DOCUMENT RESUME

ED 333 820 HE 024 669

AUTHOR Altonji, Joseph G.
TITLE Accounting for Uncertain Educational Outcomes in Estimating the Return to Education. Technical Paper No. 15.
INSTITUTION National Center on Education and Employment, New York, NY.; Rand Corp., Santa Monica, Calif.
SPONS AGENCY Northwestern Univ., Evanston, Ill. Center for Urban Affairs.; Office of Educational Research and Improvement (ED), Washington, DC.; Spencer Foundation, Chicago, Ill.
PUB DATE Feb 90
CONTRACT G008690008
NOTE 16p.
AVAILABLE FROM National Center on Education and Employment, Box 174, Teachers College, Columbia University, New York, NY 10027.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Blacks; *College Outcomes Assessment; College Role; *Educational Economics; Family Characteristics; Higher Education; High School Graduates; High Schools; *Income; *Models; Racial Factors; Salaries; Sex Differences; Socioeconomic Status; Wages
IDENTIFIERS National Longitudinal Study High School Class 1972

ABSTRACT

A prototype model was developed to estimate the effect of various variables on the internal rate of return to starting college. The model arose out of observations that there is a nonlinear relationship between education and earnings, that many students who begin college never finish, that there are large differences across fields of specialization in the earnings differential between college and high school graduates, and that many high school and family background characteristics appear to affect earnings. The model was then applied to data from the National Longitudinal Survey of the High School Class of 1972. Results found that black males have a lower rate of return on one year of college than white males (5.4 percent versus 5.9 percent) which appears, however, to be due not to race but to other characteristics. Females also had a lower rate of return than men. It was noted that none of these differences are explained by differences in the observed characteristics of males and females. Other research using this model has shown a higher return on college attendance for women over men. It is suggested that further research and refinement of the model is necessary before strong conclusions can be developed. The study includes one table displaying results and nine references. (JB)

National Center on Education and Employment

NCEE

Technical Paper Series

ACCOUNTING FOR UNCERTAIN EDUCATIONAL OUTCOMES IN ESTIMATING THE RETURN TO EDUCATION

Joseph G. Altonji

Center for Urban Affairs and Policy Research
and Department of Economics,
Northwestern University

Technical Paper No. 15

ACCOUNTING FOR UNCERTAIN EDUCATIONAL OUTCOMES IN ESTIMATING THE RETURN TO EDUCATION

Joseph G. Altonji

Center for Urban Affairs and Policy Research and Department of Economics, Northwestern University

Technical Paper No. 15

February 1990

This publication is based on work sponsored by the National Center on Education and Employment under grant number G008690008 from the Office of Educational Research and Improvement, U.S. Department of Education, by the Spencer Foundation, and by the Center for Urban Affairs and Policy Research at Northwestern University.

PREFACE

The National Center on Education and Employment is funded by the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education. The Center is based at Teachers College, Columbia University, in New York City. The RAND Corporation of Santa Monica, California, and Washington, D.C., is a partner with Teachers College in this enterprise.

This publication appears in the Center's Technical Paper Series. This series is designed to communicate the interim technical results of a large-scale Center research project or the final results of a small project. The paper has been reviewed by one reader external to the project and the Center, and was approved for publication by Center leadership.

For information about ordering additional copies of this document, write or call:

National Center on Education and Employment
Box 174
Teachers College, Columbia University
New York, New York 10027

(212) 678-3091

This publication is based on work sponsored wholly or in part by the Office of Educational Research and Improvement, U.S. Department of Education, under grant number G008690008. Its contents do not necessarily reflect the views of OERI, the Department, or any other agency of the U.S. Government.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the support of the National Center on Education and Employment, the Spencer Foundation, and the Center for Urban Affairs and Policy Research. The author is grateful also for the excellent research assistance of Avner Grief, Christian Stadlinger, and James Spletzer. The author is solely responsible for any errors or omissions.

INTRODUCTION

This paper discusses a method for accounting for uncertainty about educational outcomes and for nonlinearity in the relationship between years of education and earnings when estimating the expected return to a year of school, and the effects of high school and family background characteristics on the expected return. Four facts motivate the analysis. The first is that the relationship between education and earnings is nonlinear. The effect of completion of the final year of education on earnings is larger than the previous three. For example, for the subsample from the National Longitudinal Survey of the High School Class of 1972 (NLSHS72) used in this paper, I find that attending college for more than two years but leaving without a degree raises the wage above that of a high school graduate by only about 12 percent, while obtaining a college degree raises the wage by about 29 percent. The nonlinearity is especially dramatic for men.

The second fact is that many individuals attend but never finish college, even though they report that they plan to complete college or graduate work at the time that they begin college. In the NLSHS72 sample, 89 percent of high school seniors who plan to complete college or graduate school start college but only 58.1 have completed college by 1979.[1]

The third fact is that there are large differences across fields of specialization in the earnings differential between college and high school graduates. Engineers, for example, earn higher wages than English majors. The fourth fact is that many high school and family background characteristics appear to have larger effects on earnings through educational attainment and area of specialization than through earnings conditional on educational attainment.

Most of the economics literature on education choice and the returns to education following the work of Becker (1975) and Mincer (1974) has been based on the assumptions that workers choose education to maximize present value of lifetime wealth and that there is no uncertainty about the

---

[1] This result is for a sample of 9,032 for whom valid data on educational attainment in 1979 and educational plans as of senior year of high school are available. The small fraction does not appear to reflect unreliability of the responses to the question about education plans: 93.5 percent of the people who obtained a college or graduate degree indicated in 1972 that they planned to obtain a college or graduate degree.

amount of education the individual will successfully complete. The literature has not dealt systematically with the implications of the fact that (1) the returns to education are highly nonlinear and (2) students are uncertain at each stage in the education process about the number of grades they will complete and about the field they will specialize in. When the relationship between education and log earnings is nonlinear, and the educational outcomes are uncertain, the return to the first year of college is not the earnings differential between individuals with twelve and thirteen years of schooling who are the same in other dimensions that affect earnings. Rather, the return is the difference between the earnings of the person who stops at twelve and the expected earnings, net of education costs, of a person who attends the first year of college, where the expectation is taken across earnings associated with fourteen years, fifteen years, sixteen years, and higher education levels, weighted by the probability that the individual who has completed year thirteen will successfully complete those higher levels.

The coefficients on (a) years of education and (b) interaction terms involving years of education and school or background characteristics in a conventional earnings regression may provide misleading estimates of both (1) value ex ante of the additional year of schooling and (2) the effect of school characteristics and high school curriculum on that value. This is because such regressions condition the ultimate educational outcome, which is uncertain at the time that the decision about whether to pursue an additional year of schooling is made. This simple point was made explicitly by Weisbrod (1962). A related one underlies the matching models of occupational choice analyzed by Miller (1984) and Shaw (1985). Bamberger (1986) formulates and estimates a model of the choice of major in which students base decisions about their majors on both the expected returns if they successfully complete the major and the probability that they will do so given past educational choices and outcomes. However, the point has been ignored in empirical studies of the returns to education.[2]

In this paper, I provide a simple model that may be used to account for the fact that school and family characteristics may affect earnings by altering the probability that the individual will complete the levels and types of postsecondary education that have the largest effect on earnings. I illustrate the approach using the NLSHS72 data set and present some estimates of the expected internal

[2] After completing the empirical work leading up to this paper I became aware of research in progress by Charles Manski that addresses similar issues.

rate of return to starting college for various groups. The preliminary results indicate that ability and achievement raise the internal rate of return to starting college. High school courses in math, science, and foreign languages also raise the internal rate of return. The results also indicate that the ex ante rate of return to starting college is higher for those who actually start college than for those who do not.

The paper proceeds as follows: The first section presents a simple model that may be used to analyze the returns to education and the effects of school characteristics and background on earnings. The next section presents some initial empirical results that serve to illustrate the approach. The final section summarizes the paper and discusses a number of extensions for future research.

A Framework for Measuring the Returns to Education

In this section I present a simple model of educational attainment and earnings. I begin by deriving an equation for the expected value of future earnings as a function of an individual's ability and knowledge, determined as of grade twelve and of the individual's most recent year of education and field of study. Let the log earnings of a person with postsecondary education level s with field c as the final field of specialization in school and with t years of experience be

$$(1) \qquad \ln w_{sct} = \sum X_k B_k + r_{sc} + \phi(t)$$

where $\phi(t)$ is the experience profile of earnings and where r_{sc} is the difference in the log wage of a person with s years of schooling relative to a person with a high school degree, with s=0 for the person with a high school degree. Equation (1) is based on the assumption that the percentage effect of education on wages depends only on the total years of schooling and the most recent field of study. This is a simplification that is consistent with the wage specification used in the empirical work below, but is inconsistent with evidence that the relationship between high school courses and ability and achievement measures depends on education outcomes.[3] It is important to note that even with the

[3] More generally, the wage equation is inconsistent with wage specifications of the general class discussed in Willis and Rosen (1979) and Willis (1986), in which the personal characteristics shift the percentage effect of education on wage rates. These models assume that educational attainment is

assumption, early choices of field during the education process do matter, since they may alter the probability that one can attend school in year s in field c. This is made explicit below.

Let ρ denote the interest rate at which earnings are discounted. Then the present value of future wages W_{sc} is

$$(2) \qquad W_{sc} = Q\exp^{\sum X_i\beta_i + r_{sc} - \rho s}$$

$$\text{where } Q = \int_{t=0}^{T} \exp^{\phi(t) - \rho t}.$$

Equation (2) is based on the assumption that the length of a career is T years and does not depend on s and c.[4]

Now consider the expected present value of earnings conditional on attending school level s + 1 in field c' for a person who has completed s years of schooling with c as the most recent field of specialization. The present value is

$$(3) \qquad PV(s+1,c'|s,X) = \sum_{j=s+2}\sum_{k=c} P(X)_{s+1c'jk} W_0(X)\exp[r_{jk} - \rho j]$$
$$+ P(X)_{s+1,c',s+1,c'} \quad W_0(X)\exp^{[r_{s+1c'} - \rho(s+1)]}$$

---

based on a once-and-for-all decision, with certainty about the probability of successfully completing the program chosen. There is no difficulty in modifying the framework below so that the r_{jk} depend on personal characteristics, and I intend to do so in future work. However, the Willis and Rosen model implies a relationship between the unobservables affecting education choice (the P(X) functions) and the unobservables affecting the expected wage associated with a particular education outcome. It will be very difficult to correct for selectivity bias in the coefficients of the wage and probability functions that such a relationship would induce.

[4] Mincer (1974) presents evidence that this is a reasonable approximation.

where $W_0(X)$ is the present value of leaving school after high school, and $P(X)_{s+1, c', j, k}$ is the probability that a person with characteristics X will end up with schooling level j in field k, given that they currently are at schooling level s+1 in field c'. I view the probabilities as the outcomes of sequential decisions that are made after each year of schooling based upon information about performance in school, grade and course requirements associated with particular programs of study, the wages associated with different educational outcomes, and preferences for particular fields of study and work. However, for present purposes it is not necessary to write down a specific structural model of education demand that generates the probabilities.

Assuming that there is no specialization until after the first year of postsecondary school, the expected present value of earnings for attending the first year of college for a person who has completed high school (s=0) and has characteristics X is

$$(4) \qquad PV(1,. \mid 0,X) = \sum_{j=2} \sum_{k=c}^{P(X)} 1,jk \; W_0(X) \exp^{[r_{1.} - \rho j]} + P(X)_{1.1.} W_0(X) \exp^{[r_{1.} - \rho(1)]}$$

The expected internal rate of return to attending the first year of college is the value of $\exp(\rho)$ that equates the right hand side of equation (4) with $W_0(X)$. It may differ from $\exp(r_{1.})$ unless $\exp(r_{jk}/j) = \exp(r_{nm}/n)$ for all j,n greater than 1 and all fields k and m.

In Table 1 I report estimates of the internal rate of return to attending the first years of college conditional on completing high school. I also report estimates of the ratio $R(1, 0, X, \rho)$ of the present value of attending the first year of college to the present value of leaving school after high school.

Estimates of the Relative Present Value and Internal Rate of Return to the First Year of College

The data for the study are from the NLSHS72 panel data set, and are discussed in detail in Altonji (1988). To implement the above approach, one must have estimates of the $P(X)_{1,jk}$ functions and the wage differentials r_{jk}. The estimates of the r_{jk}, and of the effect on the log wage of each of the educational outcomes are based on an equation that includes controls for race, Hispanic background, sex, family background, geographic area, high school program, aptitude and achievement, labor market

experience and experience squared, the year and year squared, and dummy variables for whether the person's highest postsecondary education level was vocational education (VOC79), less than two years of college (SOCO1479), more than two years of college but no degree (SOCO1579), a degree in one of twelve fields of study, or an advanced degree in one of twelve fields of study. The equation was estimated by taking deviations from the mean values over all students from a given high school in a given year.[5] I will not discuss these coefficients here, although the negative signs on science semester hours, hours spent on homework, and vocabulary and reading test scores come as a surprise. The coefficients on the education outcomes are generally sensible.

To illustrate the approach, I simply estimate linear probability models by ordinary least squares for each of the educational outcomes using the sample of persons who started college. (Persons who attended vocational school but who never attended college are excluded from the estimation.) That is, I approximate the probability function $P(X)_{i,jk}$ with an equation of the following form for each jk pair:

$$P(X)_{i,jk} = X_{ih}\theta_{i,jk}.$$

The parameter vector $\theta_{i,jk}$ for a particular education outcome was estimated by regressing the dummy variables for that outcome against X_{ih} by least squares. The variables are the same as the variables in the wage equations, although the education outcome equations are used as the dependent variables rather than right-hand side variables.

The relative present value of the first year of college (R) was calculated assuming a 4 percent real interest rate. The first row of Table 1 shows R evaluated at the means for the full sample. It is equal to 1.036. The internal rate of return to the first year of college is 5.6 percent. In contrast, the coefficient on SOCO 1479 in the wage equation implies an ex post return of only .93 percent (not

---

[5] The sample means for each high school/year were computed using all the available data for students from that high school, rather than simply for students who were in the effective wage sample for that year. In practice, this has very little effect on the parameter estimates. Finally, a problem with the coding of the curriculum measures was not discovered until after the computations for the paper were largely completed. The effects of curriculum on the rate of return are probably understated, but the problem is not likely to make much difference for the other results.

reported). The latter estimate, which is the difference in log wages of a high school graduate versus that of a person who starts college but leaves with less than two years, ignores the fact that educational outcomes are uncertain when the decision to begin college is made. The second panel in the table reports the relative present value of the first year of college for students who are one standard deviation above the mean in all of the ability and achievement measures. For these students, the relative present value is substantially higher (1.0527 versus 1.0357), and the internal rate of return is 6.2 percent. A one standard deviation increase in semester hours of math, science, and foreign languages raises the internal rate of return to 6 percent, with all other variables evaluated at the mean for the full sample.

The third panel of the table reports differences in the internal rate of return and the relative present value of the first year of college as a function of race, sex, and average characteristics. Black males have a lower internal rate of return than white males (5.4 percent versus 5.9 percent), but this differential is due primarily to the variations in their average characteristics. The internal rate of return for black males evaluated at the mean of the sample characteristics of white males is 6.1 percent, which exceeds the internal rate of return for white males evaluated at the sample means for white males.

Females have a lower rate of return than men. Specifically, the internal rate of return for black females is 4.8 percent, compared with an internal rate of return for black males of 5.4 percent. The internal rate of return for white females is 5.4 percent, compared with an internal rate of return for white males of 5.9 percent. Almost none of these differences are explained by differences in the observed characteristics of males and females. It should also be kept in mind that these calculations are based upon estimates of the wage equation for the combined sample. Separate estimates of the wage equation for men and women indicate that the wage premium associated with college is larger for females than for males. In research in progress, I am using estimates of the r_{jk} obtained from wage equations for each race and sex group to distinguish between the effects on the rate of return to starting college of (1) sex differences in the effect of educational outcomes on wages and (2) sex differences in the probability of specific educational outcomes. The results indicate that education raises the wages of women more than men. As a result, the ex ante return to starting college is actually higher for women than for men. It will also be important in the future work to examine the effect of sex and races differences in work hours on the return to education that may arise from labor

supply preferences or unemployment constraints. One must assume that all persons work the same number of hours to justify the use of wage rates in the rate of return comparison.

It also interesting to ask whether the ex ante rate of return to beginning college is correlated with the decision to start college. As a first pass on this, the internal rate of return to attending the first year of college was calculated using the background characteristics of those who ended their education with high school. For this group, the internal rate of return is only 4.3 percent. For students who attended vocational school but not college, the figure is 4.7 percent. These estimates compare with 5.6 percent for the full sample (which includes the high school and vocational groups). Thus, it appears that the ex ante return to starting college is in fact lower for those who choose not to start college than for those who do. A difficult but natural extension of this research is to replace the reduced form equations for the P(X) with structural equations in which the ex ante returns play a role.

Conclusions and Suggestions for Future Research

This paper develops and implements a prototype model in which the rate of return to starting college depends upon the earnings associated with each of the possible outcomes of starting college (including dropping out after a year, majoring in biology, getting a master's degree in education, etc.), weighted by the probability of the specific outcome. I use the model to estimate the effect of parental background, high school curriculum, academic ability, race, and sex on the internal rate of return to starting college. The results from the empirical example are promising, but many refinements to the model and improvements in the empirical estimation of it will have to be made before strong conclusions can be drawn about the factors that influence the ex ante return to starting college.

An obvious extension is to allow the coefficients of educational outcome in the wage equation and the coefficients of the educational outcome equation to depend on race, sex, family background, ability, and high school curriculum. A more difficult but important line for future research is to extend the section of the model describing the probability of particular education outcomes. On the theoretical level, one would like to have a structural model relating the decision to remain in school in a given field to the probability of success, that is, the monetary rewards associated with specific outcomes, given background, ability, and past educational decisions and courses. Specifically, it would be useful to examine the interplay between college courses taken and grades in decisions about field of study and whether to remain in college. On the empirical level, one could use data from the

NLHS 72 Post Secondary Transcript Survey to link college grades and courses taken to measures of family background, high school curriculum, and ability, in order to better understand how the latter variables affect who starts and successfully completes particular programs of study.

REFERENCES

Altonji, Joseph, *The Effects of Family Background and School Characteristics on Education and Labor Market Outcomes*. Mimeo, Northwestern University Press (1988).

Bamberger, Gustavo, "Occupation Choice: The Role of Undergraduate Education," unpublished paper, Graduate School of Business, University of Chicago (October 1986).

Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis*, New York: NBER (1975).

Miller, Robert, "Job Matching and Occupational Choice," *Journal of Political Economy*, 92 (December 1984): 1086-1120.

Mincer, Jacob, *Schooling, Experience, and Earnings*, New York: NBER (1974).

Shaw, Kathryn L., "Occupational Change, Employer Change, and the Transferability of Skills" (1985)

Weisbrod, Burton A., "Education and Investment in Human Capital," *Journal of Political Economy*, Supplement, LXX (October 1962): 106-123.

Willis, Robert J., and Sherwin Rosen, "Education and Self Selection," *Journal of Political Economy*, 87(2) (1979): S7-S36

Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in O. Ashenfelter and R. Layard (editors), *Handbook of Labor Economics*, Elsevier Science Publishers (1986).

Table 1

Relative Present Value and Internal Rate of Return to First Year of College

	Relative present value of first year of college (R) at 4% discount rate	Internal rate of return (i)
1. Means of full sample	1.0357	5.6
2. Deviations from mean in		
Ability	1.0527	6.2
Parents' background	1.0575	5.9
Courses taken	1.0489	6.0

3. Race and sex

Person that is a	With Average Characteristics of				
	Full sample	BM	BF	WM	WF
Black male (BM)	R: 1.0527	1.0299		1.0572	
	i: 6.0	5.4		6.1	
Black female (BF)	R:	1.0142	1.0168		1.0411
	i:	4.7	4.8		5.6
White male (WM)	R:	1.0190		1.0473	
	i:	5.1		5.9	
White female (WF)	R:			1.0317	1.0312
	i:			5.4	5.4

Ability: One standard deviation higher in high school grades, tests of vocabulary, associative memory, reading, inductive reasoning, math, and perceptual speed and accuracy, a college ability self-report, and teacher's expectations of the student, as perceived by the student.

Parents' Background: Parents discussed plans with student, influenced plans, have sixteen years of education, not low socioeconomic status, and wanted college education for student.

Courses: Student took one standard deviation more hours in high school math, science, and foreign languages.