



Munich Personal RePEc Archive

## **Disaggregating rates of return to education**

Khan, Shahrukh R. and Mohammad, Irfan

Vassar College, ILO, Pakistan Institute of Development Economics  
Islamabad Pakistan

1990

Online at <https://mpra.ub.uni-muenchen.de/39208/>  
MPRA Paper No. 39208, posted 04 Jun 2012 13:38 UTC

**DISAGGREGATING RATES OF RETURN TO EDUCATION\***

**Shahrukh R. Khan and Mohammed Irfan**

**February, 1990**

\* The authors are assistant professor, Vassar College and ~~senior~~  
~~research economist, Pakistan Institute of Development Economics.~~  
Thanks are due to Richard C. Porter and Marsha Van Wagner for  
comments on an earlier draft.

**DISAGGREGATING RATES OF RETURN TO EDUCATION**

Abstract Education is effectively used as a discontinuous variable in studies estimating the rates of return by level of education. We find that the normal procedure used for estimating the rate of return to broad aggregates such as secondary and high understates the returns to these levels and subsequently suggest a procedure for rectifying this bias. We also find, as expected, that broad aggregations that ignore sub-levels of education result in a loss of important insights about the interaction of education and the labor market.

The main focus of this paper is to identify a flaw in the measurement of private rates of return to aggregate levels of education and to suggest a method to correct this flaw. What follows in this introduction is the placing of our contribution in the context of the literature in this field.

Rates of return are the core of human capital theory, and they embody its two critical assumptions that education enhances productivity and that increased productivity enhances earnings. One can identify at least three strands in the literature. First, there are numerous empirical estimates of rates of returns for different countries accepting the premises of the human capital model and, by and large, using the standard procedures for estimating rates of return.<sup>1</sup>

Second, with in the context of the human capital model, there are studies pertaining to the appropriate form of measurement. Here one can distinguish further between the more

---

<sup>1</sup> See for example the compilation by Psacharopoulos (1985).

conceptually oriented and the methodologically oriented studies. The former studies have refined and extended the human capital model by modeling and estimating the impact of phenomena such as dropping-out, school quality differentials, innate ability, family background, sample selectivity and truncation bias, and the impact of educational expansion on earnings.<sup>2</sup> The latter studies have demonstrated methods of estimation of rates of return or a critique of these returns based on measurement problems.<sup>3</sup>

Third, there are the studies by scholars that differ with human capital theorists on the interpretation of estimated rates of return. Thus the the screening hypothesis is a competing explanation for the association of education and income that challenges the fundamental premises underlying human capital theory.<sup>4</sup>

---

<sup>2</sup> For are recent review of such literature see Schultz (1988).

<sup>3</sup> The early works include Becker (1975), Mincer (1974) and Blaug (1967). More recently Psacharopoulos (1985a) and Mingat and Tan (1988) include chapters in their books demonstrating measurements of rates of returns. In the more critical vein, Schultz (1968) discusses the appropriate form of the dependent variable and (1988) the bias resulting from excluding the self-employed from the sample. Also see Field's critique of social cost-benefit analysis in educational planning (1978).

<sup>4</sup> Scholars who ascribe to the screening hypothesis do not believe that earning differentials across different levels of education are representative of productivity differentials. For recent evidence see Berry (1980), Hungerford and Solon (1987) and

While this classification may be helpful in placing our work in the context of the literature, we do not claim that it either does justice to the depth and coverage of all studies or ensures water tight categorization.<sup>5</sup> Our contribution fits into the methodological category since we are concerned with the accuracy of measurement of rates of return to aggregate levels of education.

Before proceeding; we should mention that measuring rates of return by aggregating school years for the level of education is consistent with the screening hypothesis because it allows for possible discontinuities in the returns to education due to the earning of diplomas. However, as Hungerford and Solon (1987, p. 177) indicate, these discontinuities are not inconsistent with a human capital interpretation of the data.

Our main objective is not to provide evidence for either interpretation. We do show that when education is used as an aggregate variable, different sub-levels of education (i.e. using finer disaggregation) are more meaningfully associated with the labor market than the broader categories of secondary and high

---

Gullison (1988).

<sup>5</sup> Thus Houthakker (1959) deals with both issues of measurement and interpretation. Along with refining the human capital model, Boissiere, Knight and Sabot (1985) also find support for it.

generally used in rates of return studies. However, our main objective is to show that 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.

This paper addresses these points using a recent data set from Pakistan. First, estimates of the more disaggregate rates of return (including sub-levels) will be related to the job market in Pakistan to indicate that the story is richer and more plausible than with 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.

### **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:

$$\sum_{t=n}^N \frac{Et - Ct}{(1 + r)^{t-n}} = 0 \quad (1)$$

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).

### **Findings**

Our results are reported below in Table 1. The earning functions are not refined by taking into account issues like cognitive ability, family background and selectivity bias in the sample. We therefore arrive at crude estimates of the rates of return. However, our inability to refine the estimates, due to the lack of the relevant data, is not important since the focus of this paper is methodological.

(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 (1985) 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 sum 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 follows:

$$\sum_{t=18}^{65} \frac{Et(m) - Ct(m)}{(1+r)^{t-18}} + \sum_{t=16}^{65} \frac{Et(b) - Ct(b)}{(1+r)^{t-16}} + \sum_{t=14}^{65} \frac{Et(i) - Ct(i)}{(1+r)^{t-14}} = 0 \quad (2),$$

where for example  $Et(m)$  represents the earning differential of

the masters' over the bachelors' level.

The three components of the formula above could be weighted by the probability of completing the different levels of schooling. Instead of the formula above, the formula usually utilized to calculate the private rate of return to the high level is as follows:

$$\sum_{t=14}^{65} \frac{Et(h) - Ct(h)}{(1+r)^{t-14}} = 0 \quad (3)$$

where  $Et(h)$  represents the earnings differential of the high over the secondary level.

The general form of the alternative formula can be represented as follows:

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

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 4 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 1 was used. This occurs because the addition suggested in equation 4 increases the net earnings and because the costs foregone are lower.

For a 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 matric, 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, even the adjusted aggregate rates of return in Table 1 based on equation 2 vary inversely with 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 usual method of estimating rates of return to the aggregate levels. However, estimating returns at the 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 the particular data set used for this paper. 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.

### **Summary**

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.



## REFERENCES

- Becker, G. S., Human Capital (New York: National Bureau of Economic Research, 1975)
- Berry, A., "Education, Income, Productivity and Urban Poverty," in ed. T. King, Education and Income, Staff Working Paper No. 402 (Washington, D. C.: World Bank, July 1980).
- Blau, D. M., J. R. Behrman and B. L. Wolfe, "Schooling and Earnings Distributions with Endogenous Labour Force Participation, Martial Status and Family Size," Economica, 55, (August, 1988), 297-316.
- Blaug, M., "A Cost-Benefit Approach To Educational Planning in Developing Countries," Economics Department Report No. EC 157 (Washington D. C.: World Bank, 1967).
- Boissiere, M., J. B. Knight and R. H. Sabot, "Earnings, Schooling, Ability and Cognitive Skills," American Economic Review, 75, (December 1985) 1016-1030.
- Fields, G. S., "Assesing Educational Progress and Commitment," Report for the U. S. Agency for International Development (Ithaca: Cornell, 1978).
- Gullison, E. T., "New Evidence Supporting the Screening Hypothesis," Atlantic Economic Journal, 17, (June 1989), 80.
- Houthakker, H. S., "Education and Income," The Review of Economics and Statistics, 51, (February, 1959), 24-28.
- Hungerford, T., and G. Salon, "Sheepskin Effects in the Returns to Education," The Review of Economics and Statistics, 69, (February, 1987), 175-177.
- Jimenez, E. and J. Tan, "Educational Development in Pakistan," Discussion Paper No. EDT 16, (Washington, D. C.: World Bank, December 1985).
- Kazi, S. and Z. A. Sathar, "Differences in Household Characteristics by Income Distribution in Pakistan," Pakistan Development Review, 24, (Autumn-Winter 1985), 657-667.
- Knight, J. B., and R. H. Sabot, "The Rate of Return to Educational Expansion," Economics of Education Review, 6, (Fall, 1987), 255-62.
- Mincer, J., Schooling Experience and Earnings (New York: National Bureau of Economic Research, 1974)
- Mingat, A., and J. Tan, Analytical Tools for Sector Work in Education (Baltimore: The Johns Hopkins University Press, 1988).
- 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, 20, (Winter, 1985), 583-604.

Psacharopoulos, G, Education for Development: An Analysis of Investment Choices (New York: Oxford University Press, 1985a).

Schultz, T. P., "Education Investments and Returns," in eds. H. Chinery and T. N. Srinivasan, Handbook of Development Economics, Vol 1. (New York: Elsevier Science Publishers, 1988).

Schultz, T. P., "Returns to Education in Bogota, Columbia," Report No. RM-5645 (Santa Monica: Rand Corporation, 1968).

## APPENDIX

## EARNINGS FUNCTIONS BY EDUCATION LEVEL

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

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.

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			
Matric	45.05	18.88	33.10	Secondary
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 (Rs.)

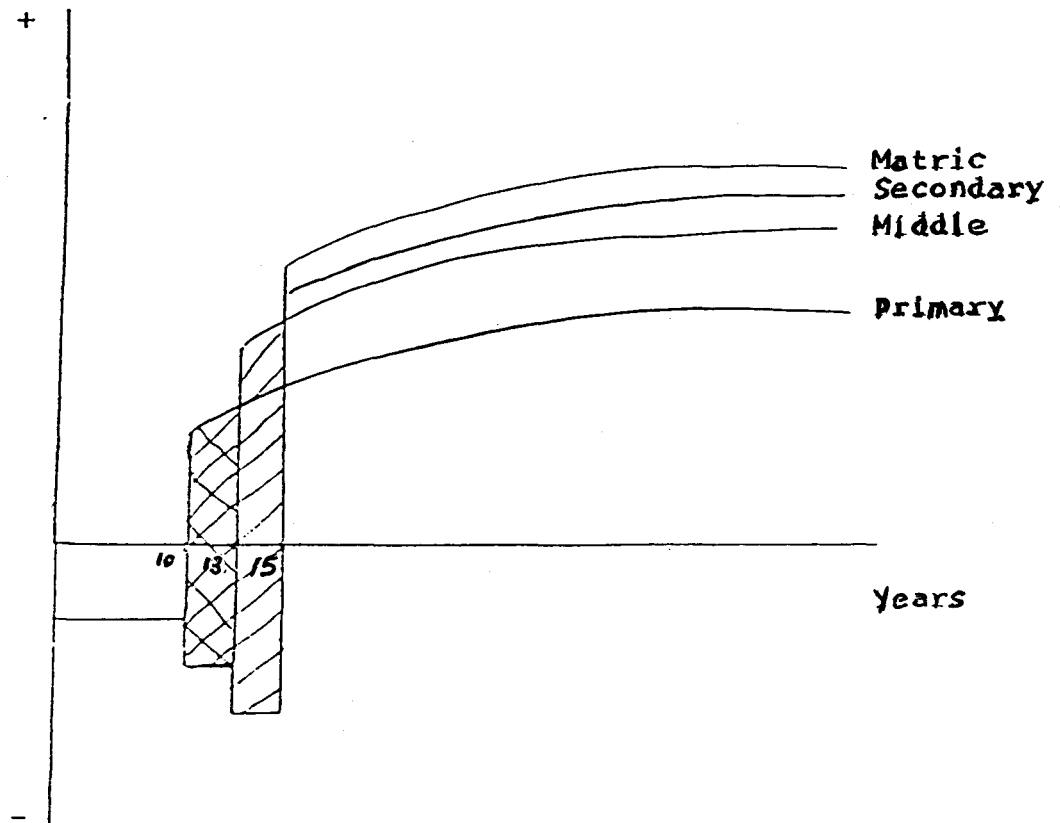
Yrs.	NED:INT.	NED:B	NED:M	NED:H	NED:HSum
5.					
6.					
7.					
8.					
9.					
10.					
11.					
12.					
13.					
14.					
15.	-1196.8			-1196.8	-1196.8
16.	-1231.3			-1231.3	-1231.3
17.	902.2	-1407.2		-1266.5	-505.0
18.	959.6	-1464.6		-1302.3	-505.0
19.	1019.1	1216.0	-4175.4	-3792.7	-1939.9
20.	1080.5	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

Earnings,  
Costs



Shaded areas represent the direct and opportunity costs.



direct and opportunity costs for middle



direct and opportunity costs for matric