Graduate Education, Ability, and Earnings
Author(s): Orley Ashenfelter and Joseph D. Mooney
Source: *The Review of Economics and Statistics*, Vol. 50, No. 1 (Feb., 1968), pp. 78-86
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/1927058
Accessed: 07-10-2016 18:40 UTC

GRADUATE EDUCATION, ABILITY, AND EARNINGS

Orley Ashenfelter and Joseph D. Mooney *

SUBSTANTIAL interest in recent years has centered on the relationship between personal earnings and a myriad of education related variables.[1] In this paper we present estimates of the impact on earnings of schooling, an index of ability, and a set of other relevant variables for a cohort of recent entrants to the labor market who have had some graduate education in the arts and sciences.

Aside from a purely intellectual curiosity, there are several other reasons for investigating the annual earnings for a group of this sort. First, estimates of an earnings function are necessary for calculation of rates of return to various quantities of educational investment. To date, very little work of an economic nature has been done in the growing field of graduate education.[2] It is hoped that the estimates presented in this paper may be viewed as an exploratory attempt to come to grips with problems in this important and neglected area.

Secondly, we have explicitly attempted the specification of an earnings relationship which allows the differential impact of schooling related variables to depend on the values of other relevant explanatory variables.[3] The existence of such interactive effects is interesting in itself and has important implications both for the rate of return analyses already available in the literature and for any further work on graduate education which may be attempted.

Finally, there is substantial interest in the specification of an earnings relationship which explicitly attempts to deal with the slippery concept of ability to earn income. We would like to know: (a) What sort of ability index is relevant in the context of highly educated persons, (b) the quantitative importance of an ability index, and (c) how parameter estimates of schooling related variables are changed by the inclusion of an ability variable. We do not hope to provide definitive statements on these issues, but our results should be of some interest to those working on related problems in the economics of education.

The plan of the paper is as follows: Section I outlines the nature of the data, variables, and methods used in estimation. Sections II and III present the results of the additive and interactive models. Section IV contains a few concluding remarks.

I The Sample, Variables, and Methods

Our sample consists of a cohort of recent Woodrow Wilson Fellows — all males and all elected from 1958 to 1960. Woodrow Wilson Fellowships are awarded to first year graduate students in the arts and sciences — approximately 75 per cent of all awards given to graduate students in the Humanities and Social Sciences and the remainder to those in the Natural Sciences. The selection procedure is carried out

* Mr. Mooney is a Senior Staff Economist with the Council of Economic Advisers. Mr. Ashenfelter is an Instructor in the Economics Department at Princeton University.

The authors would like to thank the members of the Seminar on Research in Progress of the Economics Department at Princeton University for their helpful comments on an earlier version of this paper. William G. Bowen, James Morgan, Stanley Black, and the referees offered helpful comments on an earlier draft. A special thanks is owed to Dr. Hans Rosenhaupt, National Director of the Woodrow Wilson National Fellowship Foundation, for his help in all phases of this study. This work made use of computer facilities supported in part by National Science Foundation Grant NSF-GP579. All errors remain the authors' sole responsibility.

[1] For example, see Gary Becker, *Human Capital* (New York: Columbia University Press, 1964) and J. N. Morgan and M. H. David, "Education and Income," *Quarterly Journal of Economics*, Aug. 1963, 423–437.

[2] The three studies which have come to our attention are: Shane Hunt, "Income Determinants for College Graduates and the Return to Educational Investment," *Yale Economic Essays* (Fall, 1963), 305–357; Giora Hanoch, "Personal Earnings and Investment in Schooling," unpublished Ph.D. dissertation, University of Chicago (1965), and Irene H. Butter, "Economics of Graduate Education: An Exploratory Study," *U. S. Department of Health, Education, and Welfare, Office of Education*, Nov. 1966. As skillful as these studies may be, they cast almost no light on the quantitative importance of different variables in the determination of earnings for people with graduate education. For example, Hunt uses data collected twenty years ago; Hanoch is not able to separate people with the Ph.D. degree from those with the M.A., nor is he able to distinguish different types of graduate education (e.g., Arts and Sciences from Medicine or Law or Business, etc.). Because she uses salary data cross-classified only once, Butter is unable to estimate simultaneously the effects on earnings of a number of relevant control variables.

[3] Stephen Merrett, "The Rate of Return to Education: A Critique," *Oxford Economic Papers*, Nov. 1966, 289–303, has recently criticized studies of the rate of return to education for failing even to consider this particular problem.

nationwide (regional quotas being explicitly used) and aside from the obvious academic qualifications, candidates need only express an interest in college teaching as a future career in order to qualify for a fellowship. In mid-1966, all male Woodrow Wilson Fellows from the three years noted above were asked via a mail questionnaire their first and present annual salaries, i.e., their earnings for the full year. After a number of unusable responses were deleted, 1322 people became the sample group on which the following analyses were conducted.[4]

It should be obvious that our sample is rather homogeneous. All members possess excellent college academic credentials; in addition they were of the same sex, generally the same age, and relatively recent entrants to the labor force. Therefore, since we have implicit controls for those variables, we do not have to control for them explicitly. Homogeneity, and the fact that we are dealing with survey data where individual tastes and preferences may cause large random fluctuations, make it unlikely that a reasonably well specified model will explain a large proportion of the variation in income levels. In view of these considerations, our equations do a very good job of explaining income differentials.

Although both Present Annual Salary (Y_{PSAL}) and First Annual Salary were requested of the respondents, for reasons of space, only the results for Present Annual Salary are presented here.[5] Present Annual Salary refers to the respondent's mid-66 estimate of his total earned income for all of 1966. First Annual Salary was defined as the respondent's total earned income during his first full year of employment. The respondents were also asked to list the dates of their first full year of employment. In the results presented below, the determinants of Present Annual Salary are analyzed in two different ways, — first by means of a general additive model and secondly by means of an interactive model.

The independent variables, all but one of which are expressed in "dummy variable" form, are: Field of Graduate Study (Humanities, Social Sciences, and Natural Sciences),[6] the number of Years of Graduate Schooling (S_1 . . . . . S_5), the Highest Degree held (B.A., Master's, Ph.D., and other, e.g., Divinity or Law degree), Profession (college teacher or other), the number of Years Working (W_0 . . . . . W_5), and the respondents' Mathematics Aptitude Scores on their Scholastic Aptitude Tests.[7]

Glossary of Definitions of Variables Used in the Multiple Regression Analysis [8]

Dependent Variable

	Y_{PSAL}	= present annual salary (expressed in dollars)

Independent Variables

(*R*)	*HUM*	Set of dummies for Field of Graduate Study
	SOCS	= *HUM* = Humanities, *SOCS* = Social Sciences, and
	NATS	*NATS* = Natural Sciences

[4] A few words about the non-respondents are required at this point. Overall, the response rate to the salary questionnaire was 63.1 per cent. Of those who did not respond (939 out of 2550), there were 359 for whom we had little or no information regarding their occupation, degree level, or years in graduate school. In addition there were 218 full-time graduate students among the nonrespondents. Those full-time graduate students who responded to the salary questionnaire and who usually listed their fellowship stipend, were deleted from the sample. Lack of finances prevented the authors from doing a thorough followup investigation of the nonrespondents. However, on the basis of the information at our disposal, we have no reason to believe that any significant biases are introduced by the lack of salary data on the non-respondents.

[5] Regressions similar to those run for Present Annual Salary were run for First Annual Salary. All results for First Annual Salary, with an explanatory text, are available from the authors on request.

[6] The Humanities include the following disciplines: English, Modern Languages, Classics, Philosophy, Musicology, Fine Arts, Speech, American Studies; the Social Sciences include History, Political Science, Economics, Sociology, Psychology, Religion, Area Studies and Anthropology; the Natural Sciences include Mathematics, Physics, Chemistry, Biology, Geology, Astronomy, and Statistics.

[7] The Scholastic Aptitude Test is administered by the Educational Testing Service of Princeton, New Jersey. This test is given on a nationwide basis, and the results are used extensively by American colleges and universities for the purposes of determining the intellectual capability of undergraduate applicants. The test makes no pretensions to being a perfect I.Q. test. Nevertheless, because of its wide use and general acceptability as a measure of intellectual ability, it was considered an ideal variable for our purposes. Aptitude test scores for 1022 male Woodrow Wilson Fellows elected from 1958 to 1960 were made available by the Educational Testing Service. The remaining Fellows had not taken this examination. After matching scores with the Fellows for whom we also had salary data, a sample of 694 Fellows resulted.

[8] The (*R*) in front of a category within each set of dummy variables indicates the reference category.

(R) S_1, S_2, S_3, S_4, S_5 = Set of dummies for Years of Graduate Schooling Completed, with subscripts indicating actual number of years of graduate schooling completed.

(R) D_{BA}, $D_{MAST.}$, $D_{Ph.D.}$, D_{OTHER} = Set of dummies for Degree Level, with subscripts B.A., Mast., Ph.D., and Other, e.g., Law Degree indicating the highest degree held by people in respective categories.

(R) P_{OTHER}, $P_{C.T.}$ = Set of dummies for Profession with subscript C.T. indicating college teacher and subscript other indicating other profession, e.g., law, government work, business, etc.

(R) W_0, W_1, W_2, W_3, W_4, W_5 = Set of dummies for the number of Years Working with subscripts indicating actual number of years working. W_0 (working zero years) indicates a person who started to work in 1966.

$MAPT.$ = Mathematics Aptitude Score on a numerical scale ranging from the lowest possible score of 200 to the highest possible score of 800.

Why were these particular variables included in the estimation of an earnings function? Since one of our aims was to isolate the influence of ability on income differentials from other variables, the use of an ability variable is self-explanatory. We were hypothesizing that a person's ability, as measured by his score on this test, would be positively related to his salary level. Our education variables, Years of Graduate Schooling, and Degree Level, would also be thought to have a positive effect on salary level, i.e., the more Years of Graduate Schooling completed and/or the higher the Degree Level other things being constant, the higher one's salary. However, there were two reasons why we were not overly optimistic about a steady positive net relationship between schooling and income. First, it seems reasonable to suppose that payments for labor are more closely related to degree level than to actual number of years of school completed, especially if we are considering the present salary of those who have had some experience in the labor market.[9] Second, in a sample such as ours, there is bound to be an inverse relationship between years of schooling and years of work and hence some multicollinearity.

Aside from the ability and education variables, a variable depicting Profession was included in our earnings function. This category is subdivided only once, between college teachers and all others. The "other" professional categories include occupations such as lawyer, minister, government civil servant, and research worker outside of a university. These categories simply held too few observations to be included separately. Within the context of our model, the variable, College Teacher, was expected to have a strong negative relationship with income level for reasons which have to do with the well-known economic plight of college teachers relative to their nonteaching peers. As far as Field of Graduate Study is concerned, we expected that natural scientists would fare better income-wise than their counterparts in the Humanities and Social Sciences. One reason for this expectation stems from the fact that a natural scientist usually has more job opportunities outside the university than a humanist or social scientist. Furthermore, these jobs often pay more than a comparable university position. Finally, a variable for the Number of Years Working is included in the Present Salary regressions to take account of labor market experience and the upward drift of salaries through time.

The general model we use for the estimation of the earning's function is a multiple linear regression of the form:

$$Y_j = \sum_{i=0}^{I} b_i X_{ij} + u_j \tag{1}$$

where j stands for an individual; i for a variable; y is the earnings variable; X_0 a variable which always takes the value 1 and X_i ($i = 1, \ldots\ldots, I$) are I explanatory variables (only one of which is continuous); the b_i are parameters to be estimated; u_i is a random distur-

[9] Acquisition of the Ph.D., for example, may really serve as a screening device or as a "union card" for employers. Hence, the actual number of years spent in school may be irrelevant from the employer's point of view. Worse yet, it is quite conceivable that, after holding degree level and age constant, there is some relatively low number of years of graduate school for which income reaches an absolute maximum.

bance with zero mean and constant (unknown) variance. Since each set of dummies is mutually exclusive, we omit one from each set in the estimation (denoted by an R in the glossary). The regression coefficients on the dummy variables can thus be interpreted as the net differences in income due to a person's being in a particular category rather than in the category denoted by an R, after holding other variables constant in the usual multiple regression sense.[10]

In estimating an equation like (1), it is often interesting to examine the intercorrelations among the explanatory variables. When the explanatory variables are continuous, the usual procedure is to inspect the simple correlation coefficients between the variables. When the explanatory variables are dichotomous, as in this study, the above procedure seems neither feasible nor desirable. Although there may be a large number of correlation coefficients to examine, it is not clear how they should be interpreted.

Our resolution of the problem of gauging the magnitude of the intercorrelations among the explanatory variables is to calculate the canonical correlations among the sets of dummy explanatory variables used in our equation.[11] Briefly, this amounts to finding the linear combinations of variables in each set of dummy variables that have maximum correlation. This maximum correlation is called the canonical correlation between the two explanatory factors and is our rough measure of intercorrelations.[12] If one of the two sets of variables being compared consisted of only one variable, the canonical correlation found would be identical with the ordinary multiple correlation coefficient. Thus it may be useful to interpret the relative magnitudes of the canonical correlations in the way one interprets the relative magnitudes of multiple correlation coefficients.

II The Additive Earnings Function

Results for Present Annual Salary: As a first specification we inserted the variables listed in the glossary into equation (1) and estimated regression coefficients for them under the assumption that there were no interactive effects of the explanatory variables on annual salary. Table 1 presents the results of a regression for the entire sample, and table 2 the canonical correlations among the explanatory factors. The ability variable was not included in this run. We wanted to test for the influence of the other variables with as many observations as possible. There is a total of 15 variables and 11 of the coefficients are significant at the 0.01 or 0.05 level. The category with the largest coefficient and also with the highest "t" value is Natural Sciences within the set of variables representing Field of Graduate Study. On the average, a natural scientist receives $1,606 more than a humanist and $728 more than a social scientist.

The coefficients for each of the Years of Schooling categories are significant but it should be clear that caution must be used in interpreting them. A glance at table 2 suggests that the Years of Schooling variables are rather highly correlated with the Degree, Profession, and Years of Work variables. Not unexpectedly the highest canonical correlation in the table, 0.53, is between the Schooling and Work Experience variables. It is satisfying that all those in the schooling categories, greater than one year, receive significantly more than those in the S_1 category. We are inclined at present to attribute the decline in the positive-valued coefficients in the table to the high intercorrela-

[10] For a discussion of various dummy regression techniques, see Emmanuel Melichar "Least-Squares Analysis of Economic Survey Data," *1965 Proceedings of the Business and Economic Statistics Section,* American Statistical Association and Arthur S. Goldberger. Econometric Theory (John Wiley and Sons, 1964), 173–177, 218–231, 248–255.

[11] J. Morgan and J. Sonquist have suggested the use of canonical correlations in this situation in "Problems in the Analysis of Survey Data, and a Proposal," *Journal of the American Statistical Association,* June 1963, 416.

[12] The usual formulation of the canonical correlation problem (although other formulations are perhaps more useful in the context of this paper, they lead to similar results) is as follows: Let X_1 and X_2 be random column vectors whose components are the sets of variables we wish to correlate and α and β corresponding vectors of parameters to be estimated. Then $U = \alpha X_1$ and $V = X_2\beta$ are the two canonical variates we desire. Maximizing the expression for the correlation between U and V with respect to α and β subject to normalization constraints on α and β leads to a characteristic value problem which delivers the desired canonical correlation. See T. W. Anderson, *An Introduction to Multivariate Statistical Analysis* (John Wiley and Sons, Inc., 1958), chap. 12. For a computer program similar to the one the authors used see Cooley and Lohnes, *Multivariate Procedures for the Behavioral Sciences* (John Wiley and Sons, Inc., 1962, chap. 3).

TABLE 1. — DETERMINANTS OF PRESENT ANNUAL SALARY (FOR ENTIRE SAMPLE)

	Set of Independent Variables									
	Field	"t" Value	Years of Graduate Schooling	"t" Value	Degree Level	"t" Value	Profession	"t" Value	Number of Years of Work	"t" Value
Dependent Variable:	+(R) *HUM*		+(R) S_1		+(R) D_{BA}		+ (R) $P_{oth.}$		+(R) W_0	
Y_{PSAL}	+878 *SOCS* (137)	6.4[b]	+1,023 S_2 (354)	2.9[b]	−119 $D_{MAST.}$ (214)	0.55	−1552 $P_{C.T.}$ (164)	−9.4[b]	−61 W_1 (251)	−.24
Intercept Term: 6,957	+1,606 *NATS* (157)	10.2[b]	+981 S_3 (323)	3.0[b]	+970 $D_{Ph.D.}$ (221)	4.4[b]			+265 W_2 (241)	1.1
Mean = 9,527; s = 2,373			+814 S_4 (323)	2.5[a]	−273 $D_{oth.}$ (318)	−.85			+552 W_3 (252)	2.2[a]
			+679 S_5 (324)	2.1[a]					+572 W_4 (270)	2.1[a]
Overall regression:									+1302 W_5 (273)	4.7[b]

Overall regression:
Number of observations: 1322.
Coefficient of multiple determination (R^2) = 0.25.
F value: 27.9.

[a] Significant at the 0.05 level.
[b] Significant at the 0.01 level.

TABLE 2. — CANONICAL CORRELATIONS BETWEEN SETS OF DUMMY VARIABLES: THE ENTIRE SAMPLE (N = 1,322)

Variables	Years of Schooling	Degree Level	Profession	Years Working
Field of Study	.10	.29	.14	.09
Years of Schooling		.37	.36	.53
Degree Level			.43	.32
Profession				.30

tions between the Schooling and Work Experience variables. It should also be noted that none of the differences between adjacent schooling coefficients, except the one difference noted above, is significant.

The degree variables exhibit an interesting pattern. Not unexpectedly, those with Ph.D. degrees make almost $1,000 more than those with only a B.A. degree, other things constant. The coefficient for the Master's degree holders is not significant and indeed has a negative sign. This latter result suggests that the Master's degree has little or no earning power independent of the additional years of graduate schooling involved. Of course, the time period examined herein represents the first few years of lifetime careers. Possibly, over a longer run, those with Master's degrees may begin to earn significantly higher incomes than those with a B.A. only.[13] Alas, the coefficient for College Teachers has the expected sign, is highly significant, and indicates that, on the average, college teachers earn $1,552 less than those in other professions. Finally, except for the W_1 category (working one year), the coefficients for the Years Working categories all have the right sign, and their values move in the expected directions.

At this point, we began experimentation with the explicit inclusion of a variable designed to measure ability. Actually, we had at our disposal four candidate variables to work with, and no real hypotheses about whether any or all might have a significant impact on earnings. One was a dummy variable which took on the value of one if a person had been elected to Phi Beta Kappa as an undergraduate and zero otherwise. This variable usually had the wrong sign in our regressions and its coefficient was never significantly different from zero. Two other insignificant variables were a verbal aptitude score and a mean aptitude score obtained by averaging the verbal and mathematical aptitude scores of the Scholastic Aptitude Test. As can be seen in tables 3 and 4, our fourth candidate variable, Mathematical Aptitude Score, does seem to have a significant impact on income levels. Its coefficient has a value of slightly more than two dollars and is significant at the 0.05 level.[14] Note that this represents a

[13] It should also be noted that it is a Master's Degree in the arts and sciences which is being analyzed here. A Master's Degree in education or business for example may very well have substantial earning power vis-a-vis the B.A.

[14] In an attempt to ascertain the possible existence of diminishing returns to increased aptitude, the Mathematics Aptitude Score was also tried in quadratic form, without success. (The results of the above regression and others reported only partially in this section are available from the authors on request.)

TABLE 3. — DETERMINANTS OF PRESENT ANNUAL SALARY, INCLUDING AN APTITUDE VARIABLE

	Set of Independent Variables											
	Field	"t" Value	Years of Graduate Schooling	"t" Value	Degree Level	"t" Value	Profession	"t" Value	Mathematics Aptitude Score	"t" Value	Number of Years Working	"t" Value
Dependent Variable:	+(R)HUM		+(R)$_1$		+(R)D_{BA}		+ (R) $P_{oth.}$				+(R) W_0	
Y_{PSAL}	+893 SOCS (186)	4.8[b]	+916 S_2 (570)	1.6	−239 $D_{MAST.}$ (292)	−.82	−1,542 $P_{C.T.}$ (232)	6.6[b]	+2.1 MAPT (.82)	2.6[a]	−59 W_1 (316)	−.18
Intercept Team: 6,956	+1,252 NATS (223)	5.6[b]	+1,114 S_3 (527)	2.1[a]	+728 $D_{Ph.D.}$ (295)	2.5[a]					+555 W_2 (305)	1.8
Mean = 9,429 s = 2,307			+721 S_4 (527)	1.4	−641 $D_{oth.}$ (406)	−1.6					+666 W_3 (325)	2.0[a]
			+773 S_5 (531)	1.4							+969 W_4 (348)	2.7[a]
Overall Regression: Number of observations: 694. Coefficient of multiple determination (R^2) 0.26. F Value: 14.7.											+1,623 W_5 (372)	4.3[b]

[a] Significant at the 0.05 level.
[b] Significant at the 0.01 level.

TABLE 4. — CANONICAL CORRELATIONS BETWEEN SETS OF DUMMY VARIABLES: THIS SAMPLE CONTAINS PEOPLE FOR WHOM MATHEMATICS APTITUDE SCORES WERE AVAILABLE (N = 694)

Variables	Degree Level	Years of Schooling	Profession	*MAPT.*	Years Working
Field of Study	.32	.11	.18	.31	.11
Degree Level		.36	.48	.13	.28
Years of Schooling			.37	.11	.51
Profession				.05	.26
MAPT.					.13

difference in income of two dollars for *every point difference in test score.* A difference in test scores of 300 points, for example, implies a $600 income differential — suggesting the quantitative as well as statistical relevance of the aptitude variable.

Table 5 contains the relevant coefficients of a regression using the same sample and the same variables as the regression in table 3 except for the Mathematical Aptitude Score. A comparison of the results in the two tables is interesting for the light it sheds on what happens to the other coefficients in the basic relationship when a relevant ability variable is excluded.[15] As one might expect, the coefficients on six of the seven school related variables decline with the inclusion of an ability variable. What is surprising is that they decline so minutely and that in the case of the important coefficient for Ph.D. holders an *increase* of $50 is actually registered! Even though one must regard these results with a certain caution, one wonders how far it might be possible to extend them, at least in the area of graduate education, if the proper data were available.

There is one bothersome aspect about the preceding experiments with an ability variable that seems to indicate a possibly spurious re-

[15] That is, we are in the unique position where we directly analyze a misspecification error.

TABLE 5. — DETERMINANTS OF PRESENT ANNUAL SALARY (REDUCED SAMPLE)
SET OF INDEPENDENT VARIABLES

	Field	"t" Value	Years of Graduate Schooling	"t" Value	Degree Level	"t" Value	Profession	"t" Value	Number of Years of Work	"t" Value
Dependent Variable:	(R) HUM		(R) S_1		(R) D_{BA}		(R) $P_{oth.}$		(R) W_0	
Y_{PSAL}	+927 SOCS	5.0[b]	+981 S_2	1.7	−300 D_{MAST}	−1.0	−1508 $P_{C.T.}$	−6.5[b]	−81 W_1	−.3
Intercept Term: 8,210	+1412 NATS	6.6[b]	+1183 S_3	2.2[a]	+678 $D_{Ph.D.}$	2.3[b]			+541 W_2	1.8
			+784 S_4	1.5	−677 $D_{oth.}$	−1.66			+621 W_3	1.9
			+802 S_5	1.5					+894 W_4	2.6[b]
Overall Regression: Number of observations: 694 R^2: .25 F value: 15.2									+1535 W_5	4.1[b]

[a] Significant at the 0.05 level.
[b] Significant at the 0.01 level.

sult. Why does *only* the mathematical aptitude score seem to have a significant impact on earnings? One possibility is that there is some unaccounted for interaction between the Mathematics Aptitude Score and a person's Field of Study. Two indications that this is true were present. First, table 4 shows that the (canonical) correlation between the ability variable and a person's graduate Field of Study is higher than any of the others.[16] Second, the one coefficient which does seem to be altered significantly by inclusion of the Mathematical Aptitude variable is that of the Natural Sciences variable. In order to account for possible interaction, the regression reported in table 3 was run for the humanists and social scientists only. The coefficient for the Mathematics Aptitude variable rose from 2.1 to 2.9 and its "*t*" value rose from 2.6 to 3.5. Not only is the aptitude variable still significant but its coefficient is larger and more significant for the humanists and social scientists together than it was for the entire group. Needless to say, our curiosity was aroused by this result and we proceeded to subdivide the sample once more.[17] Exactly the same test was conducted for the humanists and social scientists separately. An interesting result emerges. The coefficient for the Mathematics Aptitude Score of humanists equals 1.3 and has a "*t*" value of 1.6 (significant at the 0.1 level). The same coefficient for the social scientists equals 4.3 and has a "*t*" value of 3.1 (significant at the 0.01 level). Thus, it does not seem likely that the neglect of interaction terms is responsible for a spurious result. We conclude this section by noting: (1) Even when several candidate variables for measuring ability are available, it is not at all obvious which should be selected, and (2) our analysis of specification error shows that omitting a relevant ability variable in the earnings model for this group does not cause serious errors in estimating the impact of education related variables on earnings. Although fully aware of the limited nature of our results, we hope that they may provide some comfort to other economists working in the education area.

[16] At this point, it is instructive to note that although there is only a small increase in the total variance explained by the regression in table 3 over that in table 5, table 4 shows that it is unlikely that this is due to significant intercorrelations between ability and the other variables in the regression.

[17] At this point, of course, what we have done is equivalent to introducing complete interaction between the three variables representing Field of Study and all the other variables depicted in table 3.

III The Interactive Earnings Function

Is the economic value of a Ph.D. really independent of the profession which the Ph.D. holder enters and the field of graduate study which he pursues? Are the salary differences between non-college teachers and college teachers independent of graduate degree and field of study? Are the salary differences for those in different fields of graduate study independent of their professions and degree levels? The additive model employed above assumes that the answers to the above questions are in the affirmative. An interactive model, on the other hand, permits us to group variables in such a way as to allow for possible interactions amongst the independent variables. Such a model was employed by the authors in this study and a brief description follows.

We allowed interaction to take place among the three categories — Field of Study, Degree Level, and Profession. The actual regression included the Schooling and Years of Work variables in additive form and 23 of the 24 cross-product variables formed from the three Field of Study variables, the four Degree Level variables, and the two Profession variables. The dependent variable was Present Annual Salary, and the cross-product dummy variable dropped (reference category) referred to those who held only a B.A., had studied in the Humanities, and were *not* presently College Teachers.

The basic defect, evident in this type of analysis, is the small number of observations in certain cells even when the overall sample equals 1322 observations (with Present Annual Salary as the dependent variable). The statistical significance of the results suffers accordingly.

Table 6 presents selected results obtained from our interactive regression analysis. A few words of explanation are necessary. According to the additive model in table 1, the difference in present salaries for those with Ph.D.'s over those with B.A.'s was $970. This result is included once again in table 6. The table in section (A) of table 6 indicates the salary differentials between those with Ph.D.'s and those with B.A.'s when the Profession and Field of Graduate Study variables are allowed to interact with the Degree Level variable. Now it can be seen that the positive and significant salary difference between Ph.D.'s and B.A.'s reported in the additive model (= $970) is, in fact, the result of the large salary differences enjoyed by Ph.D. social scientists and natural scientists *who are not college teachers* and Ph.D. natural scientists *who are college teachers* — over those with B.A.'s in the corresponding Field and Profession categories.

In turn, the positive and significant salary difference between non-college teachers and college teachers reported in the additive model (= $1,552) is largely the result of the large, positive salary differentials enjoyed by Ph.D. social scientists and Ph.D. natural scientists *who were not college teachers* and B.A. natural scientists *who were not college teachers* — over

TABLE 6. — PRESENT SALARY DIFFERENTIALS FOR SELECTED GROUPS; THE RESULTS OF THE INTERACTIVE MODEL [a]

(A) SALARY DIFFERENTIALS FOR PH.D.'S OVER B.A.'S (PH.D.'S-B.A.'S)

Field	Interactive Model[b] ($) Profession: Non-College Teachers	College Teachers	Additive Model (see table 1)
			+970[d]
Humanities	−395 (1,121)[b]	577 (423)[a]	
Social Sciences	2,616[d] (644)	495 (362)	
Natural Sciences	2,245[d] (678)	2,376[d] (1,026)	

(B) SALARY DIFFERENTIALS FOR NON-COLLEGE TEACHERS OVER COLLEGE TEACHERS ($P_{Oth.}$-$P_{C.T.}$)

Field	Interactive Model[b] Degree Level: B.A.	Ph.D.	Additive Model (see table 1)
			+1,552[d]
Humanities	623 (621)	−349 (1,023)	
Social Sciences	727 (521)	2,848[d] (524)	
Natural Sciences	2,871 (1,174)[d]	2,749 (332)[d]	

(C) SALARY DIFFERENTIALS FOR SOCIAL SCIENTISTS OVER HUMANISTS (*SOCS-HUM*)

Profession	Interactive Model[b] ($) Degree Level: B.A.	Ph.D.	Additive Model (See table 1)
			+878[d]
Non-College Teachers	1,098 (586)	4,109[d] (1,132)	
College Teachers	994 (518)	912[d] (199)	

$N = 1322$
$R^2 = .29$
$F = 16.6$

[a] See text for variables included in the regression underlying these tables.
[b] Estimated standard errors, reported in parentheses, are obtained from the appropriate combinations of the terms in the variance — covariance matrix of the regression coefficients for the equation actually estimated.
[c] Significant at the 0.05 level.
[d] Significant at the 0.01 level.

college teachers in the same Field of Graduate Study and at the same Degree Level.

Finally, it is interesting to note that the salary difference of $878 between social scientists and humanists reported in table 1 is almost uniformly accurate regardless of Degree Level and Profession. Not unexpectedly, the one exception to this statement is the much larger Social Sciences-Humanities differential for the Ph.D. who is not engaged in college teaching.

The results of the interactive model did not *contradict or negate* any of the results obtained from the additive model. The two unexpected negative coefficients for salary differentials between Ph.D.'s and B.A.'s and between non-college teachers and college teachers, listed in table 6, are not significantly different from zero. Thus, accurate answers to general questions concerning salary differentials between Ph.D.'s and B.A.'s and between non-college teachers and college teachers are given by the additive expression. However, if detailed questions are posed which call for the consideration of the effects on salaries of a variety of different variables simultaneously, then some version of an interactive model should be used.

IV Conclusions

At the outset, it was stated that one of our objectives was to indicate the relative importance of factors other than years of education on the earnings stream of a cohort of recent, former graduate students. We feel that this objective has been accomplished. Regardless of the particular model or subsample used, variables such as Profession, Degree Level, and Field of Graduate Study, always explained more of the variance in earnings than Years of Graduate Study. For this particular sample, this result can hardly be considered surprising. An employer, whether a university or not, uses graduate degree and/or field of graduate study as his criteria for establishing salary levels rather than the number of formal years of graduate education. What does this imply for the problem of calculating rates of return to graduate education? On the basis of our results, we conclude that the application of traditional "rate of return" analysis to the area of graduate education in which years of graduate education would be used as the sole education-related variable would be highly misleading.[18]

Another objective of this study was to allow for the possibility of interactive effects amongst the independent variables. Interactions among Field of Graduate Study, Degree Level, and Profession were permitted and evidence of significant interactive effects among these variables was obtained. However, it should be reiterated that the results of the interactive model were consistent with those of the additive model. Thus, we conclude that the importance of taking into account the possibility of interactions is directly related to the specificity of the questions which the analyst wants to answer. For example, if an investigator wants to compute the rate of return for a Ph.D. in the Social Sciences over a B.A., he should distinguish between Ph.D. social scientists who are college teachers and those who are not (see table 6, section A). A final note of caution is called for at this point. Unless one's sample is extremely large, the possibility of reducing the observations in each cell to a statistically useless number becomes a real likelihood as soon as we permit interactions amongst a number of independent variables.

Finally, what can one conclude about the effects of ability on the earnings level of the individuals in this sample? In the first place, it was found that an index of mathematics aptitude was the only candidate variable to have any significant relationship with earnings. Although many possible explanations for such a result came quickly to mind, we have decided not to speculate freely. The reason for this caution is that we are unsure about the meaningfulness of this result in the light of the failure of other relevant ability variables to explain any significant portion of the variation in earnings. Secondly, we found that the inclusion of an ability variable affected the estimates of the coefficients for the other education-related variables only in a very marginal fashion. One tentative conclusion which we draw from this result is that investigators who work with a sample of highly educated people for which there are a number of relevant control variables need not worry unduly about the lack of an ability variable in constructing rates of return. The misspecifications, caused by the absence of an ability variable, seem to be quite small indeed.

[18] Along these lines, the authors are currently undertaking a study of the returns to graduate education which will explicitly control for a number of the variables included in the above analysis. In addition, the implications of equilibrium and dynamic adjustment in particular labor markets will be considered for their relevance to the rate of return type analysis.