

**Cardiff Economics
Working Papers**

Qaisar Abbasa and James Foreman-Peck

*The Mincer Human Capital Model in Pakistan: Implications
for Education Policy*

E2007/24

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
www.cardiff.ac.uk/carbs

ISSN 1749-6101
August 2007

The Mincer Human Capital Model in Pakistan: Implications for Education Policy

Qaisar Abbas^{a, *}, James Foreman-Peck^b

^a Comsats Institute of Information Technology, Islamabad, Pakistan & Cardiff Business School, Cardiff University, UK

^b Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive Cardiff, CF10 3EU UK, <http://www.cardiff.ac.uk/carbs/econ/foreman-peckj/>

Email: Foreman-PeckJ@cardiff.ac.uk

Phone: +44(0)29 2087 6395

Fax: +44 (0)29 2087 4419

Abstract

This paper estimates and interprets returns to education for three sub-sectors of labour market by gender in Pakistan, using the most recent data set of Pakistan Social and Living Standards Measurement (PSLM) Survey 2004-05. The results show two distinctive features of Pakistani education, the high apparent returns to female education outside agriculture, and the remarkable increase of returns with successive levels of education, are to be explained primarily by two departures from the basic Mincer model; generally poor quality primary schooling and family unwillingness to invest in female education because of lack of earning opportunities. There is some signaling in Pakistani education investment but mainly the education is productivity-enhancing investment in human capital, according to a comparison of self-employed and paid employed earnings equations. Returns to public spending of education are extremely high, suggesting very considerable state underinvestment. The policy challenge is in the low wages and high education in the female paid employment sector, and the low participation rate.

JEL Classification: J16, J18, J24.

Keywords : Rates of return, gender, occupation, Pakistan

* Corresponding Author: Qaisar Abbas, Cardiff Business School, Aberconway Building, Room T31B, Colum Drive Cardiff CF10 3EU, Wales, UK.

Email: abbasq1@cardiff.ac.uk

Tel +44 (0)29 2087 5505, (till August 30, 2007)

Department of Management Sciences, Comsats Institute of Information Technology, Plot no.30, H-8/I, Islamabad, Pakistan

Email: qaisar@comsats.edu.pk

Tel: 0092-51-9258481-83, Fax: 0092-51-4442805

(after August 30, 2007)

August 2007

The Mincer Human Capital Model in Pakistan: Implications for Education Policy

Introduction

Empirical estimates of returns to investment in human capital have long been dominated by the demand side formulation of Jacob Mincer (1958) (Rosen 1992). A rational individual with perfect foresight chooses schooling up to the point where the wage gains of an extra year's schooling are offset by the additional cost (assumed to be mainly the earnings foregone). Then as long as the direct costs of schooling are small and the period of earning is long, the proportionate increase in wages from an extra year of schooling equals the rate of return.

The Mincer model has provided the framework for many estimates of the rate of returns to education in Pakistan (see Aslam (2007) for a review). All but one of these studies has estimated returns to education only for paid employment. Yet little more than half of the total labour force is engaged in this sector. The remainder are self-employed (non-agricultural) or working in agriculture. So far only one study (Kingdon et al., 2007) in Pakistan has estimated returns to education in all these three sectors of labour market using Pakistan Integrated Household Survey (PIHS) 1998-99 data. Moreover, they have divided the sample into young (16-30 years olds) men and women and old (31-70 years old) men and women. In the present study an attempt has been made to estimate and to interpret returns to education for three sub-sectors of labour market by gender in Pakistan, using the most recent data set of Pakistan Social and Living Standards Measurement (PSLM) Survey 2004-05.

If earnings returns are persistently high in one sector we need to know why. For in a market economy, high returns in any activity attract more resources, which drive down rewards in that sector. Persistent excess returns point to some barrier to the flow of resources. A much discussed hindrance in the Mincer literature is ability. If high ability is scarce, and high ability is necessary to take full advantage of a large number of years of schooling, some of the higher earnings associated with schooling will be properly attributable to ability rather than to the schooling. Consequently, increasing schooling for lower ability individuals will not generate the same returns.

A number of other possible departures from the simplest Mincer model assumptions may influence the interpretation of the earnings/schooling correlations. Preferences or choices in this demand model may be conflated with other constraints, particularly those of finance and those imposed by the family, often the decision-taker in practice. Pure time preference may be responsible for choosing fewer, rather than more, years of schooling, but so too may be the

inability to borrow money for upkeep in more schooling, because of lack of credit. Family decision takers, regarding younger family members as sources of collective income, can be critical. The future earnings of these members could be discounted heavily on the grounds that they will become independent of the family and no longer contribute. In some societies this could be particularly true of female children.

On the supply side, schooling available is typically provided by a political process, which may not always create the ideal education options. For example primary school age children cannot be expected to travel far, so their early education is constrained by local availability of schooling. The cumulative nature of education means that the number of years of schooling chosen may be influenced by the quality of earlier education. If the quality of schooling supplied is not homogenous, then some children, receiving low quality education, may be prevented from proceeding to the higher levels of education that their abilities and preferences warrant.

In this paper we contend that two distinctive features of Pakistani education, the high apparent returns to female education outside agriculture, and the remarkable increase of returns with successive levels of education, are to be explained primarily by two departures from the basic Mincer model; generally poor quality primary schooling and family unwillingness to invest in female education because of lack of earning opportunities.

We show that the earnings returns to public spending on education are colossal, holding constant years of schooling - because so little is spent on primary education. We go on to show that the strong nonlinearity of the years of schooling function stems from the highly variable quality of early schooling and low spending upon it. More spending in earlier years to raise average quality there so that returns at least match world averages would be far more effective than increasing HE expenditure.

High measured returns to female education we show stem from a combination of much lower workforce participation than males and fewer average years in the labour force. Actual returns are much lower. Apparently higher returns are necessary to compensate for the shorter period, and lower probability, of earning, that are not factored into the basic model. If the education supply side were the principal constraint then educated female earnings would be higher than those of educated males, whereas the reverse is the case. The constraint we suggest is not on the supply of educated females but on the opportunities that create the demand for them. Increasing these opportunities is more fundamental than increasing investment. However we concede that this argument does not

take into account other important reasons for boosting investment in female education.

The paper is structured as follows. Section 1 presents the model of the paper. Section 2 discusses data and its descriptive statistics. Section 3 highlights the specifications and estimation procedures. In section 4, the estimated results are presented. And section 5 concludes the paper.

1. The Model

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.

T-s

$$\sum_{t=1}^{T-s} (w_s - w_{s-1}) / (1+r)^t = w_{s-1} + cs \quad \dots (1)$$

where T is the maximum potential number of years in the labour force, s is years of schooling, w is earnings where the s subscripts indicates the number of years of schooling that are associated with the earnings and cs is the direct net cost of schooling (negative if the pleasure outweighs the pain and private financial outlays). r is the private internal rate of return to schooling investment. Where female labour force participation is lower than male, in effect T is lower and therefore so are returns for given earnings.

To derive the Mincer equation T must be large for all individuals and cs must be small. Then

$$r \sim (w_s - w_{s-1}) / w_{s-1} = \log w_s - \log w_{s-1} \quad \dots (2)$$

The rate of return to another year of schooling is the proportionate increase in earnings consequent upon taking that extra year. If T is not large then the formula (2) must be modified to

$$r / k = (w_s - w_{s-1}) / w_{s-1} = \log w_s - \log w_{s-1} \quad \dots (2a)$$

where $k = 1 - (1/(1+r))^t$

The standard Mincerian human capital earnings function also includes experience after schooling, to which returns rise at a diminishing rate. A range of other controls are often included as well.

In the present paper our basic equation is that from (2) a one year increase in schooling times the internal rate of return to schooling, controlling for other influences, gives the proportionate change in earnings.

$$\text{Log } w_i = \ln y_{mi} = f(S_i, X_i, Z_i, u_i) \quad \dots(3)$$

where, $\ln y_{mi}$ is the log of monthly earnings for an individual i , S is years of schooling completed, X_i is a vector of personal characteristics such as experience, experience squared, and gender. In addition we allow that individuals differ in their leisure preferences and labour supplies by including a 'days worked' measure. In the supply of education (in Z), State expenditure per school age person determines the average quality of years of (public) schooling¹, although the spending (but not the quality) is ignored in the private investment decision modelled by the Mincer equation. Since other factors influencing earnings may differ with the same areas over which state spending varies, controls for these factors must also be included in Z . u_i is the random error term measuring the impact of unobserved variables.

A modification to (3) is that if the workers in question do not work for 'a long period', from (2a) the coefficient on schooling is the (return/ k) rather than simply the rate of return. When $r=0.1$ and $t=10$, $k=0.61$; the rate of return is three fifths of the coefficient on schooling.

Conventionally the log of earnings is assumed linear in years of schooling. This assumption is here tested because of the interest in the cumulative nature of education. Linearity presupposes an adequate supply side. If the quality of investment in schooling on the supply side is on average low and variable then only a small proportion of graduates from primary education will be in a position to take advantage of secondary education (and similarly for tertiary education). Consequently the returns to successive layers of education will be higher because of the increasingly restricted pool of individuals who can benefit.

We employ multiple treatment functions to test this possibility. In the single treatment model, schooling is defined as years of education completed assuming returns are the same at every level. In the multiple treatment models the rate of return to education is calculated separately for the different education levels or qualifications.

$$\ln y_{mi} = \beta_0 + \beta_p \text{ prim} + \beta_m \text{ mid} + \beta_s \text{ sec} + \beta_{hs} \text{ hsec} + \beta_r \text{ ter} + \beta_x X_i + \beta_z Z_i + u_i \quad \dots(4)$$

¹ Actual current expenditure incurred on education by the provinces for the year 2002—03.

where 'prim' is a dummy for primary (5 years) being the highest level of education completed, 'mid' is highest level completed being middle school (6-8 years), 'sec' is highest level secondary (9-10 years), 'hsec' is higher secondary (11-12 years) and 'ter' is tertiary (13+ years). The coefficients associated with prim, mid, sec, hsec and ter in equation (4) indicates the increase in earnings with education at the respective levels.

Linearity implies the β_j coefficients are proportional to the years of schooling at each level; $\beta_m = (8/5) \beta_p$, for instance². The average rate of return r_j per year of schooling for the j th level of education can be computed as $r_j = (\beta_j - \beta_{j-1})/Y_j$, where Y_j is the number of years of schooling at j th level of education. For example if $\beta_p = \beta_m$, the return to five years of primary schooling is the same as that for three years of middle schooling on top of the primary schooling. In this case there would be no return to middle schooling. If $\beta_p = 0.2$ and $\beta_m = 0.5$, the rate of return to middle schooling is 10 percent per year ($0.3/3$), compared with a 4 percent ($0.2/5$) return to a primary year.

Another supply side influence on earnings is the curriculum. A curriculum for example that emphasises Arabic when Arabic is not spoken or much used in trade might not enhance human capital but persons who were good at Arabic might nevertheless be in high demand by employers if such achievement is a signal of particular ability³. Criticisms of the Pakistan state school curriculum make tests of this signalling hypothesis, compared with the human capital concept, especially important (for example Nayyar and Salim 2003). Schooling that does not add to human capital can nonetheless increase wages if it identifies the ability of individuals.

The self-employed do not need to signal ability to themselves. They have no reason to pay themselves more if their extra schooling does not increase their

² Assuming the same age at retirement, a person who enters the workforce at age 11 will work for 10 more years than someone who enters at age 21. Consequently the person with only primary education will have ten more years of earning and failure to take this into account will bias the rate of return comparison of primary education with tertiary in favour of the second. With a seven percent internal rate of return and 50 years of working life for the primary educated individual and 40 years for the tertiary educated person, the ratio of the primary schooling coefficient to the tertiary coefficient should be multiplied by about 1.035 to obtain the true ratio of returns. The bias is quite small.

³ So judged Thomas Babington Macaulay in his 1833 speech 'If, instead of learning Greek, we learned the Cherokee, the man who understood the Cherokee best, who made the most correct and melodious Cherokee verses--who comprehended most accurately the effect of the Cherokee particles--would generally be a superior man to him who was destitute of these accomplishments.' Berriedale Keith (1922 226-265).

productivity. So a comparison of employed and self-employed earnings is a test of the value of the curriculum and the human capital model. If the two sectors have divergent earnings/schooling relations, there is evidence of signalling. A caveat is that if earning ability is perfectly matched with years of schooling then the test would not discriminate, but this seems unlikely.

2. Data and Descriptive Statistics

The data set used in this analysis is the Pakistan Social and Living Standards Measurement Survey (PSLM) 2004-05. It consists of nationally representative sample of 80,000 households and 500,635 individuals. The survey was conducted using a two-stage stratified random sampling approach. The survey collected the information on the years of education of each member of the household, cognitive ability, employment type, occupation and industry.

Table 1: Labour force participation rate

Age groups	Male	Female
15-30	68.1	13.0
31-45	95.5	17.0
46 and above	81.0	14.0
Total	77.7	13.9

A distinctive characteristic of the Pakistani labour market, according to this sample, is the very low participation of women (Table 1). Moreover, as Table 2 shows, more than half of the female participants are effectively unemployed (unpaid family workers). Those that do find paid work receive on average half the average monthly earnings of males in the full sample (table 2), consistent with very limited labour market opportunities for females.

There are significant differences in average monthly earnings across the three sub-sectors of the labour market, but it is not entirely clear whether the figures are comparable. Earnings in agriculture should not include returns to land and self-employed earnings should not include returns to capital if they are to be on the same basis. Otherwise they will overstate true labour returns. In addition agricultural income is likely to be highly seasonal and therefore monthly earnings could include a substantial element of estimation. Subject to these caveats, self-employed and agricultural male worker's earnings are greater than the workers in paid employment sector. By contrast earnings of female workers are highest in paid employment, but still well below male earnings. Less than half the sample who could have reported earnings in fact do so.

Average years of schooling of paid- and self-employed workers are higher than those of agricultural workers. In agriculture the average is 7.2 (male) and 5.8 (female) whereas for the paid employment and the self-employed average years of schooling are 9.1 (male) and 10.9 (female), and 8.4 (male) and 7.8 (female), respectively. Females have almost two more years of schooling than males in paid employment, even though they are much lower paid.

Table 2: Descriptive Statistics (persons aged 15-65)

Variables	Total		Paid employed		Self-employed		Agriculture		Unpaid family helper	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Monthly earnings	5585 (11809)	2757 (4836)	4859 (11024)	2980 (5008)	7057 (12580)	2160 (4091)	5676 (15604)	1853 (3688)	0	0
Log earnings	8.26 (0.77)	7.335 (1.02)	8.16 (0.71)	7.44 (1.01)	8.48 (0.79)	7.12 (0.77)	8.10 (1.03)	7.07 (0.97)	0	0
Years of Schooling	8.53 (3.45)	8.38 (3.39)	9.01 (3.81)	10.87 (3.97)	8.41 (3.45)	7.82 (3.63)	7.19 (3.16)	5.84 (2.63)	7.77 (3.03)	6.42 (2.95)
Age	32.46 (14.08)	32.18 (13.44)	33.19 (11.94)	31.80 (11.65)	37.18 (12.61)	31.86 (11.77)	41.48 (13.05)	37.62 (11.31)	23.25 (8.31)	32.34 (12.69)
Exp	16.44 (13.26)	13.13 (11.45)	18.42 (11.39)	13.32 (9.68)	22.43 (12.66)	15.32 (10.47)	26.01 (13.90)	21.97 (12.42)	10.02 (6.92)	12.80 (9.48)
Maths skills	0.73 (0.44)	0.45 (0.50)	0.74 (0.44)	0.59 (0.49)	0.76 (0.43)	0.57 (0.50)	0.54 (0.50)	0.28 (0.45)	0.72 (0.44)	0.30 (0.46)
Literacy	0.64 (0.48)	0.34 (0.47)	0.65 (0.48)	0.47 (0.50)	0.69 (0.46)	0.43 (0.49)	0.38 (0.49)	0.12 (0.33)	0.61 (0.49)	0.13 (0.33)
Rural	0.62 (0.49)	0.63 (0.48)	0.55 (0.50)	0.56 (0.50)	0.47 (0.50)	0.51 (0.50)	0.94 (0.24)	0.95 (0.21)	0.76 (0.42)	0.93 (0.25)
Married	0.57 (0.49)	0.68 (0.47)	0.67 (0.47)	0.58 (0.49)	0.78 (0.41)	0.56 (0.50)	0.87 (0.34)	0.79 (0.40)	0.31 (0.46)	0.76 (0.43)
Punjab	0.42 (0.49)	0.44 (0.50)	0.41 (0.49)	0.71 (0.45)	0.52 (0.50)	0.72 (0.45)	0.36 (0.48)	0.68 (0.47)	0.39 (0.49)	0.49 (0.50)
Sindh	0.26 (0.44)	0.24 (0.43)	0.27 (0.44)	0.16 (0.36)	0.23 (0.42)	0.12 (0.33)	0.30 (0.46)	0.04 (0.19)	0.29 (0.45)	0.24 (0.43)
NWFP	0.18 (0.38)	0.20 (0.40)	0.17 (0.38)	0.09 (0.29)	0.16 (0.37)	0.13 (0.34)	0.15 (0.35)	0.26 (0.44)	0.13 (0.34)	0.11 (0.31)
Balochistan	0.15 (0.35)	0.13 (0.33)	0.14 (0.35)	0.04 (0.20)	0.09 (0.29)	0.02 (0.15)	0.19 (0.40)	0.03 (0.17)	0.19 (0.39)	0.16 (0.37)
Observations	138750	134401	47193	6237	23754	1734	20774	1322	16155	9357
Earnings observations	73195	8230	46686	6107	23106	1670	3278	325	0	0

Source: Pakistan Social Living and Standards Measurement Survey, 2004-05
Standard deviations are reported in parentheses. Paid employed includes all defined as paid employee, self-employed includes those who have their own business (non-agriculture) or they are entrepreneurs, agricultural workers are defined as own cultivator, tenant, sharecropper and livestock owners.

It is worth noting that the average years of schooling amongst unemployed workers is greater than for agricultural workers, even though agricultural workers earn more than those in paid employment. Patterns of literacy and numeracy are similar to schooling. Percentages of literates and those with basic maths skills are higher in self-employment and paid employment than in the agriculture sector.

These data indicate that the main difference in earnings for male workers is between self-employed and agricultural, while for female workers it is between self-employed and paid-employment. In contrast to earnings, the principal divergence for schooling is between the self-employed and paid employed for both sexes on one hand, and agricultural workers and unpaid family helpers on the other. This suggests that schooling influences in which occupational groups an individual ends up in. Since unemployed male individuals possess the mean schooling levels of self- and wage-employed workers, they are apparently in a queue for suitable job opportunities in the labour market.

Table 3 presents the distribution of educational attainment by gender and by the three main labour markets sectors in Pakistan. As hinted at in Table 2 wage earners with tertiary education are concentrated in the paid employment sector. More than one third of female paid employees are graduates, but female paid employees in the sample amount to less than one tenth of males. The sample of Table 3 is well under half the size of Table 2.

Table 3: Level of Education Attained by gender and Occupations (Percentage)

Education levels	Paid Employment			Self-employed			Agriculture		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
No & less than 5 years Schooling	8.98	9.21	6.42	9.62	9.47	12.56	14.82	14.67	21.91
Primary	16.39	16.96	10.46	20.00	19.65	28.01	30.33	29.97	46.45
Middle	18.50	19.47	8.39	21.60	21.75	18.46	20.66	20.81	13.66
Secondary	27.20	27.52	23.86	29.82	30.13	23.04	24.75	24.98	14.75
Higher Secondary	8.30	7.78	13.67	7.01	7.00	7.20	4.28	4.31	2.73
Tertiary	20.63	19.08	37.20	11.95	12.00	10.73	5.16	5.26	0.5
Observations	34273	31281	2992	17352	16588	764	8395	8212	183

Source: Pakistan Social and Living Standards Measurement Survey (PSLM), 2004-05

The percentage of workers having only middle and secondary education is highest in the self-employed sector. Agriculture has the highest percentage of workers with only primary education. Almost 50 percent of the female agricultural labour force received only primary education and 22 percent have none or less than 5 years of education. This female sample is small (183), even compared with the 1322 female agricultural workers of Table 2.

3. Specification and Estimation

As noted in the previous section, Pakistan's labour markets are segmented between male and female workers and between paid employment, self-employment and agriculture. We therefore estimate earnings functions separately for each, taking OLS as a benchmark. However, OLS estimates potentially suffer from sample selectivity and endogeneity biases.

In view of the large number of non-completions, possibly those that answered the earnings question are not a random draw from the population. If so the earnings equations could be biased. Analogously the small proportion of females of working age who are employed or self-employed in market activities may not constitute a random sample of all Pakistani women. Consequently inferences about market returns to female schooling estimated from uncorrected equations could be misleading – if for example those working are more productive than those not. A third selection bias might arise because Mincerian earnings functions are typically estimated for sub-samples such as wage earners. In developing countries paid employees may not be representative of the majority or substantial minority of the population who are self-employed. If education determines whether a person is in the paid employment sector or not, an OLS estimate of the education earnings relation in paid employment will be biased upwards.

We use Heckman's correction for sample selectivity for each of the three possible biases. Household demographic variables such as 'head' and 'married', that will influence reporting of earnings, female labour force participation and occupational choice, are here used to satisfy the exclusion restriction in the first stage probit

The original Mincerian equation did not include any measure of ability, motivation or taste for education. Omitted variables bias estimates, when correlated with other independent variables. Omitting 'ability' will cause OLS coefficients to overestimate the returns to schooling if individuals with higher income-generating capacities are also individuals who choose more education (Ichino and Winter-Ember, 1999). The typical solution is the instrumental

variables (IV) methodology. This procedure at the same time addresses measurement error.

We use as instruments variables problems/reasons for not studying. Education is instrumented using six variables: *dtoexp* ('education too expensive'), *dtoofar* ('schools are too far away'), *dwah* ('had to help at home'), *dhtw* ('had to help with work'), *dpna* ('parents did not approve') and *dcnw* ('child not willing to attend school').

We also attempt to address the contribution of family decision making and characteristics to schooling and earnings by estimating a family fixed effects regression of earnings. To the extent that unobserved traits, such as ability, motivation, or taste for education, are shared within the family, their effects will be netted out in a family differenced model. It is unlikely that unobserved traits are identical across family members but they are likely to be more similar within a family than between families. Thus, family fixed effects estimation of earnings functions gives rates of return to schooling with reduced bias (Kingdon et al., 2007).

For the wage equations, observations were dropped if the respondent was less than 15 years or above 65 years of age. Earnings functions are estimated on subsamples of paid employees (47193 males and 6237 females), self-employed (23754 males and 1734 females) and agricultural workers (20774 males and 1322 females).

The dependent variable in earnings functions is the natural log of monthly earnings ($\ln y_m$). The definitions of the variables used in the earning function are given in table A1 in the Appendix. The education variable has been specified in two ways, as years of completed education (*school*) and as education dummy variables representing various levels of education (*prim*, *mid*, *sec*, *hsec* and *ter*). The reference category for the dummy variables is individuals with zero and less than five years of education.

The earning equations include experience and its quadratic term (*EXP* and EXP^2). This variable is computed as $(age - \text{years of education} - 5)$ on the grounds that individuals start schooling at the age of 5 and enter the labour market after education is completed. The vector of other variables in the earning equation includes dummy variables for the regions and provinces.

4. Estimating Earning Functions and Rates of Return

On the grounds of their robustness we first consider OLS estimates of equation 3 by occupation and gender in table 4.

4.1 The Basic Estimates: OLS

A table 4 shows that returns to an additional year of schooling are apparently greater for females than males in paid employment; the returns to an additional year of schooling are 9.2 percent and 14 percent for men and women respectively. These returns are higher than the estimates of Kingdon et al., 2007, (which were 3.3 percent for men and 14.9 percent for young women in the same sector), possibly because of their more restricted age range. This large and significant gender differences in Pakistan cannot be attributed to the scarcity of educated women; otherwise their wages would not be lower, as observed in section 2. Instead as the labour force participation rates also noted in section 2 indicate, the low probability or short duration of employment requires a high apparent rate of return compared with males, to provide an equivalent actual return.

With female employment at 10.5 percent of male, this could be accounted for by females having only a 10.5 percent chance of being in the employed workforce. In which case, actual returns for those in the workforce must be adjusted to take account of those outside, by multiplying by the probability. This leads to a far lower financial return to female schooling than to male. Alternatively females could be assumed to be in the labour force for the first 10.5 percent of the years of the average male, say four years. Because female earnings with this method are assumed to occur sooner, the discrepancy between males and female returns is smaller but still much larger than necessary to account for the observed difference. Earnings peak at 31 years and 37 years of experience respectively for males and females in the paid employment sector.

In the self employment sector, the returns to an additional year of education are 8.3 percent for men and 10.3 percent for women⁴. That the returns are closer might suggest that women are crowded in to the self-employed sector because of inadequate opportunities in paid employment and in agriculture. However, the small sample (745 in self-employment compared with 2950 in paid employment) contradicts the crowding interpretation- if sampling procedures are appropriate. A more plausible possibility is that self-employed women work for more years and in this respect are more comparable with men than in paid employment.

The returns to education in agriculture sector are 7.1 percent for males and 4.3 percent for females (a very small sample, reflecting very few opportunities). The coefficient of experience indicates substantial increase in wages with each additional year for males in each sector of the labour market but not for females outside paid employment.

⁴ These returns are greater for men than the estimates of Kingdon et al., 2007, which were 4.8 percent and 10.5 percent for young men and women respectively.

A greater labour supply at the individual level is a significant contributor to monthly earnings, the largest effect being for self-employed females, consistent with this sector being the only one where females are not discriminated against. An extra day worked by self-employed females increases their earnings by 4.5 percent, whereas for self-employed males, an extra day's work boosts earnings by only 0.8 percent. The coefficients on NWD (number of days worked during last month) are positive and statistically significant for all the sectors of labour market for both males and females, excepting females in agriculture.

Table 4: Earnings and years of schooling (OLS estimates) – dependent variable $\ln y_m$

Variables	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
School	0.092* (101.52)	0.140* (35.58)	0.083* (49.07)	0.103* (10.11)	0.071* (7.74)	0.043 (0.75)
Exp	0.062* (59.92)	0.072* (15.93)	0.056* (33.23)	0.021** (1.92)	0.058* (7.25)	0.019 (0.44)
Exp²	-0.001* (-38.41)	-0.001* (-8.59)	-0.001* (-22.43)	0.0001 (0.35)	-0.001* (-5.88)	0.00004 (0.06)
NDW	0.014* (17.15)	0.034* (11.57)	0.008* (5.43)	0.045* (8.40)	0.020* (3.40)	0.048 (1.38)
EDUEXP	0.001* (14.22)	0.002* (7.51)	0.001* (8.97)	0.004** (3.61)	0.001 (1.50)	0.003 (0.70)
Constant	6.047* (170.58)	3.781* (24.45)	6.577* (93.75)	2.521* (3.59)	6.020* (16.78)	3.190 (0.90)
R²	0.40	0.47	0.27	0.29	0.10	0.17
F-statistic	2551.89	323.45	731.50	37.43	19.21	1.31
Observations	30852	2950	16080	745	1404	52
Chow test (F-test)	889.33			10.40		
p-value	0.00			0.00		

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. T-statistics is in parentheses. Dummy variables for provinces and regional dummies also included in all regressions.

Provincial educational expenditures per school age population (EDUEXP) are also positive and statistically significant for both male and female in all sectors of the labour market, except for agriculture, where it is positive but not significant.

Since the variable is measured in money units (rupees), the average rate of return to this state investment in the quality of education (holding years of schooling constant) can be computed directly. Taking the average monthly earnings from table 2 and the coefficients of provincial education expenditure per school age person from table 4, the average returns for a rupee increase per head are;

paid employed (male) = $0.001 \times 4859 = 4.9$,

paid employed (female) = $0.002 \times 2980 = 6.0$,

self employed (male) = $0.001 \times 7057 = 7.1$

self employed (female) = $0.004 \times 2160 = 8.6$

agriculture (male) = $0.001 \times 5657 = 5.7$ (not significantly greater than zero)

agriculture (female) = $0.003 \times 1853 = 5.6$ (not significantly greater than zero)

These returns as percentages are colossal (4.9=490 percent) and can only be possible if the existing quality of education is extremely low. Interestingly they are greater for the self-employed than the employed.

According to human capital theory, education directly increases individual productivity by augmenting skills. Signalling theory by contrast represents education as a sign of productivity. To test these theories, we compare the coefficients of schooling estimated for the paid employed with those for the self-employed. Since the coefficients of schooling for the self-employed are lower than for the paid employment there is apparently a small element of signalling. But the bulk of the return to schooling is a consequence of enhanced human capital. Although Chow tests (table 4) reject the null hypothesis of equality of regression coefficients of schooling in the two sectors, consistent with the signalling hypothesis, the coefficients are sufficiently similar to support the position the schooling primarily raises human capital.

The linear schooling hypothesis, underpinning the estimates of Table 4, is tested in the first instance with OLS estimates of equation 4, and presented in tables 5 and 6. Increasing and significant marginal returns to each level of schooling identify a distinctive feature of Pakistani human capital supply. The coefficients on almost all education levels are positive and progressively increasing with higher levels of schooling in all sectors. The exceptions are for females in paid employment at primary level and in agriculture at secondary and higher secondary levels.

The coefficients at all education levels across all sectors of the labor market tend to be higher for females than for males. The apparent returns to an additional year of schooling at various levels (table 6) are substantially greater for women than men in both wage employed and self-employed sectors, though not in agriculture, except for the primary level. For agricultural sector males, the

returns to education increase until the higher secondary level and decline afterwards while for females the returns are high at the primary level.

These estimates are similar to the returns at different levels of Jamal et al., (2003) in Pakistan for the total paid employed workers, which were 3, 4, 16, 11 and 13 for primary, secondary, higher secondary, tertiary general and tertiary technical. The present study, however, uses a different classification for educational attainment. Our interpretation is that high returns reflect the small number of qualified persons because of shortcomings in the lowest levels of the education system. Therefore simply expanding the higher levels of education to take advantage of the higher returns would not dig down to the root of the problem.

Subject to two caveats, the observation that male monthly earnings in agriculture exceed those in paid employment is not consistent with an under-supply of educated manpower. The qualifications are about imputed returns to land and a possible agricultural premium for non-pecuniary disadvantages and/or greater costs in agriculture. The apparent high return for females is likely to be spurious for reasons already discussed. Even so the increasing rates for males are striking.

Table 5: Estimated returns at different levels of schooling by gender and occupation (OLS estimates)

Levels of education	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
Primary	0.057* (4.23)	-0.073 (-1.06)	0.042** (1.93)	0.063 (0.55)	0.087 (0.94)	0.558** (2.11)
Middle	0.176* (13.36)	0.237* (3.25)	0.218* (10.23)	0.204*** (1.65)	0.252* (2.73)	0.706 (1.49)
Secondary	0.392* (30.97)	0.569* (8.97)	0.385* (18.66)	0.494* (4.09)	0.344* (3.75)	0.247 (0.52)
Higher Secondary	0.635* (39.17)	0.851* (12.38)	0.652* (23.55)	0.762* (4.69)	0.816* (5.29)	-0.267 (-0.26)
Tertiary	1.042* (76.79)	1.426* (22.62)	0.978* (40.05)	1.255* (8.61)	0.938* (6.75)	-
NDW	0.014* (18.33)	0.035* (11.85)	0.008* (5.55)	0.044* (8.03)	0.020* (3.45)	0.037 (1.04)
EDUEXP	0.001* (13.94)	0.002* (7.65)	0.001* (9.12)	0.004* (3.46)	0.001 (1.36)	0.001 (0.22)
R²	0.41	0.47	0.28	0.30	0.11	0.27
F-statistic	1762.04	214.49	511.06	25.85	13.80	1.51
Observations	30852	2950	16080	745	1404	52

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. T-statistics is in parentheses. (-) indicates no observations. No education and less than 5 years of schooling is the reference category for education. Exp, Exp², regional and provincial dummy variables are included in all regressions.

Table 6: Rates of returns to additional year of education by level of education, gender and occupation

Level of education	Rates of return (%)					
	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
Primary	1.1	-1.5	0.8	1.3	1.7	11.2
Middle	4.0	10.3	5.9	4.7	5.5	4.9
Secondary	10.8	16.6	8.4	14.5	4.6	-23.0
Higher Secondary	12.2	14.1	13.4	13.4	23.6	-25.7
Tertiary	20.1	28.8	16.3	24.6	6.1	-

Source: Regression results (table 4). (-) indicates no observations.

The magnitude of increase in coefficients with education levels is much greater for females than males. Returns as high as 20 percent or more can persist only because shortcomings at lower levels of education restrict the supply of qualified students for higher levels⁵. Since these returns are higher than earlier estimates they may reflect a disequilibrium, where labour supply lags behind expanding demand. Alternatively they may stem from estimation bias.

4.2 Sample Selectivity Correction

OLS estimates of returns to education potentially suffer from sample selectivity bias. Those working in paid employment are more likely to have received substantial schooling and chosen to be in that sector than in agriculture perhaps. Those in agriculture will have received less schooling. In this case the marginal return to education will have two parts, one due to its influence increasing the probability of the individual moving in to the paid employment sector and the other due to its influence on earnings in paid employment. So the regression coefficient overstates the education effect in paid employment and understates it for agriculture.

We offset the sample selectivity bias by employing the Heckman two-step procedure and incorporating 'lambda' - the selection term - into earnings function estimates. The selectivity corrected earnings functions reported in table A2, A3 and A4 (appendix) include the standard variables – schooling, experience and its square, provincial educational expenditure per school age pupil, number of days

⁵ A schooling squared term added to the basic specification of table 4 is positive and significant, consistent with the results of table 5. However the estimates of the other coefficients are not substantially changed by addition of this term.

worked during the month and the provincial and regional dummy variables. Household demographic variables such relationship with the head of household (head) and marital status (married) are used as exclusion restrictions. These variables determine participation in the relevant occupation but do not directly affect workers' earnings. Most of them are statistically significant.

The selectivity term is large and statistically significant in most equations (table A2, A3 and A4) for males and females (in both years and levels of schooling specifications). A comparison of table 4 and A2 shows the effect of correcting for sample selection. Inclusion of the selectivity term in the earnings functions slightly increases the point estimates on years of schooling for males and females in the paid employed, and self-employed sectors. In the levels specifications, the inclusion of selectivity term has also slightly reduced the coefficients in most equations. These findings suggest that, because of movements between sectors, OLS underestimate the return to education specifically for females in the years of education specification (though for other reasons there may be different biases).

Experience and its square have a fairly standard relationship with earnings across all occupations for both male and female workers. For example, earnings peak at 30 years for males and 38 years for females in the paid employed sector of Pakistani labour market.

We have also estimated Heckman two step procedure based on the selection of reported income and labour force participation for both genders and the results are reported in table A5 (appendix). The selectivity term λ is significant in 3 out of 4 earnings equations. These results indicate that female returns to an additional year of schooling are apparently higher as compared to males for both selections. Moreover, the results are very close to OLS estimates for the paid employed sector. These results are consistent with some studies of Pakistan (Aslam 2007; Kingdon et al., 2007). Moreover, the convexity result is also similar to other recent studies of other countries (Kingdon and Unni, 2001 on India; Belzil and Hansen, 2002 for USA and Soderbom et al., 2005 in Kenya and Tanzania).

4.3 Endogenous Schooling: Instrumental Variable (IV) Estimates

The household choice of years of schooling for an individual cannot necessarily be regarded as independent of the expected earnings of that person. Aside from questions of ability, if earnings are extremely variable, then current earnings may dominate future earnings and the household may choose to reduce schooling when labour market prospects are buoyant. The perfect foresight underpinning equation (1) may not be appropriate. In this case the coefficient on schooling would be a downward biased estimator of the true return.

We approach this endogeneity issue by applying the IV procedure across all occupations, reported in table 7 and 8. The instruments are the problems of education/schooling (too expensive education; schools are too far away; had to help at home; had to help with work; poor schooling; parents did not approve and child are not willing to attend school)⁶. Schooling is specified as a continuous variable⁷.

The summary statistics reveal R^2 ranging from 0.10 to 0.37 for males and 0.07 to 0.41 for females across the occupations. The rates of return to an additional year of education for males are 12.8 percent, 11.3 percent and 7.0 percent in paid employment, self employment and agriculture sectors respectively. While for females they are 20.9 percent, 18.7 percent and -7.4 percent in paid employment, self employment and agriculture sectors respectively. The rate of return to education for females in agriculture occupation is negative probably because of the small number of observations.

The findings confirm the previous section results that the apparent rate of return to education is higher for females as compared to males in both the paid- and self-employed occupations. Moreover, the results are consistent with other studies (Aslam, 2007) that IV estimates are higher as compared to OLS in some cases. In that there is at least a one and a half percentage point difference between employed and self-employed sectors, the IV result supports the signaling theory for both males and females.

4.4 Ability Bias: Household Fixed Effects (FE) Estimates

The 'nature and nurture' bias of schooling return estimates is addressed by estimating a household fixed effects earnings function across sectors and genders. Identification of the effect of education on earnings comes only from within family variation among members in earnings and in education. The estimates are based on sub-samples of at least one worker within household who is related in any way (for example, father-daughter, mother-son, or husband-wife) or are siblings (only brother-sisters pairs). The results are reported in table A6, A7 and A8 (appendix). For male paid employees there are almost 5000 households and an average of about six male paid employees per household. On the other hand for self-employed females there are only 622 households with a total of 745 cases.

⁶ Arrazola et al. (2003) used instrument variables conveying information about differences in opportunities to estimates returns to education in Spain.

⁷ With few instruments for each sample, the 'levels' specification cannot be used in IV procedure.

Focusing first on the schooling years specification, the fixed effects point estimates are lower than the OLS estimates in table 4 for all sectors (for males the returns fall from 9.2 percent to 8.5 percent, and from 8.3 percent to 5.6 percent in the paid- and self-employed sectors respectively, while slightly increase from 7.1 percent to 7.7 percent in the agriculture sector).

Table 7: Earnings and years of schooling for males (IV estimates) – dependent variable lnym

Variables	Paid employed		Self employed		Agriculture	
	IV	First stage	IV	First stage	IV	First stage
School	0.128* (37.81)	-	0.113* (14.28)	-	0.070*** (1.72)	-
Exp	0.061* (56.78)	-0.017* (-2.72)	0.057* (31.86)	-0.080* (-10.28)	0.056* (6.07)	-0.111* (-4.59)
Exp²	-0.001* (-33.27)	-0.001* (-9.24)	-0.001* (-21.30)	-0.00002 (-0.19)	-0.001* (-5.40)	0.0006 (1.54)
NDW	0.010* (9.35)	0.107* (22.71)	0.008* (4.97)	0.011*** (1.62)	0.021* (3.54)	0.010 (0.37)
EDUEXP	0.001* (10.64)	0.001* (5.94)	0.001* (9.10)	-0.002* (-4.33)	0.001 (1.44)	-0.0001 (-0.06)
dtooxp	-	-2.232* (-29.28)	-	-1.654* (-16.23)	-	-0.684** (-2.40)
dtoofar	-	-1.879* (-9.19)	-	-0.967* (-3.79)	-	-1.588* (-3.09)
dwah	-	-1.692* (-9.98)	-	-1.314* (-7.17)	-	-1.214* (-2.64)
dhtw	-	-2.537* (-21.72)	-	-1.315* (-10.62)	-	-1.003* (-2.80)
dpna	-	-2.491* (-11.00)	-	-1.707* (-6.55)	-	-2.200** (-2.78)
dcnw	-	-3.092* (-38.08)	-	-2.145* (-21.63)	-	-1.578* (-5.90)
Constant	5.915* (152.88)	6.851* (30.97)	6.277* (59.31)	11.810* (37.05)	6.032* (10.61)	11.134* (10.65)
R²	0.37	0.17	0.25	0.16	0.10	0.14
F-statistic	1362.04	460.53	437.01	235.17	10.10	17.02
Observations	30469	30469	15910	15910	1379	1379
Chow test (F-test)	903.79					
p-value	0.00					

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. T-statistics is in parentheses. Dummy variables for provinces and regional dummies also included in all regressions.

Table 8: Earnings and years of schooling for females (IV estimates) – dependent variable $\ln y_m$

Variables	Paid employed		Self employed		Agriculture	
	IV	First stage	IV	First stage	IV	First stage
School	0.209* (20.26)	-	0.187* (5.68)	-	-0.074 (-0.41)	-
Exp	0.075* (14.96)	-0.099* (-4.89)	0.035* (2.68)	-0.216* (-5.60)	-0.016 (-0.24)	-0.307* (-2.96)
Exp²	-0.001* (-7.26)	-0.0002 (-0.30)	-0.00004 (-0.14)	0.002* (2.46)	0.0005 (0.49)	0.003** (1.99)
NDW	0.021* (5.71)	0.142* (11.25)	0.042* (7.05)	0.045** (2.42)	0.038 (0.96)	-0.065 (-0.66)
EDUEXP	0.001* (5.21)	0.002** (2.29)	0.003* (2.83)	0.004 (1.01)	0.003 (0.70)	-0.006 (-0.44)
dtoexp	-	-3.529* (-13.20)	-	-2.712* (-7.11)	-	1.415 (0.59)
dtoofar	-	-3.015* (-7.57)	-	-2.315* (-4.07)	-	-1.353 (-0.98)
dwah	-	-2.360* (-12.33)	-	-1.752* (-5.94)	-	0.252 (0.33)
dhtw	-	-2.910* (-7.42)	-	-1.360* (-2.05)	-	-1.425 (-0.65)
dpna	-	-4.096* (-14.33)	-	-2.288* (-5.49)	-	-1.384 (-1.38)
dcnw	-	-4.338* (-11.88)	-	-2.352* (-4.06)	-	-2.266 (-1.04)
Constant	3.565* (21.08)	8.425* (11.97)	2.129* (2.85)	8.844* (3.61)	4.504 (1.12)	15.446 (1.34)
R²	0.41	0.36	0.23	0.28	0.07	0.48
F-statistic	194.98	118.99	27.05	21.39	1.04	3.33
Observations	2830	2830	729	729	51	51
Chow test (F-test)	23.12					
p-value	0.00					

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. T-statistics is in parentheses. Dummy variables for provinces and regional dummies also included in all regressions.

These findings are consistent with the previous literature⁸. The decline in estimates may be because of an upward bias in the OLS estimates due to omitted ability and household human capital investment variables. It could also be due to

⁸ In his study using South African data, Hertz (2003) finds that whereas OLS estimates yield returns of about 13 percent and returns to education are 3 percent by using FE. Behrman and Deolalikar (1995) also find FE estimates for male and female workers to be significantly lower than corresponding OLS estimates.

schooling measurement error⁹. The estimates for the 'levels' specification present a similar picture.

The comparison of fixed effects estimates and IV estimates (table 7 and 8) reveal that the rates of return to education are higher with the IV procedure than the FE estimates for paid- and self-employed occupations for both genders. While the rates of return to education for males are higher by FE (7.7 percent) than IV (7.0 percent) in the agriculture sector.

We have also estimated household fixed effect with schooling IVs (table 9). The estimates are close to the IV estimators except for females in self-employment. The fixed effects (IV) results appear to imply that within households' variation in female schooling does not affect their earnings, but between household it does. This is because there is little variation within households and most is between in these variables, in contrast to male schooling and earnings.

⁹ Hertz (2003) corrects for measurement error in schooling using two observations of schooling on the same individual. This was possible because 13 percent of the individuals were re-surveyed in the sample to get measures of reliability of schooling. However, he still finds the return to education to be lower than OLS. The correction for ME raises the estimates over the uncorrected ones. Data constraints prevent such a correction in the current paper.

Table 9: Earnings and years of schooling: controlling for household fixed effects (with schooling IVs)- dependent variable lnym

Variables	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
School	0.120* (30.82)	0.235* (8.21)	0.094* (9.85)	-0.006 (-0.04)	0.064 (0.98)	-
Exp	0.058* (51.67)	0.070* (7.94)	0.046* (23.11)	0.0002 (0.01)	0.036* (2.51)	-
Exp²	-0.001* (-30.08)	-0.001* (-3.74)	-0.001* (-15.61)	0.0002 (0.42)	-0.001*** (-1.87)	-
NWD	0.011* (10.99)	0.016* (2.47)	0.010 (5.73)	0.041* (2.55)	0.004 (0.33)	-
EDUEXP	0.027 (0.00)	0.007 (0.00)	-	0.003 (0.00)	0.011 (0.00)	-
Rural	-	-	-	-	-	-
Constant	-14.462 (-0.00)	-1.011 (-0.00)	6.868* (66.87)	4.555* (0.00)	-0.666 (-0.00)	-
R²(within)	0.31	0.09	0.11	0.12	0.08	-
R²(between)	0.04	0.40	0.27	0.15	0.0007	-
R²(overall)	0.04	0.39	0.23	0.15	0.0004	-
Wald chi2	6555.61	187.39	3.01	16.81	7.64	-
p-value	0.00	0.00	0.00	0.01	0.18	-
Observations	30469	2830	15910	729	1379	-
No. groups	4981	1811	4560	609	1010	-
F-test that all u_i=0	1.56	1.00	1.76	2.40	1.31	-
p-value	0.00	0.51	0.00	0.00	0.00	-

Note: ** and * represent significance at the 5% and 1% levels respectively. T-statistics is in parentheses. (-) indicates no observations. Dummy variables for provinces also included in all regressions.

5. Conclusion

There is some signaling in Pakistani education investment but mainly the education is productivity-enhancing investment in human capital, according to a

comparison of self-employed and paid employed earnings equations. Returns to public spending of education are extremely high, suggesting very considerable state underinvestment. This most probably accounts for the increasing returns with higher levels of schooling (which are difficult to square with the Mincer assumptions). Inadequate primary education reduces the numbers able to take advantage of secondary education, and so on through the levels of education.

Instrumental variable (IV) correction for endogenous schooling raises estimated returns to schooling, while household fixed effects, which are presumed to control for ability, reduce the measured return. The IV effect implies a negative association of schooling with earnings, which seems contrary to the simple Mincer model.

Low labour force participation of women must be a contributor to the apparently high measured returns to female schooling. Indeed this low participation is likely to reduce the private financial returns below that of males. It would therefore be mistaken to recommend more investment in female education simply on the basis of high returns. The policy challenge is in the low wages and high education in the female paid employment sector, and the low participation rate. These reduce the private financial returns to female education and therefore the private incentive to invest.

References

1. Aslam M. (2007), "Rates of return to education by gender in Pakistan," GPRG-WPS-064, Global Poverty Research Group-WPS-064, www.gprg.org.
2. Aslam, M. and Kingdon, G. (2006), "Gender and household education expenditure in Pakistan", CSAE Working Paper Series GPRS-WPS 025, University of Oxford. Forthcoming in Applied Economics.
3. Arrazola, M., De Hevia J., Risueno M. and Sanz J.F., (2003), "Returns to education in Spain: some evidence on Endogeneity of school", Education Economics, vol. 11, No. 3, 293-304.
4. Behrman, J.R. and Deolalikar (1995), "Are there differential returns to schooling by gender? The case of Indonesian labour market," Oxford Bulletin of Economics and Statistics, 57(1), pp 97-117.
5. Belzil, C. and Hansen, J. (2002), "Unobserved ability and the return to schooling", Econometrica, 70(5), pp 2075-91.
6. Berriedale Keith, A (1922), (ed). "Speeches and Documents on Indian Policy, 1750-1921", Vol. I. London: Humphrey Milford, Oxford University Press.

7. Government of Pakistan, (2002), "Economic Surveys of Pakistan: Health section".
8. Government of Pakistan, (2003), "Economic Surveys of Pakistan: Health section".
9. Hertz, T. (2003), "Upward biased in the estimated returns to education: evidence from South Africa," *The American Economic Review*, 93(4), pp 1354-68.
10. Jamal. H., Asharf I., and Sami, F., (2003), "Private returns to education: evidence from Pakistan", State Bank of Pakistan, Research Report No. 50.
11. Kingdon, G. and Unni, J. (2001), "Education and women's labour market outcomes in India", *Education Economics*, Vol. 9(2), pp 173-195.
12. Kingdon G. and Soderman M. (2007), "Education, skills and labor market outcome: evidence from Pakistan", Department of Economics, University of Oxford.
13. Harmon C., Hogan V. and Walker I. (2003), "Dispersion in the economic return to schooling", Vol. 10, pp 205-214.
14. Harmon, Colm, Hessel Oosterbeek, and Ian Walker (2003), "The Returns to Education:Microeconomics", *Journal of Economic Surveys*, 17, 115-155.
15. Ichino A. and Winter-Ebmer R. (1999), "Lower and upper bounds of returns to schooling: An exercise in IV estimation with different instruments", *European Economic Review*, Vol. 43, pp 889-901.
16. Mincer J. (1958), "Investment in human capital and the personal income distribution", *Journal of Political Economy*, Vol. 66; 4, pp 281-302.
17. Nayyar A H and Salim A (eds) (2003), "The Subtle Subversion: The State of Curricula and Textbooks in Pakistan. Urdu, English, Social Studies and Civics", Sustainable Development Policy Institute, Islamabad.
18. Psacharopoulos George, and Harry Anthony Patrinos (2004), "Human capital and rates of return" in *International Handbook on the Economics of Education*, edited by Geraint Johnes and Jill Johnes, 1-57, Cheltenham, UK: Edward Elgar.
19. Rosen S. (1992), "Distinguished Fellow: Mincering Labor Economics", *Journal of Economic Perspectives*, 6(2),pp 157-170.
20. Soderbom, M., Teal, F., Wambugu, A. and Kahyarara, G. (2005), "The dynamics of returns to education in Kenyan and Tanzanian manufacturing", *Global Poverty Research Group Working Paper Series*, No.017.
21. Stanovink T. (1997), "The returns to education", *Economic of Education Review*, Vol.16 (4), Pergamon Publications.

Appendix

Table A1: Definition of Variables used in the Study

Variable	Description
EXP	Experience (age-schooling-5)
EXP ²	Square of experience
lnym	Log of monthly earnings from main occupation
married	Equals 1 if individual is married, 0 otherwise
rural	Equals 1 if the individual is from rural area, 0 otherwise
Punjab	Equals 1 if the individual is from Punjab province, 0 otherwise
Sindh	Equals 1 if the individual is from Sindh province, 0 otherwise
NWFP	Equals 1 if the individual is from NWFP province, 0 otherwise
Baloch	Equals 1 if the individual is from Balochistan province, 0 otherwise
school	Number of years of education acquired
No school	Equals 1 if individual reports 0 years of education
Less_primary	Equals 1 if individual has completed less than 5 years of education that class 1, 2, 3 or 4, 0 otherwise
prim	Equal 1 if individual has completed 5 years of schooling, 0 otherwise
mid	Equals 1 if individual has completed 6, 7 or 8 years of schooling, 0 otherwise
sec	Equals 1 if individual has completed 9 or 10 years of schooling, 0 otherwise
hsec	Equals 1 if individual has completed 11 or 12 years of schooling, 0 otherwise
ter	Equals 1 if individual has completed 13 years of education or more, 0 otherwise
lit	Equals 1 if individual can read or write in any language with understanding, 0 otherwise
num	Equals 1 if individual can solve simple (plus minus) sums, 0 otherwise
dtooexp	Equals 1 if education is too expensive, 0 otherwise
dtoofar	Equals 1 if schools are too far, 0 otherwise
dwah	Equals 1 if doing working at home, 0 otherwise
dhtw	Equals 1 if helping at work, 0 otherwise
dpna	Equals 1 if parent did not allow to go to school, 0 otherwise
dcnw	Equals 1 if child is not willing to go to school, 0 otherwise.
NWD	Number of days worked during last month
EDUEXP	Provincial education expenditure per school age pupil

Table A2: Earnings and years of schooling (Hickman two-step estimates) – dependent variable lnym

Variables	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
School	0.103* (70.11)	0.183* (16.15)	0.091* (37.39)	0.226* (2.97)	0.077* (7.50)	0.018 (0.29)
Exp	0.060* (45.58)	0.075* (14.28)	0.046* (17.57)	0.022 (1.51)	0.057* (6.81)	0.027 (0.66)
Exp²	-0.001* (-34.12)	-0.001* (-8.33)	-0.001* (-15.05)	0.0002 (0.48)	-0.001* (-5.56)	-0.00002 (-0.03)
NWD	0.005* (4.20)	0.039* (11.13)	-0.002 (-0.84)	0.075* (3.81)	0.020* (3.39)	0.057*** (1.63)
EDUEXP	0.001* (16.55)	0.002* (7.38)	0.001* (9.18)	0.008* (2.87)	0.001** (2.10)	0.003 (0.79)
Constant	5.418* (79.40)	2.252* (5.46)	7.293* (48.51)	0.132 (0.08)	7.359* (7.71)	2.138 (0.61)
Lambda (1)	0.815* (12.53)	0.921* (4.18)	-0.823* (-5.66)	-1.743*** (1.65)	-0.679 (-1.52)	0.271 (0.64)
Uncensored Observations	30852	2950	16080	745	1404	52
Wald chi2	14588.98	1426.40	5255.14	443.30	1193.18	8763.91
Pvalue (wald)	0.00	0.00	0.00	0.00	0.00	0.00

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. Z-statistics is in parentheses. Dummy variables for provinces and regional dummies also included in all regressions.

Table A3 Estimated returns at different levels of schooling by gender and occupation (Heckman estimates)

Levels of education	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
Primary	0.015 (0.91)	-0.010 (-0.14)	-0.010 (-0.23)	-0.320 (-0.88)	0.084 (0.88)	0.540** (2.22)
Middle	0.140* (8.56)	0.239* (3.18)	0.154* (5.64)	-0.074 (-0.22)	0.225** (2.32)	0.659 (1.44)
Secondary	0.353* (22.43)	0.441* (5.05)	0.316* (11.67)	1.085** (2.32)	0.311* (3.20)	0.192 (0.41)
Higher Secondary	0.643* (31.95)	0.651* (5.66)	0.630* (19.37)	2.137** (2.39)	0.856* (5.35)	-0.369 (-0.38)
Tertiary	1.145* (60.07)	1.179* (9.15)	1.069* (33.01)	3.309* (2.46)	1.043* (6.72)	-
NDW	0.010* (5.69)	0.033* (9.91)	-0.001 (-0.59)	0.087* (2.81)	0.021* (3.46)	0.041 (1.18)
EDUEXP	0.001* (16.23)	0.001* (3.35)	0.001* (9.31)	0.010** (2.20)	0.001** (1.93)	0.001 (0.27)
Lambda (1)	0.761* (12.02)	-0.568** (-2.24)	-0.792* (5.61)	-2.85*** (-1.58)	-0.619 (-1.41)	0.108 (0.28)
Uncensored Observations	30852	2950	16080	745	1404	52
Wald chi2	15809.63	1480.25	5880.98	423.72	1209.08	8713.54
Pvalue (wald)	0.00	0.00	0.00	0.00	0.00	0.00

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. Z-statistics is in parentheses. (-) indicates no observations. No education and less than 5 years of schooling is the reference category for education. Exp, Exp², regional and provincial dummy variables are included in all regressions.

Table A4 Rates of returns to additional year of education by level of education, gender and occupation (Heckman procedure)

Level of education	Rates of return (%)					
	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
Primary	0.3	--0.2	-0.2	-6.4	1.7	10.8
Middle	4.2	8.3	5.5	8.2	4.7	4.0
Secondary	10.7	10.1	8.1	58.0	4.3	-23.4
Higher Secondary	14.5	11.0	15.7	52.6	27.3	-28.1
Tertiary	25.1	26.4	22.3	58.6	9.4	-

Source: Regression results (table 8). (-) indicates no observations.

Table A5: Earnings and years of schooling (Hickman two-step estimates) using selection of reported income and labour force participation – dependent variable lnym

Variables	Selection for reported income		Selection for labour force participation	
	Male	Female	Male	Female
School	0.090* (98.33)	0.139* (35.72)	0.088* (102.07)	0.135* (29.85)
Exp	0.062* (67.36)	0.064* (14.86)	0.062* (65.95)	0.061* (11.71)
Exp²	-0.001* (-42.56)	-0.001* (-7.54)	-0.001* (-42.05)	-0.001* (-5.61)
NWD	0.014* (19.63)	0.045* (12.14)	0.015* (12.15)	0.026* (4.99)
EDUEXP	0.001* (13.65)	0.002* (8.24)	0.001* (10.26)	0.002* (6.96)
Constant	6.189* (182.50)	3.355* (18.95)	6.207* (172.13)	3.918* (17.94)
Lambda (1)	0.298* (5.78)	0.509** (2.34)	0.208 (0.75)	-0.998* (-3.29)
Uncensored Observations	48410	3767	48188	3737
Wald chi2	25999.01	2489	21709.56	1810.51
Pvalue (wald)	0.00	0.00	0.00	0.00

Table A6 Earnings and years of schooling: controlling for household fixed effects – dependent variable lnym

Variables	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
School	0.085* (85.76)	0.115* (16.10)	0.056* (29.95)	0.042** (2.02)	0.077* (4.76)	-
Exp	0.059* (54.43)	0.060* (8.49)	0.044* (24.69)	0.003 (0.14)	0.044* (3.20)	-
Exp²	-0.001* (-34.91)	-0.001* (-4.96)	-0.001* (-16.44)	0.0003 (0.60)	-0.001* (-2.41)	-
NWD	0.015* (17.24)	0.031* (6.29)	0.010 (6.45)	0.038* (3.29)	0.008 (0.73)	-
EDUEXP	-	-	-	-	-	-
Constant	6.411* (253.20)	5.288* (34.59)	7.211* (140.99)	5.895* (16.12)	6.783* (18.42)	-
R²(within)	0.34	0.28	0.14	0.16	0.08	-
R²(between)	0.47	0.48	0.25	0.20	0.08	-
R²(overall)	0.38	0.45	0.22	0.21	0.09	-
F-statistic	3356.41	107.75	463.04	5.44	8.40	-
Observations	30852	2950	16080	745	1404	-
No. groups	4987	1858	4571	622	1024	-
F-test that all u_i=0	1.70	1.29	1.94	2.54	1.29	-
p-value	0.00	0.00	0.00	0.00	0.00	-
Chow test (F-test)	887.66			1.80		
p-value	0.00			0.17		

Note: ** and * represent significance at the 5% and 1% levels respectively. T -statistics is in parentheses. (-) indicates no observations. Dummy variables for provinces and regional dummies also included in all regressions.

Table A7: Estimated returns at different levels of schooling by gender and occupation- controlling for household fixed effects

Levels of education	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
Primary	0.053 (3.71)	-0.099 (-0.83)	0.040** (1.68)	-0.260 (-1.44)	0.034 (0.21)	-
Middle	0.144* (10.24)	0.131 (1.05)	0.171* (7.45)	-0.115 (-0.59)	0.132 (0.85)	-
Secondary	0.353* (26.12)	0.473* (4.13)	0.298* (13.28)	0.221 (1.08)	0.388* (2.49)	-
Higher Secondary	0.576* (33.67)	0.697* (5.68)	0.466* (15.68)	-0.301 (-0.98)	0.834* (3.57)	-
Tertiary	0.951* (65.25)	1.132* (9.69)	0.657* (24.37)	0.237 (0.85)	0.821* (3.15)	-
NDW	0.015* (18.22)	0.031* (6.27)	0.011* (6.54)	0.046* (3.86)	0.010 (0.75)	-
EDUEXP	-	-	-	-	-	-
R²	0.39	0.45	0.24	0.16	0.09	-
F-statistic	1747.04	53.21	237.14	3.82	4.54	-
Observations	30852	2950	16080	745	1404	-
No. groups	4987	1858	4571	622	1024	-
F-test that all $u_i=0$	1.69	1.28	1.90	2.61	1.28	-
p-value	0.00	0.00	0.00	0.00	0.00	-

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. Z-statistics is in parentheses. (-) indicates no observations. No education and less than 5 years of schooling is the reference category for education. Exp, Exp², regional and provincial dummy variables are included in all regressions.

Table A8 Rates of returns to additional year of education by level of education, gender and occupation- controlling for household fixed effects

Level of education	Rates of return (%)					
	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
Primary	1.1	-1.9	0.8	-5.2	0.7	-
Middle	3.3	7.6	4.4	-4.8	4.9	-
Secondary	10.5	17.1	6.4	16.8	12.8	-
Higher Secondary	11.2	11.2	8.4	-26.1	22.3	-
Tertiary	18.8	21.8	9.6	26.9	-0.7	-

Source: Regression results (table13). (-) indicates no observations.

Table A9: Summary of schooling coefficients by gender and sectors for estimation procedures

Procedures	Coefficients on schooling					
	Paid employed		Self employed		Agriculture	
	Male	Female	Male	Female	Male	Female
OLS	0.092* (101.52)	0.140* (35.58)	0.083* (49.07)	0.103* (10.11)	0.071* (7.74)	0.043 (0.75)
Heckman	0.103* (70.11)	0.183* (16.15)	0.091* (37.39)	0.226* (2.97)	0.077* (7.50)	0.018 (0.29)
IV	0.128* (37.81)	0.209* (20.26)	0.113* (14.28)	0.187* (5.68)	0.070** (1.72)	-0.074 (-0.41)
FE	0.085* (85.76)	0.115* (16.10)	0.056* (29.95)	0.042** (2.02)	0.077* (4.76)	-
FE(IV)	0.120* (30.82)	0.235* (8.21)	0.094* (9.85)	-0.006 (-0.04)	0.064 (0.98)	-

Note: ***, **, and * represent significance at the 10%, 5% and 1% levels respectively. t-statistics is in parentheses.