruistic to their children (Strauss and Thomas 1995). Different studies find a significant effect of parental education on child schooling and are consistent with better educated parents valuing child quality more highly (Strauss and Thomas 1995; Ray 2000a, 2000b, 2001; Bhalotra 2003; Rosati and Rossi 2003; Ersado, 2005; Hou 2010).

$w_{p,}$ parental income or wealth in the model, is measured by household level income (also known as 'household income net of child earnings') as calculated by Bhalotra (2007)[40]. This excludes estimated income from child employment, to measure household poverty more accurately and to avoid endogeneity with child labour, but includes non-labour earnings and savings income. Expenditure is usually preferred over income as it is less volatile and reflects permanent income (Behrman et al. 1997; Bhalotra 2003). The variable has been made consistent over households using the OECD (1982) Equivalence Scale (also called Oxford scale[41], see Jenkins and Cowell 1994).

Bhalotra and Heady (2003) reported that the children of households having more land are less likely to be in school than work. Transaction costs for child labour (family and hired labour are not perfect substitutes) mean that a household with more land has an incentive to employ more of its children (Alderman et al 1996). Farm animals typically supply milk and milk products for their owners' consumption and also are a source of income from sale of the surplus. Non-farm animals usually provide power for agricultural operations[42]. The two categories of animals have different implications for child activity; child schooling is

---

Alternatively, female headship signifies a higher role of women in household decision making, which may increase/decrease child schooling/child labour if it prevails that women are more altruistic than men towards children. Patrinos and Psacharapoulos (1995) reported from Paraguay that children of female-headed households are less likely to be in school and more likely to be in work.

[40] See appendix C for detail on relevant calculation.

[41] Equivalence scales correct for the fact that the needs of each household grow with each additional member but not proportionally. Needs for space, electricity and so on are not twice as high for a household with two members as for a single individual. With equivalence scales each household member is assigned a value in proportion to their needs. The OECD equivalence scale or the Oxford scale assigns a value of 1 to the first member of the household and of 0.7, 0.5 to each additional adult and child, respectively, in the household. Specifically the formula is ; $H = 1 + \alpha(n_A - 1) + \beta n_C$, where $\alpha = 0.7$ and $\beta = 0.5$ and, $n_A$ and $n_c$ are the number of adults and children in the household.

[42] Farm animals include the cow, buffalo, goat and sheep, whereas non-farm animals include the horse, donkey, camel and bullock.

decreasing in farm animals and increasing in non-farm animals (Hou 2010). Hence, livestock ownership and possession of land[43] variables are included as possible explanatory variables for schooling.

Credit market imperfections supposedly are an important determinant of child schooling and labour decisions in developing countries (Ranjan 1999; Baland and Robinson 2000). Baland and Robinson (2000) demonstrate that when a household is credit constrained, its current period marginal utility of consumption is high and thus child labour is oversupplied. They do so by constructing an intergenerational model of two periods that relies on credit market imperfection to generate conditions under which inefficiently high level of child labour may arise.

In the present model there is no intergenerational borrowing or lending apart from that arising from investment in human capital. Nonetheless from year to year, rather than between generations, income smoothing may matter (Jacoby 1994; Jacoby and Skoufias 1997; Maitra 2001; Fuwa et al. 2009). But measurement of credit access is problematic (Ersado 2005). A contrarian view is that of Menon (2010) who argues that the widely held belief that improved access to credit has a positive impact on child school participation among poor households is not true in the case of Pakistan. Here credit constrained and unconstrained households are distinguished in a similar spirit to that of Fuwa et al. (2009) (Appendix D). But the necessary information is only available for agricultural households. So distance from a commercial bank at the community level is used as proxy for access to the formal credit market for the whole sample. A provincial (regional) dummy is included to capture spatial variation in labour demand, prices and productivity, as well as in culture.

To summarize, the estimating equation for the probability of child schooling (S) decision by the *i*th household ($S_{ci}$) is;

$\Pr(S_{ci}=1) = F$ (Child sex, child age, own farm animals, own non-farm animals, household expenditure, household head sex, household head age, mother's level of education, father's level of education, [fertility], number of children less than five years, household size, land owned, land owned squared, land ownership dummy, number of schools in the community, [student teacher ratio], proportion of schools having a playground, proportion of schools having library, credit constraint, province). (5.14)

Variables in square brackets are potentially endogenous.

In testing for the endogeneity of the student-teacher ratio in the above equation (5.14) we utilise distinctive features of the Pakistani education system. As a first approximation we can

[43] Schooling and child work may be related to parental occupation and land tenure.

assume that new teachers to a school have no experience. Teachers with a long time at school will have considerable teaching experience. Whether new teachers are appointed or not is dependent on political and other contingent features, not on school requirements. When teachers are appointed average school experience falls and so does student-teacher ratio. When they are not appointed the student-teacher ratio will be higher because a high birth rate generally ensures there are always more children coming to start school. Hence average teacher experience is correlated with new teacher appointment and the student-teacher ratio. Given newer teachers are more likely to possess qualifications than old, it follows that the probability of qualification per teacher in community schools also is a measure of new teacher appointment and student-teacher ratio.

It might be thought that qualifications and teacher experience improved Pakistani school quality and therefore school attendance. In this case the exclusion restriction for the student-teacher ratio instrumental variable would not be satisfied. We therefore tested whether parental assessment of their children's cognitive achievement varied significantly with the teacher qualification and experience and found that they did not (Table 4.9 chapter 4).

Testing whether fertility in the above equation is endogenous we used as instrumental variables the 'proportion of males of aged over 60 in the household' and 'number of females aged 16-59 in the household'. It seems plausible that the age composition of males and females in a household would affect numbers of children present in the household – particularly for females who must stop producing children earlier than males. With older heads of households former children will have grown up. There is no reason to suppose that ages of adults in the household should influence school attendance. As reported below we find the two instrumental variables are indeed correlated with fertility but insignificant in the school attendance equation.

We use a different approach (different instruments) from the two in the literature, to estimate the effect of fertility on schooling. Schultz (1997) argues there is no obvious choice of exclusion restriction to identify exogenous variation in fertility in the demand equation for child schooling. He reviews the two major approaches used in the literature to identify the impact of fertility on child school attendance. One approach was used by Rosenzweig and Wolpin (1980a, 1980b) and the other approach by Rosenzweig and Schultz (1985, 1987, and 1989) and Schultz (1990). More recently Angrist et al. (2010) empirically evaluated child

quality and quantity trade-off[44] in a quasi-experimental setup. These studies depend on distinctive characteristics of the available data for their instruments.

Rosenzweig and Wolpin (1980a) used twins on first birth to identify the effect of fertility on schooling. They argue that twin birth (within a sample of mothers) is uncorrelated with other determinants of child demand, for example, wages/prices, preferences and endowments, but twins are correlated with children ever born as some couples either want few births or want only one child and are unable to avoid perfectly subsequent births. It is argued that twins are a valid instrument to predict fertility/children ever born to a couple and enable them to estimate the effect of exogenous variation in fertility on other forms of household behaviour without simultaneous equation bias. Rosenzweig and Wolpin (1980a) used twins to identify the effect of fertility on child schooling in the family in a large Indian rural survey. The same technique was utilized by Rosenzweig and Wolpin (1980b) in the United States to analyse how unanticipated fertility shocks affect subsequent labour force participation of mothers.

Angrist et al. (2010) report empirical evidence while utilizing quasi-experimental variation due to twins on second births, household preferences for a sex composition that preferences for boys at higher order births and ethnic differences in the effects of twin births and sex composition across Israeli population. Their results show no impact of fertility on child school attendance.

The second strategy to identify the effect of fertility on schooling is used by Rosenzweig and Schultz (1985, 1987, and 1989) and Schultz (1990). They used information on unpredictable failures of specific contraceptive practices which result in unplanned births, which are then used to explain behavioural choices: fertility, schooling and contraceptive choices. The results from all these studies, except Angrist et al. (2010), show that fertility and schooling are substitutes and that an unanticipated increase in fertility reduces child schooling.

### 5.4 Child Schooling Results and Discussion

This section presents probit and instrumental variable probit estimation of the child schooling decisions model discussed earlier (equation 5.14)[45]. The focus is on the relationship between child schooling and school quality variables while controlling for other important correlates.

[44] The theoretical literature postulates that there is a negative trade-off between family size and children's schooling. It is proposed that fewer children frees up resources which are then invested in the education of the remaining offspring (Becker and Lewis 1973; Becker and Tomes 1976).
[45] Kharif season only

The probit analysis is undertaken both for the whole sample and for agricultural households only, because a special credit module is available only for the second group.

Specifically, referring to the main tables reporting results of child schooling, Table 5.2 provides probit and instrumental variable probit estimates for the whole and agricultural only samples. Their respective AMEs are reported in Table 5.3. We also undertook the same analysis in Table 5.4 and 5.5; separately for children above and below the national poverty line to test the model's implication that school quality has a different effect for both sets of children. Simulations of household expenditure or income equivalent of school quality are presented in Table 5.6. The schooling equation is also estimated separately for boys and girls. Table 5.7 reports probit and instrumental variable probit estimates for boys and girls. Estimates of AMEs for boys and girls are provided in Table 5.8. Furthermore the detailed description on estimated equations on each is presented below.

### 5.4.1 Child Schooling (Whole and Agriculture Sample)

Probit and IV Probit estimates are reported for both full and restricted (agricultural only) samples (Table 5.2). AMEs of parameters are reported for both samples in Table 5.3. Estimates of cluster robust standard errors are reported in square brackets for the same equations. We tested for zero restrictions individually for household head sex, household head age and number of children less than five years in both child-level and couple-level estimation. The results justified the exclusion of these variables from all reported equations.

The simple probit results reported for both full and restricted (agricultural only) samples show that the estimates of school quality variables were consistent over both specifications and are in accordance with the theoretical prediction of the model presented earlier, except student-teacher ratio. We instrumented the student-teacher ratio along with fertility in the schooling equation (Table 5.2) for whole and agriculture household sample. The Amemiya-Lee-Newey minimum chi-sq statistic test of over-identifying restrictions shows that our chosen instruments for the student-teacher ratio (average number of teachers per school with secondary qualification and average teacher experience per school in the community) and for fertility (proportion of males of age above 60 in the household and number of females aged 16-59 in the household) are valid (p-value > 0.05). The Wald test of exogeneity show that both variables 'student-teacher ratio and 'fertility' are endogenous to child school attendance as the null hypothesis 'Ho-both variables are exogenous' is rejected as $p < 0.05$ (equations 2 and 4). The results (equations 2 and 4 in Table 5.2) indicate that the previous positive

coefficient on student-teacher ratio in the simple probit reverses in sign in the IV equation and significantly negatively affects child school participation as expected. The parameter estimate of student-teacher ratio, for the whole sample, is significant at 5 percent level of significance with cluster robust estimate of the standard error (equation 2). However, the same is not true in the case of agriculture only sample. However, the previous negative sign on the fertility estimate in the simple probit schooling equation remains unchanged and significant in the instrumental variable model.

The parameter estimate of other school quality variable 'proportion of schools with a library' is positive (as expected) and significant (equations 2 and 4). The statistical significance of library proportion is 0.1 percent with default estimate of the standard error. However, its statistical significance falls to 5 percent with cluster robust estimate of the standard error.

The measure of school accessibility or availability; number of schools in the community, significantly positively affect the likelihood of child school participation with default estimate of the standard error. However, its parameter estimate is insignificant with cluster robust estimate of the standard error (p-value > 0.05 or 0.10) (equations 2 and 4).

The parameter estimates of school quality correlates show that 'proportion of schools with a library' significantly positively influences the likelihood of child schooling participation in the child schooling equation (equations 3 and 4), while the variable proxying the school crowding effect (student-teacher ratio) significantly negatively affects child school participation (equation 3 only).

Table 5.3 equation 1 shows that being female decreases the chances of attending school by 12.8 percent in the sampled area. The impact of age in child schooling equation is positive but insignificant (when accounting for statistically significant non-linearity in age in the sample). The probability of child school participation significantly increases with household income. The relationship holds in both full and restricted (agricultural only) samples.

Maternal education substantially improves the probability of child school participation, although the estimated effect of parental education is insignificant with cluster robust estimate of the standard error. The marginal impact of mother's education and father's education on child school participation is 8.4 percent and 2.4 percent, respectively in the school attendance equation (whole sample IV equation 1 Table 5.3). Child school attendance significantly increases with plot size and the land ownership dummy shows that non-landless

households have 4.8 percent higher chance of child school participation than landless households (IV equation 1 Table 5.3).

The variable supposedly measuring access to credit (distance to the nearest bank) has a significant negative impact on child school participation with default estimate of the standard error but does not have the same relationship with cluster robust estimate of the standard error (equation 1, Table 5.3). The estimates of the credit constraint parameter in the agriculture household sample shows that being a credit constrained household does not affect child school attendance (equation 2, Table 5.3 or equation 4, Table 5.2). Hence, we conclude that there is no support for the assumption that credit constraint is the critical factor affecting child schooling.

Table 5.4 reports simple probit and Instrumental Variable (IV) probit results on child school attendance, separately for children whose household income is above or below the national poverty line. The Amemiya-Lee-Newey minimum chi-sq statistic test of over-identifying restrictions shows that the chosen instruments, as reported above, for the student-teacher ratio and for fertility are valid (p-value > 0.05). The Wald test of exogeneity show that both variables are endogenous to child school attendance as the null hypothesis is rejected at $p < 0.05$ (equations 2 and 4). Other child and household level controls in the estimated equations (3 and 4) for both set of children show that the parameter estimate of maternal education is significant only for children from richer households with cluster robust estimate of the standard error. The parameter estimate of father's education is insignificant, with cluster robust estimate of the standard error, for children of both set of households (equations 2 and 4). The probability of child school participation significantly increases with household income, for richer and poor children, after accounting for statistically significant non-linearity in household income in the sample (Table 5.5). The parameter estimate of land ownership dummy is insignificant (equations 2 and 4, Table 5.4). The estimated effect of variable measuring access to credit 'distance to the nearest bank' remains insignificant as in the case of whole and agriculture only sample analysis.

The estimates of school quality measure 'proportion of schools with a library' show that school quality significantly improves child school attendance in poorer households just as it does for the richer. This shows the model implications are relevant to poorer households who are not deterred from sending children to school by their low incomes. The parameter estimate of school quality measure 'student-teacher ratio' is insignificant with cluster robust estimate of the standard error. Most importantly school quality measure 'proportion of

schools with a library' has bigger marginal impacts on the probability of schooling of children from poor households than in the rich households (Table 5.5). These results support the claims that the reason for low school attendance in Pakistan is the lack of quality schools not poverty.

In summary the results support the model prediction that school quality is a statistically significant reason for child school attendance, while controlling for important covariates such as subsistence poverty[46], credit constraints and other household characteristics. School quality has the same relevance and significance for child school attendance in poor households as it has in the case of rich households. We now show how important school quality is to households in our sample.

[46] Another measure (Equivalent Household Expenditure) that includes child labour earnings was also estimated instead of 'household expenditure net of child earnings' while using the same specification to see the stability of the parameter estimates of the school quality variables. The school quality variables behave consistently while using the two different measures of household income. The results on specification involving 'Equivalent Household Expenditure' has not been reported as those are similar to the reported results.

**Table 5.2:** Probit and Instrumental Variable (IV) Probit Estimates of Child Schooling Decisions in Rural Pakistan (Kharif Season): With Household Income net of Child Earnings (Equivalent) (Whole and Agriculture Household Sample)

| | (1) Whole Sample Schooling | (2) Whole Sample IV. Schooling | (3) Agriculture Sample Schooling | (4) Agriculture Sample IV. Schooling |
|---|---|---|---|---|
| Child Sex (female) | -0.502<br>(0.0433) *** [ 0.0663] *** | -0.401<br>(0.0472) *** [ 0.104] *** | -0.507<br>(0.0549) *** [ 0.0760] *** | -0.389<br>(0.0661) *** [ 0.138] ** |
| Child Age | 0.676<br>(0.0501) *** [0.0448] *** | 0.554<br>(0.0578) *** [ 0.113] *** | 0.661<br>(0.0637) *** [ 0.0588] *** | 0.534<br>(0.0822) *** [ 0.154] *** |
| Child Age Squared | -0.0347<br>(0.00251) *** [ 0.00226] *** | -0.0286<br>(0.00292) *** [ 0.00574] *** | -0.0337<br>(0.00317) *** [ 0.00296] *** | -0.0277<br>(0.00413) *** [ 0.00781] *** |
| Own Farm Animals (Yes) | 0.0133<br>(0.0531)[ 0.0821] | -0.0239<br>(0.0492)[ 0.0871] | 0.0810<br>(0.0758)[ 0.105] | -0.0858<br>(0.0774)[ 0.136] |
| Own Non Farm Animals (Yes) | 0.0296<br>(0.0501)[ 0.0822] | -0.00147<br>(0.0466)[ 0.0898] | 0.0851<br>(0.0593)[ 0.0926] | -0.0232<br>(0.0602)[ 0.108] |
| Household Income net of Child Earnings | 2.624<br>(0.376) *** [ 0.494] *** | 3.125<br>(0.372) *** [ 0.547] *** | 2.442<br>(0.463) *** [ 0.643] *** | 2.693<br>(0.457) *** [ 0.660] *** |
| Household Income net of Child Earnings Squared | -2.079<br>(0.439) *** [ 0.480] *** | -2.683<br>(0.459) *** [ 0.570] *** | -1.926<br>(0.471) *** [ 0.565] *** | -2.435<br>(0.509) *** [ 0.571] *** |
| **Fertility** | -0.0515<br>(0.0114) *** [ 0.0154] *** | -0.0999<br>(0.0230) *** [ 0.0308] ** | -0.0526<br>(0.0140) *** [ 0.0183] ** | -0.116<br>(0.0308) *** [ 0.0430] ** |
| Mother's Level of Education | 0.338<br>(0.0612) *** [ 0.0756] *** | 0.263<br>(0.0520) *** [ 0.0785] *** | 0.298<br>(0.0734) *** [ 0.0859] *** | 0.214<br>(0.0664) **[ 0.0874] * |
| Father's Level of Education | 0.139<br>(0.0208) *** [ 0.0287] *** | 0.0737<br>(0.0225) **[ 0.0507] | 0.128<br>(0.0239) *** [ 0.0277] *** | 0.0577<br>(0.0293) *[ 0.0621] |
| Land owned in Kharif 2004 | 0.00923<br>(0.00342) ** [ 0.00555] + | 0.0160<br>(0.00328) *** [ 0.00478] *** | 0.00777<br>(0.00347) *[ 0.00553] | 0.0185<br>(0.00354) *** [ 0.00574] ** |
| Land owned in | -0.0000451 | -0.0000886 | -0.0000395 | -0.0000993 |

| | | | | |
|---|---|---|---|---|
| Kharif Squared 2004 | (0.0000211) * [ 0.0000359] | (0.0000267) *** [ 0.0000345] * | (0.0000213) + [ 0.0000353] | (0.0000273) *** [ 0.0000398] * |
| Land Ownership Kharif (Yes) | 0.165<br>(0.0522) ** [ 0.0788] * | 0.149<br>(0.0491) ** [ 0.0762] + | 0.282<br>(0.0696) *** [ 0.0995] ** | 0.251<br>(0.0677) *** [ 0.113] * |
| Number of schools in a Community or Village | 0.0546<br>(0.00756) *** [ 0.0186] ** | 0.0407<br>(0.00782) *** [ 0.0316] | 0.0600<br>(0.00956) *** [ 0.0218] ** | 0.0408<br>(0.0116) *** [ 0.0378] |
| **Student Teacher Ratio** | 0.00450<br>(0.00176) * [ 0.00369] | -0.0446<br>(0.00635) *** [ 0.0223] * | 0.00387<br>(0.00211) + [ 0.00419] | -0.0443<br>(0.00986) *** [ 0.0280] |
| Proportion of Schools with a Library | 0.764<br>(0.127) *** [ 0.309] * | 0.921<br>(0.118) *** [ 0.402] * | 0.868<br>(0.168) *** [ 0.392] * | 1.129<br>(0.157) *** [ 0.440] * |
| Distance to nearest bank | -0.00845<br>(0.00341) * [ 0.00763] | 0.0119<br>(0.00423) ** [ 0.0124] | | |
| Province (Punjab) | 0.0359<br>(0.0552)[ 0.126] | -0.354<br>(0.0676) *** [ 0.220] | 0.0268<br>(0.0722)[ 0.162] | -0.420<br>(0.0992) *** [ 0.249] + |
| Credit constraint (Yes) | | | -0.0803<br>(0.0865)[ 0.144] | 0.0288<br>(0.0813)[ 0.145] |
| Constant | -4.196<br>(0.273) *** [ 0.339] *** | -1.572<br>(0.480) ** [ 1.452] | -4.344<br>(0.345) *** [ 0.390] *** | -1.241<br>(0.786)[ 1.888] |
| s21<br>_cons | | 0.232<br>(0.0720) ** [ 0.0957] * | | 0.321<br>(0.0990) ** [ 0.151] * |
| s31<br>_cons | | 6.871<br>(0.844) *** [ 2.180] ** | | 7.327<br>(1.424) *** [ 3.194] * |
| s22<br>_cons | | 1.666<br>(0.0209) *** [ 0.0542] *** | | 1.713<br>(0.0298) *** [ 0.0649] *** |
| s32<br>_cons | | -0.495<br>(0.336)[ 0.685] | | -0.776<br>(0.490)[ 0.838] |
| s33<br>_cons | | 9.487<br>(0.614) *** [ 2.377] *** | | 9.787***<br>(1.081) *** [ 3.305] ** |

| | | | | |
|---|---|---|---|---|
| Pseudo R-Square | 0.154[0.154] | | 0.166[0.166] | |
| Wald chi2 (18) | 662.6[479.9] | 1121.4[564.0] | 442.5[381.5] | 787.2[449.5] |
| p | 4.96e-129[1.74e-90] | 7.57e-227[3.48e-108] | 1.23e-82[6.44e-70] | 1.67e-155[4.23e-84] |
| N | 3950[94] | 3950[94] | 2492[94] | 2492[94] |
| Wald Test of Exogeneity: | | | | |
| Chi2(2) | | 78.08[15.35] | | 42.94[13.06] |
| P | | 0.000[0.000] | | 0.000[0.001] |
| Test of over-identifying restrictions[(a)]: | | | | |
| Amemiya-Lee-Newey minimum chi-sq statistic | | | | |
| Chi-sq(2) | | 0.169 | | 0.947 |
| P | | 0.9189 | | 0.6228 |

Standard errors in parentheses, cluster standard errors in square brackets

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$

(a): Test of over-identifying restrictions based on two-step procedure of ivprobit estimation.

Note: 94 clusters in village or community

Instrumented: Student Teacher Ratio, Fertility

Instruments: Variables List + Average Number of teacher per school with secondary qualification + Average Teacher Experience per School in the Community + Proportion of Males of age Above 60 in the household + Number of Female Age 16-59 in the household

Note: $S_{ij}$ represent the correlation coefficient of error terms.

**Table 5.3:** Probit and Instrumental Variable (IV) Probit Average Marginal Effects of Child Schooling Decisions in Rural Pakistan (Kharif Season): With Household Income net of Child Earnings (Equivalent) (Whole and Agriculture Household Sample)

| | (1)<br>Whole Sample<br>IV. Schooling | (2)<br>Agriculture Sample<br>IV. Schooling |
|---|---|---|
| Child Sex (female) | -0.128<br>(0.0159) *** [ 0.0381] *** | -0.123<br>(0.0220) *** [ 0.0487] * |
| Child Age | -0.000319<br>(0.00191)[ 0.00246] | -0.00333<br>(0.00237)[ 0.00289] |
| Own Farm Animals (Yes) | -0.00759<br>(0.0156)[ 0.0275] | -0.0267<br>(0.0239)[ 0.0416] |
| Own Non Farm Animals (Yes) | -0.000467<br>(0.0148)[ 0.0286] | -0.00724<br>(0.0188)[ 0.0336] |
| Household Income net of Child Earnings | 0.733<br>(0.0834) *** [ 0.148] *** | 0.608<br>(0.105) *** [ 0.168] *** |
| **Fertility** | -0.0318<br>(0.00725) *** [ 0.0103] ** | -0.0364<br>(0.00962) *** [ 0.0143] * |
| Mother's Level of Education | 0.0838<br>(0.0168) *** [ 0.0268] ** | 0.0670<br>(0.0210) ** [ 0.0288] * |
| Father's Level of Education | 0.0234<br>(0.00727) ** [ 0.0170] | 0.0180<br>(0.00929) + [ 0.0202] |
| Land owned in Kharif 2004 | 0.00486<br>(0.000978) *** [ 0.00139] *** | 0.00540<br>(0.000987) *** [ 0.00158] *** |
| Land Ownership Kharif (Yes) | 0.0476<br>(0.0158) ** [ 0.0244] + | 0.0792<br>(0.0219) *** [ 0.0369] * |
| Number of schools in a Community or Village | 0.0129<br>(0.00252) *** [ 0.0103] | 0.0128<br>(0.00372) *** [ 0.0122] |
| **Student Teacher Ratio** | -0.0142<br>(0.00187) *** [ 0.00646] * | -0.0139<br>(0.00289) *** [ 0.00812] + |
| Proportion of Schools with a Library | 0.293<br>(0.0373) *** [ 0.124] * | 0.353<br>(0.0481) *** [ 0.133] ** |
| Distance to nearest bank | 0.00377<br>(0.00132) ** [ 0.00394] | |
| Province (Punjab) | -0.109<br>(0.0195) *** [ 0.0631] + | -0.126<br>(0.0272) *** [ 0.0686] + |
| Credit constraint (Yes) | | 0.00901<br>(0.0254)[ 0.0452] |
| *N* | 3950 | 2492 |

Marginal effects; Standard errors in parentheses, cluster standard errors in square brackets
(d) for discrete change of dummy variable from 0 to 1
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$
Note: 94 clusters in village or community

**Table 5.4:** Probit and Instrumental Variable (IV) Probit Estimates of Child Schooling Decisions in Rural Pakistan (Kharif Season): With Household Income net of Child Earnings (Equivalent) (Above and Below Poverty Line Sample)

| | Above Poverty Line: Sample | | Below Poverty Line: Sample | |
|---|---|---|---|---|
| | (1) Schooling | (2) IV. Schooling | (3) Schooling | (4) IV. Schooling |
| Child Sex (female) | -0.471 (0.0542)*** [ 0.0789]*** | -0.382 (0.0588)*** [ 0.127]** | -0.564 (0.0727)*** [ 0.101]*** | -0.497 (0.0819)*** [ 0.129]*** |
| Child Age | 0.694 (0.0616)*** [ 0.0537]*** | 0.562 (0.0741)*** [ 0.168]*** | 0.634 (0.0876)*** [ 0.0801]*** | 0.597 (0.0950)*** [ 0.106]*** |
| Child Age Squared | -0.0356 (0.00308)*** [ 0.00271]*** | -0.0291 (0.00374)*** [ 0.00853]*** | -0.0326 (0.00438)*** [ 0.00409]*** | -0.0307 (0.00480)*** [ 0.00540]*** |
| Own Farm Animals (Yes) | 0.0225 (0.0681)[ 0.106] | -0.0379 (0.0636)[ 0.114] | -0.0199 (0.0873)[ 0.121] | -0.0102 (0.0831)[ 0.120] |
| Own Non Farm Animals (Yes) | -0.0642 (0.0607)[ 0.0941] | -0.120 (0.0565)* [ 0.106] | 0.218* (0.0898)*[ 0.121]+ | 0.259 (0.0874)** [ 0.129]* |
| Household Income net of Child Earnings | 20.74 (5.273)*** [ 6.564]** | 28.99 (5.249)*** [ 6.599]*** | 15.89 (11.64)[ 16.59] | 16.18 (11.55)[ 16.86] |
| Household Income net of Child Earnings Squared | -161.5 (53.23)** [ 55.43]** | -251.8 (56.99)*** [ 63.87]*** | -79.72 (83.54)[ 119.2] | -73.68 (81.86)[ 120.7] |
| **Fertility** | -0.0472 (0.0143)*** [ 0.0196]* | -0.110 (0.0278)*** [ 0.0407]** | -0.0658 (0.0196)*** [ 0.0266]* | -0.109 (0.0450)*[ 0.0668] |
| Mother's Level of Education | 0.367 (0.0715)*** [ 0.0914]*** | 0.280 (0.0655)*** [ 0.112]* | 0.251 (0.129)+ [ 0.172] | 0.222 (0.107)* [ 0.181] |
| Father's Level of Education | 0.156 (0.0250)*** [ 0.0369]*** | 0.0838 (0.0280)**[ 0.0706] | 0.0681 (0.0400)+ [ 0.0469] | 0.0385 (0.0420)[ 0.0558] |
| Land owned in Kharif 2004 | 0.0118 (0.00420)**[ 0.00675]+ | 0.0190 (0.00399)*** [ 0.00605]** | 19.13 (9.904)+ [ 10.14]+ | 20.43*** (5.755)*** [ 8.512]* |
| Land owned in | -0.0000452 | -0.0000917 | -161.5 | -157.1 |

| | | | | |
|---|---|---|---|---|
| Kharif Squared 2004 | (0.0000235)$^{+}$ [ 0.0000403] | (0.0000299)$^{**}$ [ 0.0000360]$^{*}$ | (103.3)[ 103.9] | (48.12)$^{**}$ [ 86.60]$^{+}$ |
| Land Ownership Kharif (Yes) | 0.0793<br>(0.0644)[ 0.0877] | 0.0873<br>(0.0608)[ 0.0798] | 0.208<br>(0.115)$^{+}$ [ 0.145] | 0.173<br>(0.0978)$^{+}$ [ 0.152] |
| Number of schools in a Community or Village | 0.0554<br>(0.00933)$^{***}$ [ 0.0191]$^{**}$ | 0.0393<br>(0.0101)$^{***}$ [ 0.0352] | 0.0527<br>(0.0132)$^{***}$ [ 0.0229]$^{*}$ | 0.0494$^{***}$<br>(0.0129)$^{***}$ [ 0.0276]$^{+}$ |
| **Student Teacher Ratio** | 0.00621<br>(0.00215)$^{**}$ [ 0.00356]$^{+}$ | -0.0421<br>(0.00814)$^{***}$ [ 0.0302] | 0.000568<br>(0.00321)[ 0.00519] | -0.0360<br>(0.0140)$^{**}$ [ 0.0257] |
| Proportion of Schools with a Library | 0.712<br>(0.164)$^{***}$ [ 0.302]$^{*}$ | 0.833<br>(0.151)$^{***}$ [ 0.384]$^{*}$ | 0.928<br>(0.205)$^{***}$ [ 0.440]$^{*}$ | 1.110<br>(0.207)$^{***}$ [ 0.493]$^{*}$ |
| Distance to nearest bank | -0.00480<br>(0.00470)[ 0.00893] | 0.0107<br>(0.00516)$^{*}$ [ 0.0192] | -0.0123<br>(0.00537)$^{*}$ [ 0.00877] | 0.00649<br>(0.00912)[ 0.0155] |
| Province (Punjab) | 0.0817<br>(0.0703)[ 0.132] | -0.305<br>(0.0881)$^{***}$ [ 0.278] | -0.0128<br>(0.0933)[ 0.176] | -0.326<br>(0.130)$^{*}$ [ 0.243] |
| Constant | -4.345$^{***}$<br>(0.341)$^{***}$ [ 0.377]$^{***}$ | -1.595$^{*}$<br>(0.649)$^{*}$[ 2.167] | -4.157<br>(0.613)$^{***}$ [ 0.765]$^{***}$ | -2.598$^{**}$<br>(0.852)$^{**}$[ 1.330]$^{+}$ |
| s21<br>_cons | | 0.264<br>(0.0892)$^{**}$ [ 0.137]$^{+}$ | | 0.210<br>(0.130)[ 0.182] |
| s31<br>_cons | | 7.026<br>(1.117)$^{***}$ [ 3.364]$^{*}$ | | 4.636<br>(1.696)$^{**}$ [ 3.032] |
| s22<br>_cons | | 1.658<br>(0.0268)$^{***}$ [ 0.0761]$^{***}$ | | 1.640<br>(0.0345)$^{***}$ [ 0.0640]$^{***}$ |
| s32<br>_cons | | -0.707<br>(0.451)[ 1.140] | | -0.0107<br>(0.451)[ 0.814] |
| s33<br>_cons | | 9.658$^{***}$<br>(0.823)[ 3.280] | | 10.00<br>(0.796)$^{***}$ [ 1.753]$^{***}$ |
| Pseudo R-Square | 0.156[0.156] | | 0.132[0.132] | |
| Wald chi2 (18) | 423.8[299.4] | 695.6[403.5] | 216.6[166.1] | 284.3[180.7] |
| p | 1.00e-78[6.42e-53] | 4.84e-136[1.73e-74] | 4.58e-36[5.25e-26] | 8.19e-50[7.05e-29] |

| | | | | |
|---|---|---|---|---|
| N | 2533[94] | 2533[94] | 1417[88] | 1417[88] |
| Wald Test of Exogeneity: | | | | |
| Chi2(2) | | 45.64[5.85] | | 12.19[5.41] |
| P | | 0.0000[0.0536] | | 0.0023[0.0667] |
| Test of over-identifying restrictions[(a)]: | | | | |
| Amemiya-Lee-Newey minimum chi-sq statistic | | | | |
| Chi-sq(2) | | 5.070 | | 2.967 |
| P | | 0.0793 | | 0.2268 |

Standard errors in parentheses, cluster standard errors in square brackets

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$

(a): Test of over-identifying restrictions based on two-step procedure of ivprobit estimation.

Note: Household Income net of Child Earnings in equation 1 and 2 divided by 100,000,0 for convergence. In equation 3 and 4 it is divided by 100,000.
Land owned in Kharif 2004 in equation 3 and 4 divided by 500 for convergence.

Note: 94 clusters in village or community in eq. (1) and (2), whereas, 88 clusters in eq. (3) and (4)

Instrumented: Student Teacher Ratio, Fertility

Instruments: Variables List + Average Number of teacher per school with secondary qualification + Average Teacher Experience per School in the Community + Proportion of Males of age Above 60 in the household + Number of Female Age 16-59 in the household

**Table 5.5:** Probit and Instrumental Variable (IV) Probit Average Marginal Effects of Child Schooling Decisions in Rural Pakistan (Kharif Season): With Household Income net of Child Earnings (Equivalent) (Above and Below Poverty Line Sample)

| | Above Poverty Line: Sample (1) IV. Schooling | Below Poverty Line: Sample (2) IV. Schooling |
|---|---|---|
| Child Sex (female) | -0.121<br>(0.0196) *** [ 0.0470] ** | -0.162<br>(0.0274) *** [ 0.0440] *** |
| Child Age | -0.00120<br>(0.00242)[ 0.00300] | 0.00217<br>(0.00317)[ 0.00333] |
| Own Farm Animals (Yes) | -0.0119<br>(0.0200)[ 0.0355] | -0.00328<br>(0.0268)[ 0.0388] |
| Own Non Farm Animals (Yes) | -0.0378<br>(0.0178) * [ 0.0337] | 0.0848<br>(0.0288) ** [ 0.0432] * |
| Household Income net of Child Earnings | 5.974<br>(1.059) *** [ 1.507] *** | 1.702<br>(0.610) ** [ 0.874] + |
| **Fertility** | -0.0346<br>(0.00862) *** [ 0.0127] ** | -0.0353<br>(0.0144) * [ 0.0217] |
| Mother's Level of Education | 0.0883<br>(0.0211) *** [ 0.0389] * | 0.0717<br>(0.0344) * [ 0.0588] |
| Father's Level of Education | 0.0264<br>(0.00901) ** [ 0.0236] | 0.0124<br>(0.0136)[ 0.0182] |
| Land owned in Kharif 2004 | 0.00578<br>(0.00118) *** [ 0.00176] ** | 5.858<br>(1.651) *** [ 2.342] * |
| Land Ownership Kharif (Yes) | 0.0276<br>(0.0192)[ 0.0251] | 0.0564<br>(0.0322) + [ 0.0502] |
| Number of schools in a Community or Village | 0.0124<br>(0.00325) *** [ 0.0115] | 0.0159<br>(0.00414) *** [ 0.00881] + |
| **Student Teacher Ratio** | -0.0133<br>(0.00238) *** [ 0.00873] | -0.0116<br>(0.00437) ** [ 0.00802] |
| Proportion of Schools with a Library | 0.263<br>(0.0473) *** [ 0.119] * | 0.358<br>(0.0648) *** [ 0.155] * |
| Distance to nearest bank | 0.00339<br>(0.00160) * [ 0.00605] | 0.00209<br>(0.00292)[ 0.00497] |
| Province (Punjab) | -0.0935<br>(0.0253) *** [ 0.0793] | -0.103<br>(0.0388) ** [ 0.0728] |
| *N* | 2533 | 1417 |

Marginal effects; Standard errors in parentheses, cluster standard errors in square brackets
(d) for discrete change of dummy variable from 0 to 1
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$
Note: 94 clusters in village or community in eq. (1), whereas, 88 clusters in eq. (2)

### 5.4.2 Simulation of Household Expenditure/Income Equivalent of School Quality

We simulate household income/expenditure and school quality to calculate the household income/expenditure equivalent of school quality variations. The simulation exercise (Table 5.6), based on the whole sample Instrumental Variable child schooling equation (2) in Table (5.2), shows that a 10 percent improvement in school quality is equivalent to one half of the per capita household expenditure.

Expression 1 in Table (5.6) gives the implied probability (34 percent) of child school attendance for an average household with all correlates (including school quality variables) at the sample means. Expression 2 gives the implied probability (38 percent) of child school participation for the same average household when household income/expenditure per capita is raised by 10 percent (Rs.1598.95 increase). The 10 percent increase in income raises the chances of child school attendance by 4 percent for the same average household. Expression 3 in the same Table 5.6 gives the implied probability (51 percent) for a 10 percent improvement in the two school quality variables (a 2 percent increase in library proportion and a 3 units decrease in student-teacher ratio). The 10 percent improvement raises the chances of child school attendance by 17 percent for the same average household. In expression 4 household income is raised to the level that increases the chances of child school attendance by about the same level as achieved with 10 percent improvement in school quality in expression 3. A 50 percent increase in household expenditure is required. This is equivalent to Rs.8010.49 (half) of the per capita household expenditure/income in the sample. Hence an average household is willing to pay 50 percent/forego one half of the per capita household income for an improvement of 10 percent in school quality. The simulation exercise therefore shows that school quality is very highly valued by this sample.

Table 5.6: Simulation of School Attendance, Household Income and School Quality (Willingness to Pay for School Quality)
(based on Instrumental Variable child schooling equation 2 in Table 5.2 above, whole sample)

| | | | |
|---|---|---|---|
| Adjusted | predictions | | Observations = 3950 |
| Model VCE | : OIM | | |
| Expression | : Pr(Schooling), predict() | | |
| | Other Correlates | = | Sample Mean |
| 1._at | Household Income net of Child Earnings | = | 0.1598951 (Mean) |
| | Library = | | 0.2070 (Mean) |
| | Student Teacher Ratio = | | 33.73 (Mean) |
| | Probability of School Attendance = | | 0.343 |
| 2._at | Household Income net of Child Earnings | = | 0.1758846 (10% increase) |
| | Library | = | 0.2070 (Mean) |
| | Student Teacher Ratio | = | 33.73 (Mean) |
| | Probability of School Attendance = | | 0.381 |
| 3._at | Household Income net of Child Earnings | = | 0.1598951 (Mean) |
| | Library | = | 0.2270 (10% increase) |
| | Student Teacher Ratio | = | 30.357 (10% decrease) |
| | Probability of School Attendance = | | 0.512 |
| 4._at | Household Income net of Child Earnings | = | 0.2400000 (50% increase) |
| | Library | = | 0.2070 (Mean) |
| | Student Teacher Ratio | = | 33.73 (Mean) |
| | Probability of School Attendance = | | 0.510 |

Note: Household Income Net of Child Earnings per capita is in 100,000 Rs. 50 % increase equivalent to Rs. 8010.49

### 5.4.3 Boys' and Girls' Schooling (Whole Sample)

The whole sample schooling equation is here estimated separately, for girls and boys to calculate the magnitude of likely differences on the vector of coefficients for each control accounted in the model. Probit and IV Probit are reported in Table 5.7. AMEs of Probit and IV Probit parameters are presented in Table 5.8. It has been widely reported from Pakistan (Bhalotra and Heady 2003; Hou 2010), as well as, other countries (Ray 2000a, 2000b; Rosati and Rossi 2003; Ersado, 2005) that there are significant gender differential in child schooling and work participation on account of the different return to education or remittance propensities among boys and girls. Parents usually perceive boys as a source of old age support for them (also called old age support hypothesis), whereas, girls usually help their mothers with child care and household work. When there is a differential treatment of each

gender then there would not only be a differential enrolment and employment rates among boys and girls but also the type of activity in which they participate would be different.

The simple probit schooling results for both boys and girls (equations 1 and 3) show that the estimates of school quality variables are in accordance with the theoretical prediction of the model presented earlier, except the student-teacher ratio. The student-teacher ratio, as well as fertility is instrumented in both boys and girls schooling equations (2 and 4). The Amemiya-Lee-Newey minimum chi-sq statistic test of over-identifying restrictions shows that our chosen instruments for the student-teacher ratio (average number of teachers per school with secondary qualification and average teacher experience per school in the community) and for fertility (proportion of males of age above 60 in the household and number of females aged 16-59 in the household) are valid (p-value > 0.05). The results (equation 2 Table 5.7) indicate that the previous positive coefficient on student- teacher ratio in the simple probit reverses in sign in the IV equation and significantly negatively affect boys and girls school participation as expected. However, the parameter estimate of student-teacher ratio is significantly different from zero with cluster robust estimate of the standard error only in the girls' school participation equation (equation 4). The result is interpreted as that parents may require a higher return on the education of girls relative to boys. As girls are often involved in child care and other household work activities, their education is highly discounted when the quality of local schools are lower than some threshold level. A greater school quality premium is needed to outweigh the present period loss due to lower household consumption if girls are sent to school. The parameter estimate of variable 'proportion of schools with a library' is significant at 10 percent level of significance, with cluster robust estimate of the standard error, in the boys' school participation equation (equation 2) while significant at 5 percent level of significance in girls' school participation equation. The previous negative sign on the fertility estimate in the simple probit schooling equation remains unchanged but insignificant in the instrumental variable model for boys but significantly different from zero in simple probit, as well as in the instrumental variable model for girls. Higher fertility/number of children per couple in the household is more detrimental to the schooling of girls than boys in rural Pakistan, as foreshadowed in our earlier discussion.

The measure of school accessibility or availability; number of schools in the community, significantly positively affect the likelihood of boys' school participation. However, the parameter estimate in the girls' school participation equation is insignificant with cluster robust estimate of the standard error (p-value > 0.05 or 0.10) (equations 2 and 4). This might

be interpreted that school access is not of much relevance for girls compared with the quality of the community schools.

The school quality measure 'proportion of schools with a library' significantly positively influences the likelihood of boys and girls schooling participation, while the variable proxying the school crowding effect (student-teacher ratio) significantly negatively affects school participation only of girls (equations 2 and 4). The statistical significance of both school quality correlates are more pronounced in girls' school participation equation than boys.

Table 5.7 equations 2 and 4 show that the impact of age, in boys and girls schooling equation, is positive and significant. The probability of child school participation significantly increases with household income. The relationship remains the same for both boys and girls.

Parental education substantially improves the probability of child school participation, although maternal education matters more. The marginal impact of mother's education and father's education on boys' school participation is 8.3 percent and 3.2 percent, respectively in the school attendance equation (boys IV equation 1, Table 5.8). The impact of maternal education on girls' school participation holds the same as for boys (equation 2). However, the impact of paternal level of education is insignificant, with the cluster robust estimate of the standard error, in girls' school attendance equation (equation 2, Table 5.8). The result confirms the differential treatment of girls relative to boys in rural Pakistan. The impact of plot size on boys' school attendance is insignificant (when accounting for statistically significant non-linearity in plot size in the boys only sample). The land ownership dummy shows that boys from non-landless households have 7.4 percent higher chance of child school participation than landless households (IV equation 1 Table 5.8). Girls' school attendance is significantly increasing in plot size, which shows that households having land are more able or willing to finance girls' education. However, the land ownership dummy is not significantly different from zero. The variable measuring access to credit (distance to the nearest bank) is not significantly different from zero for both boys and girls (equations 2 and 4, Table 5.8).

The separate results on boys and girls strengthen the whole sample results in support of the model prediction that school quality is an important reason for child school attendance. These results suggest that the theoretical model presented earlier is generally appropriate for boys and girls, and more pronounced for the girls only sample, in rural Pakistan. The school

quality effect is more pronounced for girls due to the fact that parents require for them higher quality and relevance of education. If the old age support hypothesis was relevant, then school quality would not be a significant determinant for girls' school participation; but the reported results show that school quality is important for girls as for boys. However, the last result that parents might require greater school quality for girls than for boys does render weak support for the old age hypothesis. School quality could be important in the theoretical model even for girls as the available evidence shows that the presence of bride price[47] for girls act as an incentive for parents to invest in their daughter's human capital (Becker 1991). Education increases the economic value of women[48]. As the return to education increase, parents invest in the human capital of their daughters more often and bride price/dowries "are becoming an inferior way of providing brides with future wealth relative to investing in their human capital" (Anderson 2007, p.169; Iset 2013), in the current period. Pencavel (1999) shows that assortative[49] mating increases with the level of educational attainment. Behrman, Rosenzweig, and Taubman (1994) and Behrman and Rosenzweig (2002) found that more schooling is linked with better marriage outcomes, even when controls for heritable differences and a common environment are taken into account.

[47] Bride price is usually a significant amount of payment (cash or other goods e.g. livestock) made from the groom/or his family to the bride/ or her family at the time of marriage.

[48] The old age support hypothesis may not permit parents to invest in the education of their daughters as the returns accrue to someone else's parents. Now parents choose whether to keep their daughters uneducated and pay a bride price later at marriage or to educate daughter and not pay the bride price later on. Education is viewed as important from a parental point of view as it increases the economic value of women which also increases the chances of marriage relative to uneducated women.

[49] Assortative mating refer to a non-random mating phenomena, showing that individuals with similar characteristics mate with one another more often than what could be expected under a random mating phenomena.

**Table 5.7:** Probit and IV Probit Estimates of Child Schooling Decisions in Rural Pakistan (Kharif Season): With Household Income net of Child Earnings (Equivalent). (Boys and Girls)

| | (1)<br>Boys. Schooling | (2)<br>Boys. IV. Schooling | (5)<br>Girls. Schooling | (6)<br>Girls. IV. Schooling |
|---|---|---|---|---|
| Child Age | 0.814<br>(0.0680)*** [ 0.0667]*** | 0.742<br>(0.0805)[ 0.114] | 0.528<br>(0.0741)*** [ 0.0820]*** | 0.382<br>(0.0730)*** [ 0.107]*** |
| Child Age Squared | -0.0406<br>(0.00339)*** [ 0.00328]*** | -0.0371<br>(0.00400)*** [ 0.00567]*** | -0.0286<br>(0.00376)*** [ 0.00418]*** | -0.0210<br>(0.00375)*** [ 0.00562]*** |
| Own Farm Animals (Yes) | 0.0292<br>(0.0732)[ 0.0920] | -0.0187<br>(0.0723)[ 0.0983] | -0.0270<br>(0.0782)[ 0.103] | -0.0256<br>(0.0680)[ 0.0957] |
| Own Non Farm Animals (Yes) | -0.00950<br>(0.0696)[ 0.105] | -0.0330<br>(0.0678)[ 0.111] | 0.0729<br>(0.0731)[ 0.0968] | 0.0376<br>(0.0648)[ 0.0950] |
| Household Income net of Child Earnings | 2.635<br>(0.518)*** [ 0.675]*** | 3.328<br>(0.606)*** [ 0.652]*** | 3.155<br>(0.725)*** [ 0.874]*** | 3.158<br>(0.600)*** [ 0.925]*** |
| Household Income net of Child Earnings Squared | -1.406<br>(0.564)* [ 0.654]* | -2.113<br>(0.836)*[ 0.680]** | -3.727<br>(1.180)** [ 1.352]** | -3.641<br>(0.869)*** [ 1.300]** |
| **Fertility** | -0.0302<br>(0.0156)+ [ 0.0174]+ | -0.0522<br>(0.0345)[ 0.0396] | -0.0721<br>(0.0171)*** [ 0.0218]*** | -0.122<br>(0.0329)*** [ 0.0438]** |
| Mother's Level of Education | 0.299<br>(0.0896)*** [ 0.106]** | 0.257<br>(0.0875)**[ 0.103]* | 0.356<br>(0.0804)*** [ 0.0834]** | 0.265<br>(0.0645)*** [ 0.0954]** |
| Father's Level of Education | 0.133<br>(0.0314)*** [ 0.0371]*** | 0.0984<br>(0.0334)** [ 0.0551]+ | 0.150<br>(0.0284)*** [ 0.0340]*** | 0.0524<br>(0.0294)+ [ 0.0474] |
| Land owned in Kharif 2004 | 0.00406<br>(0.00617)[ 0.00977] | 0.00816<br>(0.00681)[ 0.00991] | 0.0455<br>(0.0108)*** [ 0.0164]** | 0.0430***<br>(0.00792)** [ 0.00956]*** |
| Land owned in Kharif Squared 2004 | 0.0000167<br>(0.0000654)[ 0.000101] | -0.00000603<br>(0.0000829)[ 0.000108] | -0.000642<br>(0.000216)** [ 0.000317]* | -0.000506<br>(0.000126)*** [ 0.000175]** |
| Land Ownership | 0.226 | 0.227 | -0.0303 | -0.0164 |

| | | | | |
|---|---|---|---|---|
| Kharif (Yes) | (0.0749) ** [ 0.0952] * | (0.0739) ** [ 0.0939] * | (0.0845)[ 0.133] | (0.0730)[ 0.105] |
| Number of schools in a Community or Village | 0.0634<br>(0.0106) *** [ 0.0179] *** | 0.0582<br>(0.0109) *** [ 0.0268] * | 0.0466<br>(0.0109) *** [ 0.0227] * | 0.0231<br>(0.0105) * [ 0.0324] |
| **Student Teacher Ratio** | 0.00233<br>(0.00247)[ 0.00422] | -0.0326<br>(0.0109) ** [ 0.0262] | 0.00575<br>(0.00262) * [ 0.00542] | -0.0535<br>(0.00621) *** [ 0.0174] ** |
| Proportion of Schools with a Library | 0.565<br>(0.178) ** [ 0.329] + | 0.659<br>(0.173) *** [ 0.395] + | 0.944<br>(0.179) *** [ 0.364] ** | 1.146***<br>(0.166) *** [ 0.448] * |
| Distance to nearest bank | -0.00442<br>(0.00480)[ 0.0102] | 0.0113<br>(0.00673) + [ 0.0142] | -0.0130<br>(0.00503) ** [ 0.00985] | 0.0103<br>(0.00520) * [ 0.0131] |
| Province (Punjab) | -0.0235<br>(0.0758)[ 0.137] | -0.292<br>(0.110) ** [ 0.226] | 0.133<br>(0.0814)[ 0.160] | -0.348<br>(0.0831) *** [ 0.232] |
| Constant | -4.956<br>(0.373) *** [ 0.464] *** | -3.298***<br>(0.741) *** [ 1.564] * | -3.856<br>(0.404) *** [ 0.454] *** | -0.512<br>(0.533)[ 1.181] |
| s21<br>_cons | | 0.105<br>(0.109)[ 0.126] | | 0.287<br>(0.0991) ** [ 0.132] * |
| s31<br>_cons | | 4.763<br>(1.436) *** [ 2.892] + | | 8.459<br>(0.816) *** [ 1.465] *** |
| s22<br>_cons | | 1.673<br>(0.0267) *** [ 0.0669] *** | | 1.649<br>(0.0311) *** [ 0.0521] *** |
| s32<br>_cons | | 0.200<br>(0.393)[ 0.717] | | -1.048<br>(0.500) * [ 0.697] |
| s33<br>_cons | | 10.44<br>(0.664) *** [ 2.041] *** | | 8.330<br>(0.809) *** [ 2.185] *** |
| Pseudo R-Square | 0.145[0.145] | | 0.155[0.155] | |
| Wald chi2 (17) | 321.5[360.5] | 416.2[278.2] | 309.6[233.1] | 710.6[361.3] |
| p | 4.08e-58[3.19e-66] | 7.73e-78[3.52e-49] | 1.18e-55[5.71e-40] | 4.75e-140[2.16e-66] |
| N | 2086[94] | 2086[94] | 1864[94] | 1864[94] |
| Wald Test of Exogeneity: | | | | |
| Chi2(2) | | 11.14[2.80] | | 139.09[44.14] |
| P | | 0.0038[0.2470] | | 0.0000[0.0000] |

| | | |
|---|---|---|
| Test of over-identifying restrictions[(a)]: | | |
| Amemiya-Lee-Newey minimum chi-sq statistic | | |
| Chi-sq(2) | 0.221 | 0.930 |
| P | 0.8955 | 0.6282 |

Standard errors in parentheses, cluster standard errors in square brackets

$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$

(a): Test of over-identifying restrictions based on two-step procedure of ivprobit estimation.

Note: 94 clusters in village or community

**Table 5.8:** Probit and IV Probit Average Marginal Effects of Child Schooling Decisions in Rural Pakistan (Kharif Season): With Household Income net of Child Earnings (Equivalent). (Boys and Girls)

| | (1)<br>Boys. IV. Schooling | (2)<br>Girls. IV. Schooling |
|---|---|---|
| Child Age | 0.00574<br>(0.00286) * [ 0.00361] | -0.00613<br>(0.00269) * [ 0.00263] * |
| Own Farm Animals (Yes) | -0.00605<br>(0.0234)[ 0.0317] | -0.00795<br>(0.0212)[ 0.0296] |
| Own Non Farm Animals (Yes) | -0.0107<br>(0.0220)[ 0.0362] | 0.0117<br>(0.0202)[ 0.0296] |
| Household Income net of Child Earnings | 0.876<br>(0.131) *** [ 0.156] *** | 0.636<br>(0.119) *** [ 0.196] ** |
| **Fertility** | -0.0169<br>(0.0111)[ 0.0127] | -0.0378<br>(0.0101) *** [ 0.0143] ** |
| Mother's Level of Education | 0.0832<br>(0.0286) ** [ 0.0347] * | 0.0824<br>(0.0201) *** [ 0.0310] ** |
| Father's Level of Education | 0.0319<br>(0.0110) ** [ 0.0190] + | 0.0163<br>(0.00916) + [ 0.0151] |
| Land owned in Kharif 2004 | 0.00263<br>(0.00202)[ 0.00293] | 0.0122<br>(0.00223) *** [ 0.00270] *** |
| Land Ownership Kharif (Yes) | 0.0741<br>(0.0242) ** [ 0.0309] * | -0.00509<br>(0.0227)[ 0.0325] |
| Number of schools in a Community or Village | 0.0189<br>(0.00359) *** [ 0.00905] * | 0.00718<br>(0.00328) * [ 0.0101] |
| **Student Teacher Ratio** | -0.0106<br>(0.00334) ** [ 0.00804] | -0.0166<br>(0.00180) *** [ 0.00475] *** |
| Proportion of Schools with a Library | 0.214<br>(0.0555) *** [ 0.125] + | 0.356<br>(0.0508) *** [ 0.133] ** |
| Distance to nearest bank | 0.00365<br>(0.00214) + [ 0.00454] | 0.00321<br>(0.00160) * [ 0.00407] |
| Province (Punjab) | -0.0926<br>(0.0326) ** [ 0.0667] | -0.105<br>(0.0240) *** [ 0.0669] |
| *N* | 2086 | 1864 |

Marginal effects; Standard errors in parentheses, cluster standard errors in square brackets
(d) for discrete change of dummy variable from 0 to 1
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$
Note: 94 clusters in village or community

## 5.5 Conclusion

Many dimensions of school quality are likely to be hard to measure and dependent on the valuations of parents. We tested a number of unreported indicators that did not appear to influence parental behaviour. The two quality variables on which we focussed for the simulation were student-teacher ratios and libraries. Assuming all sample households have optimised, as the model does, the per capita income equivalent of a ten percent increase in school quality was apparently equal to one half of the average household (equivalised) expenditure. Obviously the changes involved in the simulation are non-marginal and some caveats must be attached to the numbers, but there is no doubt that they are large. In view of the great variation in school quality measures, simply reducing the disparities by improving the poorest quality schools could substantially improve school attendance and human capital accumulation.

Short period credit constraints do not seem to be problematic as far as schooling is concerned. In the face of temporary shocks our cross-section evidence indicates that households need not be driven to supply child labour by reducing schooling. On the other hand credit will not be available for minors to lend to their parents on the strength of their future earnings and any form of borrowing or lending over, say, twenty year periods is likely to be limited by uncertainty. Hence, we cannot dismiss the notion of Baland and Robinson (2000) and Ranjan (1999) that long term credit constraint is a vital factor affecting child schooling decisions.

To test the hypothesis that schooling is deterred by below subsistence incomes we adopted a similar strategy to that of Ray (2000a, 2000b) but identified households' poverty status with the national poverty line. Our results show that school quality is even more important in raising school attendance for the poor than for the rich. Therefore the principal problem of school attendance in Pakistan cannot generally be one of subsistence poverty.

Separate analysis of boys and girls shows that school quality is of equal relevance for both sexes and strengthens the whole sample results in support of the model prediction that school quality is a statistically and quantitatively significant reason for child school attendance. However higher fertility/number of children per couple is more detrimental to the schooling of girls than boys in rural Pakistan; as girls usually help their mothers with child care and household work.

Child and adult wages are assumed exogenous in the present model. If the two types of labour are actually substitutes and child wages respond to the quantity of child labour, the

conclusion about school quality is reinforced. Improved school quality that reduces child labour would then raise child and adult wages, so enhancing the worth of better schools. In this case the current estimate of the value of raising school quality in Pakistan would be understated.

Why is school quality so low relative to parental aspirations? Pakistani educational provision is largely financed by the public sector, which spends 2 percent of GDP on education, less than other countries in the region. But the allocation mechanism is also at fault and explains the wide variation in school quality (HRCP and CEF 2005). To prevent (or at least identify) the diversion and inefficient use of funds by influential but partial local bodies that must be occurring, an independent national schools inspectorate therefore might be considered.

## Appendices

### Appendix A: Optimization with Respect to Schooling

**The budget constraint;**

$$C_1 = w_p + n.w_c(1-S) - n.H_0 - n.H_1 S$$

$$C_2 = n.w_u(1+\psi.S) \qquad \text{(A1)}$$

**Utility function;**

$$U = (C_1 - \alpha)^{\gamma_1}(C_2 - \beta_0 - \beta_1 n)^{\gamma_2} \qquad \text{(A2)}$$

Taking logarithms we get from (A1) the direct utility function:

$$V = \gamma_1 Ln(C_1 - \alpha) + \gamma_2 Ln(C_2 - \beta_0 - \beta_1 n) \qquad \text{(A3)}$$

With first and second order conditions stated as:

$$V_{C_1} = \gamma_1 . \frac{1}{(C_1 - \alpha)} > 0$$

$$V_{C_2} = \gamma_2 . \frac{1}{(C_2 - \beta_0 - \beta_1 n)} > 0$$

The second order conditions are:

$$V_{C_1C_1} = -\gamma_1 . \frac{1}{(C_1 - \alpha)^2} < 0$$

$$V_{C_2C_2} = -\gamma_2 . \frac{1}{(C_2 - \beta_0 - \beta_1 n)^2} < 0$$

These conditions are plausible only if $C_1 > \alpha$ and $C_2 > \beta_0 - \beta_1 n$.

**Derivation of optimal schooling equation (maximization by choosing S given n);**
Substituting for C1 and C2 in (A3) and taking derivative with respect to S;

$$\frac{\partial V}{\partial S} = \frac{\gamma_1}{(C_1 - \alpha)}.n.(-w_c - H_1) + \frac{\gamma_2}{(C_2 - \beta_0 - \beta_1 n)}.n.(\psi w_u) = 0$$

$$\frac{\partial V}{\partial S} = \frac{\gamma_1}{(C_1 - \alpha)}.n.(-w_c - H_1) = -\frac{\gamma_2}{(C_2 - \beta_0 - \beta_1 n)}.n.(\psi w_u)$$

$$\frac{-\gamma_1 n w_c - \gamma_1 n H_1}{(C_1 - \alpha)} = -\frac{\gamma_2 n \psi w_u}{(C_2 - \beta_0 - \beta_1 n)} \qquad \text{(A4)}$$

Putting the values of C1 and C2 in (A4) and solving for S;

$$\frac{-\gamma_1 nw_c - \gamma_1 nH_1}{w_p + nw_c(1-S) - nH_0 - n.H_1 S - \alpha)} = -\frac{\gamma_2 n\psi w_u}{(nw_u(1+\psi.S) - \beta_0 - \beta_1 n)} \quad \text{(A5)}$$

$$(-\gamma_1 nw_c - \gamma_1 nH_1)(nw_u(1+\psi.S) - \beta_0 - \beta_1 n) = -\gamma_2 n\psi w_u(w_p + nw_c(1-S) - nH_0 - n.H_1 S - \alpha) \quad \text{(A6)}$$

$$(-\gamma_1 nw_c - \gamma_1 nH_1)(nw_u + nw_u\psi.S - \beta_0 - \beta_1 n) = -\gamma_2 n\psi w_u(w_p + nw_c - nw_c S - nH_0 - nH_1 S - \alpha) \quad \text{(A7)}$$

$$\begin{aligned}&(-\gamma_1 nw_c - \gamma_1 nH_1)nw_u + (-\gamma_1 nw_c - \gamma_1 nH_1)nw_u\psi.S - (-\gamma_1 nw_c - \gamma_1 nH_1)\beta_0 \\ &\Leftrightarrow -(-\gamma_1 nw_c - \gamma_1 nH_1)\beta_1 n = -\gamma_2 n\psi w_u w_p - \gamma_2 n\psi w_u nw_c + \gamma_2 n\psi w_u nw_c S + \gamma_2 n\psi w_u nH_0 \\ &\Leftrightarrow +\gamma_2 n\psi w_u nH_1 S + \gamma_2 n\psi w_u \alpha)\end{aligned} \quad \text{(A8)}$$

$$\begin{aligned}&(-\gamma_1 nw_c - \gamma_1 nH_1)nw_u - (-\gamma_1 nw_c - \gamma_1 nH_1)\beta_0 \\ &\Leftrightarrow -(-\gamma_1 nw_c - \gamma_1 nH_1)\beta_1 n = -\gamma_2 n\psi w_u w_p - \gamma_2 n\psi w_u nw_c + \gamma_2 n\psi w_u nw_c S + \gamma_2 n\psi w_u nH_0 \\ &\Leftrightarrow +\gamma_2 n\psi w_u nH_1 S + \gamma_2 n\psi w_u \alpha - (-\gamma_1 nw_c - \gamma_1 nH_1)nw_u\psi.S\end{aligned} \quad \text{(A9)}$$

$$\begin{aligned}&(-\gamma_1 nw_c - \gamma_1 nH_1)nw_u - (-\gamma_1 nw_c - \gamma_1 nH_1)\beta_0 - (-\gamma_1 nw_c - \gamma_1 nH_1)\beta_1 n \\ &\Leftrightarrow +\gamma_2 n\psi w_u w_p + \gamma_2 n\psi w_u nw_c - \gamma_2 n\psi w_u nH_0 - \gamma_2 n\psi w_u \alpha = \gamma_2 n\psi w_u nw_c S \\ &\Leftrightarrow +\gamma_2 n\psi w_u nH_1 S - (-\gamma_1 nw_c - \gamma_1 nH_1)nw_u\psi.S\end{aligned} \quad \text{(A10)}$$

$$\begin{aligned}&(-\gamma_1 nw_c - \gamma_1 nH_1)nw_u - (-\gamma_1 nw_c - \gamma_1 nH_1)\beta_0 - (-\gamma_1 nw_c - \gamma_1 nH_1)\beta_1 n \\ &\Leftrightarrow +\gamma_2 n\psi w_u w_p + \gamma_2 n\psi w_u nw_c - \gamma_2 n\psi w_u nH_0 - \gamma_2 n\psi w_u \alpha = \gamma_2 n\psi w_u nw_c S \\ &\Leftrightarrow +\gamma_2 n\psi w_u nH_1 S - (-\gamma_1 nw_c - \gamma_1 nH_1)nw_u\psi.S\end{aligned} \quad \text{(A11)}$$

$$S = \frac{\gamma_1(-nw_c - nH_1)(nw_u - \beta_0 - \beta_1 n) + \gamma_2 n\psi w_u(w_p + nw_c - nH_0 - \alpha)}{\gamma_2 n\psi w_u nw_c + \gamma_2 n\psi w_u nH_1 - (-\gamma_1 nw_c - \gamma_1 nH_1)nw_u\psi} \quad \text{(A12)}$$

$$S = \frac{\gamma_1(-nw_c - nH_1)(nw_u - \beta_0 - \beta_1 n) + \gamma_2 n\psi w_u(w_p + nw_c - nH_0 - \alpha)}{\gamma_2 n^2\psi w_u(w_c + H_1) - \gamma_1 nnw_u\psi(-w_c - H_1)} \quad \text{(A13)}$$

$$S = \frac{\gamma_1(-nw_c - nH_1)(nw_u - \beta_0 - \beta_1 n) + \gamma_2 n\psi w_u(w_p + nw_c - nH_0 - \alpha)}{(\gamma_2 + \gamma_1)n^2\psi w_u(w_c + H_1)} \quad \text{(A14)}$$

$$S = \frac{-n\gamma_1(w_c + H_1)(nw_u - \beta_0 - \beta_1 n)}{(\gamma_2 + \gamma_1)n^2\psi w_u(w_c + H_1)} + \frac{\gamma_2 n\psi w_u(w_p + nw_c - nH_0 - \alpha)}{(\gamma_2 + \gamma_1)n^2\psi w_u(w_c + H_1)} \quad \text{(A15)}$$

$$S^* = -\frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u\psi} + \frac{\gamma_2(w_p + n(w_c - H_0) - \alpha)}{n.(w_c + H_1)} \quad \text{(A16)}$$

**Derivatives of optimum schooling equation (A16);**

The derivatives of optimum schooling equation with respect to wage premium (school quality), child wage, parental wage and fertility are calculated as follows;

**Derivative with respect to premium (school quality)**

$$\frac{\partial S}{\partial \psi} = \frac{\partial}{\partial \psi}\left(-\frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u.\psi}\right)$$

$$\frac{\partial S}{\partial \psi} = -\frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u}\frac{\partial}{\partial \psi}.(\psi)^{-1}$$

$$\frac{\partial S}{\partial \psi} = -\frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u}.-1.(\psi)^{-1-1}$$

$$\frac{\partial S}{\partial \psi} = \frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u}.(\psi)^{-2}$$

$$\frac{\partial S}{\partial \psi} = \frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u\psi^2}$$

OR

$$\frac{\partial S}{\partial \psi} = \frac{\gamma_1}{\psi^2} - \frac{\gamma_1(\beta_0 + \beta_1 n)}{n.w_u\psi^2} \qquad \text{(A17)}$$

**Derivative with respect to child wage**

$$\frac{\partial S}{\partial w_c} = \frac{\partial}{\partial w_c}\left(\frac{\gamma_2(w_p + n(w_c - H_0) - \alpha)}{n.(w_c + H_1)}\right)$$

Using the quotient Rule of Differentiation;

$$\frac{\partial S}{\partial w_c} = \frac{\gamma_2}{n}\left(\frac{(w_c + H_1).\frac{\partial}{\partial w_c}(w_p + n(w_c - H_0) - \alpha) - (w_p + n(w_c - H_0) - \alpha).\frac{\partial}{\partial w_c}(w_c + H_1))}{(w_c + H_1)^2}\right)$$

$$\frac{\partial S}{\partial w_c} = \frac{\gamma_2}{n}\left(\frac{(w_c + H_1).n - (w_p + n(w_c - H_0) - \alpha)}{(w_c + H_1 - b)^2}\right)$$

$$\frac{\partial S}{\partial w_c} = \frac{\gamma_2}{n}\left(\frac{(n.w_c + n.H_1 - w_p - nw_c + nH_0 + \alpha)}{(w_c + H_1)^2}\right)$$

$$\frac{\partial S}{\partial w_c} = \frac{\gamma_2}{n}\left(\frac{(-w_p + n.H_1 + nH_0 + \alpha)}{(w_c + H_1)^2}\right) \qquad \text{(A18)}$$

**Derivative with respect to parental wage**

$$\frac{\partial S}{\partial w_p} = \frac{\partial}{\partial w_p}\left(\frac{\gamma_2(w_p + n(w_c - H_0) - \alpha)}{n.(w_c + H_1)}\right)$$

$$\frac{\partial S}{\partial w_p} = \frac{\gamma_2}{n.(w_c + H_1)} \qquad \text{(A19)}$$

**Derivative with respect to fertility**

**First part (ratio) of the optimal schooling equation;**

$$\frac{\partial S}{\partial n} = \frac{\partial}{\partial n}\left( -\frac{\gamma_1(nw_u - \beta_0 - \beta_1 n)}{n.w_u.\psi} \right)$$

Using the quotient rule of Differentiation;

$$\frac{\partial S}{\partial w_u} = \frac{-\gamma_1}{w_u.\psi}\left( \frac{n.\frac{\partial}{\partial n}(nw_u - \beta_0 - \beta_1 n) - (nw_u - \beta_0 - \beta_1 n).\frac{\partial}{\partial n}.n}{n^2} \right)$$

$$\frac{\partial S}{\partial w_u} = \frac{-\gamma_1}{w_u.\psi}\left( \frac{n.(w_u - \beta_1) - (nw_u - \beta_0 - \beta_1 n)}{n^2} \right)$$

$$\frac{\partial.S}{\partial n} = -\frac{\gamma_1 \beta_0}{w_u \psi.n^2} \qquad \text{(A20)}$$

**Second part (ratio) of the optimal schooling equation;**

$$\frac{\partial S}{\partial n} = \frac{\partial}{\partial n}\left( \frac{\gamma_2(w_p + n(w_c - H_0) - \alpha)}{n.(w_c + H_1)} \right)$$

Using the Quotient rule of differentiation;

$$\frac{\partial S}{\partial n} = \frac{\gamma_2}{(w_c + H_1)}\left( \frac{n.\frac{\partial}{\partial n}((w_p + n(w_c - H_0) - \alpha) - ((w_p + n(w_c - H_0) - \alpha).\frac{\partial}{\partial n}.n}{n^2} \right)$$

$$\frac{\partial S}{\partial n} = \frac{\gamma_2}{(w_c + H_1)}\left( \frac{nw_c - nH_0 - w_p - n.w_c + nH_0 + \alpha)}{n^2} \right)$$

$$\frac{\partial.S}{\partial n} = -\frac{\gamma_2(w_p - \alpha)}{(w_c + H_1).n^2} \qquad \text{(A21)}$$

Combining both Parts (A20 and A21) of the differentiation gives;

$$\frac{\partial.S}{\partial n} = -\frac{\gamma_1 \beta_0}{w_u \psi.n^2} - \frac{\gamma_2(w_p - \alpha)}{(w_c + H_1).n^2}$$

## Appendix B

### Community-Level Exogeneity of School Quality and Availability

Community-level school quality variables are not endogenous to individual child labour decisions but might be correlated with community-level unobservables; one community might be politically and economically more active or better endowed and both school quality as well as schooling/child labour might be a response to these characteristics. One approach to identification is to instrument school quality and quantity variables, if appropriate instruments can be found. Here we demonstrate that all the most obvious instruments are not appropriate and this in itself is evidence for community–level exogeneity (following Alderman et al 1996).

**Table B1: Village-level Selected Variables (Mean and Standard Deviation)**

| | Mean | Standard Deviation | Min | Max |
|---|---|---|---|---|
| **Effective Distances (Km)[(a)]** | | | | |
| Distance to Mandi[(b)] | 11.69 | 9.06 | 0 | 50 |
| Distance to Nearest Weekly Market | 9.69 | 12.17 | 0 | 50 |
| Distance to Tehsil Capital | 15.74 | 10.77 | 2 | 50 |
| Distance to District Capital | 40.13 | 21.39 | 0 | 100 |
| **Indicators of Political Links** | | | | |
| Local Councillor | 41.5% | | 0 | 1 |
| Naib Nazim | 3.2% | | 0 | 1 |
| Nazim | 4.3% | | 0 | 1 |
| MNA | 1.1% | | 0 | 1 |
| MPA | 2.1% | | 0 | 1 |
| Numberdar | 38.3% | | 0 | 1 |
| Patwari | 11.7% | | 0 | 1 |
| Punchayet | 27.7% | | 0 | 1 |
| Wadera | 23.4% | | 0 | 1 |
| Biradary Leader | 40.4% | | 0 | 1 |
| Religious Leader | 20.2% | | 0 | 1 |
| Army/Police | 53.2% | | 0 | 1 |
| **Income and Population** | | | | |
| Mean Household Income (Rs) | 94516.35 | 37547.58 | 45379.88 | 233672.2 |
| Population (Households) | 710.31 | 999.56 | 18 | 6000 |
| Number of Villages | 94 | | | |

(a) For obtaining the effective distances, the actual distance was doubled if the entire road was unpaved, and if the condition of the road was a mixture of paved and unpaved road then multiplied by a factor of 1.5.
(b) Mandi means a big local market in rural villages.

Table B1 presents the means and standard deviations of obvious choices for instruments. The logit estimates are presented in Table B2. Table B3 gives OLS estimates of these village level

characteristics on village level school quality variables. Tobit estimates have also been provided in Table B4 for two measures of village level school quality, namely the proportion of schools with a playground and the proportion of schools with a library. The estimates based on the analysis show no relation between school availability/school quality and observed community level characteristics (except perhaps population). Apart from a resident local councillor as a measure of strength of linkages between the village and regional government, we also explored other specifications such as the presence of a Member of the Provincial Assembly (MPA) in the sample villages, a Member of National Assembly (MNA), a Biradary Leader (a head of the same clan in rural villages), a religious leader, a Punchayet (a village council in Rural Pakistan) and the presence of other government officials. None of these measures of political influence shows any relation with the quality or availability of school in the sample villages nor adds to the explanatory power of the model. Hence we assume that local school quality does not reflect effective household demand but is determined independently.

**Table B2: Logit Estimates of Impact of Village Characteristics on Primary, Secondary or Middle School Availability**

| | Presence of | | | Presence of School for | | |
|---|---|---|---|---|---|---|
| | (1) Primary | (2) Middle | (3) Secondary | (1) Boys | (2) Girls | (3) Both |
| Mean Household Income | -0.0000227 | -0.000000358 | 0.00000220 | -0.0000113 | 0.000000357 | 0.0000166 |
| | (-1.58) | (-0.05) | (0.28) | (-1.55) | (0.05) | (1.65) |
| Local Councilor | -0.150 | 0.899 | 0.817 | 0.872 | 0.882 | 1.496 |
| | (-0.09) | (1.58) | (1.59) | (1.35) | (1.87) | (1.45) |
| Distance to the Mandi | -0.00317 | 0.0206 | -0.00420 | -0.0241 | -0.000373 | -0.0449 |
| | (-0.10) | (0.95) | (-0.19) | (-0.88) | (-0.02) | (-1.13) |
| Population | -0.000333 | 0.000287 | -0.0000608 | 0.000320 | -0.000111 | 0.00336 |
| | (-1.11) | (0.80) | (-0.23) | (0.80) | (-0.51) | (0.98) |
| Constant | 6.509* | 0.532 | 0.617 | 2.585** | 0.379 | 0.206 |
| | (2.12) | (0.69) | (0.78) | (2.69) | (0.50) | (0.14) |
| Pseudo R-Square | 0.195 | 0.0440 | 0.0259 | 0.0689 | 0.0325 | 0.183 |
| Wald chi2(4) | 9.224 | 5.091 | 2.609 | 3.785 | 3.789 | 9.963 |
| p | 0.0557 | 0.278 | 0.625 | 0.436 | 0.435 | 0.0410 |
| N | 94 | 94 | 94 | 94 | 94 | 94 |

*t* statistics in parentheses (based on robust standard errors)
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

**Table B3: OLS Estimates of Impact of Village Characteristics on Village Level School Quality Variables**

| | (1) Number of schools in a Community or Village | (2) Student Teacher Ratio | (3) Proportion of Schools with a Playground | (4) Proportion of Schools with a Library |
|---|---|---|---|---|
| Mean Household Income | 0.0000193 | 0.0000620 | -0.0000504 | 0.0000922 |
| | (1.74) | (1.11) | (-0.59) | (1.27) |
| Local Councilor | 0.644 | -4.296 | 3.775 | 1.763 |
| | (1.12) | (-1.67) | (0.62) | (0.40) |
| Distance to the Mandi | -0.0403 | -0.190 | 0.257 | -0.0652 |
| | (-1.32) | (-1.48) | (1.04) | (-0.27) |
| Population | 0.000153 | -0.00121 | -0.00739*** | -0.00452** |
| | (0.68) | (-0.75) | (-4.17) | (-3.03) |
| Constant | 5.078*** | 33.10*** | 44.90*** | 16.42* |
| | (4.46) | (5.82) | (4.58) | (2.44) |
| *N* | 94 | 94 | 94 | 94 |
| $R^2$ | 0.065 | 0.082 | 0.054 | 0.012 |
| F (4, 89) | 3.175 | 2.359 | 7.490 | 2.626 |
| p | 0.017 | 0.133 | 0.000 | 0.039 |

*t* statistics in parentheses (based on robust standard errors)
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

**Table B4: Tobit Estimates of Impact of Village Characteristics on Village Level School Quality Variables**

| | (1) Proportion of Schools with a Playground | (2) Proportion of Schools with a Library |
|---|---|---|
| Mean Household Income | -0.0000773 | 0.000138 |
| | (-0.60) | (1.30) |
| Local Councilor | 6.241 | 3.920 |
| | (0.74) | (0.63) |
| Distance to the Mandi | 0.309 | -0.183 |
| | (0.98) | (-0.56) |
| Population | -0.0116* | -0.00812* |
| | (-2.56) | (-2.29) |
| Constant | 44.81** | 9.188 |
| | (3.22) | (0.91) |
| sigma | | |
| | 38.40 | 28.76 |
| Pseudo R-Square | 0.0134 | 0.00947 |
| F (4, 90) | 3.03 | 1.54 |
| P | 0.0216 | 0.198 |
| Log pseudolikelihood | -364.03 | -327.88 |
| N | 94 | 94 |

*t* statistics in parentheses (based on robust standard errors)
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Obs. summary: 21 left-censored observations at Proportion of Schools having Playground <=0
66 uncensored observations
7 right-censored observations at Proportion of Schools having Playground >=100

Obs. summary: 31 left-censored observations at Proportion of Schools having Library <=0
63 uncensored observations
0 right-censored observations

**Appendix C Household Income Net of Child Earnings**

A separate inter-temporally consistent measure of household level poverty (also known as household income net of child earnings) was also calculated in the same fashion as given in Bhalotra (2007). As child schooling and labour participation have inter-temporal concern for the typical household, let us suppose an inter-temporal version of the budget constraint, given in (C1), for the households who actually undertake the child employment decision. A further variable is also included in the budget constraint, which measures the household non-labour income represented by 'I'. The modified version of the budget constraint is given by;

$$C_t = I + w_p + w_c E_c$$

$$I = C_t - w_p - w_c E_c \qquad \text{(C1)}$$

Where, 'I' includes income from self employment and any interest earning from household assets '$A_t$'. Similarly, the household income net of child earnings '$Y$' is defined as:

$$Y = I + w_p \qquad \text{(C2)}$$

Following Bhalotra (2007), the time path of the household asset is given as:

$$A_{t+1} = (1+r)A_t + w_p + w_c E_c - C_t$$

$$rA_t + A_t - A_{t+1} = C_t - w_p - w_c E_c$$

$$rA_t - \Delta A_{t+1} = C_t - w_p - w_c E_c \qquad \text{(C3)}$$

Using equation (C1), (C2) is given as:

$$rA_t - \Delta A_{t+1} = I$$

In the absence of data on asset changes, household income net of child earnings '$Y$' from (C2) is given as the difference between consumption and child earnings as below:

$$Y = I + w_p$$

$$Y = C - w_p A - w_c E_c + w_p = C - w_c E_c$$

The 'household income net of child earnings' excludes income from child employment and has been used as exogenous household's poverty measure in the child's schooling and employment analysis. It has also been made consistent over all households in the same spirit

as above by using OECD (1982) Equivalence scale. This categorization represents household poverty as dependant on adult's wage earnings, non-labour earnings and any saving and no saving undertaken by the household (Bhalotra 2007). Bhalotra (2007) tested whether working children come from families whose subsistence needs exceed household income net of child earnings as postulated by Basu and Van's (1998). Her analysis, using Pakistani data, aimed to test the negative wage elasticity of children's wages while controlling for 'household income net of child earnings' in child labour supply equations, confining her sample only to those children who participate in full-time wage employment. The 'household income net of child earnings' measure of household level poverty is viewed as an improved measure over the total household income measure which has been used in most previous studies on child schooling and employment (Ersado 2005).

## Appendix D Credit Rationing Variables

Jacoby (1994) reported from Peru that credit constraints negatively affect children's primary schooling attainment. A study by Fuwa et al (2009) from rural India also found that credit market failure leads to enormous reallocations of a child's time in different activities such as schooling, work, leisure and households chores. The negative impact of credit constraint amounts to a 60 percent decrease in average schooling time of the rural children. Studies also exist that viewed a household's child labour decisions as a part of risk management strategy (Jacoby and Skoufias 1997) to minimize the impact of shocks such as bad harvest due to calamities or job loss, when access to credit is not available to smooth their consumption.

A number of studies (Ersado 2005) have found credit access to be very difficult to measure from LSMS (Living Standard Measurement Surveys), which only report whether a given household has a loan/bank account or not. It is not a good measure of access to credit because the households which did not testify to a loan account might have access to credit but have no borrowing needs (Ersado 2005). Some recent studies have recognized these potential issues and relied on some sort of exogenous income shocks to infer the effects of credit constraints on household behaviour. Beegle et al (2006) used self reported crop shocks to analyse the effect of credit constraint on child labour. But Fuwa et al (2009) reported flaws with this approach in that it mixes substitution and wealth effects. Productivity shocks change households' shadow prices consisting of substitution and wealth effects. A negative shock to farm productivity results in a decrease in demand for farm production inputs whereas the wealth effect increases the labour supply to farming, if outside employment opportunities are limited. According to Fuwa et al. (2009), the approach adopted by Beegle et al. (2006) may not identify the effect of a credit constraint (which is the marginal utility value of current wealth relative to future) on child labour but the joint effect of price changes and consequent wealth changes. To address these issues the data set (Pakistan Rural Household Survey) used in this study has the unique characteristic that it contains a special module on credit access, not usually available in large scale multipurpose household surveys, which allows us to distinguish between credit constrained and unconstrained households in the same way as used in Fuwa et al. (2009). But the downside is that the module is only available for agricultural households. Nevertheless, access to the commercial bank (distance) at the community level is also used as proxy measure for access to the formal credit market for the whole sample.

As discussed, the module on credit rationing was designed in such a manner to directly determine credit constrained households from unconstrained households as suggested by Fuwa et al. (2009).

For the identification of liquidity constrained households, heads were asked about household's experience with lenders over the last Kharif 2004 season prior to the survey. Whether the household had attempted or not to get a loan over the last Kharif 2004 season was used to identify credit constrained households. For those households which tried to borrow money, it was determined whether or not they were successful in getting the requested amount under the specified conditions. If the response was 'yes', the household was identified as unconstrained. Those households which were unable to borrow as much as requested or whose requests were rejected were classified as credit constrained households.

Those households that did not attempt to get a loan were further interviewed about the reasons for not trying to get one. The choice set was: (1) Inadequate collateral, (2) Past default with this lender, (3) Bad credit history, (4) Lender too far away, (5) Lender has worse terms than other available sources, (6) Lender's procedures are too cumbersome, (7) Need to pay bribes, (8) Do not borrow at all, (9) Other.

We used Fuwa et al.'s (2009) broad definition of credit constraint. Respondents who chose (1) to (7) were classified as credit constrained households. The remaining households that did not try to borrow were regarded as unconstrained. We used the broader definition of credit constrained because 92.5 percent of households indicated that they did not borrow at all (choice number 8 above), and there remained a limited number of households who could possibly be credit constrained.

## Appendix E Tabulated First-Stage Equations

**Table E1:** Whole and Agriculture Household Sample First-Stage Regressions

| | Whole Sample First-Stage | | Agriculture Sample First-Stage | |
|---|---|---|---|---|
| | (1) Fertility | (2) Student-Teacher Ratio | (1) Fertility | (2) Student-Teacher Ratio |
| Child Sex (female) | 0.0950<br>(0.0538)$^{+}$ [ 0.0521]$^{+}$ | -0.0134<br>(0.375)[ 0.454] | 0.0183<br>(0.0702)[ 0.0681] | 0.209<br>(0.494)[ 0.521] |
| Child Age | 0.210<br>(0.0608)$^{***}$ [ 0.0528]$^{***}$ | -0.216<br>(0.423)[ 0.378] | 0.191<br>(0.0793)$^{*}$ [ 0.0701]$^{**}$ | -0.263<br>(0.558)[ 0.471] |
| Child Age Squared | -0.0108<br>(0.00304)$^{***}$ [ 0.00276]$^{***}$ | 0.00813<br>(0.0212)[ 0.0187] | -0.0104<br>(0.00396)$^{**}$ [ 0.00366]$^{**}$ | 0.00720<br>(0.0278)[ 0.0234] |
| Own Farm Animals (Yes) | 0.140<br>(0.0664)$^{*}$ [ 0.106] | -1.159<br>(0.461)$^{*}$ [ 0.983] | -0.0804<br>(0.0962)[ 0.146] | -3.210<br>(0.677)$^{***}$ [ 1.799]$^{+}$ |
| Own Non Farm Animals (Yes) | 0.131<br>(0.0632)$^{*}$ [ 0.133] | -0.757<br>(0.441)$^{+}$ [ 1.206] | 0.120<br>(0.0773)[ 0.173] | -2.029<br>(0.543)$^{***}$ [ 1.567] |
| Household Income net of Child Earnings | -3.131<br>(0.485)$^{***}$ [ 0.912]$^{***}$ | 23.81<br>(3.366)$^{***}$ [ 9.295]$^{*}$ | -3.656<br>(0.607)$^{***}$ [ 1.018]$^{***}$ | 19.84$^{***}$<br>(4.281)$^{***}$ [ 7.938]$^{*}$ |
| Household Income net of Child Earnings Squared | 1.109<br>(0.606)$^{+}$ [ 1.010] | -21.99$^{***}$<br>(4.195)$^{***}$ [ 8.601]$^{*}$ | 1.184<br>(0.694)$^{+}$ [ 0.974] | -20.10$^{***}$<br>(4.921)$^{***}$ [ 7.328]$^{**}$ |
| Mother's Level of Education | -0.120<br>(0.0513)$^{*}$ [ 0.0781] | 0.149<br>(0.358)[ 0.739] | -0.204<br>(0.0653)$^{**}$ [ 0.0781]$^{**}$ | -0.0762<br>(0.458)[ 0.748] |
| Father's Level of Education | 0.0301<br>(0.0253)[ 0.0432] | -0.912$^{***}$<br>(0.176)$^{***}$ [ 0.363]$^{*}$ | 0.0391<br>(0.0306)[ 0.0552] | -1.046$^{***}$<br>(0.215)$^{***}$ [ 0.410]$^{*}$ |
| Land owned in Kharif 2004 | 0.0364<br>(0.00424)$^{***}$ [ 0.0109]$^{***}$ | 0.161<br>(0.0296)$^{***}$ [ 0.0717]$^{*}$ | 0.0350$^{***}$<br>(0.00450)$^{***}$ [ 0.0110]$^{**}$ | 0.231<br>(0.0316)$^{***}$ [ 0.0915]$^{*}$ |
| Land owned in Kharif Squared 2004 | -0.000169<br>(0.0000318)$^{***}$ [ 0.0000520]$^{**}$ | -0.00107$^{***}$<br>(0.000221)$^{***}$ [ 0.000395]$^{**}$ | -0.000155<br>(0.0000335)$^{***}$ [ 0.0000521]$^{**}$ | -0.00135<br>(0.000236)$^{***}$ [ 0.000482]$^{**}$ |

| | | | | |
|---|---|---|---|---|
| Land Ownership Kharif (Yes) | -0.0570<br>(0.0665)[ 0.136] | 0.673<br>(0.463)[ 0.934] | 0.0763<br>(0.0924)[ 0.159] | 0.984<br>(0.651)[ 1.410] |
| Number of schools in a Community or Village | 0.00756<br>(0.0101)[ 0.0279] | -0.0917<br>(0.0695)[ 0.450] | 0.0464<br>(0.0133)$^{***}$ [ 0.0370] | -0.175<br>(0.0923)$^{+}$ [ 0.439] |
| Proportion of Schools with a Library | 0.128<br>(0.168)[ 0.344] | 6.002<br>(1.154)$^{***}$ [ 6.737] | 0.344<br>(0.226)[ 0.449] | 10.10<br>(1.575)$^{***}$ [ 6.294] |
| Distance to nearest bank | -0.0110<br>(0.00423)$^{**}$ [ 0.00989] | 0.412<br>(0.0294)$^{***}$ [ 0.301] | | |
| Province (Punjab) | -0.820<br>(0.0693)$^{***}$ [ 0.166]$^{***}$ | -5.717<br>(0.482)$^{***}$ [ 3.107]$^{+}$ | -0.912<br>(0.0924)$^{***}$ [ 0.225]$^{***}$ | -7.005<br>(0.649)$^{***}$ [ 3.088]$^{*}$ |
| Proportion of Males of age Above 60 in the household | -0.269<br>(0.0611)$^{***}$ [ 0.129]$^{*}$ | 0.451<br>(0.381)[ 0.857] | -0.193<br>(0.0749)$^{**}$ [ 0.156] | 0.0690<br>(0.469)[ 0.876] |
| Number of Female Age 16-59 in the household | -0.0900<br>(0.00297)$^{***}$ [ 0.00517]$^{***}$ | 0.0189<br>(0.0205)[ 0.0423] | -0.0919<br>(0.00396)$^{***}$ [ 0.00666]$^{***}$ | 0.0109<br>(0.0279)[ 0.0418] |
| Average Number of teacher per school with secondary qualification | -0.309$^{*}$<br>(0.138)$^{*}$ [ 0.484] | -7.893<br>(0.910)$^{***}$ [ 6.115] | -0.265<br>(0.168)[ 0.564] | -4.803<br>(1.068)$^{***}$ [ 5.536] |
| Average Teacher Experience per School in the Community | -0.00339<br>(0.0114)[ 0.0314] | 0.633$^{***}$<br>(0.0740)$^{***}$ [ 0.375]$^{+}$ | 0.00755<br>(0.0141)[ 0.0373] | 0.620<br>(0.0990)$^{***}$ [ 0.414] |
| Credit constraint (Yes) | | | 0.338<br>(0.111)$^{**}$ [ 0.263] | 1.388<br>(0.782)$^{+}$ [ 1.534] |
| Constant | 6.389<br>(0.321)$^{***}$ [ 0.424]$^{***}$ | 31.93<br>(2.226)$^{***}$ [ 5.672]$^{***}$ | 6.389<br>(0.415)$^{***}$ [ 0.516]$^{***}$ | 38.39<br>(2.915)$^{***}$ [ 5.058]$^{***}$ |
| N | 3950 | 3950 | 2492 | 2492 |

Standard errors in parentheses, cluster standard errors in square brackets
$^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$

**Table E2:** Above and Below Poverty Line Sample First-Stage Regressions

| | Above Poverty Line: Sample First-Stage | | Below Poverty Line: Sample First-Stage | |
|---|---|---|---|---|
| | (1) Fertility | (2) Student-Teacher Ratio | (1) Fertility | (2) Student-Teacher Ratio |
| Child Sex (female) | 0.134 (0.0674)* [ 0.0706]+ | -0.363 (0.480)[ 0.435] | 0.0317 (0.0884)[ 0.0703] | 0.494 (0.587)[ 0.718] |
| Child Age | 0.141 (0.0748)+ [ 0.0622]* | -0.237 (0.534)[ 0.539] | 0.313 (0.103)** [ 0.0883]*** | -0.0950 (0.682)[ 0.548] |
| Child Age Squared | -0.00727 (0.00375)+ [ 0.00315]* | 0.00796 (0.0268)[ 0.0258] | -0.0160 (0.00512)** [ 0.00447]*** | 0.00417 (0.0340)[ 0.0281] |
| Own Farm Animals (Yes) | 0.116 (0.0842)[ 0.129] | -1.699 (0.599)** [ 1.189] | 0.210 (0.107)+ [ 0.200] | -0.408 (0.713)[ 1.423] |
| Own Non Farm Animals (Yes) | 0.134 (0.0767)+ [ 0.161] | -1.691** (0.547)** [ 1.568] | 0.0862 (0.113)[ 0.231] | 1.275 (0.748)+ [ 1.531] |
| Household Income net of Child Earnings | -24.45 (6.666)*** [ 12.40]* | 282.7 (47.25)*** [ 97.54]** | 17.57 (14.61)[ 24.40] | 57.91 (97.22)[ 204.9] |
| Household Income net of Child Earnings Squared | 51.34 (72.04)[ 119.8] | -2588.7*** (509.7)*** [ 902.5]** | -155.0 (103.8)[ 170.0] | -39.44 (690.3)[ 1416.4] |
| Mother's Level of Education | -0.139* (0.0562)* [ 0.0857] | 0.0392 (0.400)[ 0.713] | -0.0448 (0.131)[ 0.171] | 0.315 (0.866)[ 1.909] |
| Father's Level of Education | -0.00496 (0.0292)[ 0.0393] | -0.945 (0.209)*** [ 0.408]* | 0.106* (0.0508)* [ 0.106] | -0.910 (0.338)** [ 0.603] |
| Land owned in Kharif 2004 | 0.0399*** (0.00500)*** [ 0.0134]** | 0.162*** (0.0354)*** [ 0.0888]+ | 33.63 (7.099)*** [ 13.28]* | 113.8 (45.07)* [ 107.2] |
| Land owned in Kharif Squared 2004 | -0.000175 (0.0000340)*** [ 0.0000629]** | -0.00106 (0.000241)*** [ 0.000439]* | -202.0 (56.52)*** [ 85.72]* | -570.6 (342.3)+ [ 614.0] |

| | | | | |
|---|---|---|---|---|
| Land Ownership Kharif (Yes) | -0.209<br>(0.0818) * [ 0.148] | 1.127<br>(0.583) + [ 0.920] | 0.108<br>(0.123)[ 0.234] | -0.511<br>(0.813)[ 1.656] |
| Number of schools in a Community or Village | 0.0230<br>(0.0126) + [ 0.0321] | -0.180<br>(0.0893) * [ 0.447] | -0.0239<br>(0.0169)[ 0.0322] | 0.0905<br>(0.112)[ 0.485] |
| Proportion of Schools with a Library | 0.228<br>(0.210)[ 0.382] | 4.764<br>(1.481) ** [ 6.891] | -0.233<br>(0.280)[ 0.588] | 7.611<br>(1.857) *** [ 7.549] |
| Distance to nearest bank | -0.0142<br>(0.00589) * [ 0.0170] | 0.298<br>(0.0419) *** [ 0.413] | -0.00876<br>(0.00620)[ 0.00939] | 0.526<br>(0.0412) *** [ 0.187] ** |
| Province (Punjab) | -0.820<br>(0.0881) *** [ 0.201] *** | -5.251<br>(0.625) *** [ 3.510] | -0.772<br>(0.112) *** [ 0.248] ** | -6.152<br>(0.745) *** [ 2.908] * |
| Proportion of Males of age Above 60 in the household | -0.142<br>(0.0774) + [ 0.172] | 1.019<br>(0.496) * [ 1.220] | -0.557<br>(0.101) *** [ 0.174] ** | -0.317<br>(0.638)[ 1.324] |
| Number of Female Age 16-59 in the household | -0.0950<br>(0.00369) *** [ 0.00644] *** | 0.0473<br>(0.0263) + [ 0.0466] | -0.0836<br>(0.00499) *** [ 0.00920] *** | -0.0264<br>(0.0328)[ 0.0774] |
| Average Number of teacher per school with secondary qualification | -0.180<br>(0.182)[ 0.559] | -10.28<br>(1.321) *** [ 7.958] | -0.626**<br>(0.218) ** [ 0.536] | -5.817<br>(1.372) *** [ 5.091] |
| Average Teacher Experience per School in the Community | 0.00226<br>(0.0152)[ 0.0409] | 0.715<br>(0.0986) *** [ 0.418] + | -0.00706<br>(0.0173)[ 0.0352] | 0.638<br>(0.116) *** [ 0.354] + |
| Constant | 6.631<br>(0.398) *** [ 0.487] *** | 33.19<br>(2.828) *** [ 6.642] *** | 5.299<br>(0.730) *** [ 1.096] *** | 26.42<br>(4.858) *** [ 8.757] ** |
| N | 2533 | 2533 | 1417 | 1417 |

Standard errors in parentheses, cluster standard errors in square brackets
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

**Table E3:** Boys' Sample and Girls' Sample First-Stage Regressions

| | Boy's Sample First-Stage | | Girl's Sample First-Stage | |
|---|---|---|---|---|
| | (1) Fertility | (2) Student-Teacher Ratio | (1) Fertility | (2) Student-Teacher Ratio |
| Child Age | 0.238 (0.0832)** [ 0.0706]*** | -0.226 (0.569)[ 0.516] | 0.200 (0.0889)* [ 0.0821]* | -0.229 (0.632)[ 0.532] |
| Child Age Squared | -0.0123 (0.00414)** [ 0.00374]*** | 0.00914 (0.0283)[ 0.0252] | -0.0101 (0.00448)* [ 0.00409]* | 0.00794 (0.0318)[ 0.0272] |
| Own Farm Animals (Yes) | 0.140 (0.0914)[ 0.116] | -1.666 (0.625)** [ 1.062] | 0.180 (0.0959)+ [ 0.132] | -0.630 (0.678)[ 1.019] |
| Own Non Farm Animals (Yes) | 0.138 (0.0869)[ 0.143] | -1.079 (0.595)+ [ 1.373] | 0.124 (0.0915)[ 0.147] | -0.327 (0.650)[ 1.177] |
| Household Income net of Child Earnings | -3.607 (0.630)*** [ 0.963]*** | 28.31 (4.308)*** [ 8.954]** | -2.731 (0.778)*** [ 1.213]* | 17.99** (5.519)** [ 11.19] |
| Household Income net of Child Earnings Squared | 1.208 (0.740)[ 0.990] | -24.50*** (5.064)*** [ 7.989]** | 1.244 (1.063)[ 1.644] | -18.50* (7.525)* [ 12.80] |
| Mother's Level of Education | -0.255 (0.0719)*** [ 0.0740]*** | -0.175 (0.492)[ 0.805] | 0.0230 (0.0733)[ 0.107] | 0.521 (0.522)[ 0.917] |
| Father's Level of Education | 0.0497 (0.0365)[ 0.0531] | -0.794 (0.249)** [ 0.420]+ | 0.0180 (0.0353)[ 0.0476] | -1.007 (0.251)*** [ 0.399]* |
| Land owned in Kharif 2004 | 0.0380 (0.00545)*** [ 0.00921]*** | 0.141 (0.0373)*** [ 0.0872] | 0.0147 (0.00956)[ 0.0193] | 0.286 (0.0679)*** [ 0.105]** |
| Land owned in Kharif Squared 2004 | -0.000186 (0.0000354)*** [ 0.0000384]*** | -0.000914*** (0.000242)*** [ 0.000397]* | 0.000154 (0.000120)[ 0.000239] | -0.00283*** (0.000850)*** [ 0.00129]* |
| Land Ownership Kharif (Yes) | -0.0326 (0.0909)[ 0.154] | 0.816 (0.622)[ 0.834] | 0.00241 (0.102)[ 0.157] | 0.0276 (0.724)[ 1.349] |

| | | | | |
|---|---|---|---|---|
| Number of schools in a Community or Village | 0.0164<br>(0.0139)[ 0.0303] | 0.0236<br>(0.0946)[ 0.468] | -0.00127<br>(0.0146)[ 0.0272] | -0.218*<br>(0.102)* [ 0.432] |
| Proportion of Schools with a Library | 0.230<br>(0.230)[ 0.359] | 3.807<br>(1.565)* [ 6.654] | 0.0666<br>(0.244)[ 0.401] | 8.206<br>(1.701)*** [ 6.978] |
| Distance to nearest bank | -0.00718<br>(0.00596)[ 0.00966] | 0.445<br>(0.0407)*** [ 0.320] | -0.0145<br>(0.00598)* [ 0.0103] | 0.377***<br>(0.0425)*** [ 0.287] |
| Province (Punjab) | -0.894<br>(0.0948)*** [ 0.183]*** | -5.805<br>(0.648)*** [ 3.133]+ | -0.767<br>(0.101)*** [ 0.176]*** | -5.474<br>(0.711)*** [ 3.071]+ |
| Proportion of Males of age Above 60 in the household | -0.147<br>(0.0839)+ [ 0.155] | 1.237<br>(0.548)* [ 1.319] | -0.403<br>(0.0889)*** [ 0.126]** | -0.312<br>(0.524)[ 0.648] |
| Number of Female Age 16-59 in the household | -0.0898<br>(0.00410)*** [ 0.00521]*** | 0.0309<br>(0.0280)[ 0.0357] | -0.0901<br>(0.00427)*** [ 0.00643]*** | -0.000154<br>(0.0298)[ 0.0537] |
| Average Number of teacher per school with secondary qualification | -0.375<br>(0.191)* [ 0.427] | -8.188<br>(1.283)*** [ 6.629] | -0.179<br>(0.199)[ 0.583] | -7.729<br>(1.267)*** [ 5.460] |
| Average Teacher Experience per School in the Community | -0.00719<br>(0.0155)[ 0.0296] | 0.663<br>(0.103)*** [ 0.402]+ | -0.000456<br>(0.0168)[ 0.0370] | 0.614<br>(0.105)*** [ 0.342]+ |
| Constant | 6.370<br>(0.431)*** [ 0.478]*** | 30.86<br>(2.947)*** [ 5.758]*** | 6.359<br>(0.478)*** [ 0.541]*** | 33.32<br>(3.364)*** [ 5.796]*** |
| N | 2086 | 2086 | 1864 | 1864 |

Standard errors in parentheses, cluster standard errors in square brackets
+ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

# Chapter 6 Policy for Raising School Quality and School Attendance: Private versus Public Schools

## 6.1 Introduction

We have established that rural Pakistani families value school quality highly and that school quality is a key determinant of whether they send their children to school or not. The problem for policy then is how to improve school quality. In urban Pakistan private education has been expanding very rapidly as a response to this demand for quality, in contrast to the countryside.

Whether school quality should be improved by public or by private provision of education is a controversial matter. For many years private schools played only a marginal role in most developing countries (Lockheed and Jimenez 1994). Governments occasionally prohibit, frequently regulate, and often ignore private schooling (Kingdon 1996). The main stated reason for government's reluctance to accept private sector's role in national education policy is equity. Universal/equal access to schooling supposedly lowers earnings inequality, which reduces governments' need to undertake controversial asset or income transfers. Consequently, private schooling is not used as an instrument of mass education on the presumption that it typically serves the elite (Andrabi et al. 2006a, 2008). In Pakistan all major industries including private schooling were nationalized in 1972, but only for seven years; denationalization of private schooling was announced and undertaken in 1979.

Private schooling is often allowed on the grounds simply that the state in developing countries faces an overwhelming task in expanding educational services for a rapidly increasing population and with tight budgets. Another argument favouring private provision of education is the inefficiency of public educational expenditures; ill-timed and inadequate school supplies, building of schools in places where they are not needed and paying salaries to poorly performing teachers (Alderman et al. 2001) (Coulson (2009) reviews more than 150 studies of private and public provision of education across developed and developing countries). It is also contended that private schools, by charging fees for their services, are more accountable to parents, as well as more efficient. Competition between private and public schools may, under certain conditions, improve the efficiency of government schooling (Aslam 2009). As noted by Alderman et al. (2001), parents frequently respond to perceived inadequate public education services by enrolling their children in private schools.

Some years ago it was contended that private schools only cater to the rich in Pakistan (Jimenez and Tan 1987) but later, according to Alderman et al. (2001) even poor households were able to pay for high-quality private schools. But private schools offering education at sufficiently low fees to attract poor households might provide a poor quality service. Andrabi et al. (2002) investigate whether in Pakistan private schools that supply the poor usually exploit often illiterate parents of low income, who are unable to assess if their children are learning or not. They concluded that parents are able to infer quality variation between schools and that fees respond in predictable ways to school inputs.

The present chapter discusses the desirability of raising school quality by expanding private education. It investigates schooling cost and quality differences between public and private schools in rural Pakistan. A brief investigation is also undertaken into factors influencing the establishment of private schools in rural Pakistan. A critical element for policy, discussed in the present chapter, is the full cost of private school attendance in relation to rural incomes.

This chapter is organized as follows. Section 6.2 is an overview of private schooling in Pakistan. Section 6.3 describes private-public schooling costs in rural Pakistan (PRHS-II). Evidence on supply of private schools in the communities is presented in section 6.4. Section 6.5 presents school quality and school fee differences between public and private schools. Child cognitive achievement differences between public and private schools are presented in section 6.6, while section 6.7 concludes.

## 6.2 Overview of Private Schooling in Pakistan

Pakistani student performance is poorer than their counterparts in other countries in the region, except Iran. The major reasons for the low performance have been expressed in the Government's education policy draft (GoP 2009, p.4) as;

> Low access rates can also be attributed to the lack of confidence in the public sector schools to deliver quality education which has convinced parents either to shift their kids to private schools or absorb additional financial burden by arranging private tuitions. Where neither is affordable the households prefer to have their children drop out from school and join income earning activities. The average student of the public sector education system cannot compete in the job market. This leads to social exclusion of the already poor. The decline has primarily resulted from political interference and corruption that has permeated the entire sector. Recruitments, transfers and postings became politically driven. Absentee teachers and ghost schools have been discovered under various exercises. Cheating in examinations is a widespread phenomenon. Primary sufferers are the most poor and underprivileged in the system. Those who make it to higher education in the public sector cannot get employment due to absence of merit or poor quality of their educational abilities.

The private sector share in overall enrolment has been increasing in recent years and is partly a reflection of low quality education in public sector institutions (GoP 2009).

In Pakistan, private and public school options are broadly similar in their educational provision, but both types differ in terms of regulation and financial arrangements. The state schools are financed by the government (central and provincial). There are low nominal fees in state schools; however, low fees are always coupled with significant school related expenditures in the form of transport, books, uniform and examination fee and even admission fees.

Concerns are also raised in policy circles that the curriculum in Deeni Madrassahs[50] lies outside the domain of mainstream education (Aslam 2009). However, Deeni Madrassahs form a very small[51] component of the private institutions (Andrabi et al. 2006).

It is also suggested that private sector establishments are not always regulated properly by the state which leaves students unprotected. By law the schools are required to register with the provincial authorities but most often they do not comply with the regulations. In Pakistan legislation for regulating private schools includes specific directives regarding school inputs to be shown or evaluated at the time of registration. Registration involves recording detailed information on school facilities and other equipment like number of blackboards, cupboards, maps etc (Rose 2007). Private schools are inspected only at the time of their initial registration by the district education officer or other relevant personnel, or in the event of a complaint against them being received. Private[52] schools are for profit and most often owned and managed by an individual entrepreneur. Private schools are supposed to be registered with the education department in each province but 29 percent fails to do so (FBS 2000). Similarly, those private schools that allow their students to sit for public examinations are required to be registered and government-recognised. Often unregistered private schools avoid this obstacle by sending their students to state examinations as 'Private/home candidates[53]'. These unregistered private schools, often of poorer quality, evade large amounts of government taxation and other costs (Aslam 2009). Consequently the true size of unregistered schools remains unmeasured in government statistics. According to Aslam

[50] Deeni Madrasahs means religious institutions.
[51] We do not have any children in Deeni Madrasahs in our sample. But discussion is provided in this chapter as appropriate.
[52] These schools are secular in nature. The majority of them are private unaided schools and do not receive any funding either from the government or other private internal or external agencies. Religious schools play only a small role in Pakistan and their enrolment share is only one percent (lower than enrolment in religious institutions in the US) in a recent study by Andrabi et al. (2006).
[53] In Pakistan children can sit state examinations even if they do not formally go to school.

(2009) the Federal Bureau of Statistics' (2000) claims of interviewing and documented both unregistered and registered schools is debatable as there was no comprehensive list of the unregistered schools in the census.

The lack of discipline in recognition and registration points to a lax regulatory environment. This may have significant implications for the quality of schooling offered. By law private schools needs to charge fees approved by the educational department. Even so, the fee in private schools ranges from exorbitantly high to relatively small amounts (Alderman et al. 2001; Aslam 2009). It is believed that registered private schools often charge higher fees than those authorised by law (GoP 2009). Similarly, private schools should enrol 10 percent poor but outstanding students free of cost, but they seldom do (Aslam 2009).

Andrabi et al. (2006a, 2008) nowhere refer to the dimension of registered and unregistered schools. They tested private school availability but did not investigate the quality dimension of public and private schools. They set forth their hypothesis (lower fees on account of lower costs due to locally available female teacher) in a framework where government policy toward private schools is one of laissez faire, which is nowhere evident from the government policy guidelines about private schools in Pakistan. They need to be registered, set school fees within stated rules, provide admission concession to needy and outstanding students, and operate within stated framework of monitoring and supervision to maintain quality. Individual entrepreneurs opting for establishing private school in Pakistan would have two options in this scenario; to comply with the stated rules and guidelines from the government or not to comply with it. If the individual entrepreneur does not comply with the rules, then he would have already acquired some political leverage to avoid the unlikely circumstances if the stated rules are enforced effectively. However, political leverage might also be important, if individual entrepreneurs need help to register their school with the provincial/district level authorities. In the scenario of weak local level monitoring and supervision, variables proxying political strength in the local communities would be correlated with the supply of private schools. Private school supply in this case would also depend on the local political structure.

### 6.2.1 The Rise of Private Schooling in Pakistan

The Federal Bureau of Statistics undertook a census of private institutions in the country in 2000. The statistics from FBS 2000 show that the number of private institutions in Pakistan was 36,096 with 6.1 million children enrolled (which is not identical with attendance, the

principal concern of this thesis). Enrolment in urban areas was 69.2 percent of the private total (Table 6.1).

Table 6.1: Urban and Rural Private School Enrolment in Pakistan in 2000

| Region | Enrolment | Percent |
|---|---|---|
| Rural | 18,76,839 | 30.8 |
| Urban | 42,12,300 | 69.2 |
| Total (Pakistan) | 60,89,139 | 100 |

Source: Calculated from FBS 2000

According to Jimenez and Tan (1987) there were 3343 private institutions in the four main provinces of Pakistan in 1983. Private schools then increased to 36096 by the year 2000, a rise of more than 900 percent over the eighteen years (Table 6.2).

Table 6.2: Increase in Number of Private Institutions from 1983-2000 in Pakistan (FBS)

| Year | Private Institutions (Number) |
|---|---|
| 1983 | 3,343 |
| 2000 | 36,096 |
| Percent increase | 980 |

Much of the rapid increase of private schools has occurred in Punjab where 56 percent of the country's total population lives. Out of 36,096 private institutions in the country reported by FBS (2000), 66.4 per cent (24,325) were in Punjab. The FBS census figures show, that 61 per cent (22,004) of all private institutions are in urban areas.

After the census of private schools in 2000 the Federal Bureau of Statistics (FBS) undertook the National Education Census (NEC) for the first time in the 2005[54] to report on all public and private educational institutions. There were 76,047 private institutions in the country in 2005, catering to 12.1 million children in the country. The number of private institutions in rural and urban areas was 35,438 (46.6 percent) and 40,609 (53.4 percent) respectively. The share of children enrolment in these institutions in rural areas was 37.43 percent (Table 6.3).

[54] The FBS (2000) is the survey of private institutions only in the country while the FBS (2005) is a National Education Census of both public and private institutions in the country. Neither of these surveys are household surveys. The survey used in the paper, PRHS-II, is a household survey with children currently in school. The survey recorded school information of currently enrolled children, as well as, visited these schools for further in-depth statistics/information/school quality data on these schools. The FBS statistic or percentages on rural/urban child enrolments are out of total child enrolments in private schools/institutions. The stated figures also include enrolment shares of different level of institutions, higher than the age range adopted in PRHS-II survey. Whereas the PRHS-II percentages on child enrolment in private schools are relative to child enrolment in government schools (forthcoming). Calculating the same percentage figure from FBS (2005) census data for rural areas of the whole country, for the child age range adopted in the paper, the share of private schools enrolment relative to public schools stands at 21 percent which is largely consistent with the PRHS-II survey. The share of private schools enrolment stands at 19.05 percent in rural Punjab. There is no further bifurcation of total child enrolment over each province other than over rural and urban areas of the whole of Pakistan in FBS (2005) census data. The FBS census was undertaken in later 2005 while the PRHS-II survey was undertaken in 2004.

According to the Pakistan Economic Survey (2011-12), Pakistan's rural population is 63 percent, nearly sixty percent of which is below thirty years of age. The number of private institutions as well as enrolment in private sector has doubled (over the span of just four years time period) between 2000 and 2005; an increase of almost 110 percent.

Table 6.3: Private Institutions, Enrolment and Teaching Staff by Region (Urban and Rural) in Pakistan (2005)

| Region | Private Institutions (Number) | Percent | Enrolment | Percent | Teaching staff | | |
|---|---|---|---|---|---|---|---|
| | | | | | Male | Female | Total |
| Rural | 35,438 | 46.6 | 4,536,690 | 37.43 | 88,132 | 123,760 | 211,892 |
| Urban | 40,609 | 53.4 | 7,584,704 | 62.57 | 125,126 | 295,908 | 421,034 |
| Total (Pakistan) | 76,047 | 100 | 12,121,394 | 100 | 213,258 | 419,668 | 632,926 |

Source: Calculated from FBS 2005

Another notable point from the FBS (2005) is that the proportion of female teaching staff in private institutions is higher than male staff. The total teaching staff serving in these institutions was 632,926 (Table 6.3). The female teaching staff accounted for 66.3 percent of the total teaching staff. In rural areas the share of female teaching staff stands at 58.4 percent while the same proportion for urban areas is 70.3 percent. Andrabi et al. (2006a, 2008) argued that the rise of private schools in Pakistan shows that a greater and increasing proportion of children enrolled in these schools are from lower and middle class families. The average fee of rural private school is estimated to be less than Rs.6 per day. The key to this lower fee (but not the lower cost, as argued below), apart from lower rural incomes on the demand side, is the local availability of moderately educated female teachers who have limited alternative employment opportunities outside the local village. These teachers are hired at low wages which keeps the cost down and enables these schools to pass on savings to the parents through lower school fees.

Summarizing the discussion of private schooling phenomena in Pakistan, several notable patterns are painted in the following lines.

- Private schools are much more abundant in urban areas.
- Private schools, as well as enrolment in these establishments, have been strongly increasing in Pakistan in both urban and rural areas.
- This increase is attributable to the lack of confidence in public sector education provision on the grounds of the poor quality of education provided (GoP 2009).

When private schools are not available parents prefer to keep their children out of schools or send them to join the local labour market (GoP 2009).

**Education Provision: Defining Private Provision**

Private education provision is defined by Kitaev (1999, p.43), cited in Rose (2007, p.2) as:
all formal schools that are *not* public, and may be founded, owned, managed and financed by actors other than the state, even in cases when the state provides most of the funding and has considerable control over these schools (teachers, curriculum, accreditations etc) e.g. schools are also run by army in Pakistan for selected member's children. The definition points to the nature of complexity in distinguishing between public and private spheres in education (Rose 2007). Private sector schools are often subsidized, but not in Pakistan, by the government who usually bear the cost of teacher training, examinations, inspection and curriculum development, even if they are managed and owned by private sector or individual entrepreneur. Commonly the state regulates private sector education.

The draft Pakistani National Education Policy (2009) established an array of regulatory frameworks to be operated – but there has been no implementation. Despite these policy guidelines private schools in Pakistan operate in a lax regulatory environment (Aslam 2009).

Non-state actors in Pakistan include profit-oriented private entrepreneurs, faith based/charitable organizations and NGOs. Most often they have different motives for involvement in education sector. Some institutions run on commercial business principles but claim to be motivated by concerns for the poor communities. Similarly NGOs may operate purely on philanthropic principles but sometimes they are established as a means to acquire donor resources (Rose 2007).

### 6.3 School Choice: Schooling Cost in Rural Pakistan

The descriptive statistics in Table 6.4 show child school participation by school type. The proportion of government school children in the survey is 87.5 percent. Private schools for profit educated 10.9 percent of sampled children. A small number of children (1.6 percent) participated in a mix of other schools including those provided by NGOs and those administered by the Pakistani army. However, to conform with the literature (which, regards 'for profit' and non-profit schools run by different entities other than the state as private

schools), children attending 'other' schools have been combined with those at 'for profit' private schools[55].

Table 6.4: Child School Participation by School Type

| | | Currently in School? |
|---|---|---|
| Type of school | | Yes |
| Government | Count | 1708 |
| | % of Total | 87.50 |
| Private | Count | 212 |
| | % of Total | 10.90 |
| Other | Count | 32 |
| | % of Total | 1.60 |
| Total | Count | 1952 |
| | % of Total | 100.00 |

Source: Calculated from PRHS-II

### 6.3.1 Private Schooling Cost

Andrabi et al. (2002) showed for rural as well as urban Pakistan that the distribution of annual fees in private schools is centred on a low median. The highest median school fees (Rs.145 per month) was reported for the urban areas of Balochistan. In rural regions it drops to Rs.112 per month. The lowest private school fees were reported for the Punjab, Rs.71 per month and Rs.53 per month for urban and rural regions respectively. From household expenditure data in the Pakistan Integrated Household Survey (PIHS), in the Punjab tuition fees account for 1.7 percent of the average household expenditure in the countryside. In urban areas of Punjab fees amount to 2.1 percent of average household expenditure.

Table 6.5: Median (per Month) Fee in Rural and Urban Pakistan

| | Rural | Urban |
|---|---|---|
| Province | Median fee per Month | Median fee per Month |
| NWFP | 100 | 111 |
| Punjab | 71 | 53 |
| Sindh | 95 | 108 |
| Balochistan | 112 | 145 |

Source: Calculated/adopted from Andrabi et al. (2002)

[55] Figures also support the FBS (2000) estimates (discussed above) that there has been a stronger growth of private institutions in the Punjab than anywhere else in the country in the last two decades.

In this section we show that Alderman et al.'s (2001) and Andrabi et al.'s (2002) conclusion regarding lower private school fees generally holds in this survey. However, households that send their children to private schools have higher total household expenditure (per capita expenditure) than those that send their children to government schools. Private schooling is for relatively well-off households in rural Pakistan. Table 6.6 summarizes the distribution of private school fees in rural Pakistan.

Table 6.6: Annual Fees for Private Schools in Rural Pakistan

| Stats | School Fees (Rs) |
|---|---|
| mean | 1550.28 |
| Standard deviation | 2184.69 |
| minimum | 0 |
| maximum | 18000 |
| p25 | 300 |
| p50 (Median) | 960 |
| p75 | 1800 |
| Inter quartile range | 1500 |
| N | 244 |

Source: Calculated from PRHS-II

Annual average private school fees in rural Pakistan stand at Rs.1550.28, which amounts to Rs.172 per month. To assess the affordability of these fees, household expenditure data, as well as annual incomes of the daily wage male workers (based on wages of unskilled labourers and semi-skilled or types of skilled manual workers like masons, or carpenters) have been used from PRHS survey. In rural Pakistan the mean tuition fee is 1.63 percent of the mean household expenditure as given in Table 6.7 below. In the lower income bracket of the expenditure distribution, tuition fees are 3.3 percent of the spending of those household's that fall in the $25^{th}$ percentile of the expenditure distribution. Thus an average household in the rural areas of Pakistan with three children will have to spend 4.89 percent of their annual budget if their children are admitted to an average private school. However, this 4.89 percent figure is an underestimate because of schooling costs such as transport, uniforms and expenses incurred on stationery (Table 6.8). The total of all these sub-heads of schooling expenditures represents 7.18 percent ((6808.83÷ 94879.2) ×100=7.18) of the average household income given in Table 6.7. So a household with three children on average would need to spend 21.5 percent of their annual budget on education.

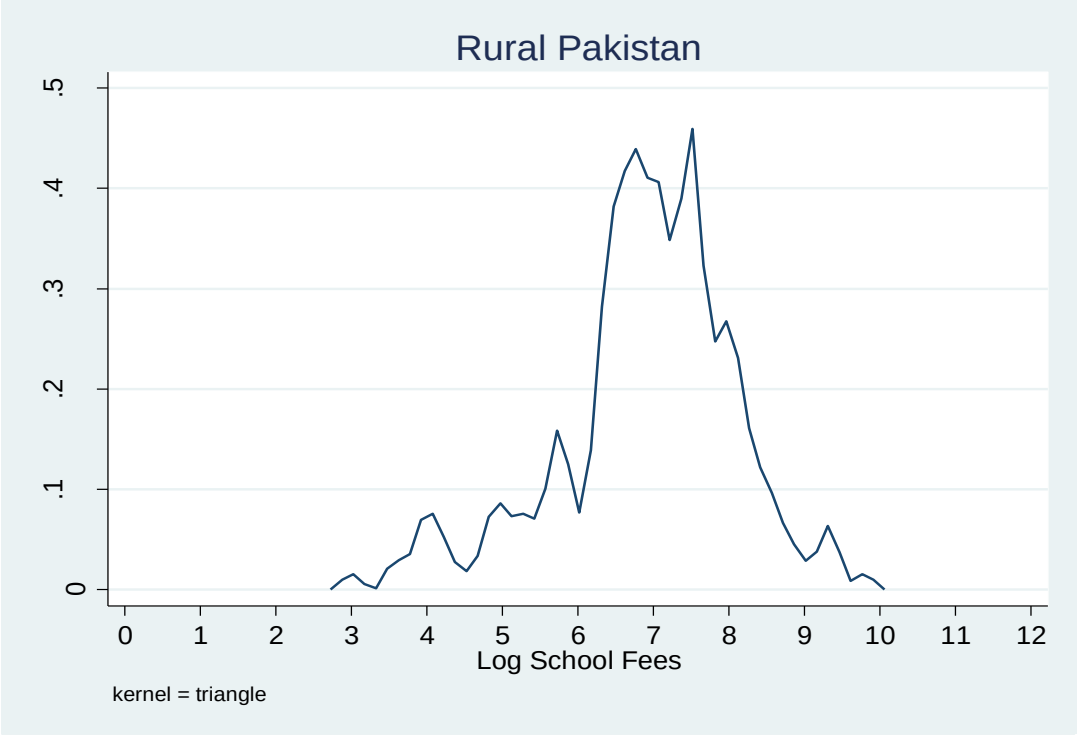
Table 6.7: School Fees as Percentage of Household Expenditure

| Total (Sample) | | | | |
|---|---|---|---|---|
| Stats | Total Household Expenditure | Household Expenditure (Per capita)[(a)] | School fee (% of Total H. Expenditure) | School fee (% of Per capita Expenditure) |
| mean | 94879.2 | 15989.5 | 1.63 | 9.7 |
| Std. Dev. | 79579.4 | 11807.5 | 1.95 | 13.13 |
| min | 8308 | 1978.1 | 18.66 | 78.37 |
| max | 771996 | 156364 | 0.2 | 0.99 |
| p5 | 24730 | 4891.87 | 6.27 | 31.69 |
| p10 | 31920 | 5899.08 | 4.86 | 26.28 |
| p50 | 74695 | 12946.3 | 2.08 | 11.97 |
| N | 3950 | 3950 | | |

(a): This variable has also been made consistent over all households using the OECD (1982) Equivalence Scale (also called Oxford scale)
Source: Calculated from PRHS-II

Another important observation on schools fees given in Table 6.6 (above) is the higher mean compared to the median, which suggests that the distribution is skewed to the right, showing that a high concentration of schools are charging lower fees consistent with Andrabi's analysis.

Figure 6.1: Log Fees Distribution of Private School Children in Rural Pakistan (only for Profit Schools).




Source: Calculated from PRHS-II

Figure 6.1 shows the kernel density of log fees of private for profit schools in rural Pakistan at the time of PRHS-II. Schools charging no fees have not been included in the kernel density construction. It is evident that there are a lot of schools in the left tail of the distribution charging lower fees. The figure also shows highest concentration of fees around the median.

The average annual income of the daily wage paid male workers in the surveyed communities is approximately 27,600 rupees given in Table 7.8 (for calculation of annual income, six days/week work has been assumed (44 weeks/ year). It is an upper bound in the prevailing economic and employment conditions in rural Pakistan). The female average annual income is about 23,000 rupees. The average private school fee, per child enrolled is 5.6 percent of male and 6.8 percent of female average annual income.

These percentages are approximately equivalent to the 5$^{th}$ and 10$^{th}$ percentile of the household expenditure distribution in Table 6.7 (above). These estimates show that the private school fee represents a small proportion of the annual average income of the daily wage labourer in rural Pakistan [56]. However, no one would deem school fees alone sufficient for school access; schooling involves significant costs in the form of transportation costs, uniform costs and

[56] Household expenditure combines earnings of more than one earner in the household and so is higher than the average earnings of skilled labourer in Pakistan.

stationary (Table 6.8). For a daily wage male worker and including all other important sub-heads, school costs per child represents 24.62 percent ((6808.83÷27653.94) ×100=24.62) of his annual average income. The estimate for three children of school age amounts to 73.86 percent (24.62 ×3=73.86) of the annual budget.

Similarly, school expenditure for one child in terms of female annual income amounts to 30 percent ((6808.83÷22690.58) ×100=30). In the event of having three children, she would need to spend 90 percent of her annual income to get them educated in an average private school in rural Pakistan.

Table 6.8: Mean and Standard Deviation of Annual Household per Child Investment in Public/Private Schools and of Annual Income of Skilled/Semiskilled Wage Labourer

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Per pupil investment (state schools)** | | | | | |
| Transport cost per student | 1708 | 951.31[(a)] | 5608.96 | 0 | 72000 |
| School fees per student | 1708 | 194.95[(b)] | 720.71 | 0 | 12000 |
| Uniform cost per student | 1708 | 324.87[(c)] | 489.23 | 0 | 6000 |
| Expenses on stationary per student | 1708 | 425.67[(d)] | 638.72 | 0 | 8000 |
| Aggregate expenditure | 1708 | 1896.8 | | | |
| **Per pupil investment (private schools)** | | | | | |
| Transport cost per student | 244 | 3841.87[(e)] | 9727.12 | 0 | 43200 |
| School fees per student | 244 | 1550.28[(f)] | 2184.7 | 0 | 18000 |
| Uniform cost per student | 244 | 581.07[(g)] | 515.57 | 0 | 2510 |
| Expenses on stationary per student | 244 | 835.61[(h)] | 714.91 | 0 | 3200 |
| Aggregate expenditure | 244 | 6808.83 | | | |
| **Annual income of unskilled/semi-skilled/skilled daily wage labourer [(i)]** | | | | | |
| Annual income (male) | 3950 | 27653.9 | 15291.7 | 12480 | 115128 |
| Annual income (female) | 3950 | 22690.6 | 14063.1 | 7800 | 115128 |
| t-value of difference in (a) and (e) | | | | | -6.73 (0.000) |
| t-value of difference in (b) and (f) | | | | | -19.33 (0.000) |
| t-value of difference in (c) and (g) | | | | | -7.59 (0.000) |
| t-value of difference in (d) and (h) | | | | | -9.23 (0.000) |

(i): Based on community level wages; includes daily wages for manual labourers, masons, carpenters. For calculation of annual income 6 days/week work has been assumed.
Source: Calculated from PRHS-II

The mean expenditure of rural households that send their children to private schools is more than one half greater than the mean expenditure of those households that send their children to government schools (Table 6.9). The distribution of household expenditure for different

income profiles (percentiles) shows the same picture. For those households that send their children to private schools and belong to the lowest income bracket ($5^{th}$ and $10^{th}$ percentile), expenditure would be nearly double that of those households that send their children to government schools and belong to the same income group ($5^{th}$ and $10^{th}$ percentile). The arrangement of children by school choice (no-schooling, government and private schooling) shows the same picture (Table 6.10). Nearly 77 percent of children who are in private schools are from the top two quintile of the rural income distribution. Children in private schools are comparatively better-off than children currently not at school or in government schools.

Table 6.9: Household Expenditure by School Type (Public and Private)

| | Government | | Private | |
|---|---|---|---|---|
| Stats | Total Household Expenditure | Household Expenditure (Per capita) [(a)] | Total Household Expenditure | Household Expenditure (Per capita) [(a)] |
| mean | 101146.80 | 17198.06 | 161315.90 | 25662.68 |
| Std. Dev. | 83630.26 | 12704.90 | 126079.40 | 16066.40 |
| min | 8308.00 | 1978.10 | 21512.00 | 4438.55 |
| max | 771996.00 | 156363.50 | 771996.00 | 107225.80 |
| p5 | 27787.00 | 5244.92 | 47328.00 | 7951.55 |
| p10 | 34717.00 | 6557.50 | 54598.00 | 10964.71 |
| p50 | 79853.00 | 13876.99 | 121484.00 | 21252.21 |
| N | 1708 | 1708 | 244 | 244 |

(a): This variable has also been made consistent over all households using the OECD (1982) Equivalence Scale (also called Oxford scale)
Source: Calculated from PRHS-II

Table 6.10: Child Enrolment by Total Household Expenditure (Percentile)

| | school choice | | | |
|---|---|---|---|---|
| Total Household Income net of Child Earnings (expenditure) (percentile) | No-schooling | Government | Private | Total |
| ≤10 % (Rs. 31920) | 254[(a)] (12.71) [(b)] | 138 (8.08) | 4 (1.64) | 396 (10.03) |
| 10% to 25% (Rs. 31921-Rs.46804) | 354 (17.72) | 230 (13.47) | 8 (3.28) | 592 (14.99) |
| 25% to 50% (Rs. 46805-Rs. 74695) | 519 (25.98) | 423 (24.77) | 45 (18.44) | 987 (24.99) |
| 50% to 75% (Rs. 74696-Rs.109901 | 500 (25.03) | 437 (25.59) | 52 (21.31) | 989 (25.04) |
| >75% (Rs.109902) | 371 (18.57) | 480 (28.10) | 135 (55.33) | 986 (24.96) |
| Total | 1998 (100.00) | 1708 (100.00) | 244 (100.00) | 3950 (100.00) |

(a) Shows number (frequency).
(b) Shows percent of total in the same column.
Source: Calculated from PRHS-II

Comparison of private school fees exclusive of other schooling costs with the total household expenditure and income of ordinary labourer in rural Pakistan supports the findings of previous studies (Alderman et al. 2001; Andrabi et al. 2002; Andrabi et al. 2011) that private schools need not supply only the rich/elite but are also accessible to the poor. But there are many other bigger costs involved and discussed above. Additionally, the opportunity cost of child labour needs to be factored in.

### 6.4 Private School Supply

In chapter 5 (appendix B) we have discussed why the availability and quality of local schools in Rural Pakistan is not a direct response to village characteristics and households demand but is determined by district and higher level decisions above the villages. In chapter 5 (section 5.4), we tested whether the quality of the available schooling was sufficiently high to cover the opportunity cost of children attending any sort of school. The school quality variables were treated as an aggregate across the schooling options available in a local community. This approach was adopted on the grounds that nearly 90 percent of children participated in government schools. However, as already highlighted, the debate on school quality often centres on the public versus the private provision of education. This section isolates factors that might be important in the establishment of private schools in rural Pakistan with regression analysis.

Alderman et al. (1996) and Behrman et al. (2008) treated school presence or school quality as exogenous to child schooling. But their argument might not be correct in the case of private school supply. As market rather than government institutions, private school availability might respond to household level and community level characteristics so that these characteristics significantly affect private school supply.

As noted earlier that the available studies on the supply of social services like; Rosenzweig and Wolpin (1986), Behrman and Birdsall (1988), Pitt et al. (1993), and Gershberg and Schuermann (2001) have emphasized that the provision of local social services, including state schools, may respond to local demands or political pressures for schooling among other services. In such a scenario, higher income villages or villages with more political power would have a higher probability of having schools. Alternatively, due to concerns about equity in social services, higher income villages might have lower school availability.

Alderman et al. (1996) assess the plausibility of the notion in rural Pakistan that whether a sample village has a primary school depends on village average household income, the presence of Union Council[57] member in the village (which reflects the strength of linkages between the village and regional government), population and other village characteristics. Their assessment confirms that governmental decisions regarding school supply is taken without reference to observed indicators of village-level demands. Similarly, Behrman et al. (2008) also assumed the availability and quality of local schools in rural Pakistan as exogenous for their estimates. Our assessment confirmed their argument at least in the case of government supply of education.

We used the same covariates for private school availability as for government schools. An educational entrepreneur would be likely to take into account the average household income in the community when deciding whether to establish a private school. Average village level household income is used. Political strength (local councillor or *biradary* leader and other indicators available in the data set) might help in the private school registration process. Or in some cases the individual entrepreneur might be the person with sufficient political influence to avoid registration and so evade taxation without penalty even if identified by the local administration. Other controls adopted in the analysis are population and other village level characteristics. The survey (PRHS-II) has data for 94 villages with which we investigate the relation between local private school availability and village characteristics in sample villages. Private schools are present in 41 out of 94 communities and a local councillor is present in 53 out of 94 communities. There are 28 communities where both private schools and local councillors are available (Table 6.11).

The probit and tobit equations in Table 6.12 have the same specifications. The probit equation (1) estimates the probability of a community having a private school; it also contains measures for mean household income, distance to the *mandi* (local market), political strength and population of the village. The tobit estimates the number of private schools in the community. Contrary to the mainly government school results presented earlier, private school availability does depend on community political strength, average household income and other community level characteristics.

The present data do not differentiate between registered and unregistered private schools; both types are included in the estimation process. It might be of interest to know whether

[57] There are three tiers of local government system in Pakistan; Union Council, Tehsil Council and District Council. Union Councils are the lowest level of the local government system.

political support is more important to obtain registration or to avoid the registration process. As argued before, registering a private school in Pakistan is complicated (Rose 2007), and so a large proportion of schools operate without formal registration (Aslam 2009). Hence it is certain that political support is important for either route to running a private school business in Pakistan.

Table 6.11: Private School and Local Councillor in the Communities

| Private School in the Community | Local Councillor in the Community | | |
|---|---|---|---|
| | no | Yes | Total |
| no | 43 | 10 | 53 |
| yes | 13 | 28 | 41 |
| Total | 56 | 38 | 94 |

Source: Calculated from PRHS-II

Table 6.12: Probit/ Tobit Estimates of Impact of Village Characteristics on Private School Availability.

| | (1)<br>Probit<br>Presence of<br>Private School | (3)<br>Tobit<br>No. Of Private School<br>In the community |
|---|---|---|
| Mean Household Income | 0.00000697 | 0.0000188* |
| | (1.71) | (2.22) |
| Local-Councilor (yes) | 0.852** | 2.014*** |
| | (3.09) | (3.74) |
| Distance to the Mandi | -0.00802 | -0.0296 |
| | (-0.47) | (-0.92) |
| Population | -0.000165 | -0.000574 |
| | (-1.12) | (-1.80) |
| Constant | -0.982* | -2.137* |
| | (-2.10) | (-2.17) |
| sigma _cons | | 2.309*** |
| | | (8.91) |
| Pseudo R-Square | 0.103 | 0.0722 |
| Wald chi2 (4) | 12.12 | |
| F (4, 90) | | 4.803 |
| p | 0.0165 | 0.00147 |
| N | 94 | 94 |

*t* statistics in parentheses for tobit equations, z statistics for probit equations; (based on robust standard errors)
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Obs. summary: 53 left-censored observations at Number of Private School in the Community<=0
41 uncensored observations
0 right-censored observations

## 6.5 Government versus Private School Fees and Quality Differences

### 6.5.1 School Fees Differences

To see how much more expensive private schools are than state schools, Table 6.8 (above) reports household expenses on child schooling in public and private schools. Annual average school fee per student in the government school stands at 194.95 (Rupees) while the same figure for private school students stands at 1550.28 (Rupees). All figures are annual. The t-test of differences shows that per student expenditure, under the sub-heads of expenditure, in private schools is significantly higher than in government schools (consistent with the Aslam (2009) finding from Pakistan Integrated Household Survey (PIHS-2002)).

### 6.5.2 School Quality Differences

Contrary to earlier work from Pakistan (Alderman et al. 2001; Andrabi et al. 2002; Andrabi et al. 2011), private schools do not cater to those households who fall in the lowest income bracket of the income distribution, once generalized for average number of children per couple in the survey. However, the finding is driven largely by other schooling costs in addition to school fees. A large fraction of private schooling costs goes on transportation charges. This is because private schools are less available in the community; being located farther from households, they cost more to access. Improved availability of these schools would certainly enable poor households in rural Pakistan to consider these schools for child education. However, even if these schools are close to households offering education at lower fees, concern also arises about the quality of these schools. They might be of poorer quality than state schools and merely exploiting the poor.

In Table (6.13), the t-tests of differences in means of the quality variables of government and private school are based on individual children who participated either in government or private schools.

Table 6.13: T-test of Differences in Means of Selected School Quality Indicators in Public and Private Schools

| **School Inputs** | **Government** | **Private** | **t-value of difference (Government –Private)** |
|---|---|---|---|
| Student teacher ratio | 35.97 | 25.72 | 11.61 (0.000) |
| Number of schools in a community or village | 5.36 | 2.98 | 14.32 (0.000) |
| Percentage of schools with a playground | 45.53 | 35.87 | 4.65 (0.000) |
| Percentage of schools with a library | 21.85 | 39.58 | -10.99 (0.000) |
| Average teacher experience | 12.3 | 5.52 | 32.31 (0.000) |
| Average number of classrooms per school | 5.68 | 7.49 | -6.98 (0.000) |
| Percentage of schools with a toilet facility | 57.77 | 90.57 | -18.05 (0.000) |
| Percentage of schools with a toilet facility for girls | 25.85 | 49.69 | -16.14 (0.000) |
| Percentage of schools with furniture for students | 67.5 | 76.5 | -4.53 (0.000) |
| Percentage of schools with blackboard/chalk | 87.43 | 93.2 | -3.94 (0.000) |
| Percentage of schools with potable drinking water facility for children | 63.26 | 87.7 | -11.37 (0.000) |
| Percentage of schools with a boundary wall | 59.95 | 81.2 | -10.62 (0.000) |
| Percentage of schools with electricity | 48.87 | 85.76 | -17.62 (0.000) |
| Percentage of schools with textbooks for students | 79.31 | 60.25 | 9.82 (0.000) |
| Teacher's (total staff) Presence at the time of visit to school (percent) | 89.41 | 94.27 | -2.39 (0.008) [(a)] |
| No of observations | 1708 | 244 | |

Note: p value is given in the brackets.
(a): Over the community t-test.
Source: Calculated from PRHS-II

The t-test shows that the mean student-teacher ratio is significantly higher in state schools. The number of government schools is significantly higher than private establishment. Except for playgrounds, textbook provision and average teacher experience, private schools compared favourably with government schools on the available school quality indicators in the data set. Provision of libraries, toilet facility, furniture availability for students, black board/chalk, potable drinking water facility, school boundary wall, electricity in the schools, average teacher presence at the time of school visit, are significantly higher in private schools than in the government schools. Government schools stood better in terms of average teacher experience than private schools- but as found earlier, teacher experience has no effect on perceived cognitive achievement. Greater teacher experience in state schools is consistent with earlier work (Andrabi et al. 2002) with census data and with the Pakistan Integrated Household Survey data (PIHS).

Private schooling is a new phenomenon in Pakistan, employing young teachers with less experience. On average government school teachers spends less time per student (and may get paid more, if experience is rewarded in the state sector) because of the higher student-teacher ratio. Also teacher absences are almost double those of the private sector (10.59 percent in government schools 5.73 percent in private schools).

### 6.6 Cognitive Achievement: Private and State Schools

Using parental assessment of child cognitive ability, Tables 6.14 and 6.16 show that private school students outperform government school children both in reading and in maths. Almost three quarters of children in private schools are able to read a simple letter or a newspaper with understanding. But the proportion of children in government schools is little over one half. For simple addition and subtraction the private advantage is greater; 81 percent of children in private schools are able to do it easily, whereas in government schools the percentage is 57.

Table 6.14: Reading Ability/Performance by School Type (Percentage)

| Is able to read a simple letter or the newspaper with understanding | | School Type | | Total |
|---|---|---|---|---|
| | | Government | Private | |
| Not at all | Percent | 27.09 | 10 | 25.03 |
| Yes, with difficulty | Percent | 18.57 | 16.67 | 18.34 |
| Yes, easily | Percent | 54.34 | 73.33 | 56.63 |
| Total | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

Table 6.15: Mathematics Ability/Performance by School Type (Percentage)

| Is able to do simple addition and subtraction? | | School Type | | Total |
|---|---|---|---|---|
| | | Government | Private | |
| Not at all | Percent | 26.94 | 7.78 | 24.63 |
| Yes, with difficulty | Percent | 16.44 | 11.11 | 15.8 |
| Yes, easily | Percent | 56.62 | 81.11 | 59.57 |
| Total | Percent | 100 | 100 | 100 |

Source: Calculated from PRHS-II

In addition to the higher cognitive achievement of children educated in private schools than their counterparts in state schools, the mean parental level of education (Table 6.16) of children in private schools is also higher. Similarly, private school children's equivalent household income net of child earnings is higher than for children in government schools and

for children not in schools. For the US and for Italy comparable associations have been documented (Epple et al. 2004; Mocetti 2012).

Table 6.16: Mean and Standard Deviation of Selected Variables by Schooling Status (Age 11-15)

| Variable | Children not in School | | Children in Government schools | | Children in private schools | |
|---|---|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| Mother's Level of Education[(a)] | 1.03 | 0.23 | 1.14 | 0.52 | 1.46 | 0.94 |
| Father's Level of Education[(a)] | 1.45 | 0.88 | 1.82 | 1.18 | 2.61 | 1.61 |
| Number of observations | 859 | | 657 | | 90 | |

(a): Level of Education is a categorical variable comprised of; No education=1, Primary=2, Middle=3, Secondary=4, Higher secondary=5 and Tertiary=6. Whereas primary stands for 5 years, being the highest level of education completed, middle stands for 6-8years, being the highest level of education completed, 9-10 years is highest level secondary, higher secondary (11-12 years) and tertiary (13+ years).
Std. Dev.: stands for standard deviation
Source: Calculated from PRHS-II

## 6.7 Conclusion

The statistics of school quality then show that private schools provide higher quality services than government schools in rural Pakistan. Consequently children in private schools have higher cognitive achievement[58]. But most importantly the full costs of private schooling are too high for a substantial proportion of the rural population.

Regarding the equity concern that private schools cater to an elite rather than to poor households in Pakistan, comparison of private school fees with total household expenditure as well as the income of the ordinary labourer in rural Pakistan, shows that if school access were permitted simply by paying school fees then the findings of the previous studies are supported (Alderman et al. 2001; Andrabi et al. 2002; Andrabi et al. 2008) - that private schools are accessible to the poor. But taking into account that households must also pay for uniforms, books and transport the evidence presented shows that an ordinary labourer, as well as those households whose income falls below the 25$^{th}$ percentile of the income distribution with three children, would be simply unable to get their children educated in an average private school in rural Pakistan. For a skilled male, such as a carpenter or mason, more than 70% of their income would be absorbed by schooling costs. In addition the opportunity cost of child work should be factored in. Merely expanding private schooling provision is not a solution to the currently poor education available to these households.

[58] Using nested logit to control for selection into private school (for example by parental education or income), private school children are estimated to have higher cognitive achievement than state school children. But the results are not tabulated in the present thesis.

In any event such expansion is not without its problems. There are typically arbitrary political or regulatory barriers to establishing private schools in many localities. Both registered and non-registered private schools are included in the sample and cannot be separately identified. Political support can be necessary to achieve registration or even to pursue the cheaper route of running an unregistered private school. We have shown that whether or not a community has a local councillor influences the chances of whether it has a private school.

These findings suggest that, both from private school expense and politically constrained supply, improving (perceived) state school quality may be the most effective strategy for improving human capital in rural Pakistan. However, further detailed consideration and evaluation is required before this recommendation is implemented. For example the competence and willingness of the state authorities to raise state school quality should be assessed and a cost-benefit evaluation of raising quality in state schools versus alternative programmes, such as distributing subsidies to low income families to pay private school fees, must be undertaken. This last policy still might not be satisfactory because children already in state schools would be worse off while getting lower school quality, since school quality is lower in government schools. Public-private partnerships could also be examined as a solution.

# Chapter 7 Concluding Discussion

This study has focussed on the effectiveness of a policy of raising school quality to increase schooling and reduce child labour. The standard indicators show that the performance of the education system is quite deplorable in Pakistan. Moreover we have demonstrated that there are great spatial disparities in educational quality and a high proportion of schools lack basic physical infrastructure. The theoretical interest of the study arises, firstly, from recent work of Basu and Van (1998), whose model's implication is that household subsistence poverty is the main cause of the child schooling or labour decision in developing countries. This implication was later on empirically tested and confirmed by Bhalotra (2007) for rural Pakistan. Secondly, Baland and Robinson's (2000) theoretical paper considered household's credit considerations and postulated that households are unable to borrow and lend optimally over the period in which their children ideally attend school, with the consequence of sub-optimal school attendance.

The present research has utilized a large household survey (PRHS-II) for rural Pakistan in 2004. This data set is unique in that it has an enormous range of indicators of school quality for schools attended by children in the surveyed households. The data has been analysed using a range of empirical models/techniques (e.g. binary outcome models, multinomial ordered models, and linear and instrumental variable regressions models) and a number of important findings have been established. Detailed discussion of the main findings from each chapter is reported below.

Chapter 2 set the school quality scene with a discussion of human capital and schooling with particular reference to Pakistan. It showed how low school quality (characteristic of Pakistan) modifies many of the conclusions of the standard human capital and schooling models, and explains some of the empirical anomalies. School attendance is widely accepted long term approach to enhancing an economy's human capital. We examined the views that poor school attendance (again, a feature of Pakistan) is due to household poverty, when child work is the opportunity cost of schooling, that credit constraints are the root cause of non-attendance and that low school quality ensures too low a return to schooling for attendance to be worthwhile.

Chapter 3 provided a detailed description of PRHS-II survey. It discusses adult literacy, child education, rural households' consumption and credit rationing. It also discusses major school quality indicators available in the survey. School quality indicators in rural Pakistan are very poor and prospects for a good education are dismal. In some villages the average student-

teacher ratio across schools is as high as 79 pupils. Females compared to males lose in terms of level of education, literacy, current school enrolment and work outside the home on non-family farms in rural Pakistan. Majority of children working outside the home belong to the upper quintile of household expenditure (equivalised). A very small proportion of households could be regarded as credit rationed in rural Pakistan. Family wellbeing (consumption expenditure) as well as child education is higher in Punjab than in Sindh.

The World Bank indicators on education show that child enrolment in Pakistan is much worse than other developing countries. Pakistan spends relatively less on education consequently the student-teacher ratio is higher than in other developing countries. The estimates of school quality and child and household characteristics presented in the chapter are consistent with other national level surveys (RHPS-2012) and government statistics (GoP 2011), as well as with education estimates from the World Bank.

Chapter 4 used the Pakistan Rural Household Survey II to show that two measures of school quality (proportion of schools in the community with a library and the average student-staff ratio) are closely associated with parental assessment of child cognitive ability.

This then allowed in chapter 5 a model of the household child schooling decision to be specified. The model supposes the decision is based on a rational cross-generational calculation of the increase in expected future earnings to the educated children when grown up compared with their contribution to family income in the present generation if they do not attend school. A prediction is that if a household chooses to send its children to school because of an improvement in school quality, it cannot have been below subsistence (without child work) before the quality increased. The two school quality measures validated in chapter 4 are strongly associated with school attendance, as predicted by the model, where higher school quality considerably boosts expected future earnings. A simulation exercise shows that a ten percent increase in school quality is equal to one half of the average household (equivalised) expenditure. Similarly for a typical household below the national poverty line, higher school quality increases school attendance. Most importantly marginal improvements in school quality increase the likelihood of child school participation by more in the poor households than in the rich households. In view of the great variation in school quality measures, simply reducing the disparities could substantially improve school attendance and human capital accumulation. We also controlled for the important credit market access in the estimation process. The results showed that short term credit

considerations do not seem to adversely affect child school attendance. Furthermore it was found that fertility adversely affects child school attendance.

Having established that school quality is poor and strongly influences school attendance we want to know how school quality is best raised. Chapter 6 therefore investigated whether school quality is better improved by the state or private provision of education in rural Pakistan. For a larger rural sample than exploited previously we confirmed that, after including suitable controls, children in private schools show better cognitive achievements than children in government schools. But even if private schools show higher quality or performance (in accordance with much of the literature for Pakistan), this is of little use for national policy if large proportions of the population cannot participate; private school attendance may well be constrained by the greater costs and by inadequate household incomes.

This matters because per student costs in private schools are higher than in government schools. Parents or households who send their children to private schools are better off and better educated compared to households that either have their children in government schools or not currently at school. Despite comparatively low private school fees an important part of both private and public school costs is the child's travel cost. Schooling expenditure also involves paying for uniform and books. Taking these heads into account the evidence presented shows that, an ordinary labourer, as well as those household whose income falls below the $25^{th}$ percentile of the income distribution with three children, would be simply unable to get their children educated in an average private school in rural Pakistan. Looking at these as a proportion of household income shows that for a skilled male, such as carpenter or mason, more than 70% of their income would be absorbed by schooling costs. In addition the opportunity cost of child work should be factored in. So, simply expanding private schooling provision is not a solution to the currently poor education available to these households. Some form of voucher system would be necessary to subsidise school fees, yet the administration of such a scheme would be problematic in Pakistan at present.

This can be seen simply in the current constraints on private schools. There are typically arbitrary political or regulatory barriers to establishing private schools in many localities. Both registered and non-registered private schools are included in the present sample and cannot be separately identified. Political support can be necessary to achieve registration or even to pursue the cheaper route of running an unregistered private school. It has been shown that whether or not a community has a local councillor influences the chances of whether

there is a private school. For these reasons it is recommended, both from private school un-affordability and politically constrained supply, that improving (perceived) state school quality may be the most effective strategy for improving human capital in rural Pakistan.

Why is school quality so low and shows such vast spatial[59] variation in Pakistan? There are two obvious reasons, based on the evidence presented. Firstly, educational provision is largely financed by the public sector, but only 2 percent of the GDP is spent on education, much less than other countries in the region. Secondly, the allocation mechanism is also at fault and explains the wide variation in school quality (HRCP and CEF 2005). This thesis recommends that an independent national schools inspectorate might be considered to prevent (or at least identify) the diversion and inefficient use of funds by influential but partial local bodies that must be occurring. Possibly a link with the national judiciary may be a way to create and preserve the essential independence.

There is insufficient evidence as to decide unambiguously about the best way to improve school quality. For this reason it would be valuable to have some empirical research that investigated the competence of, and incentives for, the state authorities to raise state school quality. This might then be compared with a credible (in a Pakistani context) programme for subsidising low income families to pay private school fees. Such a subsidy would only affect children who switched between sectors not those remaining in state schools – which would need to be taken into account in the policy comparison. Another policy worth exploring is public-private partnership in education – central control of the curriculum is likely to be an important consideration here, as well as finance. Detailed appraisal of such policies is beyond the scope of this thesis, but at least we have established an empirical foundation for such future research.

In chapter 2 we showed that the neglect of school quality measurement was likely to give rise to misleading macroeconomic results about human capital and misinterpretations of the microeconomics of investment in human capital. The finding in this thesis of the great importance that parents attach to school quality in rural Pakistan reinforces these views. A reallocation of state spending towards schooling, and arrangements to ensure the resources are actually spent effectively on schooling, would ensure a stronger growth rate of living standards.

[59] School quality variables show a wide range of variation between communities and a high average deprivation.

## Bibliography


Abbas, Q. and Foreman-Peck, J. 2008a. Human Capital and Economic Growth: Pakistan, 1960-2003. *Lahore Journal of Economics* 13(1), pp. 1-27.

Abbas, Q. and Foreman-Peck, J. 2008b. The Mincer Human Capital Model in Pakistan Implications for Education Policy. *South Asia Economic Journal* 9(2), pp. 435-462.

Alderman, H.; Behrman, J. R.; Ross, D.R; Sabot, R. 1996. Decomposing the Gender Gap in Cognitive Skills in a Poor Rural Economy. *Journal of Human Resources* 31(1), pp. 229-254.

Alderman, H.; Behrman, J. R.; Ross, D.R; Sabot, R. 2009. The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labour Market. *Oxford Bulletin of Economics and Statistics* 58(1), pp. 29-55.

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* 36(2), pp. 304-326.

Anderson, S. 2007. The Economics of Dowry and Brideprice. *Journal of Economic Perspective* 21(4), pp. 151-174.

Andrabi, T.; Das, J.; Khwaja, A. 2002. The Rise of Private Schooling in Pakistan: Catering to the Urban Elite or Educating the Rural Poor? *World Bank and Harvard University*.

Andrabi, T.; Das, J.; Khwaja, A. 2011. *Students Today, Teachers Tomorrow? Identifying Constraints on the Provision of Education*. World Bank Policy Research Working Paper No. 5674

Andrabi, T.; Das, J.; Khwaja, A. 2006a. *A Dime a Day: The Possibilities and Limits of Private Schooling in Pakistan*. World Bank Policy Research Working Paper 4066.

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

Andrabi, T.; Das, J.; Khwaja, A.; Tristan, Z. 2006. Religious School Enrolment in Pakistan: A Look at the Data. *Comparative Education Review* 50(3), pp. 446-477.

Angrist, J. D. and Krueger, A. B. 1991. Does Compulsory School Attendance Affect Schooling and Earnings? *Quarterly Journal of Economics* 106(4), pp. 979–1014.

Angrist, J.; Lavy, V.; Schlosser, A. 2010. Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labour Economics* 28(4), pp. 773-824.

Angrist, J. D. and Lavy, V. 1999. Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics* 114(2), pp. 533-575.

Arif, G. M. and Saqib, N. 2003. Production of Cognitive and Life Skills in Public, Private, and NGO Schools in Pakistan. *The Pakistan Development Review* 42(1), pp. 1-28.

Arif, G. M. and Farooq, S. 2012. *Rural Poverty Dynamics in Pakistan: Evidence from Three Waves of the Panel Survey.* Poverty and Social Dynamics Paper Series No.02. Pakistan Institute of Development Economics.

Ashenfelter, O. and Krueger, A. B. 1994. Estimates of the Economic Returns to Schooling from a New Sample of Twins. *American Economic Review* 84(5), pp. 1157-1173.

Aslam, M. 2009. The Relative Effectiveness of Government and Private Schools in Pakistan: are Girls Worse-off? *Education Economics* 17(3), pp. 329-354.

Baland, J.-M. and Robinson, J. A. 2000. Is Child Labour Inefficient? *Journal of Political Economy* 108(4), pp. 663-679.

Bardhan, P. K. and Udry, C. 1999. *Development Microeconomics.* Oxford University Press.

Barro, R. J. and Lee, J. W. 1993. International Comparisons of Educational Attainment. *Journal of monetary economics* 32(3), pp. 363-394.

Bartus, T. 2005. Estimation of Marginal Effects using margeff. *The Stata Journal* 5(3), pp. 309-329.

Basu, K. and Van, P. H. 1998. The Economics of Child Labour. *The American Economic Review* 88(3), pp. 412-427.

Becker, G. S. 1965. A Theory of the Allocation of Time. *Economic Journal* 5, pp. 493-517.

Becker, G. S. 1975. Front matter, Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, 2nd ed.* NBER, pp. -22-20.

Becker, G. S. 1991. *A Treatise on the Family*. Harvard University Press, Cambridge.

Becker, G. S. 1994. *Human capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press.

Becker, G. S. and Lewis, H. G. 1973. On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy* 81(2), pp. S279-S288.

Becker, G. S. and Tomes, N. 1976. Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy* 84(4), pp. S143-S162.

Beegle, K.; Dehejia, R. H.; Gatti, R. 2006. Child Labour and Agricultural Shocks. *Journal of Development Economics* 81(1), pp. 80-96.

Behrman, J. R. and Birdsall, N. 1983. The Quality of Schooling: Quantity Alone is Misleading. *The American Economic Review* 73(5), pp. 928-946.

Behrman, J. R. and Birdsall, N. 1988. The Equity-Productivity Trade-off: Public School Resources in Brazil *European Economic Review* 32(8), pp. 1585-1601.

Behrman, J. R.; Rosenzweig M. R.; Taubman, P. 1994. Endowments and the Allocation of Schooling in the Family and in the Marriage Market: the Twins Experiment. *Journal of Political Economy* 102(6), pp. 1131-74.

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), pp. 127-142.

Behrman, J. R. and Rosenzweig, M. R. 2002. Does Increasing Women's Schooling Raise the Schooling of the next Generation? *American Economic Review* 92(1), pp. 323-34.

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), pp. 94-104.

Benhabib, J. and Spiegel, M. M. 1994. The Role of Human Capital in Economic Development Evidence from Aggregate Cross-Country Data. *Journal of Monetary economics* 34(2), pp. 143-173.

Bhalotra, S. 2003. *Child Labour in Asia and Africa*. UNESCO.

Bhalotra, S. and Heady, C. 2003. Child Farm Labour: The Wealth Paradox. *The World Bank Economic Review* 17(2), pp. 197-227.

Bhalotra, S. R. 2007. Is Child Work Necessary?. *Oxford Bulletin of Economics and Statistics* 69(1), pp. 29-55.

Bils, M. and Klenow, P. J. 2000. Does Schooling Cause Growth? *The American Economic Review* 90(5), pp. 1160-1183.

Boissiere, M.; Knight, J.B.; Sabot, R.H. 1985. Earnings, Schooling, Ability, and Cognitive Skills. *The American Economic Review* 75(5), pp. 1016-1030.

Bonjour, D.; Cherkas, L.F.; Haskel, J.E.; Hawkes, D.D.; Spector, T.D. 2003. Returns to Education: Evidence from U.K. Twins. *The American Economic Review* 93(5), pp. 1799-1812.

Burney, N. A. and Irfan, M. 1991. Parental Characteristics, Supply of Schools, and Child School-Enrolment in Pakistan. *The Pakistan Development Review* 30(1), pp. 21-62.

Butcher, K. F. and Case, A. 1994. The Effect of Sibling Sex Composition on Women's Education and Earnings. *The Quarterly Journal of Economics* 109(3), pp. 531-563.

Cain, M. T. 1977. The Economic Activities of Children in a Village in Bangladesh. *Population and Development Review* 3(3), pp. 201-227.

Cameron, A. C. and Trivedi, P. K. 2009. *Microeconometrics using Stata*. Stata Press.

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

pp. 1-40.

Case, A. and Deaton, A. 1999. School Inputs and Educational Outcomes in South Africa. *The Quarterly Journal of Economics* 114(3), pp. 1047-1084.

Coombs, P. H. and Ahmed, M. 1974. *Attacking Rural Poverty: how non-Formal Education can Help*. Johns Hopkins University Press.

Coulson, A. J. 2009. Comparing Public, Private, and Market Schools: The International Evidence. *Journal of School Choice* 3(1), pp. 31-54.
Das, J.; Dercon, S.; Habyarimana, J.; Krishnan, P.; Muralidharan, K.; Sundararaman, V. 2011. *School Inputs, Household Substitution, and Test Scores*. National Bureau of Economic Research.

Das, J.; Pandey, P.; Zajonc, T. 2006. Learning Levels and Gaps in Pakistan. The World Bank, Policy Research Working Paper Series: 4067.

Dasgupta, P. 1995. The Population Problem: Theory and Evidence. *Journal of Economic Literature* 33(4), pp. 1879–1902.

Deaton, A. 1988. Quality, Quantity, and Spatial Variation of Price. *The American Economic Review* 78(3), pp. 418-430.

Ederer, P.; Schuler, P.; Willms, S. 2007. *The European Human Capital Index: The Challenge of Central and Eastern Europe*. The Lisbon Council.

Edmonds, E. V. 2007. *Child Labour*. NBER Working Paper 12926, National Bureau of Economic Research, Cambridge, MA.

Emerson, P. M.; Knabb, S. D. 2006. Opportunity, Inequality and the Intergenerational Transmission of Child Labour. *Economica* 73(291), pp. 413-434.

Epple, D.; Figlio, D.; Romano, R. 2004. Competition between Private and Public Schools: Testing Stratification and Pricing Predictions. *Journal of Public Economics* 88(7–8), pp. 1215-1245.

Ersado, L. 2005. Child Labour and Schooling Decision in Urban and Rural Areas: Comparative Evidence from Nepal, Peru and Zimbabwe. *World Development* 33(3), pp. 455-480.

Eswaran, M. 1996. *Fertility, Literacy and the Institution of Child Labour*. Working Paper No. 26, IRIS-India.

Federal Bureau of Statistics 2000. Census of Private Educational Institutions in Pakistan. Federal Bureau of Statistics, Statistics Division, Government of Pakistan.

Federal Bureau of Statistics 2005. National Education Census (NEC). Federal Bureau of Statistics, Statistics Division, Government of Pakistan.

Filmer, D. 2007. If you Build it, will they Come? School Availability and School Enrolment

in 21 Poor Countries. *Journal of Development Studies* 43(5), pp. 901-928.

Fordham, P. E. 1993. Informal, Non-Formal and Formal Education Programmes In: ICE301, Y.G.W.C. ed. *Lifelong Learning Unit 2*. YMCA George Williams College, London.

Fuwa, N.; Ito, S.; Kubo, K.; Kurosaki, T.; Sawada, Y. 2009. How Does Credit Access Affect Children's Time Allocation?: Evidence from Rural India. *Journal of Globalization and Development* 3(1).

Galasso, E. 2000. Intra-household Allocation and Child Labour in Indonesia. Mimeograph, Boston College.

Gao, F. and Wahid, A. N. M. 2009. The Baland-Robinson Theory of Child Labour Efficiency: A Proposed Extension. *American Economist* 54(2), pp. 72-81.

Gershberg, A. I. and Schuermann, T. 2001. The Efficiency–Equity Trade-off of Schooling Outcomes: Public Education Expenditures and Welfare in Mexico. *Economics of Education Review* 20(1), pp. 27-40.

Glewwe, P. 1996. The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment. *Journal of Development Economics* 51(2), pp. 267-290.

Greene, W. H. 2003. *Econometric Analysis*. 5th Edition ed. Prentice Hall, USR NJ 07458.

Harmon, C.; Hogan, V.; Walker, I. 2003. Dispersion in the Economic Return to Schooling. *Labour Economics* 10(2), pp. 205-214.

Harmon, C.; Oosterbeek, H.; Walker, I. 2003. The Returns to Education: Microeconomics. *Journal of Economic Surveys* 17(2), pp. 115-156.

Harmon, C. and Walker, I. 1995. Estimates of the Economic Return to Schooling for the United Kingdom. *The American Economic Review* 85(5), pp. 1278-1286.

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

Hill, R. C.; Griffiths, W.E.; Lim, G.C.; Adkins, L.C. 2011. *Principles of Econometrics International Student Version*. 4 ed. John Wiley & Sons, ISBN: 978-0-470-87372-4.

Hou, X. 2010. Wealth: Critical but Not Sufficient-Evidence from Pakistan on Economic Growth, Child Labour and Schooling. *Journal of Development Studies* 46(3), pp. 439-465.

Hoxby, C. M. 2000. The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *The Quarterly Journal of Economics* 115(4), pp. 1239-1285.

Human Rights Commission of, P. and Commonwealth Education, F. 2005. *The Education Budget in Pakistan*. HRCP in collaboration with CEF.

Infed. 2012. *Non-Formal Education* [Online]. Available at: http://www.infed.org./biblio/b-nonfor.htm [Accessed: 08-12-2012].

Iset. 2013. *Price of a Women: Economic Rationale Behind Marriage Payments in Georgia* [Online]. Available at: http:// http://www.iset.ge/blog/?p=730 [Accessed: 07-14-2013].

Jacoby, H. G. 1994. Borrowing Constraint and Progress Through School: Evidence from Peru. *The Review of Economics and Statistics* 76(1), pp. 151-160.

Jacoby, H. G. and Skoufias, E. 1997. Risk, Financial Markets, and Human Capital in a Developing Country. *The Review of Economic Studies* 64(3), pp. 311-335.

Jenkins, S. P. and Cowell, F. A. 1994. Parametric Equivalence Scales and Scale Relativities. *The Economic Journal* 104(July), pp. 891-900.

Jimenez, E. and Tan, J. P. 1987. Decentralised and Private Education: the Case of Pakistan. *Comparative Education* 23(2), pp. 173-190.

Jolliffe, D. 1998. Skills, Schooling, and Household Income in Ghana. *The World Bank Economic Review* 12(1), pp. 81-104.

Jones, P. 2001. Are Educated Workers Really More Productive? *Journal of Development Economics* 64(1), pp. 57-79.

Kane, T. J. and Rouse, C. E. 1995. Labour-Market Returns to Two- and Four-Year College. *The American Economic Review* 85(3), pp. 600-614.

Kingdon, G. G. 1996. Private Schooling in India: Size, Nature, and Equity-Effects. *Economic and Political Weekly* 31(51), pp. 3306-3314.

Kingdon, G. and Aslam, M. 2008. Gender and Household Education Expenditure in Pakistan. *Applied Economics* 40(20), pp. 2573-2591.

Kitaev, I. 1999. *Private Education in Sub-Saharan Africa: A Re-Examination of Theories and Concepts Related to Its Development and Finance. Mechanisms and Strategies of Educational Finance*. Paris: International Institute of Educational Planning, UNESCO.

Kyriacou, G. A. 1991. *Level and Growth Effects of Human Capital: A Cross-Country Study of the Convergence Hypothesis*. CV Starr Center for Applied Economics.

Lau, L. J.; Jamison, D.T.; Liu, S.C.; Rivkin, S. 1991. *Education and Productivity in Developing Countries: an Aggregate Production Function Approach*. World Bank Publications.

Leibowitz, A. 1974. Home Investments in Children. *Journal of Political Economy* 82(2), pp. S111-s131.

Lockheed, M. E. and Jimenez, E. 1994. *Public and Private Secondary Schools in Developing Countries: What are the Differences and why do they Persist?* ESP Discussion Paper Series No. 33, The World Bank.

Lucas, R. E., Jr. 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics* 22(1), pp. 3-42.

Maitra, P. 2001. Is Consumption Smooth at the Cost of Volatile Leisure? An Investigation of Rural India. *Applied Economics* 33(6), pp. 727-734.

Mankiw, N. G.; Romer, D.; Weil, D.N. 1992. A Contribution to the Empirics of Economic Growth. *The quarterly journal of economics* 107(2), pp. 407-437.

McKelvey, R. D. and Zavoina, W. 1975. A Statistical Model for the Analysis of Ordinal Level Dependent Variables. *Journal of Mathematical Sociology* 4(1), pp. 103-120.

Menon, N. 2010. Investment Credit and Child Labour. *Applied Economics* 42(12), pp. 1461-1479.

Mincer, J. 1958. Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy* 66(4), pp. 281-302.

Mincer, J. and Polachek, S. 1974. Family Investment in Human Capital: Earnings of Women. *Journal of Political Economy* 82(2), pp. S76-s108.

Mocetti, S. 2012. Educational Choices and the Selection Process: before and after Compulsory Schooling. *Education Economics* 20(2), pp. 189-209.

Nayab, D. and Arif, G. M. 2012. *Pakistan Panel Household Survey Sample Size, Attrition and Socio-demographic Dynamics.* Poverty and Social Dynamics Paper Series No.01. Pakistan Institute of Development Economics.

Nehru, V. and Dhareshwar, A. 1993. A new Database on Physical Capital Stock: Sources, Methodology and Results. *Revista de análisis económico* 8(1), pp. 37-59.

OECD. 1982. The OECD List of Social Indicators Paris.

OECD. 1998. *Human Capital Investment: An International Comparison*. Paris, Organisation for Economic Cooperation and Development.

OECD. 2007. Policy Brief: Globalization, Jobs and Wages. Organisation for Economic Cooperation and Development.

Pakistan, G. o. 2006-07. Pakistan Economic Survey. In: Finance, M.o. ed. Islamabad, Pakistan.

Pakistan, G. o. 2007-08. Pakistan Economic Survey. In: Finance, M.o. ed. Islamabad, Pakistan.

Pakistan, G. o. 2009. Draft National Education Policy 2009. In: Ministry of Education. Islamabad, P. ed.

Pakistan, G. o. 2011-12. Pakistan Economic Survey. In: Ministry of Finance. Islamabad, P.

ed.

Patrinos, H. A. and Psacharopoulos, G. 1995. Educational Performance and Child Labour in Paraguay. *International Journal of Educational Development* 15(1), pp. 47-60.

Patrinos, H. A. and Psacharopoulos, G. 1997. Family Size, Schooling and Child Labour in Peru. An Empirical Analysis. *Journal of Population Economics* 10(4), pp. 387-405.

Pencavel, John. 1999. *Assortative Mating by Schooling and the Timing of Marriage.* Working Paper, Stanford University.

Pitt, M.M.; Rosenzweig, M.R.; Gibbons, D.M. 1993. The Determinants and Consequences of the Placement of Government Programs in Indonesia. *World Bank Economic Review* 7(3), pp. 319-348.

Pritchett, L. 2001. Where has all the Education Gone? *The World Bank Economic Review* 15(3), pp. 367-391.

Psacharopoulos, G. 1985. Educational Research and Policy Development at the World Bank. *Indian Journal of Quantitative Economics* 1(2), pp. 119-124.

Ranjan, P. 1999. An Economic Analysis of Child Labour. *Economics Letters* 64(1), pp. 99-105.

Ravallion, M. and Wodon, Q. 2001. Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrolment Subsidy. *The Economic Journal* 110(462), pp. 158-175.

Ray, R. 2000a. Analysis of Child Labour in Peru and Pakistan: A Comparative Study. *Journal of Population Economics* 13(1), pp. 3-19.

Ray, R. 2000b. Child Labour, Child Schooling, and their Interaction with Adult Labour: Empirical Evidence for Peru and Pakistan. *The World Bank Economic Review* 14(2), pp. 347-367.

Ray, R. ed. 2001. *Child Labour and Child Schooling in South Asia: A Cross Country Study of their Determinants*.

Robertson, D. and Symons, J. 2003. Do Peer Groups Matter? Peer Group versus Schooling Effects on Academic Attainment. *Economica* 70(277), pp. 31-53.

Romer, P. M. 1986. Increasing Returns and Long-Run Growth. *Journal of Political Economy* 94(5), pp. 1002-1037.

Roodman, D. (2009). *Estimating Fully Observed Recursive Mixed-Process Models with cmp.* Working Paper Number 168, Center for Global Development.

Rosati, F. C. and Rossi, M. 2003. Children's Working Hours and School Enrolment: Evidence from Pakistan and Nicaragua. *The World Bank Economic Review* 17(2), pp. 283-295.

Rose, P. M. 2007. *Supporting non-State Providers in Basic Education Service Delivery*. Research Monograph No.4. University of Sussex, UK.

Rosenzweig, M. R. 1995. Why Are There Returns to Schooling? *The American Economic Review* 85(2), pp. 153-158.

Rosenzweig, M. and Evenson, R. 1977. Fertility, Schooling and the Economic Contribution of Children in Rural India: An Econometric Analysis. *Econometrica* 45(5), pp. 1065-1079.

Rosenzweig, M. R. and Paul Schultz, T. 1987. Fertility and Investments in Human Capital: Estimates of the Consequence of Imperfect Fertility Control in Malaysia. *Journal of Econometrics* 36(1–2), pp. 163-184.

Rosenzweig, M. R. and Schultz, T. P. 1985. The Demand for and Supply of Births: Fertility and its Life Cycle Consequences. *The American Economic Review* 75(5), pp. 992-1015.

Rosenzweig, M. R. and Schultz, T. P. 1989. Schooling, Information and Nonmarket Productivity: Contraceptive Use and Its Effectiveness. *International Economic Review* 30(2), pp. 457-477.

Rosenzweig, M. R. and Wolpin, K. I. 1980a. Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica* 48(1), pp. 227-240.

Rosenzweig, M. R. and Wolpin, K. I. 1980b. Life-Cycle Labour Supply and Fertility: Causal Inferences from Household Models. *Journal of Political Economy* 88(2), pp. 328-348.

Rosenzweig, M. R. and Wolpin, K. I. 1986. Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions. *The American Economic Review* 76(3), pp. 470-482.

Schuller, T. 2001. The Complementary Roles of Human and Social Capital. *Canadian Journal of Policy Research* 2(1), pp. 18-24.

Schultz, T. P. 1990. *Assessing Family Planning Cost-effectiveness: Applicability of Individual Demand-Program Supply Framework*. Economic Growth Center, Yale University.

Schultz, T. P. 1997. Demand for Children in Low Income Countries. *Handbook of population and family economics* 1, pp. 349-430.

Schultz, T. W. 1974. References and Index to" Economics of the Family: Marriage, Children, and Human Capital". *Economics of the Family: Marriage, Children, and Human Capital*. UMI, pp. 549-584.

Singh, R. D. and Schuh, G.E. 1986. The Economic Contribution of Farm Children and the Household Fertility Decisions: Evidence from a Developing Country, Brazil. *Indian Journal of Agricultural Economics* 41(1), pp. 29-40.

Solmon, L. C. 1973. The Definition and Impact of College Quality. National Bureau of Economic Research, Inc, NBER Working Papers: 0007.

Spiegel, M. M. ed. 1994. *Determinants of Long-run Labour Productivity Growth: A Selective Survey with some new Empirical Results*. Federal Reserve Bank of San Francisco.

Strauss, J. and Thomas, D. 1995. Human Resources: Empirical Modelling of Household and Family Decisions. In: Behrman, J. and Srinivasan, T.N. eds. *Handbook of Development Economics*. Vol. iii. Amsterdam: North Holland: 1882-2023.

Temple, J. 2001. *Growth Effects of Education and Social Capital in the OECD Countries*. CEPR Discussion Paper No. 2875. Available at SSRN: http://ssrn.com/abstract=276787.

Urquiola, M. 2006. Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia. *Review of Economics and Statistics* 88(1), pp. 171-177.

Wachtel, P. 1975. The Effect of School Quality on Achievement, Attainment Levels, and Lifetime Earnings. *Explorations in Economic Research, Volume 2, number 4*. NBER, pp. 502-536.

Weiner, M. 1991. *The Child and the State in India: Child Labour and Education Policy in Comparative Perspective*. Princeton, N.J.: Princeton Univ. Press.

White, H. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 48(4), pp. 817-838.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data.* MIT Press.

Wooldridge, J. M. 2009. *Introductory Econometrics: A Modern Approach*. South-Western Pub.