Faculty Research and Reports

12-1989

Disaggregating rates of return to education

Shahrukh R. Khan
Vassar College

Mohammed Irfan

DISAGGREGATING RATES OF RETURN TO EDUCATION

Shahrukh R. Khan and Mohammed Irfan*

December, 1989

DISAGGREGATING RATES OF RETURN TO EDUCATION

Education is used as a continuous or discontinuous variable in earning functions to estimate the rates of return to education.[1] When it is used as a discontinuous variable, it was be found that the different sub-levels of education are more meaningfully associated with the labor market than the broader categories of primary, secondary and high generally used in rates of return studies. Also, the aggregated rates of return for the broader categories can easily be misleading if the correct procedure is not adopted for the aggregation. The literature demonstrating the estimation of rates of return does not take note of these issues.

The objective of this paper is to address these points using

---

[1]. The major premises underlying the human capital model that education enhances productivity and that increased productivity leads to higher earnings are the subject of intense controversy. Scholars who ascribe to the credentialist view do not believe that earning differentials across different levels of education are representative of productivity differentials. This controversy concerning the rates of return to education is side-stepped here (See Berry [1980] for details).

The earning functions in this paper are not refined by taking into account issues like cognitive ability and family background and hence the issue of selectivity bias in the sample. We therefore arrive at crude estimates of the rates of return. However, since the focus of the paper is methodological, no attempt has been made to derive refined estimates.

Recent refined estimates of the determinants of earnings by Boissiere, Knight and Sabot (1985, pp. 1016-1030) that included a measure of cognitive ability supports the human capital interpretation of the wage structure.

a recent data set from Pakistan. First, estimates of disaggregate rates of return will be related to the job market in Pakistan to indicate that the story is richer and more plausible than estimates using the usual level of aggregation. Second, the relatively more disaggregate rates of return will be compared to the aggregated rates of return to show the following anomaly: All sub-level rates of return within a given level of education can be higher than the rate of return for that level as an aggregate. It is indicated that this anomaly results from the way the internal rate of return algorithm works. Finally, an alternative method of aggregation is suggested which removes this anomaly.

These points can be established by using either private or social rates of return. The former are used in this paper since they require fewer computations.

I. Data, Method and Educational Structure

Primary household data for wage earners were drawn from the computer tapes of the 1984-1985 Household Income and Expenditure Survey. Details about the nationally representative multi-phased stratified random sample are published in the survey report (pp.xix-xxii).

The standard method of computing private rates of return is utilized. This entails first the simulation of age-earning profiles for the different levels of education to determine the net stream of benefits represented by earning differentials between the level of education for which the rate of return is being computed and the prior level of education. Apart from representing the benefits, earning profiles for one level of education represent the income foregone or the opportunity cost for the succeeding level of education during the period the education is being acquired. These simulations are based on earning functions for each level of education using age and experience (age squared as a proxy) as the independent variables. Appendix Table 1 reports the age-earning profiles to the different levels and sub-levels of education.

Second, from the stream of benefits are subtracted the opportunity costs and also the direct cost of schooling. All this is summarized in the following formula:

$$\sum_{t=n}^{N} \frac{E_t - C_t}{(1+r)^{t-n}} = 0 \qquad\qquad\qquad (1)$$

Et is the earning differential between consecutive levels of education, Ct is the direct cost and earnings foregone from attaining a given level of education, n is the starting age for a particular level of schooling, N is the retirement age, and r is

the internal rate of return solved for by iteration.

Like most educational structures, formal education in Pakistan is composed of the primary, secondary and post-secondary or high levels. However, the sub-levels at the secondary and high levels make the educational structure unusual. Five years of secondary education follow five years in primary school. After three years in secondary school, middle school is completed, and after two more years, the successful completion of national board exams earns the candidate a matric degree. After each of these three initial levels, there are options to move to vocational tracks. After the matric degree, the candidate also has the option of continuing to formal college education. For the latter, there are three additional national board examinations each after two years. After the first two years the candidates appear for an intermediate examination (also referred to as F. A./F. Sc. or higher secondary). They then either opt for the professional or technical streams or continue with general college or university education and can appear for the bachelor's examination after two years and for the master's examination after an additional two years. Thus to sum up, for formal education there are either three aggregate levels (primary, secondary and high) or six disaggregate levels (primary, middle, matric, intermediate, bachelors and masters).

II. **Findings**

Our results are reported below in Table 1.

(Table 1 about here)

First some remarks are proffered about the complex picture of the interaction of education and the labor market using column 1 of Table 1. Column 1 reports the rates of return with the disaggregated education variable. Second, in column 2 we show how the findings at the level of aggregation commonly used in rates of return studies (Psacharopoulos, 1985) contradict the findings in column 1. Finally, we show how this apparent contradiction is resolved by using an alternative method of aggregation (to be developed below), the results of which are reported in column 3.

The rates of return shown in column 1 follow a cyclical pattern rather than a directly inverse pattern that could be expected from evidence reported by Psacharopoulos for other LDCs including Pakistan. Plausible explanations for the pattern in column 1 of Table 1 are as follows:

1. The middle level, while technically representing a completed level of education, may still signal a "drop-out" candidate on the market and hence lead to a lower return than the lower cost primary level.

2. A matriculation has become the common minimum qualification for most lower level public and private sector skilled and semi-skilled jobs, which command much higher pay than those acquired with primary education.

3. A lower return at the intermediate level may reflect both the higher cost and market saturation. Most successful intermediate candidates attempt entry into professional schools (eg. medical, dentistry, law, engineering); if they fail to gain admission, they may pursue general higher education, if resources permit, or else fall back on the job market. The return at this level could be lower still if it were not the entry point for a fairly large number drawn into the armed services.

4. The higher return at the bachelor's level probably reflects the much higher earnings of those who do earn professional degrees. Those with contacts enter the private sector at this stage and also command high salaries.

5. The general master's degree is often a way to postpone entry onto the job market by those who were not admitted to professional school. Many use the master's degree as a launching pad for taking the civil service exam, entering government service via competitive interviews, or seeking a graduate academic career abroad. Civil service salaries are not comparable to remuneration in the professions, the armed services

or select private sector jobs.

The unadjusted aggregated rates of returns in column 2 reveal the usual pattern of an inverse relationship of rates of return and levels of education for less developed countries. However, the discrepancy in rates of returns between the aggregated and disaggregated estimates defies common sense; for conceptually, how can a straight average be less than all the numbers being averaged? In fact, the conceptual error is inherent in the method of averaging.

Consider for example the aggregation at the high level in Table 2, which is a reproduction of the first nineteen years of the actual net earning streams used for the rate of return estimates in this paper.

(Table 2 about here)

Before the stream of benefits begins, there are six years of direct costs and foregone earnings (these six years of costs range from Rs. 1197 at age 15 to Rs. 3865 at age twenty as shown in the column titled NED:H). These costs are less heavily discounted in the computation of the internal rate of return, since they occur earlier, than the average of net earning differentials accruing later. Thus it is not surprising that the return to the high level of 12 percent in column 2 of Table 1 is much lower than a straight average of the three disaggregate

rates of return within the high level in column 1. Similarly, the aggregate return to secondary is lower than a straight average of middle and matric.

A suggested method of aggregating in estimating the returns, for example to the high level, is to <u>sum</u> the net earnings of masters over bachelors (M), bachelors over intermediate (B) and intermediate over matric (I). Thus algebraically the alternative formula for high could be represented as

$$\sum_{t=18}^{65} \frac{E_t^M - C_t^M}{(1+r)^{t-18}} + \sum_{t=16}^{65} \frac{E_t^B - C_t^B}{(1+r)^{t-16}} + \sum_{t=14}^{65} \frac{E_t^I - C_t^I}{(1+r)^{t-14}}$$

instead of

$$\sum_{t=14}^{65} \frac{E_t^H - C_t^H}{(1+r)^{t-14}} = 0 \quad \text{---------------} \quad (3),$$

where H represents high.

Thus the general form can be represented as

$$\sum_{t=n}^{N} \sum_{j=1}^{K} \frac{E_t^j - C_t^j}{(1+r)^{t-n}} = 0 \quad \text{---------------} \quad (4),$$

instead of as in equation 1. In equation 4, j represents the sub-levels within a particular level of education.

The adjusted aggregated rates of return were calculated using equation 2 and are reported in column 3 of Table 1.

These adjusted estimates remove the anomaly referred to above. The average of rates of returns to the sub-levels (adjusted estimate) lies as might be expected between the range of returns to the sub-levels rather than below it as would be the case if equation 3 was used. This occurs because the addition suggested in equation 2 increases the net earnings (the sum of intermediate, bachelor's and master's net earnings are greater than the net earning stream to high) and because the costs foregone are lower. These lower costs occur because the positive earnings now begin after the first two years of higher education, i.e. the intermediate level, rather than after the sixth year as would be the case in equation 3.

For specific examples turn to Table 2. Equation 3 is the calculation of a rate of return based on column five, labeled NED:H, which is the net earning differential of high over the secondary level. By contrast, equation 2 is the calculation based on column six labeled NED:HSUM, which is the sum of the earning differentials of columns two through four (i.e. intermediate over matic, bachelor's over intermediate and master's over bachelor's). The first two entries in column six are identical to those in column five, which are the direct and indirect cost at the intermediate level (rows 15 and 16). The costs for the third to sixth entry in column six are lower than

column five. Consider the third entry of negative Rs. 505 in

column six. This results from subtracting out the positive

monthly earnings of Rs. 902 at the intermediate level (column 2)

from the costs at the bachelor's level of Rs. 1407. In effect,

since an average return to the higher level is being calculated,

it is done acknowledging the simultaneous existence of earnings

and costs at various levels.

Finally compare the seventh entry of Rs. 4056 in column six

with the equivalent entry of Rs. 1246 in column five. The first

number is the sum of the entries in row 17 for columns two

through four. The second number is approximately an average of

the returns to the three sub-levels and therefore is much lower.

To summarize, the adjusted aggregate rates of return in

Table 1 based on equation 2 still show an inverse relationship

between the level of education, the difference between the

secondary and high level of one percent could easily be due to

statistical discrepancies. The important point is that the rate

of return to both the secondary and high level is understated to

a great extent in the normal method of estimating rates of return

to the aggregate levels. However, estimating returns at this

usual level of aggregation, even if correctly done, still results

in a loss of important insights about the relationship of

education and the labor market.

The question remains whether these results are general or whether they pertain to this particular data set. The answer depends on whether the age-earning profiles are non-intersecting (well behaved). See for example figure 1, which shows the costs and non-intersecting age earnings profiles for the primary and secondary levels.

(Figure 1 about here)

It is self evident that the aggregate return to the secondary level will be an average of middle and matric and hence will lie below the matric age-earning profile. Thus given the same costs (ignoring the different foregone earnings in the alternative methods of aggregation), the rate of return based on summing returns of matric over middle and middle over primary will exceed the rate of return based on the net earning differential of secondary over primary. Well behaved age-earning profiles are a standard assumption of human capital theory justified by substantial cross country evidence. If the age-earning profiles intersected, it may be difficult to say a priori what the outcome would be in further disaggregating different levels of education.

III. **Conclusion**

This paper makes two points. First, if an education system has sub-levels in the aggregate levels of secondary and high, the usual procedure used for estimating rates of returns understates the true return to these aggregate levels of education. An alternative approach has been suggested to avoid this understatement. Second, using finer disaggregations of the education variable to account for the sub-levels within the usual aggregates of secondary and high may lead to richer insights about the interaction of education and the labor market.

APPENDIX

EARNINGS FUNCTIONS BY EDUCATION LEVEL

Level	Constant	Age	Age2	R^2	F	N
Illiterate	5.15043	.07336	-.00078	.22	529.5	754
	(106.29)	(24.59)	(20.00)			
Primary	4.93429	.09279	-.00097	.30	322.21	1481
	(53.94)	(16.94)	(13.11)			
Middle	5.09901	.08468	-.00081	.29	162.92	788
	(35.47)	(10.13	(7.28)			
Matric	5.54052	.06237	-.00046	.29	240.99	1166
	(40.34)	(7.74)	(4.24)			
Secondary	5.28393	.07621	-.00067	.29	400.26	1954
	(53.02)	(13.10)	(8.56)			
Intermediate	5.03763	.09810	-.00088	.32	107.33	456
	(18.13)	(5.94)	(3.81)			
Bachelor's	5.29409	.09441	-.00079	.27	89.41	492
	(16.07)	(5.12)	(3.24)			
Master's	5.20339	.10564*	-.00089**	.21	26.26	187
	(6.02)	(2.28)	(1.48)			
High	4.95920	.10791	-.00094	.31	250.30	1135
	(23.22)	(8.84)	(5.71)			

Note: Numbers in parentheses are t-values. All unstarred coefficients are significant at a 1 percent level. One star represents significance at the 5 percent level and two stars significance at the 10 percent level.

REFERENCES

Berry, A., "Education, Income, Productivity, and Urban Poverty," in ed. Timothy King, Education and Income, Washington D. C.: World Bank Staff Working Paper No. 402, 1980, 153-230.

Boissiere, M., J. B. Knight and R. H. Sabot, "Earnings, Schooling, Ability and Cognitive Skills, American Economic Review, December 1985, 75, 1016-1030.

Jimenez, E. and J. Tan, Educational Development in Pakistan, Washington D. C.: World Bank Discussion Paper No. EDT 16, December 1985.

Kazi, S. and Z. A. Sathar, "Differences in Household Characteristics by Income Distribution in Pakistan," Pakistan Development Review, Autumn-Winter 1985, 24, 657-667.

Pakistan, Household Income and Expenditure Survey, Islamabad: Federal Bureau of Statistics, Statistics Division, 1979, 1983.

Pakistan, 1984-85 Household Income and Expenditure Survey, Islamabad: Federal Bureau of Statistics, Statistics Division, 1985.

Psacharopoulos, G, "Returns to Education: A Further International Update," Journal of Human Resources, Fall 1985, 20, 583-604.

Table 1

ANNUAL WAGE EMPLOYEE PRIVATE RATES OF RETURN TO EDUCATION BY LEVEL

Education level, disaggregate	Disaggregate returns	Aggregate returns, (usual)	Aggregate returns, (alternative)	Education level, aggregate
Primary	45.56	45.56	45.56	Primary
Middle	28.72			
		18.88	33.10	Secondary
Matric	45.05			
Intermediate	37.20			
Bachelor's	40.56	11.74	32.04	High
Master's	21.59			

Note: 1. The returns to the primary level assume zero labor force participation for children under ten. This assumption seems justified for Pakistan since Kazi and Sathar (1985, p. 661) found labor force participation for the 5-9 age group to be less than 4 percent.
2. The data on direct cost were taken from Jimenez and Tan (1985, p. 12). This data was for the period 1983-84, and therefore there is a one year gap between the earnings and cost data.

Table 2

Net Earning Differentials to Different Levels of Education (ks.)

Yrs.	NED:INT	NED:B	NED:M	NED:H	NED:HSum
5.					
6.					
7.					
8.					
9.					
10					
11.					
12					
13					
14.					
15.	-1196.8			-1196.8	-1196.8
16.	-1231.3			-1231.3	-1231.3
17.	902.2	-1407.2		-1266.5	-505.0
18.	959.6	-1464.6		-1302.3	-505.0
19.	1019.1	1216.0	-4175.4	-3792.7	-1939.9
20.	1080.5	1286.0	-4245.0	-3865.1	-1878.5
17.	1143.8	1357.4	1554.7	1245.5	4055.8
18.	1208.8	1430.4	1648.4	1319.8	4287.6
19.	1275.4	1505.0	1774.6	1395.8	4525.1

NED	=	Net Earning Differentials
Int	=	Intermediate
B	=	Bachelor's
M	=	Master's
H	=	High
HSum	=	Sum of net earning differentials of intermediate, bachelor's, and master's

Earning Profiles for Different Levels
of Aggregation

Figure 1
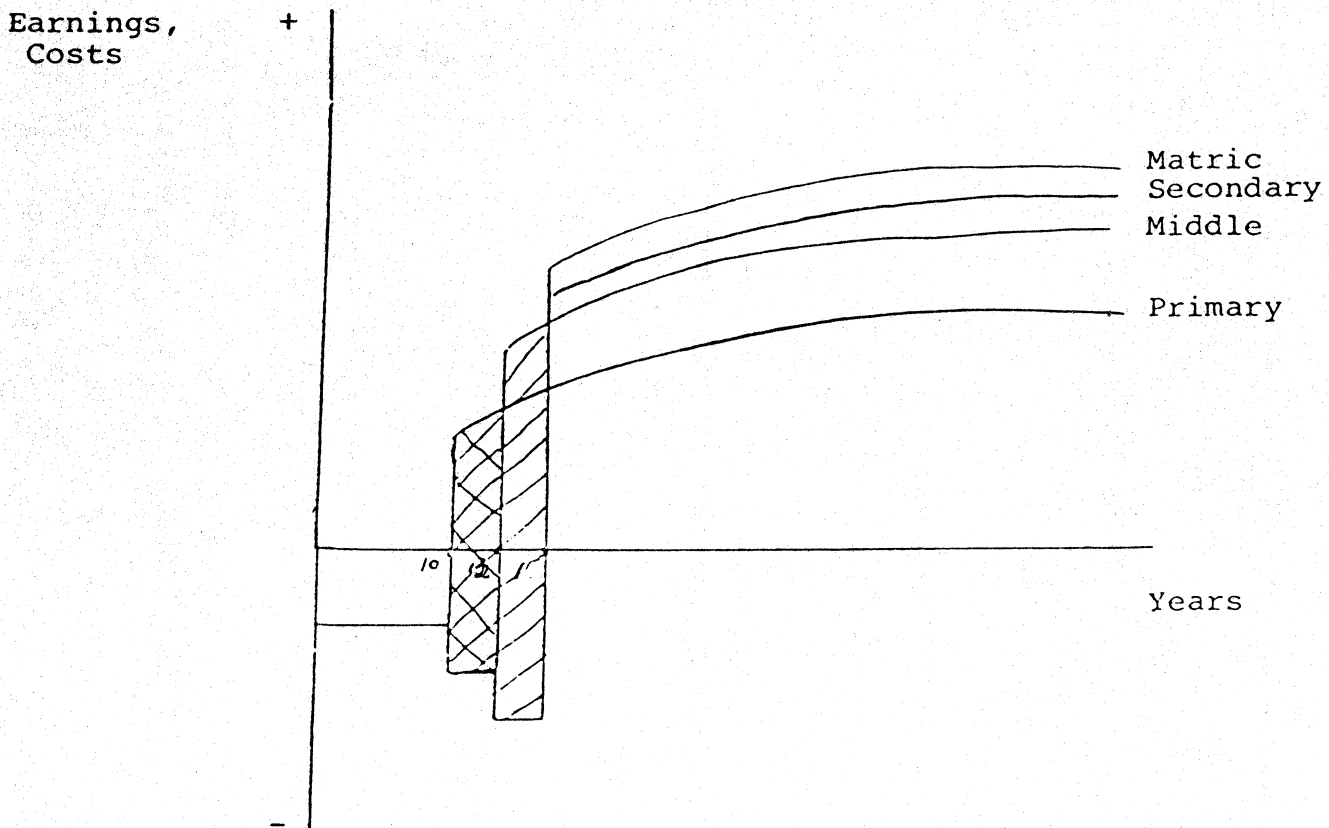
Shaded areas represent the direct and
opportunity costs.

XXXX   direct and opportunity costs for middle

////   direct and opportunity costs for matric