DOCUMENT RESUME

ED 310 167 TM 013 852

AUTHOR Kominski, Robert
TITLE Education and Earnings: Empirical Findings from Alternative Operationalizations.
PUB DATE 87
NOTE 8p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adults; *Census Figures; Degrees (Academic); *Educational Attainment; Educational Experience; Education Work Relationship; *Income; Interviews; Longitudinal Studies; *National Surveys; Regression (Statistics); Statistical Surveys
IDENTIFIERS *Survey of Income and Program Participation

ABSTRACT

Data from the third-wave interview of the 1984 panel of the Survey of Income and Program Participation (SIPP) are used to assess the empirical impact of a SIPP item concerning educational attainment on the regression of earnings on educational attainment. The SIPP is a longitudinal survey conducted by the United States Census Bureau to measure economic, social, and demographic characteristics of persons, and how these characteristics change over time. The SIPP involves a multi-stage cluster sample of all living quarters, representing all non-institutionalized residents of the United States, aged 15 years and older. Households remain in the sample for about 2.5 years, and are interviewed every 4 months. The data for this analysis were restricted to persons aged 18 to 64 years for whom data on education had not been imputed. Three different measures of education are used: (1) years of school completed; (2) a series of dummy variables based on the years of school data, representing five groups; and (3) a nine-category variable based on the highest degree completed. Analysis of the data indicates that a new measure of educational attainment, based on the highest degree attained, yields a better earnings-to-education relationship than does the existing measure of years of schooling completed. The improvement in the model appears to apply to the entire adult population; however, the addition of subgroup-specific interaction terms demonstrates the added utility of measuring degrees with respect to gender differences in returns to education. Five data tables are included. (TJH)

EDUCATION AND EARNINGS: EMPIRICAL FINDINGS FROM ALTERNATIVE OPERATIONALIZATIONS

Robert Kominski
Population Division
U.S. Census Bureau
Suitland, MD 20233

EDUCATION AND EARNINGS: EMPIRICAL FINDINGS FROM ALTERNATIVE OPERATIONALIZATIONS

Robert Kominski, Population Division, U.S. Census Bureau, Suitland, Maryland 20233

KEY WORDS: education, earnings, measurement

INTRODUCTION

The measurement of social phenomena is not an easy task, but often we identify observable indicators that we can accept as functionally practical. One measure that meets this criterion is the level of educational attainment achieved by an individual. Educational attainment is often ascertained by the use of a question which asks for the "number of years of school completed." Most of us associate the completion of a year of school as the accumulation of a "package" of knowledge. In the terms of some economists, we might say that the completion of a year of school adds something to the human capital of the person who acquires it (Mincer, 1974). Put simply, we routinely infer that higher levels of education (in years of school completed) yield higher earnings.

The years of school completed by the individual actually represent something other than time spent in the educational system, however. Along the way, defined "hurdles" are overcome, and recognition is made of this fact. The hurdles are the actual requirements one must meet in order to successfully complete not merely the grade level, but the degree or certificate that the individual is attempting to achieve. The interest for virtually all students is not merely in completing four years at a university, but in obtaining a Bachelor's degree of some sort.

As long as successful degree recipiency matches a given number of years of school completion, using years as a proxy for degrees does not matter. Most of the people who we say have completed twelve years of school in fact tell us they have a high school degree, and it is this latter point that we are truly interested in. In the past several years, however, some concern has been raised that the correspondence between these measures at levels beyond high school might be less than complete. In a series of analyses of different data sources, Kominski (1985) and Kominski and Siegel (1987), demonstrated that the relationship between years of school completed and actual earned degrees was not exact. More importantly, the data indicated that this years-to-degrees "mismatch" was becoming more serious over time (i.e., the level of mismatch was highest among younger persons). Based on the increase in the level of mismatch, as well as growing requests for actual counts of degree-holders, an attainment question measuring years at levels below high school and degrees beyond high school was tested for inclusion in the 1990 decennial census. The results of the testing indicate that the proposed question has good validity and reliability, as well as documenting that the years-degrees mismatch continues to be a problem. The Census Bureau plans to include in the 1990 census a question on educational attainment which measures "highest degree completed" at the level of high school and and beyond.

The introduction of a new education item has not occurred without some question. One area of concern is the "loss" of the interval variable, years of school completed. Education is recognized as a strong determinant of economic returns (e.g., wages), and the interval scale often facilitates discussion of the "value of each additional year of schooling." The introduction of a new educational attainment item will mean that this specific analysis will not be possible with data from the 1990 census, but it is not clear if this constitutes an analytic loss. In this paper I analyze the earnings of a group of individuals using several different operationalizations of educational attainment. The goal is to try to assess what empirical impact the new education item will have in a fairly common analysis, the regression of earnings on educational attainment.

DATA

The data for this analysis are taken from the third wave interview of the 1984 panel of the Survey of Income and Program Participation (SIPP). The SIPP is a longitudinal survey conducted by the U.S. Census Bureau, with the purpose of measuring the economic, social, and demographic characteristics of persons, and how these characteristics change over time. The sample design is a multistage cluster sample of all living quarters, representing all noninstitutionalized residents of the U.S., ages 15 and older. Households remain in sample for a period of about two and one-half years, and are interviewed every four months.

During the first interview, respondents are asked questions about basic socio-demographic statuses; e.g., race, sex, age, marital status, etc. These items are recorded on a "control card," a survey instrument which is also used to record basic information about the household (e.g., address, number of persons, etc.), and which is physically distinct from the actual survey instrument (the questions to be asked in a specific wave). Interviewers are instructed to "update" the control card each wave, that is, to physically erase and recode any control card item that has changed since the last interview. This is done prior to the actual interview. Included on the control card are the "traditional" items of highest grade attended, and whether or not the grade was completed.

In each wave, the basic questionnaire consists of recurring questions which are asked of respondents concerning their employment and economic activities during the previous four months, including their involvement in various income transfer and supplement programs. Most waves also include a series of special topic questions that are asked in a single wave only. In the third wave interview of the 1984 panel, questions on educational background were asked. Included was a question which asked respondents to report the "highest degree beyond a high school diploma that ... has earned." While interviewers were instructed to use control card

information for universe selection purposes, the coding of the highest degree question was independent of the previously-coded highest grade attended/completed item on the control card. Thus, the SIPP affords the opportunity to examine consistency between the two education items, while at the same time providing detailed earnings information collected for relatively short recall periods (monthly earnings for each of the four months are collected).

The data for this analysis have been restricted to include only persons between the ages of 18 and 64 for whom data on education has not been imputed (since these cases would have, by definition, been made consistent in the years/degree relationship.) The age restriction is imposed since most models of earnings are based only on the employed population (or full-time employed population). Three different measures of education are used: 1) the traditional measure of years of school completed, ranging from 0 to 18 ("6 or more years of college" is the highest category collected); 2) a series of dummy (0,1) variables based on the years of school data, representing five groups: high school dropout, high school graduate only, some college, bachelor's degree, college beyond bachelor's degree; 3) a nine-category variable based on the highest degree completed. All analysis is based on the unweighted sample data of the survey, and may not be specifically generalizable to the entire population. Nevertheless, the general results are likely to be indicative of the findings that would occur based on a more strictly representative sample.

ANALYSIS

Table 1 shows the crosstabulation of the years of school completed item with the question on highest completed degree. The table shows that about 9.9% of all persons with four or more years of college did not have at least a Bachelor's degree, and that two-thirds of these (6.5% of the pool) had no degree beyond high school. Beyond this, about 1 in 4 persons (26.8%) with 1 to 3 years of college have a degree of some sort (for example, associate or vocational), but the remainder have earned no degree beyond a high school diploma. In the traditional years of schooling measurement, the most we would be able to say about this clearly heterogeneous group is that they all have "some college." Most advanced degrees correspond to the years completed category of 18, because this is the upper limit. For many holders of advanced degrees a number somewhat larger than 18 would be necessary to characterize their total number of years of schooling. While the years of schooling item falls short for advanced degree holders, it serves to "overvalue" many individuals with far less impressive degree credentials; 29.4% of the cases with 18 years of school completed report their highest earned degree as a Bachelor's or less. In short, the table demonstrates that the mapping of years into degrees is not exact; while there is a high degree of correspondence between categories one cannot with certainty pick the "correct" degree level knowing years of schooling completed or vice-versa.

A simple elaboration of the relationship between education and earnings is displayed in Table 2. Here, the average monthly earnings have been computed for each of three different operationalizations of education. While the relationship between years of school and earnings is not strictly linear, the pattern of the average earnings amounts is fairly monotonic, supporting the notion of an incremental value for each year of schooling. The second measure, using of a series of dummy variables based on the years data, shows that the education-earnings relationship is well-summarized by a series of five items. The third measure based on received degrees, however, reveals differences (i.e., earnings variability) that is undetected using either of the first two measures. This is true not only for degrees beyond a bachelor's, but for associate and vocational degree holders as well. Because the proportion of persons who are high school graduates or less is so large, the overall population relationship between education and earnings is not affected nearly as much as it might be, were even more persons holders of degrees beyond high school. However, as the proportion of the population with more than a high school degree increases over time (as it has and will continue to do), the relationship of earnings and education based on years weakens.

The principal method used to assess the value of the new education measure versus the old one is a multivariate regression analysis. A simple model for estimating earnings is first constructed using some typical background variables such as age, race, sex, current level of work activity in the recent past (measured here as the number of the last four months in which the person was employed, even if only for a few days), and the number of years the person has worked at their current job. These constitute a succinct yet fairly comprehensive set of items which account for societal differences (e.g., race and sex), historical change (age), and human capital components such as experience and job tenure or seniority. The results of this regression are shown as Model I in Table 3. All of the effects estimated are significant (assessed in this case as parameters which are at least two times their standard error), and the variance explained is about 17%.

The next step is to add to this model each of the three measures of education. Model II adds the years of schooling completed variable, Model III uses the set of dummy variables based on years, and Model IV relies on the highest degree varialbe. The results, in terms of incremental addition of R2, show that the years variable explains an additional 1.9%, but that both the dummy variable measure and the degree measure add about 2.7% to the original "baseline" model. The dummy variables, however, represent 4 degrees of freedom (plus an excluded, fifth, comparison variable) while the degree status variable is a single item. The adjusted R2, accounting for this fact, implies that the degree item is better than years and as good as the multiple dummies, at least in terms of explained variance. We might choose to create dummy variables from the 9-category degree variable, yielding even better explanatory power in the model. This is not pursued since the idea

is to show that the single degree variable alone is as useful, if not more, than the traditional years of school item.

While the regressions in Table 3 lend support to the value of the degree measure, it may not be the case that degrees are the most useful education item for all segments of the population. A large proportion of all adults possess no degree beyond high school, and those who do hold postsecondary degrees tend to be differentially distributed by race, sex and age. A series of regressions, similar to those in Table 3, were run for twelve independent subpopulation groups, created by the crossclassification of age (18-34, 35-49, 50-64), race (white, nonwhite) and sex (male, female). The summarized results are shown in Table 4.

As with the regressions for the entire population, a baseline model was fit first, then the three measures of education were independently assessed for their relative addition to the R2 value. As the table indicates, the general ability to account for the variance in earnings differs substantially between these groups. This is true not only for the "background" variables, but for the effect of education as well. Nevertheless, education is always a significant predictor, and the degree variable is always more useful in explaining variance than the years variable. In about half the groups the use of a series of dummy variables based on years of schooling accounted for a larger amount of variance than did the degree item. Examination of the results indicate that no single age, race or sex dimension exhibits a consistent pattern of improvement. Thus, it does not seem that degrees differentially improve the explanation of earnings for any specific subgroup beyond the general overall improvement.

As a final test for subgroup differences, another regression equation was estimated. In this model, baseline variables and the degree variable were fit as in Model IV in Table 3. Then, years of school completed was added in order to correct for any overall population relationship with years, net of degrees. A series of fourteen possible subgroup-specific "interaction" variables were then allowed to enter the model in a stepwise fashion.

The results of this regression are shown in Table 5. Three subgroup-specific effects were added to the basic model. (Only those terms that were at least twice their standard error were accepted for inclusion.) Two of these effects, involving interactions between age and degrees, reflect the high degree attainment of persons currently in ages 35 to 49, and the fact that persons 18-34 have not yet completely finished their education (and degree completions). Thus, degrees have "extra value" for persons in their middle ages, most likely because they have had the time in their job to start to realize the economic return that the degree inevitably (for many persons) earns. For younger persons the individual has not had enough time in the job, and the degree, at least initially, is not "worth it."

The more important effect is that of female gender and degree status. With the addition of this variable to the model, the constant term changes from -141 to -245.6, and the term for female sex changes from -561.5 to -120.1. In simple terms the female/degree interaction (which has a coefficient of -115.4) allocates mush of the negative wage value of being a woman from the gender term to the constant and the sex/degree interaction term. This occurs because wages for women versus men become less equal as the degree level increases. While there is a positive effect associated with each additional earned degree, the effect is less for women than for men. It is this degree effect, not the years of schooling effect, which best captures the sex-based inequity in the relationship of education and earnings.

SUMMARY

The change to a new item for measuring educational attainment must recognize the many different uses of the concept, and what the effect of the change might be. In the case of education, one principal use of the item is in models which attempt to estimate the relationship of education and earnings. This analysis indicates that a new measure of educational attainment, based on the highest degree attained, yields a better earnings-educaiton relationship than the existing measure of years of schooling completed. The improvement in the model appears to apply to the entire adult population, however, the addition of subgroup-specific interaction terms demonstrates the added utility of measuring degrees with respect to gender differences in returns to education. The growing importance of degree attainment is also demonstrated by the inclusion of significant age/degree interactions for younger age groups. As the proportion of adults holding degrees beyond high school continues to grow, the utility of the degree item should become even more apparent.

REFERENCES

Kominski, Robert. Evaluation of the 1980 Decennial Census Education Questions. Preliminary Evaluation Results Memorandum Number 104. Bureau of the Census, U.S. Department of Commerce, Washington, D.C., 1985.

Kominski, Robert, and Paul M. Siegel. Measuring Educational Attainment in the 1990 Census. Presented at the 1987 Annual Meeting of the American Sociological Association. Chicago, Illinois, 1987.

Mincer, Jacob. Schooling, Experience and Earnings. New York: Columbia University Press, 1974.

Table 1. Years of School Completed by Highest Earned Degree

	Years of Schooling Completed								
	0-11	12	13	14	15	16	17	18	TOTAL
Professional	0	0	0	0	0	4	7	268	279
Doctorate	0	0	0	0	1	4	1	133	139
Masters	0	0	1	2	0	37	198	843	1081
Bachelors	0	4	15	41	81	2714	382	413	3650
Assoicate	0	6	57	765	246	86	14	10	1184
Voactional	0	22	196	203	93	62	7	10	593
High School, some college	0	1089	2148	1758	732	218	60	84	6089
High School	580	11265	0	0	0	0	0	0	11845
H.S. Dropout	6354	282	0	0	0	0	0	0	6636
TOTAL	6934	12668	2417	2769	1153	3125	669	1761	31496

NOTES: Values are unweighted sample counts of nonimputed cases from the 3rd wave of the 1984 panel of SIPP

Table 2. Average Monthly Earnings by Education: 3 Operationalizations (Unweighted data)

Years of School Comp.	$	Dummy Variables	$	Degree Categories	$
0	333				
1	545				
2	704				
3	502				
4	431	H.S.		H.S.	
5	489	Dropout	596	Dropout	596
6	599				
7	454				
8	632				
9	634				
10	627				
11	608				
12	942	H.S. Grad	942	H.S. Grad	942
13	965			Some col: no deg.	1025
14	1088	Some Col.	1060	Vocational	1138
15	1188			Associate	1250
16	1735	Col. Grad	1735	Bachelor's	1701
17	1777				
18	2443	Post-BA	2259	Master's	2192
				Doctorate	3196
				Professional	4073
TOTAL	1070		1070		1070

Table 3. Regressions of Average Monthly Earnings on Alternative Education Measures

Variable	Model I	Model II	Model III	Model IV
Intercept	377.2	-747.9	164.6	-65.3
Work exposure	257.8	227.3	226.5	224.1
Experience	36.5	35.9	36.6	36.6
Sex	-579.2	-580.1	-558.8	-560.2
Race	-182.3	-110.6	-125.1	-113.6
Age (35-49)	272.9	278	229.3	246.1
Age (50-64)	55.4	154.4	89.2	117.8
EDATT		94.8		
EDUM2			168.8	
EDUM3			262.3	
EDUM4			781.9	
EDUM5			1151.6	
DEGREE				178.9
R2	0.17	0.189	0.197	0.197
ADJ R2	0.17	0.189	0.196	0.197

VARIABLES:

Work exposure (0-4): Number of the last four months when individual was employed for any amount of time
Experience (0-48): Number of years working at current job
Sex: Dummy variable 1=Female
Race: Dummy variable 1=Nonwhite
Age (35-49): Dummy variable for this age category
Age (50-64): Dummy variable for this age category
EDATT (0-18): Years of schooling completed
EDUM2: Dummy variable for 12 years of school completed
EDUM3: Dummy variable for 13-15 years of school completed
EDUM4: Dummy variable for 16 years of school completed
EDUM5: Dummy variable for >16 years of school completed
DEGREE (1-9): Degree attainment variable; 1=H.S. Dropout, 9=Professional

Excluded (comparison) categories are: Male, White, ages 18-34

Table 4. Summary of Earnings Regressions for Age-Race-Sex Subpopulations, and R2 Increments, Using Three Alternative Education Measures

			Sample n	Background R2 (Adj)	R2 (Adj) Increment EDATT	EDUM	DEGREE
ALL PERSONS			31507	17.0	1.9	2.6	2.6
SEX	RACE	AGE					
M	W	Y	6246	16.9	4.2	5.9	5.8
M	W	M	3880	6.4	4.8	5.4	6.1
M	W	O	3051	9.4	1.6	1.9	2.1
M	NW	Y	1036	35.9	2.9	5.5	4.7
M	NW	M	466	31.4	8.6	10.9	11.8
M	NW	O	386	27.1	6.5	9.1	9.4
F	W	Y	6481	5.1	0.5	1.0	0.6
F	W	M	4126	36.6	2.5	3.8	4.0
F	W	O	3433	47.4	1.8	2.8	2.5
F	NW	Y	1257	10.5	0.9	3.5	1.8
F	NW	M	681	36.8	4.4	8.3	7.3
F	NW	O	464	44.0	6.9	17.9	15.9

VARIABLES: SEX (M=male, F=female); RACE (W=white, NW=nonwhite);
AGE (Y=18-34, M=35-49, O=50-64)
EDATT=Years of school completed
EDUM=Dummy variables (as specified in Table 3)
DEGREE=Highest degree earned (as shown in Table 3)

Table 5. Regression Model With Education Interactions

	Without Interactions		With Interactions	
Variable	B	s.e. B	B	s.e. B
Intercept	-141.0	-	-245.6	-
Race	-111.3	28.7	-112.2	28.6
Sex	-561.5	20.4	-120.1	36.8
Age (35-49)	248.2	23.9	1.2	43.7
Age (50-64)	123.0	27.7	-12.4	45.6
Experience	36.6	1.5	36.5	1.5
Work exposure	223.4	6.6	228.8	6.6
Degree	167.7	9.4	240.4	14.8
Edatt	8.7	5.9	8.4	5.9
Female X Degree			151.4	10.7
Age (18-34) X Degree			-41.8	14.0
Age (35-49) X Degree			35.0	14.3
R2	0.197		0.204	

NOTE: Variables are defined as in Table 3. Interactions are defined in text.