Volume Title: Higher Education and Earnings: College as an Investment and Screening Device

Volume Author/Editor: Paul J. Taubman, Terence Wales

Volume Publisher: NBER

Volume ISBN: 0-07-010121-3

Volume URL: http://www.nber.org/books/taub74-1

Publication Date: 1974

Chapter Title: Effects of Education and Mental Ability on Income: The Evidence from the Wolfe-Smith Data

Chapter Author: Paul J. Taubman, Terence Wales

Chapter URL: http://www.nber.org/chapters/c3649

Chapter pages in book: (p. 37 - 56)

3. Effects of Education and Mental Ability on Income: The Evidence from the Wolfle-Smith Data

As discussed in the preceding chapter, in order to estimate the net effect of education on earnings, it is necessary to standardize for all other determinants of earnings that are correlated with education. This is not an easy task, since innate ability, which is correlated with educational attainment and probably is a determinant of earnings, is not available in most samples. In this chapter we make use of a sample whose members span the entire IQ range to study the effects of education and the comparative usefulness of high school rank and the American College Entrance Examination. The results of this study are of limited applicability for the following reasons: only grouped data are used; the sample is drawn only from Minnesota; the respondents may be more successful than the nonrespondents, given education and ability; the individuals' educations and lives were interrupted by World War II; and earnings for only one year are available.

SUMMARY AND CONCLUSIONS

We use hitherto unpublished details of data collected by Wolfle and Smith (1956) on Minnesota high school graduates of 1938 to estimate the effect of education and ability on wages and salaries earned in 1953. Using a general nonlinear functional form, we find that, for a person with the same IQ as the average high school graduate, the extra earnings from vocational training are less than 7 percent; from attending college for less than two years, 18 percent; from attending college for more than two years but not graduating, 36 percent; from earning one degree, 47 percent for those in the first nine IQ tenths and 100 percent for those in the top tenth; and from earning two degrees, 58 percent for those in the first nine IQ tenths and 111 percent for those in the top tenth.[1] Except for the people in the top IQ tenth, the percentage increase in income falls as education grows. For those in the top IQ tenth, one college degree represents a huge 50 percent increase over not graduating.

An important feature of the sample is that because the IQ test (the ACE) was administered to all the students in their senior year in high school, subsequent differences in schooling would not affect students' scores. Besides reflecting differences in innate ability, the test scores could also depend on the quality of previous schooling and home environment. By holding constant such elements, the test scores should improve the education estimates but will overstate the effects of innate ability. We find that mental ability adds to earnings but that education is a more important determinant than IQ. The effect of mental ability on income differs for three groups: those in the lowest four IQ tenths, those in the next five tenths, and those in the top tenth.[2] While we find that high ability and high educational attainment interact strongly to produce very large income differences, ability affects income even for high school graduates.

Two important questions concern the types of mental ability that determine earnings and the best measure of a particular type of ability. In this chapter we use a general innate-mental-ability concept, two measures of which are an IQ-type test and high school rank in class.[3] For this body of data it is clear that the IQ scores are the superior measure, since they have the expected signs and are significant determinants of income differences. Coefficients on the rank-in-class variables are neither significant nor large, and do not increase with higher rank.[4] Our explanation for the poor performance of the rank-in-class measure is that, because the quality of the student body varies between schools, it is not legitimate to call all students in the

[1] The sample is Minnesota male high school graduates of 1938. The average annual salary in 1953 of those with no additional education was $4,500.

[2] Within each of these groups, variations in ability have no effect on income.

[3] Of course, both measures could also incorporate other attributes.

[4] In interpreting both measures, it must be remembered that there was a response bias on the sample, with the more educated and more able much more likely to respond. In addition, there may have been a success bias, which could have been more extreme at the lower ability and education levels.

top ranks of different schools equal or to assume that they are more able than those in the lower ranks.

For any ability measure, it is useful to calculate the bias that could be expected in samples that relate income to education without holding ability constant. We find that the bias is quite small—no more than 4 percent at the various education levels. The peculiarities of the sample used heighten the importance of this result. As indicated in our earlier paper (Taubman & Wales, 1972), the coefficient on education in the regression between ability and the percentage of students entering college is higher in the Wolfle-Smith sample than in any other studied. Thus, as long as the estimates of the effects of ability and education on income are similar in other samples, the bias from omitting ability would be larger in this sample than in any other.[5] In our analysis of the NBER-TH sample, however, we find the bias to be larger than in the Wolfle-Smith sample, especially when a different concept of ability is used.

After dividing the data into groups involving the three highest-paying (on the average) occupations and the other five, we computed our regressions within each of the two groups. For the high-paying group of professionals, semiprofessionals, and sales, we find that the education coefficients are very small and statistically insignificant as long as ability is held constant.[6] Ability, however, is statistically significant and quite large, with those in the top tenth earning 20 percent more than those in the fifth and sixth tenths and about 30 percent more than those in the bottom four tenths. For the other occupations, which have lower average wage and salary levels, we find that neither education nor ability is a significant determinant of income.[7]

[5] In samples such as the census, in which people in different age groups are studied simultaneously, the problem is more complex because the ability-education relationship shifts for different age groups. See our earlier paper (Taubman & Wales, 1972).

[6] This conclusion holds if we use the six possible education categories; if we combine the data into the three groups of no college, college dropout, college graduate; if we use the two categories of college graduate and all others; or if we use the two categories of no college and all others.

[7] See, however, the discussion in Chapter 6 on the position, steepness, and intersection of age-income profiles at various education levels.

The above regression results could occur if there were no variations in education and ability within the two occupational groups. But as shown in Appendix B, Table B-9, there are wide ranges of education and ability in the various occupations. For example, 34 percent of the people with just a high school education in the bottom four tenths of IQ are in professional, semiprofessional, and sales occupations, while more than 30 percent of high school graduates in the top six tenths are in the other occupations. However, over 90 percent of people with one or more college degrees are in the professional, semiprofessional, and sales categories in each of the bottom four tenths, middle three tenths, and top three tenths of ability.

The occupation regressions indicate that people with less education can earn the same income as more educated people with the same ability when they are given the opportunity. But there is a disproportionately low percentage of people with low education in the high-paying occupations. (In Chapter 9, we indicate a method to calculate the "proper" percentage.)

For females, we find that ability is not significant, but that one or more college degrees add substantially to income, although by lesser absolute amounts than for males.[8] Housewives are included in these data; hence, the results are partly determined by the fact that a smaller proportion of college graduates are married and/or not working full time.

While we believe the above conclusions to be important, they are subject to some qualifications. First of all, Wolfle and Smith (1956) report that there was a tendency for a greater proportion of those high school students who did not enter college not to respond to the questionnaire. It may also be true that the less successful, in terms of income, did not respond.[9]

Of more concern is the date at which the sample was taken. While chronologically the sample refers to a period 15 years after high school graduation, the interval includes World War II. Thus, lawyers and those with Ph.D.'s would have had fewer

[8] College dropouts among females may make their return through better selection of husbands.

[9] There is some evidence that the nonrespondents were those with lower scores; hence, our results with IQ held constant could be valid, although the bias calculation need not be. For the NBER-TH sample we show that there is a response bias but no success bias (see Chapter 4).

than five years in their primary occupation and M.D.'s only one or two years. Graduates with one degree who entered college in 1938 would have been on the job for as long as seven years if they were in the service; hence, their income figures should reflect promotions and phenomena other than starting salaries. However, as reported in Anderson and Berning (1941), the percentage of Minnesota high school graduates of 1938 who entered college in 1939 was one-half as large as the corresponding group who had entered by 1953, as recorded by Wolfle and Smith (1956).[10] While some of this difference reflects the response bias previously discussed, part also reflects post-World War II education. If we assume their earlier work experience to be irrelevant after college, the veterans who entered college after the war and who graduated would have been on the job about four or five years.

Besides these technical problems, there is the obvious qualification that our results apply to only one year out of an individual's lifetime experience. It is conceivable that the high school graduates who are professionals and who were making more income in 1953 than their counterparts in other jobs will have the same lifetime income as their counterparts. We doubt it, however, since census data indicate the same hierarchy of average income in all age brackets. It is also possible that the lifetime income of the more educated will be greater, even though this is not observed in this sample, because it takes time for them to overcome the advantages derived from experience on the job by their less well-educated coworkers. Both these possibilities can be tested by examining data for individuals over long time periods, as is done in Chapter 5.

The final qualification is that there may be other individual characteristics that determine income *and* are correlated with education, ability, or occupational choice, but that we have not held constant.[11] Since our sample incorporates only Minnesota high school graduates of 1938, some of these factors have been accounted for. However, such personal characteristics as drive,

[10] Wolfle has informed us that his sample was a reinterview of the people dealt with in Anderson and Berning (1941).

[11] These characteristics must be correlated with education, ability, or occupational choice in order to cause a bias.

motivation, personality, and the like have not been held constant. If these attributes determine income and are correlated with education, then the calculations of the individual return are biased because part of the credit apportioned to education belongs to these attributes. In Chapter 5, we test for the relative importance of some of these omitted variables using the NBER-TH sample.

DATA The body of data collected by Wolfle and Smith (1956) is described in detail in Taubman and Wales (1972). Summarizing briefly, the population from which the samples were drawn consists of (1) the top 20 percent of Rochester, New York, high school graduates of 1933–1938; (2) the top 60 percent of Illinois high school graduates of 1935 (excluding Chicago); and (3) all Minnesota high school graduates of 1938. Information on an individual's income in 1953, occupation in 1953, post-high school education, father's occupation, and many other sociological and economic items were collected as part of the study.

We have not been able to locate the original questionnaires or the cards on which the data were punched. However, Dr. Wolfle retained a file of extensive cross tabulations, which he graciously made available to us. Included in his files are such tables as the distribution of Minnesota males for each of ten high school class ranks, for eight classes of post-high school education, and for each of nine occupational groups.[12] Another table presents, for each tenth in class rank, the data on the distribution of wage and salary income by occupation for Minnesota males. Comparable tables for males in the other two areas and for females are available, as well as cross tabulations using IQ in place of rank in class.

Although the Wolfle-Smith sample has been extensively analyzed in the literature on the returns to education, very few of the basic data have been published. The three tables published in the original Wolfle and Smith article are presented in Appendix B, Tables B-1, B-2, and B-3. While the data contained in these tables are useful for some purposes, in general

[12] There is a misprint in the original Wolfle and Smith article in the labeling for the ability classes for the Minnesota ACE tenths. The correct classifications are given in Appendix B, Table B-2.

we believe that the information is to some extent either misleading or inadequate.[13]

Wolfle's extensive cross tabulations permit us to rectify many of the problems detailed in footnote 13 and to provide comparable information for females. Since the data are basic to the analysis and since we consider the available tables to be deficient, we present in Appendix B a more comprehensive set of tables.

RETURN TO EDUCATION: MINNESOTA MALES

For reasons given earlier, it is desirable to analyze the information for each state separately. Unfortunately, although rank-in-class data are available by tenths for all states, the IQ data are available in sufficient detail only for Minnesota. We have estimated separate equations for males and females because females may have been discriminated against in job markets and because many females in the sample were married and had dropped out of the labor force.

Using the data in Appendix B, Table B-6 (in which the measure of mental ability is the ACE score), we find[14]

[13] Specifically, the published data are deficient in the following respects:

(1) No useful mathematical operations, such as averaging, can be carried out properly using the medians available. For analysis of variance techniques or regressions, the appropriate measure of central tendency is the mean, though one problem with the mean is that an average income must be assigned to those people in the open-ended class.

(2) The published tables are not extensive enough to permit use of regression analysis, which allows other variables to be held constant and can utilize different weights for sample points with different numbers of observations.

(3) The most detailed table combines all Rochester, N.Y., Minnesota, and Illinois graduates who ranked in a given tenth of their high school class. Using combined rank-in-class data for the three states is misleading on several accounts. First, the quality of the schools or student body may differ substantially in these three areas. Second, the years in which the students graduated from high school differed in these three areas: 1933–1938 in Rochester, 1935 in Illinois, and 1938 in Minnesota. Abundant evidence exists that wages and salaries are related to time on the job. Because of this or for other reasons, the average income received in any educational and rank-in-class cell was higher in the other two states than in Minnesota. Finally, due to the design of the sample, nearly all the people in the lower ranks come from Minnesota, while most of those from the other two areas are in the top ranks.

(4) The other two tables based on IQ are more revealing, but the Rochester data do not indicate the average IQ in either of the two groups.

[14] These data have not been grouped by occupation.

$$Y = \underset{(10.3)}{4.259} + \underset{(.6)}{.302}E_V + \underset{(1.6)}{.814}E_1 + \underset{(3.4)}{1.612}E_3 + \underset{(4.8)}{2.107}E_G + \underset{(5.5)}{2.658}E_{GM}$$

$$+ \underset{(2.3)}{.483}A_{5-9} + \underset{(2.9)}{1.283}A_{10} \qquad \bar{R}^2 = .93 \quad (3\text{-}1)$$

where Y = wage and salary income in thousands of dollars

E_V = a dummy variable that equals 1 if the person attended vocational, military, or other noncollege school

E_1 = a dummy variable that equals 1 if the person attended college, but for less than two years

E_3 = a dummy variable that equals 1 if the person attended college for two or more years but did not graduate

E_G = a dummy variable that equals 1 if the person had an undergraduate degree but no graduate degree

E_{GM} = a dummy variable that equals 1 if the person had more than one degree

A_{5-9} = a dummy variable that equals 1 if the person was in the fifth through ninth IQ tenths

A_{10} = a dummy variable that equals 1 if the person was in the tenth IQ tenth (the most able students are in this tenth)

$\bar{R}^2$ = the coefficient of determination adjusted for degrees of freedom

The equations are weighted regressions in which each weight is the square root of the number of observations in the IQ group.[15]

We first tried an equation with a separate dummy variable for each possible education and ability category, which is the greatest detail available. The individual coefficients on A_2 through A_4 were practically zero, while the coefficients on A_5 through A_8 were nonzero and very similar. For some reason, the coefficient of A_9 is very low, but for convenience we combined this category with A_5 through A_8. Finally, the coefficient of A_{10} was very large, a result consistent with earlier studies on the effect

[15] Eq. (3-1) has also been estimated using $20,000 instead of $25,000 as the mean of the open-ended class. The coefficients of the variables in the same order as in Eq. (3-1) are 4.270, .250, .810, 1.475, 1.906, 2.319, .477, and 1.015.

of mental ability on income.[16] To reduce collinearity, we combined the ability variables into three groups. This aggregation does not greatly change the education coefficient.

Since the categories for no post-high school education and for the bottom IQ decile have been excluded, the coefficient on a dummy variable indicates the average additional amount of income for people in the particular category compared with those in the lowest ability tenth who had no post-high school education (the reference group). Thus in Eq. (3-1), a person in the lowest IQ tenth but with zero to two years of college earned, on the average, approximately $800 more than a person in the reference group.

Consider the magnitudes of the education dummy variables. The coefficients increase continuously with education. For any IQ level, vocational training added about $300 to earnings 15 years (including World War II) after graduation from high school. Using the *t* statistic in parenthesis, this increase—about 7 percent—is not significantly different from the earnings of a comparable person with no post-high school training.[17] A person who attended college for less than two years received $800 more than a person with no post-high school education. The $800 figure is nearly statistically significant, and the remaining education coefficients are significant and are successively larger. In terms of percentages, the first two years of college add about 18 percent to income; the next two years add about 36 percent; one degree adds about 45 percent; and more than one degree adds roughly 57 percent (as compared with no post-high school education). While the increments in these percentages decrease as we move to higher education levels, it must be kept in mind that the mean years of college education in the various categories are approximately 1, 3, 4.5, and perhaps 6.[18] Thus, while it seems clear that the largest gain in income occurs for the first year in college, the absolute differences per year for the other categories may be about the same. Even this would mean that the percentage increase in income for each additional

[16] See, *for example*, Becker (1964).

[17] The 7 percent figure is for a person with the ability level of the average person with a high school degree only. For people with higher ability, the percentage would be smaller.

[18] We do not know how many Ph.D.'s and M.D.'s are in the sample.

school year was falling, because the constant absolute differences would be divided by a continually growing base. Thus, there is some evidence that there are diminishing (percentage) returns to education for an individual.

The ability variables indicate that those in the fifth through ninth tenths earn $480 more each year, and those in the top decile $1,300 more than persons in the bottom four deciles. When the discrete variables for each ability class are replaced with a continuous variable that takes on a value equal to the particular IQ tenth (1 through 10), the coefficient on the variable is approximately .1, while the coefficients of the education variables are unchanged. Although statistically significant and numerically important, ability explains much less of the range of income differences than does education. The income difference between the top and bottom ability group is less than the income difference between those with two and four years of college.

There is one other aspect of this equation that merits discussion. Unlike most other studies of the determinants of income, our equations do not contain an age or time-on-the-job variable. An age variable is not needed because all the people in our sample graduated from high school at the same time and would have been about the same age. Time on the job, however, would vary between individuals in part because of differences of time spent in the military and in college. We do not know the military experience of each individual. However, there should be little variation in length of service after averaging to obtain our grouped data.

Length of time in education obviously varies by amount of educational attainment. Suppose we write time on the job as equal to age minus years of education (S). Further, suppose that S is the proper measure of education in the income equation

$$y = aS + b(\text{age} - S)$$

This can be written as

$$y = (a - b)\ S + (b \times \text{age})$$

The last equation is the type we have estimated, though we have omitted age because it is constant, and we have partitioned S in a set of dummy variables. From this equation we see that the coefficients on the education variables represent the net effects of two different mechanisms. That is, a gives the impact of education on income for people with a given amount of time on the job, while b gives the effect of changes in time on the job for people with a given amount of education. Obtaining additional education automatically alters both variables; hence, the net impact is given by $a - b$. With this sample we cannot estimate b, whose magnitude is of some interest. For many purposes, however, the net impact $a - b$ is the required piece of information. It is worth noting that b could vary by education level; hence, the relative size of the various education coefficients could change with age.[19]

INTERACTION OF ABILITY AND EDUCATION

In the above equations it is assumed that the effect of education on income is the same regardless of the level of ability. Many people, however, have hypothesized that education is more important, or only important, for the most able students. This possibility can be tested by including various ability-education product terms in the regression equations. For example, to test the effect of high ability, we use variables defined as $A_{10}(E_G + E_{GM})$, $A_{10}(E_1 + E_3)$, and $A_{10}(E_V + E_H)$, where E_H is high school education only. To test the effect of higher education, we use variables defined as $A_{10}(E_G + E_{GM})$, $A_{5-9}(E_G + E_{GM})$, and $A_{1-4}(E_G + E_{GM})$, and for low education, we replace $E_G + E_{GM}$ by $E_V + E_H$. The only interaction term that appears to be significantly related to earnings is the product of A_{10} and $E_G + E_{GM}$. This variable represents high ability together with high educational attainment. The following equation is the same as Eq. (3-1) except that it contains this additional term.[20]

$$\underset{(11)}{Y = 4.307} + \underset{(.6)}{.276E_V} + \underset{(1.7)}{.786E_1} + \underset{(3.6)}{1.642E_3} + \underset{(4.8)}{2.017E_G} + \underset{(5.4)}{2.494E_{GM}}$$

$$+ \underset{(2.5)}{.499A_{5-9}} - \underset{(.7)}{.593A_{10}} + \underset{(2.6)}{2.460A_{10}(E_G + E_{GM})} \qquad \bar{R}^2 = .93 \qquad (3\text{-}2)$$

[19] See Chapters 5 and 6.

[20] Eq. (3-2) has also been estimated with an open-end mean of $20,000. The corresponding coefficients are 4.307, .228, .788, 1.499, 1.834, 2.185, .490, −.509, and 2.000.

The result may be interpreted in the following ways. First, the additional income attributable to one college degree is $2,017 for those in the first nine ability classes but is $4,477 ($2,017 + $2,460) for those in the top tenth. For more than one degree, the corresponding values are approximately $2,500 and $4,900. An alternative interpretation suggests that no additional income accrues to individuals in the top tenth without a college degree, since the coefficient of A_{10} is insignificant. This result should be treated with caution, since there are only 15 people without college degrees in the top tenth. Individuals in the fifth through ninth tenths, on the other hand, earn a modest amount of added income—about $500—regardless of their education.[21]

Another proposition that we have tested is that ability differences have no effect on income for those with just high school or vocational training.[22] This was tested by including in Eq. (3-3) a set of additional variables defined as $A_{1-4}(E_{H+V})$, $A_{5-9}(E_{H+V})$, and $A_{10}(E_{H+V})$, where E_{H+V} is a dummy variable representing all those with just high school or vocational training. In order to interpret the results of including these variables, it should be noted that the ability variables also appear separately, that A_{5-9} is significant, and that (when interactions are used) A_{10} is not significant except when included as an interaction with $E_G + E_{GM}$. Consequently, if it were true that ability had no effect at low education levels, the coefficient of $A_{5-9}(E_{H+V})$ would have to be significant and negative, and of a magnitude sufficient to negate the effect of the separate ability variable A_{5-9}. This is not the case, as all three variables are insignificant, with t values less than .5. These data therefore do not support the hypothesis that ability differences contribute to income differences only for those with high educational attainment.[23]

THE BIAS FROM OMITTING IQ

Many previous studies of the returns to education have been based on census or other data sources that contain no ability

[21] We tested a variable defined as $A_{5-9}(E_G + E_{GM})$, and although the coefficient was positive (.304), its t value was less than 1.

[22] Becker (1964) reaches this conclusion.

[23] Evidence given below strongly suggests that ability adds to income only if people are employed as professionals, semiprofessionals, or salesmen. Since most high school graduates are not in these professions, there is some truth to the proposition.

measure. As the authors of those studies and others have noted, if ability and education are positively related, then the omission of ability will result in attributing too much of the earnings differential to education. To observe the extent of this bias when IQ is the measure of ability, we have reestimated the equation omitting the ability variables.[24]

$$\underset{(10)}{Y = 4.493} + \underset{(.5)}{.273E_V} + \underset{(1.5)}{.808E_1} + \underset{(3.3)}{1.673E_3} + \underset{(5.8)}{2.218E_G} + \underset{(5.8)}{2.914E_{GM}} \qquad \bar{R}^2 = .92 \qquad (3\text{-}3)$$

For the significant education variables in Eq. (3-2), we calculate the extent to which each coefficient in Eq. (3-3) exceeds the corresponding one in Eq. (3-2) and express this as a percentage of the latter. This is a measure of the upward bias in Eq. (3-3) attributable to omitting the ability measure. For E_3 this value is approximately 2 percent. The additional income due to one college degree from Eq. (3-3) is $2,017 for the first nine ability classes, and $4,477 for the top ability class. The weighted average of these two values (with weights equal to the numbers in the two groups) is $2,151. The difference between this and the value of $2,218 from Eq. (3-3) is approximately 3 percent. The corresponding estimate for those with more than one degree is −1 percent.

Since these percentages are all very small (less than 4 percent), the additional income from education can, with these data, be estimated fairly accurately without including an ability measure. This result is of great interest in view of our discussion in Taubman and Wales (1972), in which we traced the pattern of the relationship between ability and educational attainment for various samples. The conclusion reached there was that the slope of such a relation for Minnesota males, using the ACE decile measure, was *steeper* than in any other time period. The implication of these two results taken together is that the bias in the education coefficients due to omitting IQ will in general be very small *provided* the relative importance of the education and ability coefficients in determining income is as in

[24] The corresponding coefficients of Eq. (3-3) using an open-end mean of $20,000 are 4.493, .225, .808, 1.531, 2.009, and 2.532.

this sample.[25] (However, in Chapter 5, in which the NBER-TH sample is discussed, the bias based on a different ability measure is larger.)

The relatively small bias due to omitting ability is also of interest, because it is derived from the data source that was used in summary form by Becker (1964) and Denison (1964) in attempts to answer the same question. Becker concluded that an increase in ability has a negligible effect on the earnings of high school graduates and a 15 to 20 percent effect among college graduates. Denison concluded that about one-third of the income differentials between individuals with different educational attainment was not due to education. Our estimate of this differential is about 4 percent.[26] This does not imply that ability differences are unimportant, since as mentioned above, college graduates in the top decile earn approximately 30 percent more income than college graduates in the bottom four deciles. But as indicated in footnote 25, the relative bias on the education coefficient α is equal to $\beta\gamma/\alpha$, where β is the effect of income on education and γ is the *reciprocal* of the marginal effect of IQ on educational attainment. While both β and α vary by ability and educational level, their ratio would be no smaller than one-fourth. Hence, the relative bias is less than 4 percent because (in the units used in this study) γ is small. However, a small γ implies that the marginal effect of IQ on educational attainment is large.

RANK IN CLASS The above equations use IQ as a measure of mental ability. It has been suggested that high school rank in class is a more appropriate measure. The arguments for rank in class are that it is a more accurate measure of mental ability and that it accounts for such factors as drive and motivation, which are important in determining income. On the other hand, as suggested earlier, the rank-in-class data can be deficient because the quality of the

[25] Of course, the only data we have here are for individuals 15 years (including World War II) out of high school. The bias could be greater at different ages. This aspect is discussed in more detail in Chapter 5.

If the true equation is $y = \alpha S + \beta A$, then the bias from omitting A is equal to $\beta\gamma$ where γ is found by estimating the equation $A = \gamma S$. In Taubman and Wales (1972), we show that γ has varied.

[26] However, Denison's one-third also allowed for differences in family background (Denison, 1964).

student body varies from one school to the next. Since data are available for rank in class as well as IQ, it is interesting to compare the effect upon income of these two ability measures.

The following equation, in which the ability measure is rank in class (R), can be compared with Eq. (3-2).

$$
\begin{aligned}
Y = \underset{(10.5)}{4.545} + \underset{(.5)}{.227}E_V + \underset{(1.7)}{.793}E_1 + \underset{(3.5)}{1.565}E_3 + \underset{(4.9)}{2.034}E_G + \underset{(5.6)}{2.565}E_{GM} \\
- \underset{(.4)}{.112}R_{5-9} + \underset{(1.1)}{.352}R_{10} \qquad \bar{R}^2 = .92 \qquad (3\text{-}4)
\end{aligned}
$$

The most noticeable difference between the two equations is that in Eq. (3-4) the ability variables are no longer significant. In addition, the point estimates here are much lower; in fact, the coefficient of R_{5-9} is negative. For the equation in which R_1 through R_{10} are included separately, no ability t value exceeds 1.5 and there is no apparent pattern in magnitude or sign to the coefficients. Clearly, the rank-in-class ability variable is inferior to the IQ ability variable, in that the latter measures an attribute that is significantly related to income and the former does not. If data were available on the quality of the various schools, inclusion of such a variable would yield more meaningful results.

OCCUPATIONAL REGRESSIONS

According to the material to be discussed in Chapter 9, it is possible to test the hypothesis that education is used as a screening device by determining if a disproportionately lower percentage of people at lower education levels are in the high-paying occupations open to them.[27] The test requires that we estimate the income that could be earned in each occupation by a person with a given set of characteristics. To accomplish this, we estimate equations within occupation groups. Such equations are also of interest for other reasons.

The Wolfle-Smith data are available by the nine broad census occupational categories.[28] However, we have only the cross tabulations of the basic data and at most ten observations (corresponding to the ability measure) on income and education

[27] See pp. 158–163

[28] These are professional (1), semiprofessional-managerial (2), clerical (3), sales (4), service (5), skilled (6), farm (7), unskilled (8), and housewife (9).

within any occupation. The possibility, therefore, of testing various education and ability variables is severely limited. For example, the six educational categories and ten ability categories would have to be combined into fewer than ten variables in order to estimate the equation. In order to preserve degrees of freedom, we first combined the data for the professional, semiprofessional, and sales occupations.[29] In these equations we included a dummy intercept variable for two of the occupations. This method constrains the slope coefficients to be the same, while permitting the average income levels to vary among the three occupations. The number of observations in this group of occupations (30) permits estimation of the type of income-education-ability relationship discussed above. The other occupations (3, 5, 6, 7, and 8) were combined in the same way and estimated as a group with dummy constant terms for some or all of the occupations.

Examples of the types of relationships with which we have experimented are:

Occupations 1, 2, 4

$$Y = -\underset{(.1)}{.377E_{V+1+3}} - \underset{(.1)}{.373E_{G}} - \underset{(.2)}{.725E_{GM}} + \underset{(2.8)}{.693A_{5-9}}$$

$$+ \underset{(2.7)}{1.895A_{10}} - \underset{(1.6)}{.797D_{2+4}} + \underset{(1.9)}{6.582} \qquad \bar{R}^2 = .97 \quad (3\text{-}5)$$

Occupations 3, 5, 6, 7, 8

$$Y = \underset{(.7)}{.594E_{V+1+3}} + \underset{(1.6)}{2.037E_{G}} + \underset{(.7)}{1.279E_{GM}} + \underset{(.2)}{.049A_{5-9}} - \underset{(.3)}{.199A_{10}}$$

$$+ \underset{(.5)}{.182D_{3+5}} + \underset{(2.6)}{.942D_{6}} + \underset{(4.8)}{3.194} \qquad \bar{R}^2 = .91 \quad (3\text{-}6)$$

where E_{V+1+3} is a dummy variable with a value of 1 if the individual has either vocational training or any college education short of a degree, and D_i is an occupational dummy with a value of 1 for those in occupation i, and a value of zero otherwise.

In no case does the t statistic for an education variable

[29] This division was made on the basis of preliminary equations using all the occupations. In these equations we allowed the constant term to vary by using a dummy variable. These equations consistently pointed to this division which, moreover, is in accord with intuitive impressions and census data rankings.

suggest a coefficient significantly different from zero at the 5 percent level.[30] The point estimates of the education parameters from Eq. (3-5) are negative, whereas those from Eq. (3-6), although positive, are not significant, nor do they increase with educational attainment. The ability variables have significant effect on income in the former case but not the latter. Regressions with different education and ability groupings yielded similar results—in no case did education appear to influence income in a systematic manner.[31] From these regressions we reach the somewhat surprising conclusions that, at an average age of 33, there is no (determinable) significant effect of education on income within occupations.

The two equations also reveal some other interesting information. In Eq. (3-5), those in the semiprofessional and sales occupations earn about $800 a year less than those in the professional occupations.[32] This difference amounts to about $7^1/_2$ percent of the professional income. Ability is statistically significant and numerically important, with those in the tenth ability class earning nearly $1,900 more than those in the bottom four tenths. Indeed, the ability coefficients in Eq. (3-5) are larger than those in Eq. (3-1). Thus it appears that within the professional occupations, income is determined by the suboccupation that a person enters and by his IQ score, but not by his education.

In Eq. (3-6), the only significant variables are the constant term and the dummy variable for skilled workers. Thus, for unskilled, clerical, farm, skilled, and service workers, neither higher education nor mental ability significantly determines income. It certainly is not surprising to find that higher education is not important here, since such education is not generally focused toward the skills used in these areas. Similar arguments probably also apply to mental ability.

In summary, it is not astonishing to find that mental ability is only important in certain types of occupations, but it is surprising to find that education is not all important in either of the

[30] The high-ability–high-education interaction term is also insignificant when included in the equations.

[31] As noted earlier, these results, which differ from those in census data, may represent the particular set of ages involved.

[32] We tried separate dummy variables for occupations 2 and 4, but the coefficients were nearly identical.

two occupational groups. There is, however, a possible technical explanation for this result, which we consider now. If, in fact, there is little or no variation in education within occupations, then the finding that income is not affected by education is of little importance, since it follows necessarily from the data. But the fraction of people in each occupation at each IQ and education level—given in Appendix B, Tables B-9 and B-12—indicates a wide range of education and ability within the various occupations. In Table B-9 it can be seen that, at all ability levels, substantial numbers of persons with either high school or some college education are working in all the occupational groups. Those with one or more degrees are mainly employed in occupations 1, 2, and 4 (and especially 1). For any level of education, the occupational distribution of people is almost the same at various ability levels.[33]

We consider now the difference between these results and those from similar regressions using census data. In general, the latter yield significant education coefficients (particularly for college graduates versus noncollege graduates) after standardizing for such influences as age, race, weeks worked in the year, and so on. Of course, there is no standardization for ability in the equations from census data. One suspects, therefore, that the education variables in equations such as Eq. (3-5) and Eq. (3-6) might be significant if the ability variable were omitted. We have estimated the following equations to test this hypothesis; the first is for occupations 1, 2, and 4, and the second is for occupations 3, 5, 6, 7, and 8.

$$\underset{(6.2)}{Y = 4.637} + \underset{(2.7)}{2.831}(E_G + E_{GM}) + \underset{(.5)}{.249}D_{2+4} \qquad \bar{R}^2 = .96 \quad (3\text{-}7)$$

[33] Another question that arises in connection with these results concerns the occupational breakdown. Admittedly, the nine occupations are rather broad and, consequently, heterogeneous. Even if a finer classification were used, however, it seems unlikely that the results would differ, although it is conceivable. For example, if the professionals were divided into four or five groups, then it would be possible to have education and income positively related in one group and to have, in the others, either education and income negatively related (which is unlikely) or virtually no variation in education *within* them, but variation *between* them, with income higher in the occupation with lower average education. It seems unlikely, however, that these relationships would be such as to cancel out when the data were combined. In addition, an obvious difficulty in defining occupations less broadly is that the variation in education within any ability class is likely to be smaller.

$$Y = \underset{(12.4)}{3.607} + \underset{(1.7)}{1.301}(E_G + E_{GM}) + \underset{(.9)}{.296}D_{3+5} + \underset{(3.3)}{1.065}D_6$$

$$\bar{R}^2 = .92 \quad (3\text{-}8)$$

These equations differ from Eq. (3-5) and Eq. (3-6) in that the ability variables are omitted and the education coefficients have been collapsed to represent college graduates and all others (with at least a high school education).[34] Eq. (3-7), for the professional, semiprofessional, and sales occupations, is in accord with general census results in which college graduates are found to have a significantly higher income. The Wolfle-Smith data suggest, however, that this higher income, once entrance into any of these occupations is achieved, is entirely due to ability. For the remaining occupations, on the other hand, education appears to be more important than ability in determining income, although the education variable is not significant even when ability is omitted.[35]

One final comment is in order. We showed earlier that the education coefficient we estimate is net of the value of work experience forgone while obtaining education. Making allowance for time spent in the military, the people in the sample with a college degree had a maximum of seven years' experience. At an older age, greater impacts of education might be found. Indeed, this is the case with the NBER-TH sample as analyzed in Chapter 9.

MINNESOTA FEMALES

We have also used the Wolfle-Smith data on females ranked by ACE decile. The following equation is analogous to Eq. (3-1) for males.

$$Y = \underset{(17.8)}{2.138} + \underset{(.7)}{.108}E_V + \underset{(1.0)}{.175}E_1 + \underset{(1.0)}{.158}E_3 + \underset{(4.9)}{.777}E_G + \underset{(7.0)}{2.184}E_{GM}$$

$$+ \underset{(1.2)}{.124}A_{5-9} - \underset{(1.2)}{.171}A_{10} \qquad \bar{R}^2 = .83 \quad (3\text{-}9)$$

[34] With the ability variables included in Eq. (3-7), the coefficient of $(E_G + E_{GM})$ has a t value of .2, while inclusion in Eq. (3-8) yields a t value on the education variable of 1.5.

[35] Morgan and David (1963) find that education is significant when dummy variables for five occupations are included. Their ability measure, however, does not seem to be very appropriate.

The income measure here is average earnings of all females in the sample regardless of whether they were working. Approximately 70 percent of the females in the sample listed their occupation as housewife, although some also worked part time.

There are several interesting conclusions to be drawn from Eq. (3-9). First, the ability variables are not significant, nor were they significant when included in continuous form, as 10 decile tenths, or as part of a high-ability–high-education interaction term. Second, only the education variables representing one college degree and post-bachelor's education are significant—and these are highly significant. On the other hand, the magnitudes differ considerably from those for males—particularly the coefficient of E_G. One college degree adds only \$777 to the average female's income, but \$2,100 to the average male's income; the corresponding values for more than one degree are \$2,184 and \$2,658. The relatively better performance for women with more than one degree and the lack of significance of ability may be because teachers, who are not necessarily drawn from the high IQ groups but who plan to be in the labor force, obtain master's degrees for certification and higher pay. Thus, people who are not in the labor force—and who did not plan to be—would have less education.

FINAL COMMENT Unfortunately, without the original observations, we cannot proceed much further with this data set in exploring questions that need to be answered. Much additional information is available in another body of data, which regrettably contains people from the top half of the IQ distribution only. Most of this book will be concerned with this latter data set.