NEW EVIDENCE ON RATES OF RETURN TO EDUCATION IN PAKISTAN

Javed Ashraf, University of St. Thomas

ABSTRACT

Relatively few articles have examined the rates of return to education in Pakistan in recent years. This paper uses data from the 2001-02 Pakistan Integrated Household Survey to compute returns from different levels of education. Returns to education are found to be higher for females than for males at the lowest level of schooling. The results presented here support the findings of other researchers. In addition the result have significant policy implications for educators and policymakers.

JEL: J24; J31

KEYWORDS: Education, Returns, Pakistan

INTRODUCTION

After a flurry of studies in the 1970s and 1980s, there appears to have been a barren period in which little attention was directed towards computing rates of return to education in Pakistan. Recently however, some studies, both published and unpublished have investigated this important area. This article updates, and provides comparisons with the results of some of these studies. Most of these studies, including this one, determine that the returns to females are higher than they are for males. This is a finding with considerable policy significance.

Pakistan is estimated to have a population of approximately 160 million. The labor force participation rate is 46%, reflecting the large number of children in the population who cannot work. The participation rate is only 19% for women. This is not surprising, given the social and cultural factors which often prevent women from engaging in paid employment. Agriculture is the mainstay of the economy accounting for 43% of all employment. The unemployment rate of 6.2% reported officially, is widely regarded as being significantly understated. As much as 73% of employment outside agriculture is in the informal sector, which is largely unregulated and untaxed. The literacy rate is 53%, but is only 40% for women.

This paper uses data from the *2001-2002 Pakistan Integrated Household Survey* to compute reported results. While generally supportive of results from previous studies, the use of new data reinforces these earlier findings. The paper provides a survey of some of the more recent studies. It uses a modification of Oaxaca's model suggested by Cotton (1988) and Neumark (1988) to estimate results. Empirical findings are reported thereafter.

The remainder of the paper is organized as follows. In the following section, a survey of the literature is provided. Next, the methodology and data used in the study are discussed. The empirical results are presented and discussed. The paper closes with some concluding comments.

LITERATURE REVIEW

Hamdani (1977) used data from the *1975 Socio-Economic Survey of Rawalpindi* carried out by the *Pakistan Institute of Development Economics*. The same data was used by Haque (1977) and Guisinger *et. al.* (1984) in calculating rates of returns to education. A fourth study by Khan and Irfan (1985) used

the *Population, Labor Force and Migration (PLM) Survey.* The Survey was a joint project of the Pakistan Institute of Development Economics and ILO-UNFPA. Another study from this time period was Pasha and Wasti (1995).

Comparison between the results of these studies is difficult. Khan and Irfan's results show the percentage gains from each educational level relative to a common base (which is "Less than primary level of education."). However, Hamdani, Haque, Guisinger *et. al.,* and Pasha and Wasti report incremental returns to each educational level, so that returns relative to a base level would have to be calculated separately. Guisinger *et. al.* (like Hamdani and Haque) based their study on data from the *1975 Socioeconomic Survey of Rawalpindi*. They reported the surprising finding that the returns to education were low. They also found, contrary to their expectations, that instead of falling, the rate of return to schooling rose with higher educational levels. They partly attributed their results to the differences in their methodology compared to Hamdani, who had found higher returns to education.

Among the more recent studies, Nasir and Nazli (2000) used data from the *Pakistan Integrated Household Survey, 1995-96,* which covered 12,622 households, and more than 84,000 individuals. They examined the effect of education, technical training, and school quality on the earnings of wage earners and salaried individuals. Their study differed from previous ones in that they were able to estimate the increase in earnings resulting from an additional year of education at different educational levels. An important finding was that each year of education raises salary by approximately 7%.

Nazli (2004) used data from the *Pakistan Socio-Economic Survey (PSES) 1998-99*. The survey used a two-stage stratified random sampling design to select a sample of 3,564 households. She examined the effect of education, experience and occupation on individual earnings for wage earners and salaried individuals. She found that the education-experience interaction had a positive and significant impact on earnings. When she stratified the earnings functions according to experience groups, she found the returns to education declined as experience increased.

In an excellent, article, Aslam (2007) used four statistical methods to estimate rates of return to males and females. She found that the estimated return to additional years of education ranged between 7-11% for men, and between 13-18% for women. She also found the education-earnings profile to be sharply convex for both males and females, and provided explanations for this pattern. Aslam's data were from the *Pakistan Integrated Household Survey, 2002.*

Hyder (2007) defined seven levels of education, and computed the rate of return to each, relative to the preceding level. The gains ranged from 1.5% for primary education to 9.23% for professional education. Much of her paper dealt with differences between the public and private sector. Her data were from the *Labor Force Survey, 2001-02.*

Abbas and Foreman-Peck (2007) used data from the *Pakistan Social and Living Standards Measurement Survey 2004-05*. Consistent with other studies, they found the rates of return to be consistently higher for females than for males. Among paid employed workers, they found the returns for males to range from 5.7% for primary education to 63.5% for higher secondary education. Returns for females were higher, ranging from -7.3% at the primary education level to 142% for tertiary level. In both cases, the reference group was workers with less than five years of schooling.

METHODOLOGY AND DATA

The data used for this study were drawn from the *2001-02 Pakistan Integrated Household Survey (PIHS),* and were collected by the Government of Pakistan's Federal Bureau of Statistics. The *PIHS* is a large sample survey covering 16,182 households. The occupation variables were determined using the

Pakistan Standard Classification of Occupations (PSCO), 1994 as revised by the *International Standard Classification of Occupations (ISCO),* 1988.

The wage equation used for the estimates in this study is:

$$\text{Log Monthly Salary} = \alpha + \sum_{i=1}\beta X_i + \sum_{i=1}\delta_i OC_i + \sum_{i=1}\mu_i IC_i \quad \text{equation} \qquad (1)$$

where the X_i represent eleven characteristics of working individuals that impact earnings, OC_i are eight occupational categories, and IC_i represent nine industrial categories. In addition to AGE (which was a proxy for experience), and AGE-SQUARED (which captures the concavity of the age-earnings profile), we included the following variables: Marital status represented by a dummy variable. This variable was included to test the hypothesis that a correlation exists between marital status and earnings, as is suggested in the literature.

English instruction was a dummy variable applied to survey respondents who received their education in English. We are not aware of a previous study using Pakistani data that has included this variable in an earnings equation. The rationale for its use was straightforward. Schools where the medium of instruction is English are generally viewed as elitist, and cater to the more affluent segments of the society. The quality of instruction is regarded as superior to those schools with mediums of instruction in other languages.

URDU instruction. This dummy variable applied to survey respondents who received their education in Urdu. A number of schools in Pakistan impart instruction in a regional language. Our hypothesis is that while schools where students are taught in Urdu are superior to those in the regional languages, they are generally not as good as those with English as the medium of instruction.

We used five dummy variables for different levels of education: Middle, Matric, Intermediate, BA/BSc., and Masters or other advanced degree. The missing base variable was education below Middle level. The *a priori* expectation was that earnings would rise monotonically as the level of education rises.

Dummy variables were used for eight different occupational categories: legislators, senior officials, and managers; professionals; technicians and associate professionals; clerks; service workers, shop and market sales workers; skilled agricultural and fishery workers; and plant and machine operators and assemblers. The missing base variable for the occupational categories was *Elementary Occupations.*

Nine industrial classifications were also recognized, and dummy variables were used in the regressions to identify each of them. These industries were mining and quarrying; manufacturing; electricity, gas and water; construction; trade hotels and restaurants; transport and storage; finance and real estate; community services; other activities not defined. The missing base variable for industrial classifications was *Agriculture, Forestry, Hunting and Fishing.*

It is very likely the case that working women in a less-developed country like Pakistan do not constitute a random sample of all females in the population. This is the well-known selectivity-bias problem expressed by Heckman (1979), following whom variables to correct for samples selectivity bias were developed, and used as an additional explanatory variable in the wage equations. These variables, or 'Inverse-Mills ratios' have been seldom employed in computations of the gender earnings gap, but have been widely used in the literature on union-nonunion wage differentials. They are $[f(EMP_i/F(EMP_i]$ and $[-f(EMP_i/\{1-F(EMP_i]\}$ where F and f are the cumulative and density functions of a standard normal variable. EMP_i is the predicted employment status of an individual, obtained from probit estimates of the reduced-form equation determining employment status. Given the likelihood of different factors not

influencing the work decision of males and females in the same manner, the selectivity variables were computed from separate probits for each sex. Compactly stated then, joint determination of participation and earnings is given by:

$$\ln W_{mi} = \beta_{m0} + \beta_{m1}X_{mi} + \beta_{m2}[-f(EMP_i/F(EMP_i] + u_{mi} \quad (2)$$
$$Emp_{mi} = Z'\delta_{mi} + \varepsilon_{mi} \quad (3)$$

for males, and:

$$\ln W_{mi} = \beta_{m0} + \beta_{m1}X_{mi} + \beta_{m2}[-f(EMP_i/F(EMP_i] + u_{mi} \quad (4)$$
$$Emp_{mi} = Z'\delta_{mi} + \varepsilon_{mi} \quad (5)$$

for females. X_{mi} and X_{fi} represent characteristics of males and females respectively, as represented in equation (1), and Z is a subset of X, and represents those worker characteristics that are instrumental in determining whether an individual will be in the workforce. In equations (2) and (4), the selectivity variables proxy for the probability of male and female participation in the workforce. The use of these variables leads to consistent estimation of the coefficients of the equation.

EMPIRICAL RESULTS

Only individuals between the ages of 16 and 65 were used in this study. The mean age of 37 indicates that the sample is relatively young. Eighty percent of the survey respondents were married. The sample is heavily biased toward men, with 96% being male. We identified English, Urdu, and Regional Languages as the mediums in which instruction is imparted in Pakistan. Only 3% of the respondents attended English-medium institutions, 50% went to Urdu-Medium institutions, and the remaining 47% received their education in regional languages. Forty-eight percent had a level of education below Middle. The percentage of those with higher levels of education was Middle (12%), Matric (17%), Intermediate (8%), Bachelors (7%), and Masters or other advanced degree (4%). Table 1 lists the coefficient estimates from three different regressions: the first for the entire sample, and the other two for males and for females separately.

AGE was used as a proxy for experience. The data did not contain a more precise variable to capture the effect of work experience As expected, we found AGE to be positively correlated to earnings. However, both the size of the coefficient estimate as well as the statistical significance of the variable was considerably higher for men than they were for women. The negative coefficient for AGE-SQUARED confirmed the concavity of the age-earnings profile, although it was statistically insignificant for females. The coefficient estimates indicate that earnings rises by 2.3% for every year that a male advances in age. The percentage increase for women, was lower, at 1.2%.

It has been suggested in the literature that married individuals tend to have higher earnings than those unmarried. The suggestion is made that higher-earning men tend to self-select into the marriage market. It is further postulated that as a result of social conditioning, women often seek out jobs that are traditionally lower paid than those pursued by their male counterparts. Our finding of a positive and statistically significant coefficient for MALE reinforces the findings of other researchers. Our coefficient estimates indicate that married men earn approximately 8% more than their unmarried counterparts, while the gain for married women is 6%.

Our earnings equation used a variable that we have not seen hitherto, in previous studies computing rates of return to education. It is well known that the most elite academic institutions with the highest quality of instruction in Pakistan are generally those which impart education in English. Many institutions also

impart education in the national language, Urdu. Though the quality of instruction is not considered at par with English-medium institutions, it is nonetheless superior to that at institutions which teach using regional languages. We therefore used a dummy variable for respondents who had had their education at English medium institutions as well as a dummy variable for those who were taught in Urdu. The missing base variable in the regression equation consisted of the regional languages.

As we expected, the coefficient estimate for instruction in the English language was positive and highly significant for both males and females. The coefficient estimate suggests that individuals from English-medium institutions earn approximately 35% more than those from institutions using regional languages. The impact was very similar for both males and females. The coefficient estimates for URDU was also positive and statistically significant, but less so than for ENGLISH. The coefficient estimates translate to percentage earnings gains of 13% and 9% for males and females respectively, relative to those schooled in the regional languages.

Table 1: Coefficient Estimates for Entire Sample, PIHS, 2001-2002

Coefficient	**Entire Sample**		**Males**		**Females**	
Intercept	7.11	(152.88)	7.26	(166.29)	7.34	(39.64)
Age	2.29	(10.24)	2.31	(10.04)	1.22	(1.19)
Age-Squared	1.76	(-6.52)	1.79	(-6.45)	0.58	(-.43)
Married	0.08	(6.74)	0.08	(6.57)	0.06	(1.70)
Male	0.15	(7.86)				
English Language Instruction	0.33	(13.60)	0.33	(12.79)	0.32	(4.47)
Urdu Language instruction	0.12	(11.97)	0.12	(11.88)	0.09	(1.83)
Middle Education	0.05	(3.69)	0.05	(3.82)	0.13	(-1.33)
Matric Education	0.14	(10.95)	0.14	(10.69)	0.13	(1.95)
Intermediate Education	0.26	(15.15)	0.27	(15.29)	0.25	(1.28)
BA/BS Education	0.41	(22.46)	0.42	(21.64)	0.33	(4.86)
Masters and Other Advanced Degrees	0.59	(25.30)	0.58	(23.32)	0.60	(8.14)
Legislators, Senior Officials and Managers	0.53	(20.83)	0.52	(20.12)	0.89	(6.79)
Professionals	0.23	(11.26)	0.24	(10.77)	0.17	(2.40)
Technicians and Associate Professionals	0.19	(9.95)	0.20	(9.70)	0.12	(1.65)
Clerks	0.05	(2.55)	0.05	(2.35)	0.19	(1.58)
Service Workers, Shop & Market Sales Workers	0.16	(11.54)	0.16	(11.30)	0.12	(1.60)
Skilled Agricultural and Fishery Workers	0.23	(11.55)	0.23	(11.17)	0.30	(3.47)
Craft and Related Trades Workers	0.14	(9.04)	0.14	(8.91)	0.12	(1.45)
Plant and Machine Operators and Assemblers	0.12	(7.19)	0.13	(7.23)	0.08	(-.42)
Industry: Other Activities Not Defined	0.14	(-2.39)	0.15	(-2.42)	0.17	(.47)
Industry: Mining and Quarrying	0.08	(1.57)	0.09	(1.64)	0.03	(-.12)
Industry: Manufacturing	0.05	(2.21)	0.05	(2.22)	0.03	(-.26)
Industry: Electricity Gas and Water	0.03	(1.01)	0.03	(.91)		
Industry: Construction	0.02	(-1.06)	0.02	(-1.07)	0.02	(.08)
Industry: Trade, Hotels and Restaurants	0.13	(6.16)	0.13	(5.99)	0.02	(-.14)
Industry: Transport and Storage	0.09	(3.74)	0.08	(3.52)	0.66	(2.45)
Industry: Finance and Real Estate	0.33	(9.42)	0.33	(9.29)	0.04	(.17)
Industry: Community Services	0.10	(-5.11)	0.11	(-5.19)	0.03	(.31)
Selectivity Variable			0.52	(5.11)	- 0.19	(-0.62)
Note: Figures in parentheses are t-statistics						

Table 1 lists the coefficient estimates for the equation Log Monthly Salary = $\alpha + \Sigma\beta X_i + \Sigma\delta OC_i + \Sigma\mu_i IC_i$ The first column iss based on results from the entire sample, the second for the male sample, and the third column reports results for females. Figures in parentheses are t-statistics.

A number of interpretations could be suggested for these findings. *Prima facie,* the inference to draw would be that instruction in English is superior to that in Urdu, while the latter trumps the regional languages. However, as is clear to any Pakistani observer, the level of resources devoted to education (financial, manpower, and capital) is highest in the case of English-medium institutions, followed by Urdu-medium institutions, and finally, the schools that impart education in the regional languages.

English continues to be the language of the powerful elite, who ensure the provision of ample resources to the institutions where their offspring are nurtured. With higher compensation levels, the English-medium schools get first pick of incoming teachers. Buildings including laboratories are acceptable, sometimes even by Western standards at the best of the English-medium institutions, but are often non-existent in the regional language schools. The resultant difference in the quality of education imparted, and the outcome of significant differentials in earnings, is a reflection of the biases and power plays inherent in our society. The data, without getting into a social commentary, clearly uphold the benefits of receiving education in the English language. For those not as privileged, the Urdu language is to be preferred to the regional languages.

It must be stressed that it is very likely not the medium of instruction itself, but the attendant resources that generally accompany English medium instruction, and to a lesser degree, Urdu medium instruction (relative to the regional languages) --- that are likely the reason for the earnings premiums reported in this article.

Nine different occupations were identified in the data. Our regressions included eight of these, with Elementary Occupations as the missing base variable. The coefficient estimates associated with each of these occupations (listed in Table 3) indicate the earnings premium relative to respondents employed in the elementary occupations ranged from 5 percent to 70 percent. Given the broad definitions of the occupational categories, it is difficult to provide much meaning to these estimates.

Ten industrial categories were identified in the data. Nine of them were included in the regressions, with Agriculture, Forestry, Hunting and Fishing being the missing base variable. As with the occupational categories, the industrial categories were very broadly defined. The only industry with a considerable premium relative to the omitted base was Finance and Real Estate, where the earnings premium was 39% above those from the omitted base industry. Our coefficient estimate suggests that males earn 15% more than females. The gender earnings gap is well documented, and we will therefore refer the interested reader to these other studies. They include Ashraf and Ashraf (1993).

The principal objective of this study was to determine the rates of return to different levels of education in Pakistan. Five levels of education were identified in the earnings equation: Middle (eight years of schooling), Matric (ten years of schooling), Intermediate (twelve years of schooling), Bachelors (fourteen years of schooling), and Masters and Other Advanced (sixteen or more years of schooling). Less than Middle level of education was the missing base variable in the regression equation. Consistent with *a priori* expectations, coefficient estimates rose monotonically with the level of education. The percentage gain to men and women from different levels of education (relative to the base of "less than middle level of education") is shown below in Table 2. These results show the considerable enhancement in earnings from receiving progressively higher levels of education. The differences across gender are relatively minor, underlining the benefits that accrue to both men and women with formal academic instruction. As a clear recommended policy measure, it reinforces the need to provide education as a means to combat poverty in the country. Even the attainment of Middle level of education would, in the case of females, enhance earnings by 13%. Given distressing levels of economic well-being, such an enhancement of income would be welcome.

Table 3 lays out the returns to education across the four provinces. It appears that education at all levels leads to higher returns in the NWFP than it does in the other provinces. It is noteworthy that even a middle level of education leads to a considerable increase in earnings (except in Sindh, where the coefficient estimate was statistically insignificant). As a policy prescription, it suggests the need for allocating resources to lower levels of education. As we observed earlier, the results are even stronger in the case of females.

Table 2: Percentage Earnings Gains from Different Levels of Education

	All	Male	Female
Middle	5%	5%	13%
Matric	15%	15%	14%
Intermediate	30%	31%	28%
Bachelors	51%	52%	39%
Master/Advanced	80%	79%	82%

Table 2 reports percentage earnings gains from different levels of education. The first column is based on results from the entire sample, the second is for males only, and the third reports results for females. The percentage earnings gains were computed as the exponent of the coefficient estimates of the dummy variables for each of the different levels of education computed from the equation: Log Monthly Salary $= \alpha + \Sigma\beta X_i + \Sigma\delta_i OC_i + \Sigma\mu_i IC_i$

Table 3: Rates of Return to Education by Province

Education Level	Punjab	Sindh	NWFP	Baluchistan
Middle	0.068 (3.20)	-0.002 (0.007)	0.137 (3.91)	0.109 (2.505)
Matric	0.171 (8.48)	0.079 (3.38)	0.238 (6.93)	0.192 (4.83)
Intermediate	0.325 (11.41)	0.226 (7.77)	0.259 (5.65)	0.218 (4.48)
BA/BS	0.553 (16.84)	0.330 (10.75)	0.427 (8.31)	0.304 (5.56)
Masters/Advanced Degree	0.707 (17.69)	0.418 (10.06)	0.726 (12.13)	0.494 (7.74)
R-Squared	0.34	0.37	0.35	0.26
No. of Observations	5,182	3,796	1,976	2,355

Table 3 is an extension of Table 2. It reports percentage earnings gains from different levels of education. The first column is based on results from the entire sample, the second is for males only, and the third reports results for females. The percentage earnings gains were computed as the exponent of the coefficient estimates of the dummy variables for each of the different levels of education computed from the equation: Log Monthly Salary $= \alpha + \Sigma\beta X_i + \Sigma\delta_i OC_i + \Sigma\mu_i IC_i$

CONCLUSIONS

This paper updated earlier estimations of the returns to education in Pakistan. New data from the Pakistan Integrated Household Survey, 2001-2002 were used for this purpose. We used a modified version of the well-known Oaxaca model suggested by Cotton and Neumark for our estimations.

One of our most important findings was that returns for females at 13% is considerably higher than for males (5%) at the Middle level of education. This suggests that policy makers should devote more resources toward female education, in a country where large numbers of women go without any formal education. These results are consistent with the findings of earlier researchers, and serves to reinforce those results.

Because of the relatively small number of observations for females in the sample, results have to be viewed with caution. Also, precise data on work-experience is difficult to get, since females in Pakistan start their education at different ages. It is hoped that as better data becomes available, estimations of the returns from education will be more reliable. We expect that future research will explore other determinants earnings that were not available in the data set used for this article. With all of the limitations of the data, we nonetheless feel that the basic conclusions are valid, and can serve as a basis for further exploration of this important issue.

REFERENCES

Abbas, Q. and Foreman-Peck, J. (2007) "The Mincer Human Capital Model in Pakistan: Implications for Education Policy," (Unpublished Manuscript), Cardiff Business School, United Kingdom.

Aslam, M. (2007), "Rate of Return to Education by Gender in Pakistan," *Research Consortium on Educational Outcomes and Poverty,* Working Paper No. 1. p. 1-41.

Ashraf, J., Birjees Ashraf (1993) " Estimating the Gender Wage Gap in Rawalpindi City," *Journal of Development Studies,* vol. 29 (1), p. 365-376.

Cotton, J. (1988) "On the Decomposition of Wage Differentials," *The Review of Economics and Statistics,* vol. 70 (2), p. 236-43.

Guisinger, S. J. W. Henderson, and G. W. Scully (1984) "Earnings, Rates of Return to Education, and the Earnings Distribution in Pakistan," *Economics of Education Review*, vol. 3 (2), p. 257-267.

Hamdani, K. (1977) "Education and the Income Differential: An Estimation for Rawalpindi City," *The Pakistan Development Review*, vol. 16 (1), p. 144-164.

Haque, N. (1977) "Economic Analysis of Personal Earnings in Rawalpindi City," *Pakistan Development Review*, vol. 16 (4), p. 353-382.

Heckman, J.,(1979) "Sample Selection Bias as a Specification Error," *Econometrica,* vol. 47 (1), p. 153-61.

Hyder, A. (2007) "Wage Differentials, Rate of Return to Education, and Occupational Wage Share in the Labour Market of Pakistan." *PIDE Working Papers, 2007*. Pakistan Institute of Development Economics, Islamabad.

Khan, S. and Irfan, M. (1985) "Rates of Return to Education and the Determinants of Earnings in Pakistan," *The Pakistan Development Review*, 24 (4), p. 671-683

Nasir, Z. and Nazli, H. (2000) "Education and Earnings in Pakistan," *Research Report No. 177*, Pakistan Institute of Development Economics.

Nazli, H. (2004) "The Effect of Education, Experience and Occupation on Earnings: Evidence from Pakistan," *The Lahore Journal of Economics* vol. 9 (2), p. 1-30.

Neumark, D. (1988) "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *Journal of Human Resources*, vol. 23 (3), p. 279-295.

Oaxaca, R. (1973) "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review,* vol. 14 (3), p. 693-709.

Pasha, H. and Wasti, S. (1989) "Unemployment and Rates of Return to Education," *Singapore Economic Review*, vol. 34 (2), p. 64-77.

BIOGRAPHY

Javed Ashraf is Professor of Economics at the University of St. Thomas in Houston (USA). He has previously held faculty appointments at the University of West Florida, the University of Hartford, and the University of Wisconsin. He is a former dean of the Cameron School of Business at his present institution, where he also held the Cullen Endowed Chair in Economics for four years.