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: The BVER-TH Sample Regression Results

Chapter Author: Paul J. Taubman, Terence Wales

Chapter URL: http://www.nber.org/chapters/c3651

Chapter pages in book: (p. 69 - 112)

5. The NBER-TH Sample Regression Results

In this chapter and its appendixes (Appendixes H and I), we will examine the relationship of earnings to education, ability, and several other variables, using the NBER-TH sample. An extensive discussion of how this sample was drawn, the accuracy of the data, the extent of response, and the derivation of the four major measures of ability can be found in Chapter 4. The reader may especially wish to refer to Chapter 4 for a discussion of alternative procedures and methods that could have been used in the derivation (and interpretation) of the ability measures and, to a lesser extent, elsewhere. To aid recollection, we present a very brief summary of the important information about the sample.

In 1943, some 75,000 men who had volunteered to undertake pilot, bombardier, or navigator training in the Army Air Corps, and who had scored in the top half of the Armed Forces Qualification Test, were given a battery of some 17 tests. In 1955, Thorndike and Hagen (1959) surveyed a random sample of about 17,000 men and obtained information from about 10,000 civilians and 2,000 who were still in the military. In 1969 the NBER surveyed the 10,000 civilians. About 7,500 people were contacted, and 5,200 responded. Some 800 questionnaires were received too late to be included in this study. While the 4,400 people we used in the study are brighter and more educated than the 10,000 people who responded in 1955, the respondents within each ability-education cell appear to be a random drawing from the same population.[1]

[1] The other people who responded in 1955 are not included in any of the analyses because of the lack of a crucial piece of education data, as explained in Chapter 4.

The major purpose of this chapter is to develop the information necessary to answer the following questions at various points in the age-income profile: What is the effect of education on income once other determinants of income are held constant? What types of ability are important determinants of income? Are there significantly different effects of education at different ability levels? What is the bias on the education coefficients from omitting ability and other variables? Are there some other unmeasured variables whose effects on earnings persist over time? In Chapter 4 we argued that a stratified sample drawn from only a portion of the population will yield results that can be generalized to a random sample of the truncated population (except for the bias problem). The results need not be capable of generalization to the whole population since the functional form may be different in the portion of the population not sampled.

SUMMARY OF RESULTS

Since this is a particularly long and, at times, technical chapter, we present first a summary of what we have found to be the major determinants of wage, salary, and unincorporated-business income. The variables that are important in some or all years in our equations for males are education, mathematical ability, being a high school or elementary school teacher, personal biography, health, marital status, father's education, and age. In addition, we find that the residual in each year consists in part of personal effects that persist over long periods of time. When we test the variables that represent coordination, verbal IQ, and spatial-perception abilities, however, there is little evidence that these variables add to earnings in this sample, which was drawn from the top half of the ability distribution.[2] Except, perhaps, for the types of abilities that were not significant, the list contains few surprises, since other investigators have often found the same types of variables (as well as a few we have not tried) to be important. There are many surprises, however, in the magnitude of the effects. We consider first the earnings differentials due to education.

[2]Other variables that explain income, but are not presented in the results because they are inappropriate for calculating the returns to education, include occupational dummy variables such as "business proprietor." The teaching dummy variable is included because we think it represents the nonpecuniary returns prevalent in this occupation.

The reader is once again reminded that there is some dispute about the interpretation of the factors we call mathematical ability and IQ. See Chapter 4.

The extra earnings from education are given in Table 5-1 for persons with the characteristics of the average high school graduate. In 1955, when the average age in the sample was 33, the extra earnings from education were 10 to 15 percent of those obtainable with a high school degree at all educational levels —except for the highest graduate degrees or for those whose occupation was elementary school or high school teacher. Although not shown in the table, the differentials for B.A. and B.S. holders are the same. The income differential was 70 percent for M.D.'s, 2 percent for Ph.D.'s, and 20 percent for LL.B.'s. (The Ph.D. and M.D. estimates are based on fewer than 100 and 50 people, respectively.)

In 1969, those with some college earned about 17 percent more than high school graduates, while those with an undergraduate degree, some graduate work, and a master's degree received 25 to 30 percent more.[3] Ph.D.'s, LL.B.'s, and M.D.'s earned about 25, 85, and 105 percent more, respectively, than high school graduates of the same ability level. These increases would also hold for hourly wage rates, since average hours worked are the same at all educational levels except for the combination of Ph.D., LL.B., and M.D., in which hours are 8 percent greater than those of high school graduates. In 1969, the income differentials were higher at all education levels, with the greatest percentage increase accruing to the most highly educated. Since most of the post-high school education was obtained after

[3] The questionnaire was somewhat vague as to whether "most recent earnings" referred to 1969 or 1968, but as explained at a later point, there is evidence that 1969 annual earnings were given.

TABLE 5-1 ***Increases in earnings for the average high school graduate, by education level, 1955 and 1969 (as a percentage of high school earnings)***

Education	*1955*	*1969*
Some college	11	17
*Undergraduate degree**	12	31
*Some graduate work**	15	26
*Master's**	10	32
Ph.D.	2	27
M.D.	72	106
LL.B.	19	84

* For those who do not teach.

SOURCES: Tables 5-4 and 5-8.

the tests were taken in 1943, and since evidence presented below indicates that pretest education differences had little effect on the test score, the education coefficients measure both the cognitive and affective benefits of higher education.

In 1969, we can identify the people who own their own business.[4] The earnings reported for these people probably include some return to financial investment and some compensation for extra risk. If we include a zero-one dummy variable to remove such effects, we find that the 1969 income differentials for non-business owners are 17 percent and 39 percent for those with some college and those with an undergraduate degree, respectively. A bachelor's-degree holder's performance is improved because the business owners, who have high earnings, on the average, are less heavily concentrated in the bachelor's than in the high school group.

There is one other important result in Table 5-1. In all the years studied, the returns to education display an erratic behavior in that each additional year or plateau does not add as much income as the previous one. Thus, in 1955, there is essentially no increase in income beyond the some-college level until the huge jump for M.D.'s. The increases for M.D.'s could be justified as a monopoly return, but the lack of additional income for undergraduates and those in other categories requires some explanation. This is discussed in more detail in Chapter 6, but a brief description follows.

According to Mincer (1970), people pay for general on-the-job training and higher future earnings by receiving lower wages while being trained. If, on the average, the more educated invest more in such training, differences in average starting salary by education level will be less than lifetime differences (and may even be in reverse order). The earnings profiles of those who invest more will increase faster with experience. In no more than $1/r$ years (where r is the rate of return on the on-the-job training), the returns from such investment will outweigh the costs of any current investments and people's current earnings will be above the "no investment" lifetime earnings flow. After this, the earnings profiles by education level will continue to diverge if investment continues.

Mincer's formulation also emphasizes that it is work ex-

[4] This category does not include self-employed professionals.

perience and not age that determines earnings. We indicated in Chapter 3 that when age is held constant, the education coefficients measure the net effect of education, that is, the increase in productivity minus the loss in work experience. The combination of these two ideas and information in Table 5-1 suggests that in 1955 those with some college and those with a bachelor's degree were near the point on their profiles where observed earnings equaled their "no investment" earnings, while high school graduates, who had more experience, were above their constant earnings level. But if those who are more educated invest more, then Mincer's model would imply that earnings of the more educated should in 1969 be even further from their constant flow than earnings of high school graduates. That is, his model implies that the 1955 results understate, and the 1969 results overstate, the effect of education on lifetime earnings.

While Mincer's theory can help explain some of our results, it is necessary to note that there exists little direct evidence on which either to accept or to refute the wage-adjustment process he predicts. Thus, it is interesting to find in this sample that, beginning in the late 1940s, the starting salary of high school graduates (in a given year) is nearly the same as that of college graduates, that graduate students receive less than college graduates, and, finally, that those with some college may earn more than those with a college degree. (These starting-salaries figures are subject to recall error since they were recorded in 1969 and are less reliable than the 1955 or 1969 contemporaneous estimates).

On the other hand, from 1955 to 1969 the growth rates in income of those with a college degree, some graduate work, and a master's degree were essentially the same. This would suggest that the difference in investment in on-the-job training was not very large at these levels. This would also reinforce the idea that the returns to education do not follow a smooth pattern as education increases.

Further evidence of the erratic nature of the returns to education is evident in the 1969 equations, in which we replace the some-college dummy variable with three separate ones for those who had one, two, and three years of college. The coefficient for completing one year of college is essentially equal to that of the some-college variable discussed above, while the coefficients for completing the second and third years of college indicate no further increase in earnings.

There is one further peculiarity to report on the returns to education. When we included a variable to represent vocational education, we obtained significant but negative coefficients. Our explanation is that those who took such courses had no intention, or chance, or perhaps lacked sufficient drive, to enter the higher-paying managerial and professional occupations.[5] Under this interpretation, the variable does not represent causation, and we exclude it from the analysis.

The returns to education given above are calculated with ability and other characteristics held constant. To derive measures of different abilities, the 1943 test scores were subjected to factor analysis, as described in detail in Chapter 4. We used the factors representing mathematical ability, coordination, verbal IQ, and spatial perception. (As explained in Chapter 4, Professor Thorndike believes that the first factor is a closer approximation to a standard IQ test than the third factor, which he contends to be more of a heterogeneous mix.) Of these four factors, only coordination is a physical skill; the others are mental. To allow for nonlinear effects of ability and changes in scores due to maturation, we divided each factor into fifths (based on the whole 1955 sample)—which for the verbal IQ factor and to a lesser degree for the mathematical factor are equivalent to tenths, since only those in the top half of the mental-ability distribution were allowed into the test program.

We find that, of these ability measures, only mathematical ability was a significant determinant of income.[6] The extra incomes above the bottom fifth are given in Table 5-2 for 1955 and 1969. The second through fourth fifths may be subject to mild diminishing returns, but the top fifth yields substantially more than the fourth.[7] The difference between the top and bottom fifth of $1,000 per year in 1955 is more important than all the education variables (except M.D.), and in 1969 the $3,350 per

[5] If the variable represents drive, it could be argued that those high school graduates who did not have vocational training are more like college attenders. Since the average income of high school graduates without vocational training is higher, the returns to education would be less.

[6] The second fifth was not significant, but the other three were.

[7] For 1955, a linear, continuous ability measure would be almost as appropriate as these dummy variables in the regression analysis.

TABLE 5-2
Extra income per month for those above the bottom fifth in mathematical ability, 1955 and 1969 (in dollars)

Year	*Mathematical-ability fifth* Q_2	Q_3	Q_4	Q_5
1955	23	33	50	84
1969	69	107	143	279

NOTE: The second fifth was not significant, but the other three were. Rank Q_5 is the highest.
SOURCES: Equation 5, Tables 5-3 and 5-7.

year difference is nearly as important as an undergraduate degree. These results are of considerable interest for those interested in the relative importance of ability and education in determining the distribution of income.[8] Since our sample was drawn from only the top half of the ability distribution, it is almost certain that for those with at least a high school education, ability is a more important determinant of the range of distribution than higher education. (The range is a more interesting measure since the sample is stratified.)

An individual's score on any test at time *t* depends upon inherited intelligence, the quantity and quality of schooling obtained prior to *t*, and the individual's home environment. Evidence presented below indicates that differences in the quantity and quality of pretest education had little if any effect on scores on the particular test used. Neither environment, which presumably has several dimensions, nor school quality has been held constant in our analysis. However, since both of these variables are related to family background, which is held constant, the ability coefficients primarily reflect differences in inherited intelligence.

In the above calculations, the ability effects were constrained to be the same at all levels of education studied. Two related questions of some importance are: If this constraint is removed, do the effects of ability persist at low levels of education? What interactions exist between ability and education? We have examined these questions in some detail. For 1955, we find practically no evidence that the effect of education was different at the various ability levels, while even for high school graduates there was a return to ability. For 1969, however, we found some evidence that those who were in the fourth (and, to some ex-

[8] See Mincer (1970) for an excellent summary of the issues.

tent, fifth) fifth and who had graduate training received extra income. We still find that ability was an important determinant of income even for high school graduates. Finally, in our study of initial salaries, we find that ability had no effect on income except for those with graduate education. This discussion and the entries in Table 5-2 suggest that ability has little effect on earnings initially, but that over time this effect grows and perhaps grows more rapidly for those with graduate training and high ability.[9]

A basic criticism that has been made of other studies of the returns to education is that the returns obtained are biased upward because relevant abilities have not been held constant. While we have (some) measures of ability, it is possible for us to estimate the equations omitting ability and then to calculate the percentage change in the education coefficients. Because in some studies the effects of education have been estimated holding constant sociodemographic background information, it is also useful to calculate the bias on the same basis. However, as explained in more detail below, one of our important variables is a mixture of background and ability; thus we can only calculate the upper and lower bound of the bias from omitting ability. For simplicity, in this summary we use the average of these bounds.

We have calculated the bias from assuming both that each factor was the only type of ability to be measured and that all abilities should be included in our equations. In either instance we find that only the omission of mathematical ability leads to bias of any magnitude. In 1955 the bias on the education coefficients from omitting mathematical ability is about 25 percent, varying from a low of 15 percent for the some-college category to a high of 31 percent for the master's degree category.[10] In 1969 the biases are somewhat smaller, averaging about 15 percent and ranging from 10 to 19 percent. The decline in the bias over time occurs because the coefficients on ability did not grow as

[9] The greater importance of ability over time was also found in a study of American Telephone and Telegraph employees. See Weisbrod and Karpoff (1968).

[10] The 15 percent bias for the some-college category is higher than in other studies. This may be due to our use of mathematical ability rather than IQ.

rapidly between 1955 and 1969 as those on education.[11] In some studies, rates of return have been calculated using differences in average incomes between education groups for various age groups. In this sample, such a procedure would overstate the returns to higher education by 35 and 20 percent in 1955 and 1969, respectively. However, the biases calculated in this sample may not apply to the population, because the relationships between education and either ability or background may be different in this special sample than in the population as a whole.

We have also examined the role of various sociodemographic background factors. Several such variables are significant and important determinants of income. For example, the difference between excellent and poor health in 1969 is worth $7,000 a year, and the nearly 100 people who are single earn $2,800 a year less than others.[12] Those whose fathers' education was at least the ninth grade earn about $1,200 a year more in 1969 and $300 more in 1955 than those whose fathers did not enter high school. The other background information is contained in a biography variable that includes information on hobbies, pre-World War II family income, education prior to 1946, and mathematical ability. This measure is divided into a set of dummy variables for the fifths. We consistently found the fourth and fifth and either the second or third fifths of the biography variable to be significant and to have about the same effect as the mathematical ability.

In our analysis we also included an age variable. Since this variable only spans eight to nine years in each of our cross sections, we did not look for nonlinearities in age effects or for interactions with education and other variables.[13] We found the age variable to be quite significant and large numerically in

[11] The bias is also determined by δ in the equation $A = \alpha + \delta\, ED$, but because this equation would involve the sample in 1955 and 1969, and because ED changed only slightly, δ would be virtually unchanged in the two years.

[12] The health variable used was linear with 1 for "excellent" through 4 for "poor" answers. When separate dummies are used for the various categories, the coefficients confirm that the effects are linear.

[13] Some interactions can be estimated from the coefficients of the different cross sections. See Chapter 6.

1955. The difference in income between those 30 and 39 years old was about the same as the difference in income between those with a Ph.D. and those with a high school diploma. In 1969, however, the age effect was negative and insignificant. Thus, we have evidence of the familiar sharply rising age-income profile reaching a peak after the age of 40.

Our method for allowing for age effects in any cross section constrains the profiles for all educational levels to be parallel. However, the separate results for 1955 and 1969 indicate that the high school profile is less steep than the others, since the effects of education are greater in 1969. Analysis we conducted separately for each education group would indicate that there were no significant differences in age effects between the different education groups except for the graduate group in 1955, which displayed no discernible age effects.

The results discussed above were obtained from analysis of separate cross sections. It is possible to develop a combined measure of motivation, drive, personality, and whatever other characteristics persist over long periods of time by using the residuals generated in one cross section in the equations in another cross section. In principle, better results can be obtained by grouping the data. Unfortunately, because of multicollinearity, we were not able to obtain reasonable results by grouping. We used the residuals, denoted as Q, of 1955 in 1969, and vice versa. In each year, Q raised the $\bar{R}^2$ from about .1 to .33 and reduced the standard error of estimate by 15 percent.[14] However, since least-square residuals are uncorrelated with independent variables, the inclusion of Q did not alter the values of the other coefficients. The results using the Q variable are both encouraging and discouraging. It is encouraging to find that we can substantially improve the efficiency of our es-

[14] Mincer (1970) has argued that earnings data reflect not only investments in school but also investments in on-the-job training. According to his analysis, the latter investment takes the form of a reduction in earnings in the early years of working, while its payoff is in terms of a constant flow of extra earnings. Thus, those people investing more would have negative residuals in the early working years and positive residuals later. After the passage of the number of years no more than the reciprocal of the rate of return on such investment (that is, about a decade) Mincer suggests that the returns on previous investments would offset new investments. Thus, there should be no correlation between residuals calculated at this point and ones later (or earlier). Since our income differentials in 1955 are small, very little of the correlation can be attributed to Mincer's type of investment.

timates by incorporating other information on the individual. It is also encouraging that certain effects that we believe reflect personality, and so forth, persist over time.[15] The discouraging part is that the proportion of the residual that persists over time is very small. The remainder of the residual represents random events, such as luck, and/or changes in underlying characteristics.

We next turn to a detailed examination of the results. We begin with a brief discussion of the variables.

DEFINITION OF THE VARIABLES

We have divided our factors into fifths as described in Chapter 4.[16] Since the individuals in our sample are within the top half of the IQ scale, these fifths approximately correspond to tenths for the population as a whole. The dummy-variable method was used to allow for nonlinearities in the effect of each ability measure.

As described earlier, the first factor, we believe, reflects a mathematical, or numerical, aptitude. The second factor represents coordinating abilities; the third factor, a verbal IQ; and the fourth factor, spatial aptitude. In the regression analysis we refer to these factors by number and by the general type of ability that they reflect.

In our analysis we have used a variable that is intended to reflect background characteristics of individuals. The variable, which was constructed by the Army Air Corps and retained by Thorndike, is obtained from the individuals' scores on navigator and pilot biography keys.[17] These keys—described in Thorndike and Hagen (1959, pp. 38–49)—were calculated from background-information items using weights that predicted success in navigator or pilot training schools.[18] In our analysis we refer to this as the biography variable. Because it reflects

[15] However, it may be possible to interpret our results as evidence of a Markov scheme such as estimated in Solow (1951), Cutright (1969), and elsewhere.

[16] In order to have a reasonable sample size within each ability cell, finer divisions were not attempted.

[17] We weighted the navigator scores twice as heavily as the pilot scores to form the biography variable referred to in the analysis. These weights were suggested by preliminary analysis of the data.

[18] The background-information categories were general family and personal background, major subject in college (pre-1943), school subjects studied and done well, sports participated in and done well, activities done a number of times, hobbies and free-time activities, work experience (pre-1943), and reason for choosing cadet training.

ability, education, and socioeconomic background characteristics of the individuals, it is not a pure background measure. As we did for the ability variables, we have calculated the sample fifths for this variable to allow for nonlinearities.

The education categories used in the regressions are high school, some college (one to three years), undergraduate degree, some graduate work, master's degree, and three-year graduate degrees.[19] These are represented in the form of dummy variables, with a value of 1 if the individual is in a particular category and zero otherwise. The high school dummy variable is omitted in the regressions. In addition, since at one time we were not sure if there would be enough observations to study M.D.'s and LL.B.'s separately, we included them with Ph.D. and then created one dummy variable to represent M.D.'s—defined as 1 if the individual is in the medical profession and zero otherwise—and another for LL.B.'s.[20] The dummy variable for elementary school and high school teachers (as of 1955) is defined analogously.

Other variables included in some of the regressions are age (in years); health (measured in 1969), with a value of 4 for those in poor health, 3 for those in fair health, 2 for those in good health, and 1 for those in excellent health; marital status, as a dummy with a value of 1 for single people and zero otherwise; and father's education, in the form of two dummies reflecting any high school and any college education. (When there was no answer for a variable, we either inserted the modal response, as with age, or eliminated the observations, as with father's education.)

The dependent variable in the regressions is earnings, which equal wage and salary income plus unincorporated-business income. This is more appropriate than total income for measuring the returns to education.

RESULTS FOR 1955

The dummy variables for education and ability allow for all nonlinear effects within each variable, but no interactions between ability and education. We have tested for interactions be-

[19] We experimented with separate variables for B.A. and for other undergraduate degrees and for one, two, and three years of college. The results are discussed in the text. The three-year graduate degrees are Ph.D. or equivalent, M.D., and LL.B.

[20] Thus, both the LL.B. and M.D. coefficients must be combined with the Ph.D. coefficients to find the full effect of these educational categories.

tween education and ability by the following method. We estimated separate equations of the form

$$Y_i = a_i + \Sigma b_{ij} ED_j + \Sigma d_{ij} Z_j$$

where *ED* represents the various education variables and *Z* all the other variables for each of the five ability levels (*i*) of factor 1.[21]

If there were no interactions between ability and education, the estimates of b_{ij} should be the same at each of the *i* ability levels. We can test for the equality of the b_{ij} and d_{ij} by using analysis of covariance. The reader should note that the interactions we are testing for are second-order effects that are equivalent to $\delta^2 Y/\delta ED \delta A$.

On the basis of the analysis of covariance, we cannot conclude that the coefficients in the various ability equations in 1955 are significantly different at the 5 percent level.[22] However, it appeared that in 1955 the only possible significant differences in the effect of the various abilities on income occurred in the comparison of graduate education levels with all other levels. We therefore decided to estimate equations using all the data but with separate interaction terms for graduate students at each ability level. Such equations are given as numbers 14 and 15 in Table 5-3.

All the interaction terms are insignificant with only the Q_4 interaction being positive. Thus, there is very little evidence of an education-ability interaction in 1955. Consequently, for this year we can omit the interaction variables while studying the returns to education and the bias from omitting ability.

In equation 5, Table 5-3, which can be used to study the education-earnings differentials, all the education coefficients are significant, but their magnitudes are surprisingly small. In Table 5-4 we summarize the effects of education on income after holding constant ability, background factors, and age. In this table we have calculated the percent by which average earnings exceed (1) those of high school graduates ($6,000 per year) and (2) the average earnings in the immediately preceding educa-

[21] As explained below, the first factor is the only significant one.

[22] We made both tests in pairs of the equations by ability level and a joint test for all five ability levels. All the *F* ratios were less than 1.5.

TABLE 5-3 Regressions for salary, 1955 (in dollars per month)

		Constant	Some college	Undergraduate degree	Some graduate work	Master's	Ph.D. and LL.B.*	M.D.	Teacher
(1)	Y_{55}	$499.6	$75.4	$ 88.0	$ 83.3	$ 28.4	$153.3		
		(5.4)	(5.9)	(7.2)	(3.9)	(1.5)	(7.1)		
(2)	Y_{55}	243.9	77.1	96.8	90.3	33.7	161.2		
		(4.2)	(6.1)	(7.8)	(4.2)	(1.8)	(7.4)		
(3)	Y_{55}	229.6	80.8	103.0	115.4	96.8	109.5	$292.7	$−181.1
		(3.6)	(6.2)	(8.2)	(5.2)	(4.7)	(4.7)	(6.6)	(8.0)
(4)	Y_{55}	260.1	58.6	72.6	90.9	68.1	80.8	297.3	−169.1
		(4.1)	(4.5)	(5.7)	(4.1)	(3.3)	(3.5)	(6.8)	(7.6)
(5)	Y_{55}	229.3	54.0	57.8	75.0	51.2	61.2	299.8	−162.4
		(3.6)	(4.1)	(4.4)	(3.4)	(2.5)	(2.6)	(6.9)	(7.3)
(6)	Y_{55}	258.4	59.1	73.2	91.1	69.8	82.9	296.5	−168.2
		(4.0)	(4.5)	(5.7)	(4.1)	(3.4)	(3.6)	(6.8)	(7.5)
(7)	Y_{55}	251.8	57.9	67.9	85.6	63.0	76.3	299.9	−165.7
		(3.9)	(4.4)	(5.3)	(3.9)	(3.1)	(3.3)	(6.9)	(7.4)
(8)	Y_{55}	227.9	57.5	69.4	86.2	64.6	78.1	299.0	−164.0
		(3.5)	(4.4)	(5.4)	(3.9)	(3.2)	(3.4)	(6.9)	(7.4)
(9)	Y_{55}	247.6	70.3	88.5	100.9	81.0	94.9	295.9	−173.3
		(3.9)	(5.4)	(7.0)	(4.6)	(4.0)	(4.1)	(6.8)	(7.7)
(10)	Y_{55}	195.1	70.5	82.3	91.8	77.0	86.1	284.0	−173.4
		(3.2)	(5.7)	(6.6)	(4.3)	(3.8)	(3.7)	(6.7)	(7.9)
(11)	Y_{55}	220.1	77.5	102.2	112.2	100.7	114.3	281.2	−180.1
		(3.6)	(6.2)	(8.4)	(5.2)	(5.0)	(5.0)	(6.6)	(8.2)
(12)	Y_{55}	216.3	74.7	92.8	102.1	89.0	101.9	285.4	−176.3
			(6.0)	(7.6)	(4.7)	(4.4)	(4.5)	(6.7)	(8.0)
(13)	Y_{55}	188.3	73.5	94.6	104.2	92.0	105.4	283.2	−174.7
			(5.9)	(7.8)	(4.9)	(4.6)	(4.6)	(6.7)	(7.9)
(14)	Y_{55}	230.4	53.7	56.8	81.1	56.9	69.1	299.0	−162.2
		(3.6)	(4.1)	(4.4)	(3.5)	(2.6)	(2.7)	(6.9)	(7.3)
(15)	Y_{55}	229.5	53.8	57.5	84.2	60.4	73.3	299.0	−160.5
		(3.6)	(4.1)	(4.4)	(3.5)	(2.5)	(2.6)	(6.9)	(7.1)

	Ability			
Age	Q_2	Q_3	Q_4	Q_5
$7.7				
(4.2)				
8.1				
(4.3)				
7.8				
(4.1)				
7.8	$ 23.4	$33.4	$50.1	$83.7
(4.1)	(1.5)	(2.2)	(3.4)	(5.7)
7.6	− 2.0	4.0	17.7	20.6
(4.2)	(0.1)	(0.3)	(1.2)	(1.4)
7.7	5.4	8.1	12.6	40.9
(4.1)	(0.4)	(0.5)	(0.9)	(2.8)
8.2	9.4	12.2	25.2	55.7
(4.3)	(0.6)	(0.8)	(1.7)	(3.8)
9.0				
(4.8)				
8.0	33.9	40.4	58.7	99.3
(4.5)	(2.3)	(2.8)	(4.1)	(6.9)
7.7	11.5	24.9	36.3	40.4
(4.3)	(0.8)	(1.8)	(2.6)	(2.9)
8.0	5.7	18.7	25.0	59.4
(4.5)	(.04)	(1.3)	(1.8)	(4.2)
8.5	18.4	23.8	38.9	76.0
(4.8)	(1.3)	(1.7)	(2.8)	(5.4)
7.8	23.3	33.1	45.8	95.9
(4.3)	(1.5)	(2.2)	(2.9)	(6.0)
7.8	25.4	33.9	46.1	96.0
(4.1)	(1.6)	(2.1)	(2.9)	(6.0)

TABLE 5-3 (continued)

	Interactions of graduate education with:						
	Q_2	Q_3	Q_4	Q_5	*Health*	*Single*	*Father attended high school*
(1)							
(2)							
(3)							
(4)					$−34.6	$−117.5	$27.7
					(4.6)	(3.5)	(2.7)
(5)					−33.1	−121.9	26.0
					(4.4)	(3.7)	(2.5)
(6)					−33.9	−118.9	27.5
					(4.5)	(3.5)	(2.7)
(7)					−33.8	−120.4	27.4
					(4.4)	(3.6)	(2.7)
(8)					−33.0	−119.8	25.5
					(4.3)	(3.6)	(2.5)
(9)					−37.5	−126.6	36.7
					(4.9)	(3.7)	(3.6)
(10)							
(11)							
(12)							
(13)							
(14)			$15.3	−40.5	−33.7	−125.0	26.3
			(0.6)	(1.8)	(4.4)	(3.7)	(2.6)
(15)	$−12.8	$−5.2	12.4	−43.1	−33.9	−125.2	26.2
	(0.4)	(0.2)	(0.5)	(1.8)	(4.5)	(3.7)	(2.5)

* M.D.'s are also included.

NOTE: Figures in parentheses are *t* statistics.

Father attended college	*Biography* Q₂	Q₃	Q₄	Q₅	*Ability factor*	$\bar{R}^2$/S.E.
						.021
						274
						.026
						273
						.058
						269
$26.3	$ 2.1	$31.9	$66.1	$90.2		.089
(2.1)	(0.1)	(2.2)	(4.5)	(6.2)		265
21.4	0.7	30.0	63.1	81.1	1	.098
(1.7)	(0.0)	(2.0)	(4.3)	(5.6)		264
25.9	0.7	30.2	64.6	86.9	2	.090
(2.0)	(0.0)	(2.0)	(4.4)	(5.9)		265
23.2	−0.7	28.5	62.6	83.0	3	.091
(1.8)	(0.0)	(1.9)	(4.3)	(5.6)		265
24.2	−0.6	28.0	61.0	81.6	4	.094
(1.9)	(0.0)	(1.9)	(4.1)	(5.6)		265
44.3						.070
(3.5)						267
					1	.066
						268
					2	.056
						269
					3	.059
						269
					4	.062
						268
21.9	1.1	30.1	63.2	81.0	1	.099
(1.7)	(0.1)	(2.0)	(4.3)	(5.5)		264
21.9	0.9	30.1	63.0	81.0	1	.099
(1.7)	(0.1)	(2.0)	(4.3)	(5.5)		264

TABLE 5-4 Percentage increases in earnings, by education level, 1955

Education level	Percentage by which earnings in each education class exceed: Earnings of average high school graduate	Earnings of average member of preceding education level
Some college	11	11
Undergraduate degree (not teacher)	12	1
Some graduate work	15	3
Master's (not teacher)	10	−4
Ph.D.	2	−8
M.D.	72	
LL.B.	19	
Undergraduate degree (teacher)	−21	
Master's (teacher)	−22	

NOTE: Average age is 33.

SOURCE: Earnings of average person in each education class from equation 3, Table 5-3. Absolute increases from equation 5, Table 5-3.

tion class.[23] It should be noted that we have separate estimates for Ph.D.'s and LL.B.'s in this table but not in equation 5, Table 5-3. The equation with the separate categories yields approximately the same coefficients on the other variables as does equation 5, but to save space, it is not shown.

Table 5-4 indicates that those who continue their education (except for LL.B.'s, M.D.'s, and teachers) receive 10 to 15 percent more income at all education levels than the average high school graduate. Thus, in the second column (which compares income with the preceding educational level), there are sharply diminishing returns to education after the first three years of college, except, of course, for medical training. M.D.'s and LL.B.'s earn about 70 percent and 20 percent more than high school graduates, respectively, while teachers earn about 20 percent less.[24] However, the reader is reminded that these results are valid only for individuals around 33 years of age in 1955.

[23] We use equation 3 in Table 5-3 to establish the earnings level of the average high school graduate for (1) and the average person in the other education levels for (2).

[24] We consider the lower salary of teachers to be offset by a nonpecuniary reward. Thus, in our discussion we will use the salary of nonteachers.

In equation 5, Table 5-3, our ability measure is factor 1, which we interpret as a mathematical skill.[25] In the equation, the top three fifths are significant and monotonically increasing.[26] These ability effects are important since the earnings difference (per month) between the lowest and highest fifth exceeds the largest education coefficient of $75 for some graduates. We compare the results using this factor with the other ability factors later.

The biography variable is also an important determinant of income—the third through fifth fifths are again significant, with the coefficients being very similar in magnitude to those of the mathematical factor. As mentioned above, this variable reflects a mixture of mathematical ability, education, and background characteristics. Since we are holding education and ability (and some background) constant in the regression equations, the coefficients on the biography variable reflect such other background factors as parents' socioeconomic position and home environment.[27]

The other variables in equation 5, Table 5-3, are all significant except for the variable representing father's attendance at college. We consider first the age variable. Because the ages of those in the sample range from about 30 to 39 in 1955, we would expect a large impact of age on income. Thus, the highly significant coefficient, which indicates an annual earnings increase of about $95, or about $1^1/_2$ percent, is not surprising.

While testing for ability-education interactions, we automatically obtained estimates of the age effect at each education level. The coefficients of age with t values in parentheses are 12.4 (4), 6.4 (1.8), 10 (2.5), and −2 (.4) for the high school, some-college, B.A., and graduate-education categories respectively. Differences in the effect of age on income are not significant except perhaps at the graduate level.[28]

The only measure of health available is the respondents' own

[25] See, however, the discussion of the factors in Chapter 4.

[26] See Appendix I for a discussion of the appropriateness of the t tests given in the tables.

[27] When the variables for father's education are omitted, the coefficients on the biography variables are larger and more significant.

[28] The graduate result occurs because lawyers, who graduate sooner and who are younger on the average, receive more income. When a time-on-the-job variable is used in the graduate group, it is positive, significant, and greater than the one obtained for the whole sample.

evaluations in 1969. In view of this, the significance and magnitude of the variable in 1955 are surprising, suggesting that poor health and perhaps disability persist over time.[29] The variable for marital status, which is also taken from 1969 records, is significant, and its coefficient is larger than the effect of any of the education or ability variables. This type of variable is usually interpreted as a proxy for motivation, as well as the need for higher income for a family.[30] We have included the two father's education variables (attended high school and attended college) to represent the socioeconomic standing of the individuals' families. Since our dependent variable does not include unearned income except possibly for the earnings of the self-employed, the effect of inheritance will not be reflected in the coefficients.[31] The father's education variables will reflect, however, various types of training and motivations inculcated by parents as well as possible business contacts. The coefficients on these high school and college variables are approximately the same, although only the high school variable is significant.[32] As noted earlier, when these variables are omitted, the biography variables are more significant. We have tried including dummy variables for nine of the ten regions in which people went to high school, but none of these variables were significant. Hence, the interregional differences in the quality of high school education are not significant. Some results on the quality of college education are given in Chapter 1; more detailed analyses of high school and college quality are currently being carried out by Solmon and Wachtel.

Despite the presence of a large number of significant determinants of earnings, the fraction of explained variance in our

[29] Health is a much more significant determinant of income in 1969.

[30] A variable representing divorced individuals was insignificant.

[31] When the self-employed category was analyzed, as in Chapter 8, father's education was not important.

[32] Duncan, Featherman, and Duncan (1968), as well as others, have found that the socioeconomic standing of parents affects income of people with the same education. Sewell, Haller, and Portes (1969) have found that both father's education and occupation are significant determinants of occupational attainment. Subsequent work with our sample has confirmed this finding. In addition, we have found that religion and father-in-law's status likewise have a significant relationship to earnings. These new findings only change the education coefficients by a small amount and are not used, since they would necessitate recalculating all rates of return for a small refinement in the figures.

equations is only about 10 percent. While this may seem small compared with other studies of the determinants of income, there are several aspects of our sample that should be considered. First, the age variation in our study is very small, and since age increases the total variance of income without being associated with a corresponding increase in variance of the residual, samples with little age variation will have smaller $\bar{R}^2$s.[33] A comparable argument can be made because of the truncation of education and ability. Second, some studies standardize for occupations, thereby introducing more explanatory variables, a procedure that is not appropriate here.

The variables examined above explain, in some sense, part of the distribution of earnings. The distribution of the residual (the unexplained portion of earnings) is important for various aspects of human-capital and statistical theory. Appendix I contains an examination of the residuals for 1955 and 1969, as well as a study of the relative performances of equations using log earnings rather than earnings.

THE BIAS FROM OMITTING ABILITY, 1955

A primary concern of this study is the bias that occurs in the calculations of the extra income due to education when ability is omitted from the regressions. To calculate the bias from omitting ability, we compare the coefficients on the education dummies before and after including ability. To facilitate comparisons with other studies that do not have ability data, we hold constant variables that have been included in other studies. The equation without ability is called the comparison equation. The biases, calculated as the percentage differences between the corresponding coefficients in the two equations, are presented in Table 5-5.

Before examining these results, we shall explain one complication. As noted above, the biography variable contains information on mathematical ability, hobbies, and personal and family background. Since this variable remained significant after the introduction of the ability and father's education variables, it contains some information not accounted for by these two variables. If this other information is used in census studies such as Hanoch's and if we want to apply our bias corrections to

[33] In Hansen, Weisbrod, and Scanlon (1970), who study about the same age span, the $\bar{R}^2$s are comparable to ours.

TABLE 5-5 Percentage biases at various education levels from omitting different types of ability, 1955

Ability omitted	Bias at education level of: Some college	Undergraduate degree	Some graduate work	Master's	Ph.D and LL.B.
A. Bias from omitting ability; age, background, and biography held constant					
1. Factor 1	7.8	20.4	17.5	24.8	24.3
2. Factor 2	−0.9	−0.8	−0.2	−2.5	−2.6
3. Factor 3	1.2	6.5	5.8	7.5	5.6
4. Factor 4	1.9	4.4	5.2	5.1	3.3
B. Bias from omitting ability; age held constant					
5. Factor 1	9.3	19.8	18.9	22.9	22.8
6. Factor 2	0.3	0.4	0.9	−0.8	−2.4
7. Factor 3	3.9	9.6	9.8	10.9	8.7
8. Factor 4	5.4	7.8	8.0	8.0	5.6
C. Bias from omitting biography and ability; age and background held constant					
9. Factor 1	23.2	34.7	25.7	36.8	35.5
10. Factor 2	15.9	17.3	9.7	36.8	12.6
11. Factor 3	17.6	23.3	15.2	22.2	19.6
12. Factor 4	18.2	15.2	14.6	20.2	17.7

SOURCE: Table 5-3. Comparison equation for section A is equation 4. Row 1 from equation 5; row 2 from equation 6; row 3 from equation 7; row 4 from equation 8. Comparison equation for section B is equation 3. Row 5 from equation 10; row 6 from equation 11; row 7 from equation 12; row 8 from equation 13. Comparison equation for section C is equation 9. Row 9 from equation 5; row 10 from equation 6; row 11 from equation 7; row 12 from equation 8.

such studies, then we should include the biography variable in the comparison equation. If we do so, the biography variable—which includes ability—will to some extent hold ability constant. Thus, the bias from omitting ability calculated from this equation will underestimate the actual bias. On the other hand, if we omit biography from the comparison equation while including it in the equation with ability, we will overstate the bias due solely to ability. Since these two alternatives bracket the desired result, we present both. Of course, if the content of the biography variable is not contained in the census data, then this latter bias calculation is the appropriate one for correcting for the effect of omitting both sets of variables.

We have calculated the bias from omitting each of the factors separately. We have also computed the bias from omitting all four factors, but since the results are quite close to those obtained from omitting just mathematical ability, they are not presented. Section A of Table 5-5 presents the bias when the age, background, and biography variables are included in the comparison equation, while section C contains the results when biography is not included. The biases from omitting mathematical ability, which appear in row 1, are substantial, ranging from 8 to 25 percent with a mean of about 20 percent. Subject to a problem discussed below, these can be considered lower bounds to the bias from omitting mathematical ability in census studies. The upper bounds, which appear in section C, average out about 50 percent greater but with a narrower range.

Table 5-5 also contains estimates of the bias from omitting different types of ability, as represented by factors 2 through 4. There appears to be no bias at all from omitting factor 2, which represents coordination. The maximum bias from omitting either of factors 3 and 4, which represent IQ and spatial perception, ranges from 15 to 23 percent with an average of somewhat less than 20 percent. In Table 5-3 only the top fifth for factors 3 and 4 is significantly different from zero in equations 6, 7, and 8. All these variables are insignificant when included in an equation with the first factor. Thus, we conclude that of the four ability measures tested, only the mathematical one is a significant determinant of income, and it is the only one whose omission results in a substantial bias.

In evaluating these bias results, the reader is reminded that one component of the bias calculation is the coefficient in the linear relationship between the omitted variable and education. But this relationship may not apply to the population as a whole because the sample is stratified, and because the war and the GI Bill may have changed the demand for college by people with various ability levels and family backgrounds.

The significance and magnitude of the mathematical-ability variable are important in terms of the calculations of the returns to education and in terms of the design of future studies. Before accepting these results, however, we must consider the possibility that the coefficient on the mathematics factor is reflecting part of the effect of education. This could occur if the scores on some of the component tests used in calculating the mathematics factor depend on differences in mathematical course work or education obtained prior to 1943 and if post-World War II education is related to mental ability. Suppose pretest education affects the test scores. Then, considering all people with graduate degrees, for example, we would expect to find higher average test scores for relatively older individuals, since it is more likely that they would have received more education prior to taking these tests. For high school graduates, of course, we would not expect to find any pattern by age, since everyone in the sample had at least a high school education when taking the tests.

In Table 5-6 we present average scores on the arithmetic-reasoning (Math B) test for various age-education cells. A quick glance at the upper education categories of some graduate work, master's, and Ph.D. indicates that there is no tendency for older people to score higher than younger people. In fact, if anything, it would appear that the reverse holds true.[34] This suggests that the Math B ability test score is not affected to any significant degree by the amount of the individual's pretest schooling. We have calculated tables similar to Table 5-6 for the mathematics (Math A) and biography scores, and again there is no evidence to suggest that schooling affects scores.

There are several other reasons for expecting that the effect of prior education on ability has little impact on the ability and education coefficients. First, the tables on initial salary (to be presented shortly) indicate that most people with such prior education had attended college only a year or two; hence the effect of differential education on the test scores should be small, especially when the data are converted to fifths. Second, the biography measure, which incorporates data on pre-1943 education, partially holds this variable constant. Third, at least half the weight in the mathematical factor is attached to tests, such as numerical operations, that are not related to college education. Finally, when we computed regressions for those with just a high school education—that is, those whose pre-1943 education was the same—the effect of ability was not statistically different from the estimates using all education groups. We conclude, therefore, that in this sample mathematical ability is a much more important variable than IQ in studying the returns to education and that the estimates of the bias from omitting this abil-

[34] The decrease with age may mean that recent familiarity with tests or with items stressed on tests improves scores. Because the magnitude of the age effect is the same at all education levels, however, our conclusions would be unaffected by an age adjustment. In the regressions, such a familiarity bias would be eliminated by our age variable.

TABLE 5-6 ***Average scores on Mathematics B test, by age and education***

	Education					
Age in 1969	*High school*	*Some college*	*Undergraduate degree*	*Some graduate work*	*Master's*	*Ph.D. and LL.B.*
43	14.0		28.3		25.7	
	(2)	*	(3)	*	(3)	*
44	18.5	23.4	31.6	31.1	32.9	33.2
	(95)	(115)	(183)	(34)	(37)	(45)
45	17.2	24.1	32.3	31.3	32.7	31.5
	(193)	(230)	(328)	(68)	(88)	(81)
46	16.9	23.4	28.5	30.9	29.0	34.2
	(195)	(263)	(286)	(49)	(102)	(73)
47	16.6	21.2	29.7	32.5	31.7	32.6
	(155)	(151)	(198)	(38)	(59)	(58)
48	16.3	21.5	30.1	31.6	29.6	28.4
	(91)	(122)	(111)	(19)	(56)	(39)
49	15.4	21.1	27.7	26.6	24.1	29.5
	(80)	(68)	(43)	(14)	(24)	(16)
50	14.7	19.4	32.7	35.9	26.1	26.6
	(81)	(55)	(35)	(7)	(21)	(8)
51	16.3	21.8	27.3	23.3	34.0	23.2
	(61)	(49)	(29)	(6)	(7)	(9)
52	15.6	17.9	27.2	26.8	32.2	19.0
	(73)	(48)	(24)	(5)	(6)	(2)
53	13.8	14.8	29.3	22.0	20.8	49.0
	(22)	(19)	(8)	(2)	(4)	(2)
Average age	47.3	46.9	46.3	46.3	46.6	46.3

*No observations.

NOTE: Sample sizes are given in parentheses.

ity are substantial. This conclusion must be qualified to the extent that Thorndike's comment that factor 1 is primarily an IQ test is correct.

COMPARISON OF RESULTS WITH THE WOLFLE-SMITH DATA

Since the average age of those in the Wolfle-Smith data is approximately the same (when sampled) as the age of those in the NBER-TH data of 1955 and since the sample dates are only two years apart, it is useful to compare the results. The bias from omitting ability in the Wolfle-Smith data was about 4 percent. But this calculation was made with only IQ and age held constant, while in the above analysis we also hold background factors constant. In section B of Table 5-5 we present, for purposes of comparison, estimates of the bias from omitting each ability separately when no account is taken of background.[35] When what we call the IQ factor is omitted, the biases range from 4 to 11 percent with an average of around 9 percent, which is fairly close to the Wolfle-Smith result. When the mathematical factor is omitted, however, the bias ranges from 9 to 23 percent with an average bias of about 20 percent. (As noted in Chapter 4, there is a possibility that what we call mathematical ability is a better approximation to IQ.) Thus, the results from the two samples do not differ substantially when analyzed on a comparable basis. This suggests that mathematical ability may be more important for determining earnings than IQ. However, the Wolfle-Smith and NBER-TH tests differ in measurement as well as in concept, and we cannot be sure which is the cause of the difference in results.

The extra income from education can also be compared in the two samples. In the Wolfle-Smith study, after holding ability constant—as measured by ACE scores—we found education to be more important than in the NBER-TH data; that is, the first two years of college add 18 percent to income, an undergraduate degree adds 45 percent, and two or more degrees add 57 percent. Some of the extra returns arise because a different ability measure is used and because additional variables are included in the equation on which Table 5-4 is based. If only age and IQ are held constant, as in equation 1, Table 5-3, the income differentials of education in the NBER-TH sample are substantially greater. For example, the return to some college rises to 18 percent, which is identical to the Wolfle-Smith result, while the return to an undergraduate degree rises to about 20 percent.[36] It is not completely clear why there is such a difference between the two samples for returns to a college education. One possibility

[35] Age is held constant for comparability, since all people in the Wolfle-Smith sample graduated from high school in the same year.

[36] M.D. and teacher dummy variables are included in these calculations, but not in the Wolfle-Smith calculation. The discussion in this section assumes that factor 3 and not factor 1 represents IQ. If this is not the case, there is less agreement between the samples.

lies in the nature of the samples. The NBER-TH sample is drawn from a group more homogeneous with respect to such unmeasured variables as personality and drive than the Wolfle-Smith sample, because the population of the latter was composed of all high school graduates from Minnesota.[37] Second, the Wolfe-Smith sample may be more affected by response bias (see Chapters 3 and 4).

The above discussion exhausts the implications of the data for 1955 when treated as a single or isolated cross section. Later, we study the implications of using information from other years.

EARNINGS DIFFERENTIALS IN 1969

When comparing the 1955 and 1969 results, it is important to keep in mind the slight differences in definition for the income and education variables. Since nearly 10 percent of our sample completed their education after 1955, for 1969 we use education as reported in 1969, but as pointed out earlier, about 10 percent of the questionnaires contain inconsistent replies on education in the two years. The 1969 earnings variable differs from that of 1955 because in 1969 respondents were asked to report earnings from their main job only, and in 1955 they were asked to report total earnings.[38] About 10 percent of the sample indicated that they held more than one position in 1969, a fact that we ignore in our analysis.[39]

The first question to consider is whether there are any important interactions between education and ability.[40] Once again,

[37] In addition, if the returns to education interact significantly with experience and are very concave with respect to age, then the average return from the sample with variable age will be less than that from the sample with a fixed age, but this should only explain a small portion of the difference.

[38] We omit from the analysis the 50 people not reporting income in 1969.

[39] We have calculated the income change from 1968 to 1969 for those with more than one job in 1969. Assuming that they also had two jobs in 1968, we would expect them to have a relatively smaller growth in income than others. This is the case at all education levels. We can adjust their 1969 earnings to have the same average growth as single-job holders in the same education groups. When this is done, average incomes rise by 0.5 to 1.5 percent in the various education levels. Since such small differences would not affect our results significantly, we have not corrected the 1969 data. Note that these results strongly imply that the "current" earnings figures are for 1969 rather than 1968.

[40] As was the case with the 1955 data, preliminary analysis indicated that factor 1 was the only important ability factor. Factor 3—an IQ measure—was occasionally significant and negative.

the *F* statistics used in the analysis of covariance are less than 1.5; hence, we cannot reject the hypothesis that all the coefficients in the individual education regressions are equal to the corresponding coefficients in the overall regression. There is, however, some tendency for the graduate education coefficients to be higher at the high ability levels. Thus, we also computed some equations with interaction between graduate education and the high ability fifths. In equations 14 and 15 in Table 5-7 we present the equations with interactions. For each ability fifth, the effect of ability is greater for graduate students than for other students; however, only in the case of Q_4 is the difference significant, although Q_5 has a *t* value greater than 1.0.[41] For graduate students the effect of ability can be determined by adding together the ability coefficients such as Q_2, with the coefficient on the interaction term involving Q_2. Thus, at the graduate level, compared to the bottom fifth, the effect of ability is about $150, $210, $350, and $415 per month for the second to fifth fifths, respectively. Although the effect in the top fifth is larger than that in the fourth fifth, the interaction coefficient is not significant because the top fifth has a much bigger impact than the fourth fifth—$204 versus $94—at the other education levels. Since the interactions here are fairly weak, we conduct our analysis of the returns to education and the bias from omitting ability using equations without interactions.

We consider now equation 5 in Table 5-7, which contains the mathematical-ability, biography, and background variables. All the education coefficients, the top three fifths of the mathematical factor, the second, fourth, and fifth fifths of the biography variable, the background variables, and the M.D. and teacher dummies are significant.[42] Most of these variables, as well as

[41] Because there are 20 possible education-ability cells, 5 percent of the time, or once, we would find one of the interaction terms to have a significant *t* value even if its true value were zero. However, we would not expect to find all four coefficients positive.

[42] The M.D. and teacher variables are not quite correct, since they are measured as those who were M.D.'s and teachers in 1955. When we added a separate variable for lawyers, both the Ph.D. and lawyer coefficients were significant, while the other coefficients were unaffected. In a weighted regression that attempts to correct for heteroscedasticity, the coefficients on the some-graduate-work and master's variables change by about 15 percent, while the ability variables increase by about $60. There were some noticeable changes in a few of the other variables.

TABLE 5-7 Regressions for salary, 1969 (in dollars per month)

	Constant	*Some college*	*Undergraduate degree*	*Some graduate work*	*Master's*	*Ph.D. and LL.B.**	*M.D.*	*Teacher*	*Age*
(1) Y_{69}	$1,164.7	$276.4	$499.4	$376.2	$318.0	$926.6			
	(37.2)	(6.4)	(11.9)	(5.3)	(5.4)	(14.8)			
(2) Y_{69}	1,360.2	274.8	493.2	370.8	313.9	920.6			$−5.9
	(6.5)	(6.4)	(11.6)	(5.2)	(5.9)	(14.6)			(1.0)
(3) Y_{69}	1,321.9	276.9	498.3	421.6	527.6	857.7	$508.9	$−522.5	−4.7
	(6.4)	(6.5)	(11.8)	(5.9)	(7.9)	(12.9)	(4.0)	(6.9)	(0.8)
(4) Y_{69}	1,454.2	206.4	388.9	336.4	412.1	728.3	486.1	−524.9	−1.9
	(4.7)	(4.6)	(8.6)	(4.6)	(6.0)	(10.5)	(3.7)	(6.4)	(0.3)
(5) Y_{69}	1,355.7	192.9	340.8	287.1	351.0	670.0	493.7	−496.2	−1.7
	(4.4)	(4.3)	(7.5)	(3.9)	(5.1)	(9.6)	(3.8)	(6.1)	(0.3)
(6) Y_{69}	1,461.9	208.0	391.2	337.5	410.9	734.3	474.5	−516.5	−2.3
	(4.8)	(4.6)	(8.7)	(4.6)	(6.0)	(10.6)	(3.6)	(6.3)	(0.4)
(7) Y_{69}	1,425.8	204.9	382.9	330.6	403.1	723.7	485.5	−517.9	−1.8
	(4.6)	(4.5)	(8.4)	(4.5)	(5.8)	(10.5)	(3.7)	(6.2)	(0.3)
(8) Y_{69}	1,367.5	205.1	383.5	331.1	339.8	723.4	488.1	−511.2	−0.9
	(4.4)	(4.5)	(8.5)	(4.5)	(4.9)	(10.5)	(3.7)	(6.3)	(0.1)
(9) Y_{69}	1,416.5	234.5	425.0	371.3	440.9	757.1	495.6	−522.5	1.0
	(4.6)	(5.2)	(9.6)	(5.1)	(6.4)	(11.0)	(3.8)	(6.4)	(0.2)
(10) Y_{69}	1,215.7	252.9	430.2	353.1	443.6	776.7	514.9	−515.8	−4.8
	(5.9)	(5.9)	(10.0)	(4.9)	(6.6)	(11.5)	(4.1)	(6.4)	(0.8)
(11) Y_{69}	1,300.7	278.0	498.3	424.4	526.6	864.1	499.6	−543.8	−5.5
	(6.2)	(6.5)	(11.8)	(6.0)	(7.9)	(13.0)	(3.9)	(6.8)	(0.9)
(12) Y_{69}	1,294.0	270.1	478.8	403.6	504.1	838.6	510.2	−540.0	−4.8
	(6.1)	(6.2)	(11.2)	(5.6)	(7.5)	(12.5)	(4.0)	(6.7)	(0.8)
(13) Y_{69}	1,211.7	267.1	477.9	404.2	500.2	840.3	509.6	−531.3	−3.5
	(5.8)	(6.2)	(11.3)	(5.7)	(7.5)	(12.6)	(4.0)	(6.6)	(0.6)
(14) Y_{69}	1,358.8	195.3	346.9	229.9	282.5	602.2	490.7	−485.4	−1.7
	(4.4)	(4.3)	(7.5)	(2.8)	(3.5)	(7.6)	(3.7)	(5.9)	(0.3)
(15) Y_{69}	1,373.6	196.1	348.7	150.3	204.8	521.4	489.0	−487.4	−1.8
	(4.5)	(4.3)	(7.6)	(1.3)	(1.8)	(4.4)	(3.7)	(6.0)	(0.3)

TABLE 5-7 ***(continued)***

	Ability				*Interaction of graduate education with:*					
	Q_2	Q_3	Q_4	Q_5	Q_2	Q_3	Q_4	Q_5	*Health*	*Single*
(1)										
(2)										
(3)										
(4)									$−208.6	$−225.5
									(−7.9)	(2.1)
(5)	$ 68.8	$ 107.1	$143.7	$ 278.9					−205.6	−236.8
	(1.3)	(2.1)	(2.9)	(5.5)					(7.8)	(2.2)
(6)	−16.7	−38.5	72.7	−8.1					−208.1	−230.2
	(0.3)	(0.8)	(1.5)	(0.2)					(7.9)	(2.1)
(7)	14.9	41.9	26.7	57.5					−208.0	−225.1
	(0.3)	(0.8)	(0.5)	(1.1)					(7.9)	(2.0)
(8)	33.0	32.0	36.0	117.4					−205.7	−225.2
	(0.7)	(0.6)	(0.7)	(2.4)					(7.8)	(2.0)
(9)									−215.5	−250.1
									(8.2)	(2.3)
(10)	81.5	121.5	176.2	318.2						
	(1.6)	(2.5)	(3.6)	(6.5)						
(11)	35.1	8.7	121.5	57.5						
	(0.7)	(0.2)	(2.6)	(1.2)						
(12)	−9.0	53.6	54.0	102.2						
	(0.2)	(1.1)	(1.1)	(2.1)						
(13)	53.1	73.7	73.7	192.2						
	(1.1)	(1.5)	(1.5)	(4.0)						
(14)	70.3	110.7	106.0	261.3			$179.8	$ 87.2	−204.7	−242.4
	(1.4)	(2.2)	(2.0)	(4.7)			(1.9)	(1.0)	(7.8)	(2.2)
(15)	55.2	89.5	93.6	248.8	$96.1	$121.5	260.6	168.0	−204.5	−241.8
	(1.0)	(1.6)	(1.7)	(4.4)	(0.7)	(0.9)	(2.0)	(1.4)	(7.8)	(2.2)

*M.D.'s are also included.

NOTE: Figures in parentheses are *t* statistics.

Father attended high school	*Father attended college*	*Biography*				*Ability factor*	$\bar{R}^2$/S.E.
		Q_2	Q_3	Q_4	Q_5		
							.062
							954
							.062
							954
							.077
							947
$115.0	$112.4	$125.4	$101.1	$178.3	$236.5		.100
(3.3)	(2.6)	(2.5)	(2.0)	(3.6)	(4.8)		938
107.9	96.8	119.2	91.8	167.0	205.9	1	.112
(3.1)	(2.2)	(2.4)	(1.8)	(3.4)	(4.2)		934
113.7	112.7	126.6	103.1	178.9	236.9	2	.105
(3.2)	(2.6)	(2.6)	(2.0)	(3.6)	(4.7)		938
114.4	108.6	122.0	97.1	173.4	228.2	3	.104
(3.3)	(2.5)	(2.5)	(1.9)	(3.5)	(4.6)		938
110.9	108.8	120.3	92.4	168.3	219.8	4	.105
(3.1)	(2.5)	(2.4)	(1.8)	(3.4)	(4.4)		938
137.3	152.9						.095
(3.9)	(3.6)						941
						1	.087
							942
						2	.078
							947
						3	.078
							946
						4	.080
							946
107.8	96.4	120.1	92.4	167.7	205.8	1	.113
(3.1)	(2.2)	(2.4)	(1.8)	(3.4)	(4.1)		934
107.8	97.3	120.7	92.9	167.9	204.5	1	.113
(3.1)	(2.2)	(2.4)	(1.9)	(3.4)	(4.1)		934

age, were significant in 1955. The age result for 1969 is discussed in more detail below.

The percentage increases in earnings due to increases in education implied by this equation are presented in Table 5-8, for a person with the ability and background of (1) the average high school graduate and (2) the average member of the preceding education level. As compared to just going to high school, earnings increased by 17 percent for some college attendance, 25 to 30 percent for obtaining an undergraduate through a Ph.D. degree; 84 percent for obtaining a law degree, and 160 percent for completing a medical degree. Each of these numbers is larger than the corresponding entry for 1955 (Table 5-4). The greater importance of education with the passage of 14 more years can be explained either by the proposition that the age-income profile is steeper for the more educated in this range or by the proposition that because of shifts in supply and demand, the relative wage rate of the more educated was greater in 1969

TABLE 5-8 Percentage increases in earnings, by education level, 1969

	Percentage by which earnings in each education class exceed:	
	Earnings of average high school graduate	*Earnings of average member of preceding education level*
Some college	17	17
Undergraduate degree (not teacher)	31	11
Some graduate work	26	−3
Master's (not teacher)	32	4
Ph.D.	27	−8
M.D	106	
LL.B.	84	
Undergraduate degree (teacher)	−14	
Master's (teacher)	−13	

NOTE: Average age is 47.

SOURCE: Earnings of average person in each education class from equation 5, Table 5-7. Absolute increases from equation 3, Table 5-7.

than in 1955. We discuss this question in more detail in Chapter 6.

Several comments about the pattern of results in Table 5-8 are in order. First, in an equation not shown, we have replaced the some-college dummy variable with separate variables for one, two, and three years of college. We find that the absolute increase in earnings (over high school) for the first year is the same as for the some-college variable and that neither of the next two years adds anything to income.[43] Second, as shown in the second column of Table 5-8, an undergraduate degree causes another jump in income, although the increase is smaller than that for the first year of college. The erratic nature of the returns to education is made more evident by the decrease in earnings from an undergraduate degree to some graduate education and the large increase for three-year-graduate-degree holders. Third, when a dummy variable for business owners is included, the constant term and all the various education coefficients fall, but the decreases in the high school and some-college categories are larger. If calculated for non-business owners, the income differentials in Table 5-8 would be 17 and 39 percent for some college and a bachelor's degree, respectively.

We next examine the effects of ability and background. As noted earlier, only factor 1, the mathematical-ability factor, has significant coefficients and displays the expected monotonic patterns. From equation 5, Table 5-7, it is clear that the impact of this mathematical skill is important, with those in the top fifth receiving about $3,300 a year more income than those in the bottom fifth—an increase midway between that of the some-college and the undergraduate-degree variables. The biography variable has significant coefficients, which (except for a small insignificant drop in the third fifth) increase monotonically. Although the magnitudes are not quite as large as those of mathematical ability, the effect of the top fifth in biography is as important as that of some college education.

We now examine the effect of age on earnings. Before discussing this subject in detail, we remind the reader that the age variation in this sample is only eight to nine years. Because of this small range, no attempt was made to find nonlinear age ef-

[43] The information currently available from the 1955 questionnaire does not allow us to break down the some-college category to investigate this question.

fects or to discover interactions with education in 1969.[44] Hence, our results should not be extrapolated to other ages.[45] Whenever age is included in an equation, its coefficient ranges from −2 to −6, but it is always insignificant. (In 1955 it was about +7 and significant.) A common finding in studies of age-income profiles is that the peak earnings occur in the late forties or early fifties of a person's lifetime.[46] Since the average age of people in this sample was 47 years in 1969, a negative coefficient is not surprising, and the insignificance could arise because the peak is near the mean of 47.[47] There is, however, one other result to report. A good argument could be made that time on the job (after education is completed) is more important than age, and that such work experience is not completely collinear with age and education because of delays in entering or finishing education. We attempted to measure such a concept by a variable defined as 1969 minus the year of first job after completing school. (This year was determined from the job-income history.)[48] When this time-on-the job variable is used, we always find a significant coefficient of about +10, even if age is added to the equation. Since the education coefficients are not much affected by this variable, however, we shall not pursue these results any further at this time.

As was the case in 1955, the background and individual characteristics are important. The health and single coefficients are both significant, negative, and in excess of $200. Each of the two father's education coefficients are about $100 and significant. Thus, the health and single effects are as great as that of the

[44] We tested for age-education interactions by running separate regressions for the various education categories. The age variable was never significant.

[45] It is possible to study these questions by combining the 1955 and 1969 cross sections.

[46] See, for example, Becker (1964), Miller (1960), and Mincer (1970). Some recent work at the NBER by Fuchs suggests that when earnings are converted to wage rates per hour, this peak is not discernible.

[47] Moreover, in the next chapter we will present evidence that from 1968 to 1969 average income (after adjustment for secular changes in wage rates) declined in this sample.

[48] We should note that for some people the "first job" entry appears to refer to the date they began in their current job.

some-college variable, while the influence of father's education is equal to that of the third fifth of ability. As was the case in 1955, the quality of high school education, as measured by a set of dummy variables, did not determine earnings. The $\overline{R}^2$ values in 1969 of about .11 are slightly greater than those in 1955.

THE BIAS FROM OMITTING ABILITY, 1969

We turn now to the question of the bias from omitting ability. The information obtained on this question is very important, since this is the only large sample with information on income, ability, and education for people in their late forties and early fifties. In Table 5-9, we present the bias from omitting separately each of the ability factors, using a variety of assumptions

TABLE 5-9 ***Percentage biases at various educational levels from omitting various types of ability, 1969***

		Bias at education level of:				
	Ability omitted	*Some college*	*Under-graduate degree*	*Some graduate work*	*Master's*	*Ph.D. and LL.B.*
A.	**Bias from omitting ability; age, background, and biography held constant**					
1.	*Factor 1*	6.5	12.4	14.6	14.8	8.0
2.	*Factor 2*	−0.8	−0.6	0.3	0.3	−0.8
3.	*Factor 3*	0.7	1.5	1.7	2.2	0.6
4.	*Factor 4*	0.6	1.4	1.6	3.0	0.7
B.	**Bias from omitting ability; age held constant**					
5.	*Factor 1*	8.7	13.7	16.2	15.9	9.4
6.	*Factor 2*	−0.4	0.0	−0.7	0.2	−0.7
7.	*Factor 3*	2.5	3.9	4.3	4.5	2.2
8.	*Factor 4*	3.5	4.1	4.1	5.2	2.0
C.	**Bias from omitting biography and ability; age and background held constant**					
9.	*Factor 1*	17.8	19.7	22.8	20.6	11.5
10.	*Factor 2*	11.4	7.8	9.3	7.0	3.0
11.	*Factor 3*	12.7	9.8	11.1	8.8	4.4
12.	*Factor 4*	12.6	9.6	11.0	9.5	4.4

SOURCE: Table 5-7. Comparison equation for section A is equation 4. Row 1 from equation 5; row 2 from equation 6; row 3 from equation 7; row 4 from equation 8. Comparison equation for section B is equation 3. Row 5 from equation 10; row 6 from equation 11; row 7 from equation 12; row 8 from equation 13. Comparison equation for section C is equation 9. Row 9 from equation 5; row 10 from equation 6; row 11 from equation 7; row 12 from equation 8.

about the other included variables. Sections A and C yield lower and upper bounds to the percentage bias when background factors are held constant. When mathematical ability is omitted (section A), the biases at the various education levels range from 7 to 15 percent, with an average of about 10 percent. In section C, the upper bounds of such biases range from 18 to 23 percent, with an average of about 20 percent. Both the lower and upper bounds of the bias are substantially less than the corresponding estimates for 1955, because the coefficients on education grew more rapidly than those on the ability, biography, or background variables.

Mathematical ability is only one of the many types of ability that could affect income and whose omission could bias the education coefficients. In Table 5-9 we also present the biases from omitting physical coordination, IQ, and spatial-perception measures. The biases from omitting any of these average about 2 percent or less in section A and 10 percent in section C. Both these upper and lower bounds are much smaller than when mathematical ability is the omitted variable. (The biases using an equation in which all the ability variables are included were nearly the same as in the mathematical-ability rows.) Since the upper bound on the bias from omitting these three abilities does not exceed 10 percent, we conclude that in 1969, as well as in 1955, the returns to education are not greatly affected by the omission of coordination, IQ, and spatial-perception abilities. It would still be possible, however, for any of these abilities to be significant determinants of income. But in equations 6, 7, and 8 in Table 5-7, the only significant coefficient for the ability dummies is the top fifth of factor 4. When factors 2, 3, and 4 are added to equation 5, no additional coefficients are significant.

We have also made some calculations of the bias when only age is held constant. The results are given in section B of Table 5-9. Once again, omitting mathematical ability yields a large bias—about 12 percent—while omitting the other abilities implies biases of 4 percent. In 1955 the comparable biases were twice as large.

INDIVIDUAL EFFECTS PERSISTING OVER TIME

In the preceding analysis we have treated the two cross sections separately. It is possible to make use of the continuous-cross-section nature of the data to obtain more efficient estimates of the coefficients and to allow for unmeasured personal characteristics that persist over time.[49]

It can be demonstrated that if we were to regress $Y_{69} - Y_{55}$ on our independent variables (and include a new one for the change in education), the estimates would equal the difference between the coefficients estimated in 1969 and 1955. There is, however, an advantage in using the individual's residual in one time period in the regression for the other. That is, the residual consists of $\delta_t P_i + u_i$, where P_i represents the ith person's various unmeasured characteristics, such as personality, that persist over time; u_i is a random element; and δ_t is the effect on income in period t of the P_i. Let $q_{it} = \delta_t P_i + u_i$. If we include q_{i55} in our 1969 regressions, or vice versa, then we incorporate an imperfectly measured estimate of P in our analysis. Measurement error will generally lead to biased estimates of the coefficient on q and all variables with which q is correlated. But since q_t is constructed to be orthogonal to all the independent variables in the equation in year t and since these variables are unchanged from 1955 to 1969 (except for a few changes in education), each q is approximately orthogonal to the other variables. Hence, the other coefficients are practically unchanged when q is included.

We do not present the entire equation, but concentrate instead on the coefficient of q and the change in the explained variance. When we use the 1955 residual in 1969, the coefficient of q is 1.7 with a t value of 35. The $\bar{R}^2$ rises from .11 to .34, and the standard error of the equation declines from \$934 to \$785. When the 1969 residual is used in the 1955 equation, the coefficient is .15 with a t value of 35. The $\bar{R}^2$ rises from .10 to .33 and the standard error declines from \$264 to \$227.[50]

[49] In principle, it is possible to obtain, with certain grouping techniques, unbiased estimates of the coefficients of the measured and unmeasured variables even when they and income are correlated with the unmeasured personal characteristics. However, this procedure proved to be infeasible because of multicollinearity.

[50] The comparisons are with equation 5, Table 5-7, and equation 5, Table 5-3. Since the coefficients of the q variables should equal the ratio of the coefficients of the unmeasured characteristic in the two years, it might be expected that their product would be 1. However, because in both cases personality and so on are measured with error, the q coefficients are biased downward, and their product is less than 1.

The residual in any year can be partitioned into p and u.[51] It is interesting to calculate the extent to which the variance of the residual is due to the variables represented by p. Since p and u are assumed to be uncorrelated, var p + var u = var q. An estimate of var q is the (squared) standard error of the regression equation when q is included, while an estimate of var u is the corresponding statistic when q is not included. Using this information, we calculate var p expressed as a fraction of var p + var u to be 30 percent for 1969 and 26 percent for 1955.

Thus, based on observations 14 years apart, we estimate that about 30 percent of the original unexplained variance (after eliminating the influences of the measured variables) is due to variables whose effects persist over time. The other 70 percent of the variance is due to random events such as luck and changes in unmeasured variables.[52] This conclusion is discouraging in two regards. First, there apparently are some systematic but unmeasured determinants of income that are more important than all the education, ability, and background variables that we have studied. Second, random events (and changes in the unmeasured but systematic variables) are the most important determinants of income within a given age cohort. Finally, as demonstrated earlier, very little of the correlation between the residuals should be attributable to investments in on-the-job experience as formulated by Mincer (1970), since 1955 should be close to what he calls the overtaking point, where a person's current and discounted lifetime earnings are equal.

EARNINGS ON INITIAL JOB

The last earnings to be examined are those obtained on the person's first job after completing his education. There are several difficulties in using this information. First, there is the possibility of error because of the long time between event and recall in 1969. As demonstrated in Appendix H, we judge this memory lapse to be important for estimates of 1958 earnings, and the initial job occurred earlier than 1958 for nearly all the people in our sample. A possible offset to this greater-time-lapse aspect is

[51] Earlier we used δ, but since δ is just a scaling vector, it can be ignored.

[52] However, since both our education and ability measures cover limited ranges, the importance of ability and education relative to the unmeasured variables is understated compared with the population as a whole.

that the first position is a much more important event to the individual than other ones.[53]

Another problem with the initial-job data is that there is a correlation between education and the starting date. Because of inflation and technical change, wages increase over time; hence the average initial salaries of the more educated would tend to be overstated relative to those of high school graduates.[54] The adjustment for inflation and technical change given in footnote 53 assumes that the coefficients on the education variables change proportionately with the passage of time, but for reasons explained in Chapter 2, the relative wages at different education levels (at given age or experience levels) can shift over time. While such a structural change would hardly be expected to be important over the space of five or maybe even ten normal years, we do not have the luxury of five or ten normal years, because the period before 1943 is far removed from that of the late 1940s and early 1950s.

For these reasons, we resorted to simpler methods in analyzing the data. In Table 5-10, we present (for those instances in

[53] There is some partial evidence on the reliability of the earnings-on-initial-job estimates. Thorndike and Hagen asked the same question in 1955. If there are memory lapses, it seems likely that the longer the recall period, the greater the errors should be. Unfortunately, Thorndike and Hagen had not transcribed the initial-earnings data to the IBM sheets, and the NBER has not yet coded these data from the original questionnaires. However, Dr. E. Mantell, who has used the NBER-TH sample to write a dissertation on engineers, has collected the data for some 300 people for whom the year of initial job was the same in the 1955 and 1969 responses. He finds that the mean income reported in 1955 was $3,905 with a standard error of $2,039, while that reported in 1969 was $4,094 with a standard error of $1,473. This difference, which is only one-seventh of a standard error, is not significant at the 5 percent level. It might be noted that if people do not recognize the amount of inflation since the 1940s, they may tend to overstate their initial-job earnings in 1969. In a regression of the two responses, Mantell obtains

$$Y_{I,69} = \underset{(4.9)}{1{,}378} + \underset{(10.3)}{.070}\,Y_{I,55} \qquad \bar{R}^2 = .25$$

Thus while the mean estimates of two estimates are reasonably close, the individual estimates are not strongly correlated. Hence, the differences in the two responses do not solely represent an equally poor recall or inflation.

[54] This difficulty can be overcome through such devices as deflating starting salaries by a wage index and by including a time trend. Once the equation is estimated, it is easy to convert back to money wages at various education levels for each year.

TABLE 5-10 Initial annual salaries for selected years, by mathematical ability and education (mean income in dollars)

Year	Ability fifths 1	2	3	4	5	Total
High school						
1939	$1,878	$2,096	$1,769	$ 900	$1,359	$1,671
	(10)	(14)	(10)	(6)	(14)	(54)
1940	1,781	1,857	1,103	1,985	1,228	1,659
	(14)	(24)	(9)	(14)	(14)	(75)
1941	2,734	1,642	1,785	2,335	1,233	1,975
	(22)	(30)	(28)	(19)	(11)	(110)
1942	2,325	2,024	2,443	1,889	1,850	2,190
	(18)	(14)	(14)	(11)	(1)	(58)
1945	2,822	3,767	4,989	4,077	2,401	3,818
	(13)	(23)	(13)	(12)	(3)	(64)
1946	3,528	3,431	3,240	3,259	3,376	3,392
	(44)	(31)	(22)	(27)	(13)	(137)
1947	3,233	3,736	3,392	2,760	3,500	3,328
	(12)	(11)	(13)	(8)	(2)	(46)
Some college						
1940	2,708	1,815	1,594	1,323	1,422	1,624
	(4)	(8)	(14)	(8)	(15)	(49)
1941	2,104	2,392	2,519	1,727	1,554	2,012
	(11)	(6)	(10)	(10)	(12)	(49)
1945	4,229	4,432	3,713	4,027	4,711	4,180
	(18)	(10)	(12)	(17)	(9)	(66)
1946	3,560	3,517	3,267	3,321	3,246	3,361
	(24)	(24)	(32)	(41)	(33)	(154)
1947	4,172	3,572	4,112	4,896	3,544	4,089
	(22)	(22)	(24)	(24)	(19)	(111)
1948	5,326	3,998	4,194	4,181	4,206	4,377
	(22)	(23)	(21)	(26)	(19)	(111)
1949	4,668	3,293	4,211	3,951	3,743	3,894
	(10)	(18)	(11)	(12)	(7)	(58)

NOTE: Dollar entries are mean income. Sample sizes are given in parentheses below dollar amounts. Rank 1 is the lowest ability fifth.

TABLE 5-10 (continued)

	Ability fifths					
	1	*2*	*3*	*4*	*5*	*Total*
Undergraduate degree						
1946	3,370	3,375	3,444	3,233	3,543	3,443
	(5)	(8)	(16)	(11)	(32)	(72)
1947	4,050	3,408	3,917	3,427	3,230	3,464
	(6)	(13)	(20)	(23)	(48)	(110)
1948	3,654	3,615	3,399	3,845	3,936	3,755
	(20)	(30)	(32)	(47)	(74)	(203)
1949	3,783	3,649	3,320	3,402	3,590	3,518
	(24)	(38)	(44)	(59)	(59)	(224)
1950	4,788	4,100	3,120	3,576	3,976	3,808
	(17)	(25)	(32)	(41)	(43)	(158)
1951	4,118	4,329	3,920	4,222	4,600	4,297
	(11)	(17)	(14)	(24)	(28)	(94)
Graduate study						
1946	3,408	2,841	2,600	4,188	2,671	3,081
	(6)	(8)	(8)	(9)	(15)	(46)
1947	2,240	3,075	3,070	2,411	4,111	3,460
	(2)	(8)	(15)	(9)	(32)	(66)
1948	2,987	3,147	3,053	3,244	3,482	3,242
	(14)	(17)	(20)	(31)	(37)	(119)
1949	2,824	2,988	3,066	3,615	3,492	3,298
	(16)	(24)	(16)	(26)	(54)	(136)
1950	3,107	3,515	3,467	4,208	3,290	3,579
	(15)	(13)	(18)	(36)	(44)	(126)
1951	3,583	3,781	3,490	4,306	5,207	4,223
	(12)	(13)	(20)	(18)	(27)	(90)
1952	3,028	4,650	4,581	5,211	5,462	4,738
	(8)	(10)	(14)	(11)	(16)	(59)
1953	4,137	5,198	4,711	3,560	5,454	4,715
	(4)	(12)	(13)	(10)	(11)	(50)

which we have reasonable sample sizes) average starting salaries by mathematical-ability fifths, education, and year of starting salary.[55]

Consider first the high school graduates. Although it is not shown, about one-half had earned a living before 1943, but of the rest a majority started work in 1946.[56] In 1946 there is no evidence of a positive effect of ability on salary, and indeed there could even be a negative effect. None of the other years has enough observations to analyze the separate fifths.[57]

About one-third of those with some college had their first job before 1946. Fairly large groups of people began to work in each of the years between 1946 and 1948. It is encouraging to note that their starting salaries increased from 1946 to 1949 and that the salaries in 1946 were about double those before the war. In the years 1946 to 1948, we can compare the average income of the bottom two fifths with that of the top two fifths. We find the former to be larger by $250 in 1946, smaller by $420 in 1947, and larger by $550 in 1948. Using all five fifths of those averages, there is no evidence of incomes increasing with ability.

Very few people had finished their undergraduate studies and started to work before 1946.[58] From 1946 through 1951, there are large numbers of people with undergraduate degrees starting to work. The average starting salary rises in each year except for 1949.[59] In comparing the average starting income in the top two fifths with the income in the bottom two fifths, we find that

[55] Because we suspected that many people were reporting salaries other than their first, we constructed a comparable set of tables in which individuals were included only if the date of their starting salaries minus the date they terminated schooling was less than three years. In those instances in which we had large samples, for example, undergraduates around 1950 and high school graduates in 1946, the two sets of estimates were close; hence we decided to ignore the restriction and use the larger sample (1,000 more observations).

[56] The relatively large numbers in 1947, 1948, and 1949 presumably reflect later discharge dates from the military.

[57] Although education in these tables is defined from the data in 1969, when the corrected 1955 education is used, essentially the same results are achieved. The 1969 education is used because the question in 1969 asked for first salary after completion of education as reported in 1969.

[58] Including those with graduate education, there were only 4 percent. However, some graduates who began working in the late 1940s and early 1950s had completed their undergraduate work earlier.

[59] In 1949, the starting salary of those with some college as well as various national figures such as wage income declined.

in half the years the bottom fifths have more earnings, and in the other half the top fifths have more. Thus, the effect of ability on earnings still is not evident.

It is useful at this point to compare the starting salaries at the various educational levels. Since we have found that ability does not affect starting salaries, we can use the average in each year. In 1946 average incomes are \$3,392, \$3,361, and \$3,443 for those with a high school degree, some college, and an undergraduate degree, respectively. The differences in income are not significant.

It is interesting to speculate on why people coming on the market in 1946 with different amounts of education and ability received the same starting salary. One possibility is that, though ability and education add to skills, employers had no way of knowing who was better qualified and, on the average, offered the same wage to all these education and ability levels.[60] If this is correct, then as the workers' performances are monitored over time, we should find the effects of ability and education becoming more important, which is indeed what we have found to be the case. This is the opposite of the role of ability hypothesized by Lydall (1969). It also seems to run counter to the human-capital theory, since those with more training do not earn more when coming on the market. There is, however, a way to reconcile these results with the human-capital theory. As explained in detail in Mincer (1970) and Becker (1964), those with more education could be undertaking more on-the-job training and "paying" for the investment with lower current wages. In addition, those with some college education could have been tempted by very good job offers while the others finished their degrees.

The average starting salaries in 1947 and 1948 give some credence to these ideas. In these two years, those with some college received between \$500 and \$1,000 more in starting salary than those with undergraduate degrees. Also during these two years, high school graduates received slightly smaller starting salaries than undergraduate-degree holders. Finally, it might be

[60] Weisbrod and Karpoff (1968) have presented some evidence that when professional and mangerial employees are separated by quality of school and performance in school, starting salaries are the same, but salaries after 15 years differ significantly.

noted that in 1949 and 1950, when the some-college samples are smaller, it is still true that the starting salary of the college dropouts exceeds that of those with an undergraduate degree.

In order to obtain a large enough sample to permit analysis at the graduate level, we combined the three categories of graduate training. Since the average salaries of the three groups were quite close in each year, there is little danger of aggregation bias. The majority of these students finished their education between 1947 and 1951, although about 25 percent—many of whom were Ph.D.'s—began to work after 1951. In this graduate education category, we do find some evidence of the effect of ability on income. The bottom two fifths fall short of the top two in every year from 1946 to 1952. That ability is significant here and not at other education levels can be explained by the proposition that only at this education level is academic performance—which is related to ability—considered an indicator of employee quality, thus influencing starting salaries.

Consider next the difference in average starting salary of graduate students and college graduates. In each year from 1946 through 1954 (except 1953) college graduates earned as much on their first job as graduate students or earned more.[61] While the large percentage of pre-college teachers in the graduate group presumably holds down the average, the starting salaries for Ph.D.'s and LL.B.'s were about the same as the average for all graduates. Thus, once more we find that starting salaries do not rise, and may even fall, with education.

[61] The samples in both education groups are small in 1953.