THE SOCIAL ORIGINS OF SCIENTISTS IN DIFFERENT FIELDS

Yu Xie

ABSTRACT

Departing from mainstream social stratification and intergenerational mobility research, this paper studies the social origins of incumbents of a particular group of occupations—scientific occupations. With data from two large national surveys, it examines the social origins of scientists in four major fields (physical, biological, mathematical, and social) and compares these to those of nonscientists. Within the log-linear analytical framework, the paper explicitly models the difference in social origin between scientists and nonscientists as well as the interfield variation in social origin among scientists. It is shown that the social origins of scientists are largely homogeneous across different fields after the strong association between farm origin and biological science is controlled. Direct effect of social origin net of higher education is found to be nonmonotonic along the status hierarchy of social origin and thus cannot be attributed to a unidimensional notion of father's socioeconomic status.

Research in Social Stratification and Mobility,
Volume 11, pages 259-279.

ISBN: 1-55938-491-3

INTRODUCTION

From Status Attainment to Occupation Attainment

Researchers of social stratification have long been concerned with how social origin determines social status (e.g., Lipset and Bendix 1964; Blau and Duncan 1967; Sewell and Hauser 1975; Featherman and Hauser 1978). A society is *open* if the association between social origin and social status is weak and *closed* if the association is strong. Typically, social origin is measured by parent's occupation, and social status by one's current occupation.[1] Almost invariably, all studies of industrial societies have found considerable intergenerational mobility (for a review, see Treiman and Ganzeboom 1990). The classical explanation is that the efficiency needs of modern industrialization call for social mobility through the principle of recruiting the most qualified for the most important positions (Lipset and Bendix 1964, Chapter 2; Blau and Duncan 1967, pp. 425-431). That is, industrialization makes universalism necessary and functional.

By universalism I mean the predominant importance of achieved characteristics over ascribed characteristics in determining one's social status. This definition of universalism has been widely used by social stratification researchers. For example, in commenting on the relationship between universalism and social stratification, Blau and Duncan (1967, p. 430) remarked:

> Heightened universalism has profound implications for the stratification system. The achieved status of a man, what he has accomplished in terms of some objective criteria, becomes more important than his ascribed status, who he is in the sense of what family he comes from. This does not mean that family background no longer influences careers. What it does imply is that superior status cannot any more be directly inherited but must be legitimated by actual achievements that are socially acknowledged. Education assumes increasing significance for social status in general and for the transmission of social standing from fathers to sons in particular.

This tradition of separating the mediating effect of education and the direct effect of social origin has persisted in the social stratification literature. The role of education is understood to be twofold. First, it transmits the advantages and disadvantages of parental status. Second, it provides opportunities for children from low-status families to move upward (Hauser 1971, p. 144). In actual mobility processes, the dual role of education is always present. However, the importance of education varies across occupational groups (Yamaguchi 1983). As Hout (1988, p. 1381) put it, "some occupations are almost certainly more universalistic in their recruitment criteria than are others."

Unfortunately, with a few exceptions (e.g., Yamaguchi 1983; Xie 1989b), studies of intergenerational mobility have ignored the varying degrees of universalism in different occupations. In the methodological tradition of path

analysis and structural equation models, as exemplified in Blau and Duncan (1967), Duncan, Featherman, and Duncan (1972), and Sewell and Hauser (1975), occupation is treated as a continuous variable by using Duncan's (1961) Socioeconomic Index (SEI) scores. The typical concern in these studies is with vertical mobility, and the problem of "'channels' of mobility" is ignored, as Blau and Duncan (1967, p. 117) acknowledged in their pioneering study of American occupations. This body of literature may be labelled *status attainment*. In the status attainment literature, to borrow Duncan's (1979, p. 793) characterization, "dependence of destination on origin is represented, say, by a simple or partial regression coefficient." Hence, particularities of individual occupations are disregarded if they have the same or similar SEI scores.

This limitation of path analysis and structural equation models has spurred a renewed interest in the analysis of mobility tables in the framework of log-linear models (Featherman and Hauser 1978; Goodman 1978; Duncan 1979; Hauser 1979; Goodman 1984; Sobel, Hout, and Duncan 1985). In Hauser's (1978, p. 821) words,

> In short, mobility tables are useful because they encourage a direct and detailed examination of movements in the stratification system. Within a given classification they tell us where in the social structure opportunities for movement or barriers to movement are greater or less, and in so doing provide clues about stratification processes which are no less important, if different in kind, from those uncovered by multivariate causal models.

The shift in interest from causal models with Duncan's SEI scores as the dependent variable to log-linear models of mobility tables reflects a growing consensus among social stratification researchers that occupation cannot always be treated as a continuous variable. For certain research interests, occupation should be treated as a discrete variable. Unlike structural equation models, log-linear models of mobility tables take account of the discreteness of occupations and thus render possible the study of mobility processes (Hauser 1978; Duncan 1979) and mobility channels (Yamaguchi 1983).

Theoretically speaking, discrete coding of detailed occupations could be retained in log-linear analyses. In practice, only a few major occupational categories are used because a too detailed classification of occupations makes modeling unfeasible and interpretation difficult. As a consequence, researchers collapse detailed occupations into a few internally homogeneous categories before analytical tools are used; and this process always results in a loss of information, because particularities of detailed occupations are suppressed in such collapsings of occupations. For example, the unique social origins of scientists cannot be revealed from intergenerational mobility tables based on broad occupational categories because scientists are grouped into "white-collar" workers in a 3 × 3 table (Grusky and Hauser 1984), "upper-nonmanual" workers in a 5 × 5 table (Featherman and Hauser 1978, Chapter 4), "professionals and high administrative workers" in a 8 × 8 table (Duncan 1979), and "self-employed professionals" or "salaried professionals" in a 17 × 17 table (Blau and Duncan 1967, Chapter 2).

In this research, I take a different approach, treating detailed occupations as discrete positions taken by individuals. It is recognized that detailed occupations are intrinsically interesting because they differ in more than just one dimension—socioeconomic status. As it is not possible to examine all detailed occupations in one study, the research is limited to scientific occupations as a case study. While social origin is treated in the same way as in earlier studies of mobility tables, occupational destination is highly skewed in amplifying the details of scientific occupations while lumping together all other occupations. That is, the study is based on an asymmetric mobility table, with interest focussed on the social origins of scientists in different fields. I label this type of social stratification research *occupation attainment*. Even though the current study is limited to the social origins of scientists only, the same methodology can be readily extended to studying the social origins of incumbents of other occupations.

Universalism In Science

There are three reasons why scientists are an interesting case to study as an example of occupation attainment. First, scientists enjoy such high social status in modern society (e.g., Hodge, Siegel, and Rossi 1964) that becoming a scientist may be a desired "channel" of upward mobility (Xie 1989a). Second, the interest in the social origins of scientists has a long intellectual history that can be traced back to Galton (1874). Third, in science the universalism thesis in social stratification coincides with Merton's (1942) universalism hypothesis. In brief, Merton's hypothesis states that universalistic (or impersonal) criteria are normally used to evaluate a scientist's performance (Merton 1942). Through the elaboration of Cole and Cole (1973, pp. 65-68), it more specifically means that functionally irrelevant factors should not have a role in the evaluation process in order for science to work efficiently.[2] Merton's universalism hypothesis has two important implications: (1) "truth-claims, whatever their source, are to be subjected to *preestablished impersonal criteria*" (Merton 1942, p. 118, emphasis in original); and (2) "[u]niversalism finds further expression in the demand that careers be open to talents" (Merton 1942, p. 119).

In the past two decades, the sociology of science literature has mostly concentrated on the first implication of Merton's universalism hypothesis and in the meantime neglected the second. Recognition, often measured by citations and awards, "is the functional equivalent to property" (Cole and Cole 1973, p. 45); and institutional backgrounds are taken as analogous to social origin, representing particularistic factors (e.g., Long, Allison, and McGinnis 1979; Hargens and Hagstrom 1982). In this tradition, studies examining Merton's universalism hypothesis are restricted to scientists because they are concerned with the reward system within science. In contrast, the present research is concerned with the second implication of Merton's universalism hypothesis, that is, achieved rather than ascribed characteristics determine social status. It is evident that, unlike earlier studies of the reward system in science, any investigation of recruitment of scientists necessarily requires data on scientists as well as on nonscientists. I hasten to add, however, that results from this study have no direct bearing on the issue of whether the reward system is universalistic or not. It is possible that recruitment into science is universalistic even though the scientific reward system is not so universalistic, and vice versa.[3]

HYPOTHESES

This study attempts to test the following two hypotheses. First, there are no significant differences in social origin between scientists and nonscientists net of education. Second, the social origins of scientists are homogeneous across the different fields. In formulating the first hypothesis, I assume education to be a universalistic factor, either because education truly increases one's working capacity (human capital) or because it is merely used as a set of formal credentials. From the perspective of social relations, formal education is universalistic in the sense that it is the same regardless of involved actors (Parsons 1951, pp. 62-63).[4] It is well known, however, that social background has a strong effect on educational attainment (e.g., Blau and Duncan 1967; Jencks et al. 1972; Featherman and Hauser 1978; Mare 1980). Since educational attainment is a strong predictor of one's likelihood of being a scientist (Xie 1989a), social background should, as previous studies have shown (Visher 1947; Roe 1952; Zuckerman 1977), affect the process of becoming a scientist.

The interest of the present research centers on whether there is an additional advantage in having come from high-status families after controlling for education, which is known to be affected by family background. In other words, like many other studies in social stratification (e.g., Blau and Duncan 1967; Althauser, Spivack, and Amsel 1975) this research is interested in decomposing the total effect of social origin on one's likelihood of being a scientist into the indirect effect mediated by education and the direct effect net of education. If recruitment of scientists is indeed universalistic, all family advantages must be transmitted to education before they can affect one's likelihood of being a scientist. When children from low-status families break the educational barrier, they have equal access to science. If this is the case, science can be said to constitute an ideal "channel of mobility" for the socially disadvantaged who, for whatever reason, have attained higher education.[5]

The underlying theoretical basis for studying interfield variations has traditionally been the distinction between the soft and hard sciences. Hardness means not only a more frequent use of mathematical tools but also a greater degree of impersonality (Storer 1967). Numerous studies (e.g., Zuckerman and Merton 1973; Hargens and Hagstrom 1982; Hargens 1988) have suggested that the universalistic norm might be less true in the soft fields of science than in the hard fields. It is thus plausible that entry into the harder fields of science is more universalistic controlling for education than that into the softer ones, as formal education may matter more in the harder fields than it does in the softer fields.

Empirical evidence in support of this hypothesis is sparse and inconclusive. Knapp (1963) and Davis (1964) report that college students in social science majors come from families of higher socioeconomic status than those in natural science majors.[6] The contrary is found in an earlier study of 800 gifted men (Terman 1954). The present research attempts to provide a more conclusive answer to this question with national survey data and updated statistical techniques.

DATA AND METHODS

Scientists are defined as incumbents of scientific occupations with at least 16 years of formal education.[7] This definition includes natural scientists, social scientists, and college professors and instructors of the natural and social sciences. See the Appendix for occupational titles and occupational codes used by the 1960 U.S. Census.

Empirical research on the social origins of scientists needs to overcome the difficulty that there is only a very small proportion of scientists in the population. Large, representative samples of scientists that are required for rigorous statistical analysis are difficult to obtain. To cope with this problem, Galton (1874) conducted a survey among all established scientists in England of his time. Later researchers have relied on biographic data reported in such reference works as *American Men of Science* (Visher 1947) and *The Dictionary of National Biography* (Merton 1970).

The above two strategies share one limitation: they can be applied only to studies restricted to the population of scientists. This limitation makes these strategies futile in research on the social origins of scientists, where a comparison of scientists with nonscientists is imperative. This in part explains the paucity of research on this topic in recent years, with the exception of works by Pearson and his associates (Pearson and Earle 1984; Pearson 1985), which investigated race-gender variations in the social origins of scientists.

By contrast, the data to be analyzed in this paper come from large national surveys. One is the 1962 postcensal survey on scientists and other technical personnel, called the "1962 Postcensal Survey of Professional and Technical Manpower" (PTM). PTM was conducted by the Bureau of the Census for the National Science Foundation (U.S. Bureau of the Census 1969). The sampling frame was those who were in the 25 percent sample of the 1960 Census and who reported in their Census returns as having scientific and technical occupations. The original sample size was 49,082. The present study, however, is limited to 20- to 64-year-old male respondents who were scientists at the time of the survey and who provided complete data on their occupation, education, and father's occupation. After this restriction, there are 10,234 cases.[8]

Also in 1962, the survey of Occupational Changes in a Generation (OCG) was conducted by the Bureau of the Census as a supplement to the Current Population Survey (Blau and Duncan 1967). OCG and PTM are merged together to form a combined data set (OCG-PTM). Since OCG covered only the U.S. male population, females in PTM are excluded. The exclusion of scientists from the OCG sample is necessary for a convenient treatment and analysis of the OCG-PTM data. There were 139 scientists in a total of 17,618 OCG respondents before the exclusion.

OCG-PTM can be seen as a sample stratified on the discrete variable of outcome: nonscientists and scientists in diverse fields, where scientists were sampled at higher sampling ratios.[9] This is a case of what biometricians call "retrospective" (Farewell 1979) or "case-control" (Prentice and Pyke 1979) samples and what econometricians call "choice-based" (Manski and McFadden 1981) or "response-based" (Xie and Manski 1989) samples.

Analysis based on case control methods can be justified within the log-linear and logit approach (Bishop, Fienberg, and Holland 1975). Successful applications of this technique have already been made in sociology (e.g., Gortmaker 1979). Its desirability lies in the fact that a large number of cases for a particular event can be obtained even though the event under study is rare in the population. Clearly, being a scientist is a rare event. A nationally representative sample covering the entire U.S. population would be an inefficient way to study scientists—few scientists would be found in it. By combining two samples, one for scientists and another one for nonscientists, it becomes possible to effectively study the social origins of scientists in comparison with those of nonscientists.

My analytical method is log-linear analysis of cross-classified tables. For technical details, the reader is referred to consult other sources (e.g., Bishop, Fienberg, and Holland 1975; Goodman 1978, 1984; Duncan 1979; Hauser 1978, 1979; Fienberg 1980; Hout 1983; Agresti 1990).

RESULTS

The respondents from the OCG survey are grouped into two categories of the response variable: nonscientists with less than 16 years of education and nonscientists with 16 or more years of education. Scientists from the PTM survey are aggregated into four major fields of science: physical science, biological science, mathematical science, and social science. It is assumed that scientists within each of the major fields are homogeneous with respect to the analytical models to be reported. These four fields of science are sufficiently dissimilar for variations among fields of science, if any, to be observed. The classification of the major fields of science follows the practice of the National Science Foundation (U.S. Bureau of the Census 1969). The only departure is that professors and instructors of medical science are included here as biological scientists. The 1960 Census occupational titles for scientists in each of the fields are given in the Appendix.

Following earlier research on intergenerational mobility (e.g., Blau and Duncan 1967; Featherman and Hauser 1978; Hauser 1979; Duncan 1979; Yamaguchi 1983), I measure social origin by father's occupation when the respondent was 16 years old. Six major occupational categories are aggregated from three-digit 1960 Census occupational codes. This six-category classification was largely derived from Featherman and Hauser (1978, Chapter 3).[10]

The observed associations between origin and destination are displayed in Table 1. For ease of interpretation, Table 1 shows the inflow rates to the six destinations.[11] The reason for segregating nonscientists into two groups by education (less than 16 years of education versus 16 or more years of education) is to control education. It is well known that origin affects education, and education in turn affects destination. Since the definition of scientists necessarily excludes anyone without 16 years of schooling, it is of interest to separately compare scientists to nonscientists with or without 16 years of education.

It is evident from Table 1 that scientists in all the four fields are far more likely to have come from high-status families than nonscientists without 16 years of education. It is not clear, however, whether scientists have higher social origins than do nonscientists with at least 16 years of education: scientists are more likely to have fathers who were professionals, but less likely to have fathers who were managers, officials, and proprietors, than nonscientists with 16 or more years of education. As scientists are by definition professionals, this effect can be viewed as inheritance in a broad sense. Altogether, about half of scientists (except for biological scientists) come from families with fathers who have nonmanual occupations, as compared to 18.2 percent for nonscientists without 16 years of education and 56.6 percent for nonscientists with at least 16 years of education. Hence, it appears that among respondents with 16 or more years of education the effect of father's occupation on

Table 1. Crude Inflow Rates from Father's Occupation to Destination, 1962

	Destination					
	Nonscience		*Science (Edu. ≥ 16)*			
Father's Occupation[a]	*Edu. < 16*	*Edu. ≥ 16*	*Physical*	*Biological*	*Mathematical*	*Social*
(1) Professional	2.9%	16.4%	21.5%	17.8%	19.5%	21.3%
(2) Managerial	9.0	25.6	19.0	15.3	20.5	22.4
(3) Other Nonmanual	6.3	14.6	11.7	7.8	11.7	12.1
(4) Upper Manual	19.8	15.1	17.2	14.0	15.6	14.2
(5) Lower Manual	28.5	15.6	20.6	18.6	23.3	20.5
(6) Farm	33.4	12.6	10.1	26.5	9.4	9.5
Total[b]	100.0	100.0	100.0	100.0	100.0	100.0
N	15,287	2,192	4,180	2,603	1,055	2,396

Note: [a] Father's occupation is aggregated into six major categories: (1) professional and kindred workers; (2) managers, officials, and proprietors; (3) sales workers, and clerical and kindred workers; (4) craftsmen, foremen, and kindred workers; (5) service workers, operatives and kindred workers, and laborers, except farm; (6) farmers and farm managers, farm laborers and foreman.
[b] Percentages may not add to 100 because of rounding.

Source: OCG-PTM combined file.

destination is weak. This result is consistent with earlier studies that found similar decreases in the effects of family background on educational continuation transitions (Mare 1980) and on occupational placement (Hout 1988) as educational attainment increases. The fact that scientists are less likely to have fathers who were managers, officials, and proprietors than are nonscientists with 16 or more years of education may reflect a difference in value orientation between families with professional fathers and those with managerial fathers. The second group may place more emphasis on "money making" than the first group (Davis 1964).

Table 1 shows that biological science recruits heavily from farmers and farm laborers. The proportion of scientists of farm origin in biological science is about three times as large as in the other sciences. In fact, farm and farm laborers are the modal supplier of biological scientists. The reason for this is that the current definition of biological science includes agricultural science. In the data, there are 806 agricultural scientists among a total of 2,603 biological scientists. Of the agricultural scientists, 52.1 percent are of farm origin. If the 806 agricultural scientists are excluded, the inflow rate from farm origin to biological science drops dramatically from 26.5 percent to 15.1 percent. Therefore, the high inflow rate of biological scientists from farm origin is mainly due to the affinity between farm origin and agricultural science, which represents a kind of inheritance.

As hinted before, Table 1 is a "skewed" intergenerational mobility table. The table can be viewed as a contingency table deriving from a "product-multinomial" sampling scheme (Fienberg 1980, p. 15). Hence, log-linear models

Table 2. Design Matrices

	Destination					
	Nonscience		Science (Edu. ≥ 16)			
Father's Occupation[a]	Edu. < 16	Edu. ≥ 16	Physical	Biological	Mathematical	Social
EDUCATION						
Professional	0	1	1	1	1	1
Managerial	0	1	1	1	1	1
Other Nonmanual	0	1	1	1	1	1
Upper Manual	0	1	1	1	1	1
Lower Manual	0	1	1	1	1	1
Farm	0	1	1	1	1	1
SCIENCE						
Professional	0	0	1	1	1	1
Managerial	0	0	1	1	1	1
Other Nonmanual	0	0	1	1	1	1
Upper Manual	0	0	1	1	1	1
Lower Manual	0	0	1	1	1	1
Farm	0	0	1	1	1	1
BIOLOGICAL SCIENTISTS FROM FARM ORIGIN (BF)						
Professional	0	0	0	0	0	0
Managerial	0	0	0	0	0	0
Other Nonmanual	0	0	0	0	0	0
Upper Manual	0	0	0	0	0	0
Lower Manual	0	0	0	0	0	0
Farm	0	0	0	1	0	0

Note: See Table 1 for a more precise definition of father's occupation.

are appropriate tools for the data so long as the marginal effects of destination are included. The interest here is on the difference in social origin between scientists and nonscientists net of education and the interfield variation in social origin among scientists. To effectively and parsimoniously model the data, I create three design matrices as displayed in Table 2.

The first matrix, Education, is designed to capture the effect of social origin on educational attainment. The second matrix, Science, demarcates scientists from nonscientists while retaining interfield homogeneity among scientists. The third matrix, BF, identifies biological scientists from farm origin as a special group.

Two sets (Panels A and B) of hierarchical log-linear models are presented in Table 3. The models are compared according to two guidelines: (1) the difference between two nested models in their log-likelihood ratio L^2 statistics,

Table 3. Loglinear Models for the Association of Father's Occupation and Destination Controlling for Education

Models Description	L^2	DF	BIC	D
PANEL A				
A1 Origin + Destination	4,752.20	25	4496.26	16.58%
A2 Model A1 + Origin × Education	536.65	20	332.05	3.27%
A3 Model A2 + Origin × Science	439.38	15	285.93	2.49%
PANEL B				
B1 Origin + Destination + BF	4,746.68	24	4501.16	16.48%
B2 Model B1 + Origin × Education	144.84	19	−49.53	1.86%
B3 Model B2 + Origin × Science	37.00	14	−106.21	0.87%

Note: The models are based on data in Table 1. For definitions of the variables, see Tables 1 and 2. L^2 is log-likelihood ratio chi-square statistic with the degrees of freedom reported in colum DF. BIC = L^2 − (DF)log(N), where N is the sample size (27,713). D is the Index of Dissimilarity.

since the difference asymptotically follows a chi-square distribution with the degrees of freedom equal to the difference in degrees of freedom; and (2) the Baysian BIC criterion (Raftery 1986) in order to adjust for a large sample size. The rule is to select the model with the lowest BIC value. When BIC is negative, the null hypothesis is preferred relative to the saturated model. I also use the Index of Dissimilarity (denoted as D) as a purely descriptive measure of misclassification under a model.

As the baseline model, Model A1 assumes the independence between origin and destination. The model fits the data poorly with L^2 = 4,752.20 for 25 degrees of freedom, BIC = 4496.46, and D = 16.58 percent. Model A2 drastically improves goodness-of-fit by allowing origin to affect the attainment of 16 years of education: L^2 is reduced by 4,215.55 for only five degrees of freedom; BIC and D are reduced respectively to 332.05 and 3.27 percent. The improvement in goodness-of-fit of Model A2 over Model A1 demonstrates the indubitable significance of social origin in determining the attainment of 16 years of education. In Model A3, origin is further assumed to affect the destination of being in science (regardless of field). The improvement in goodness-of-fit from Model A2 to Model A3 is not as drastic as it is from Model A1 to Model A2, but nonetheless conspicuous: L^2 is reduced from 536.65 to 439.38 for 5 degrees of freedom; BIC from 332.05 to 285.93; and D from 3.27 to 2.49 percent. However, Model A3 is not an adequate model in describing the data, as it is not preferred to the saturated model (for which L^2 = BIC = 0). It should be noted that the only constraints of Model A3 are concerning scientists. The lack-of-fit of Model A3 suggests the rejection of the hypothesis that the four fields of science are homogeneous with respect to the direct effect of origin.

Further examination of residuals from these models reveals that a large portion of the L^2 statistic is contributed by the strong association between farm origin and biological science. To block out this association, I add the interaction matrix BF in Panel B. The reductions of L^2 statistics from Models A2 and A3 in Panel A to Models B2 and B3 in Panel B are respectively 391.81 and 402.38 for one degree of freedom, highly significant. The BIC statistic falls to −49.53 for Model B2 and −106.21 for Model B3; and similar reductions are evident with D. The above comparisons illustrate the great significance of farm origin in determining one's likelihood of being a biological scientist (or more precisely, an agricultural scientist).

The L^2 statistic is 144.84 for 19 degrees of freedom in Model B2 and 37.00 for 14 degrees of freedom in Model B3. Although they are significant by conventional chi-square tests, these statistics are not overwhelmingly large considering the sample size (27,713). The BIC value is −49.53 in Model B2 and −106.21 in Model B3; thus by the BIC criterion, both Models B2 and B3 are better than the saturated model. Between Models B2 and B3, B3 is preferred. The modified homogeneity hypothesis specified by Model B3 is therefore retained: after the association between farm origin and being a biological scientist is blocked out in Panel B, the four fields of science are shown to be homogeneous with respect to the direct effect of father's occupation.

The estimated interaction parameters measuring the association between origin and destination for Model B3 are presented in Table 4. The first and second columns, each of which has four nonredundant parameters, respectively report the effects of father's occupation on the attainment of higher education and the likelihood of being a scientist. There is one meaningful parameter in the third column contrasting biological scientists from farm origin with other cells in the origin-destination classification. Since the parameters are estimated from a log-linear model, the most straightforward way to interpret them is through log-odds-ratios (LORs). In the current case, an LOR typically refers to the following quantity:

$$\log\{[P(D_j|O_i)/P(D_j'|O_i)]/[P(D_j|O_i')/P(D_j'|O_i')]\},$$

where O_i and O_i' are any two possible origin categories, and D_j and D_j' any two possible destination categories. The ratio of two conditional probabilities, for example, $P(D_j|O_i)/P(D_j'|O_i)$ is called the odds. For example, the LOR of attaining 16 years of education between respondents with professional fathers and respondents from farm origin is 2.698. This estimate means that the odds of attaining 16 years of education for respondents with professional fathers are exp(2.698)=14.84 times those from farm origin. Likewise, the odds of being a scientist for respondents with professional fathers are exp (0.514 − 0.002) = 1.67 times those with managerial fathers.

Table 4. Estimated Parameters for the Association between Father's Occupation and Destination (Model B3, Table 3)

	Destinatiopn					
	Education		*Science*		*Biological Sci. (BF)*	
Father's Occupation	*Parameter*	*(S.E.)*	*Parameter*	*(S.E.)*	*Parameter*	*(S.E.)*
Professional	2.698	(0.094)	0.514	(0.091)	0.000*	
Managerial	2.021	(0.079)	0.002	(0.085)	0.000*	
Other Nonmanual	1.811	(0.089)	−0.004	(0.095)	0.000*	
Upper Manual	0.704	(0.085)	0.323	(0.093)	0.000*	
Lower Manual	0.372	(0.083)	0.562	(0.091)	0.000*	
Farm	0.000*		0.000*		1.199	(0.059)

Note: * Indicates reference groups in normalizing each column of interaction parameters. See Tables 1 and 2 for definitions of the variables and Table 3 for the goodness-of-fit statistics of the model.

From the third column of Table 4, it is calculated that being of farm origin increases the odds of being a biological scientist by a factor of exp(1.199) = 3.32. This is additional to the general effect of social origin on being a scientist (the second column). The estimated BF parameter is large and statistically significant. And it establishes the strong association between farm origin and biological science in such a way that sons of farm origin have much higher odds of being biological scientists than sons of other origins.

The arrangement of the six categories of father's occupation as in Table 1 forms an ordinal scale, with socioeconomic status (Duncan 1961) being the highest for professionals to the lowest for farmers and farm laborers (Duncan 1979). The effect of father's occupation on attaining 16 years of education is estimated to be monotonic along this scale: the odds of attaining 16 years of education are the highest for those with professional fathers and decline gradually as father's occupation declines in socioeconomic status. In contrast, the effect of father's occupation on being a scientist does not exhibit a systematic pattern along the socioeconomic scale. Sons of professional fathers have the highest odds of being scientists, followed by sons of upper- and lower-manual workers. Surprisingly, sons of fathers in managerial and other nonmanual occupations have similarly small odds of being scientists, as do sons of fathers in farming occupations (except that the last group is more likely to be biological scientists). Clearly, the direct effect of father's occupation on one's likelihood of being a scientist is nonmonotonic along the status hierarchy of social origin and thus cannot be attributed to a unidimensional notion of father's socioeconomic status. This finding challenges Blau and Duncan's (1967, p. 74) general statement that "socioeconomic status is a fundamental dimension of the social distance between occupational groups."

It should be emphasized that the nonmonotonic effect of father's occupation on son's likelihood of being a scientist reported in Table 4 is conditional on

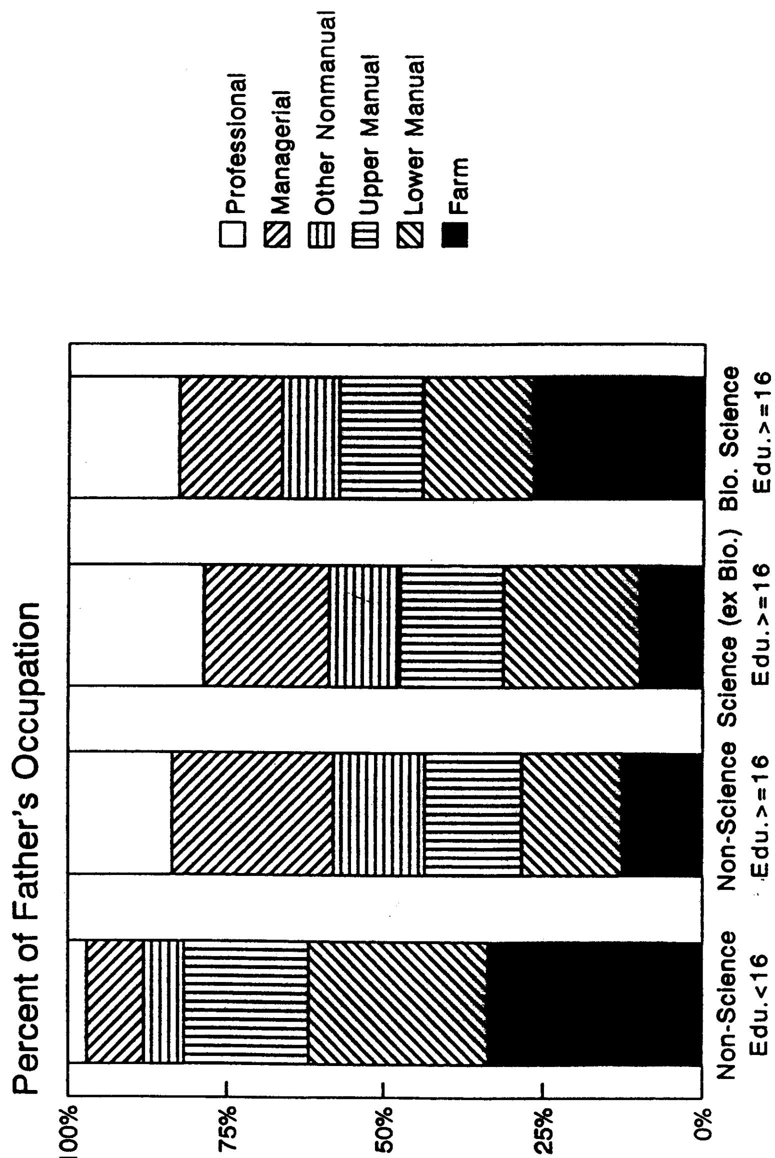

Figure 1. Smoothed Inflow Rates From Father's Occupation to Destination, Based on Model B3, Table 2

son's attainment of 16 or more years of education. It is evident from Table 1, however, that scientists are more likely to have fathers in managerial and other nonmanual occupations than are nonscientists without 16 years of education. In other words, higher social origin is conducive to scientific careers only insofar as it raises children's educational attainment. Given 16 or more years of education, a higher social origin is not necessarily associated with a higher likelihood of a scientific career.

The large and persistent effect of father being a professional net of the mediation of education suggests that entries into science may be related to factors that may be best described as "taste" or "preference." Scientists are by definition professionals. Sons of professional fathers may be more likely to be scientists through many possible channels of influences: (1) an early exposure to scientific careers; (2) an appreciation of scientific and intellectual work through socialization of values; and (3) at the cognitive level, a better understanding of science through interacting with and learning from parents. I believe that the same explanation applies to the strong association between farm origin and biological science.

How different are the social origins of scientists from those of nonscientists? To answer this question, I present the predicted compositions of father's occupation for all destinations under Model B3 in Figure 1. The predictions, easier to interpret than log-linear parameters, should be viewed as translations of the estimated parameters of the model, some of which are reported in Table 4. It should be noted that the predicted percentages for the two destinations of nonscience are the same as the observed ones reported in Table 1, as Model B3 essentially puts no constraints on nonscientists. The difference between the biological science and other fields of science is due to the BF parameter representing the strong relationship between biological science and farm origin. Comparing the different bars in Figure 1 gives a visual image of the heterogeneity in social origin across different modes of destination. It is seen that the difference between scientists and nonscientists with at least 16 years of education is much smaller than the difference between nonscientists without 16 years of education and the other groups all with 16 years of education.

DISCUSSION AND CONCLUSIONS

I have shown that social origin affects the recruitment process into science largely through the indirect effect of social origin on educational attainment. For respondents with 16 or more years of schooling, the direct effect of social origin is relatively small and cannot be attributed to father's socioeconomic status. The log-linear analysis lends partial support to the hypothesis that science is equally open to people from all kinds of family background if they complete a college education. Science is universalistic or meritocratic in that it allows individual performance (such as educational attainment) to be the sole criterion after a threshold is reached. The disadvantages of having come from low-status families are mediated by education. This is not so much the case with other high-status occupations, such as managers, officials, and proprietors (Yamaguchi 1983; Xie 1989b).

The above conclusion is subject to an alternative interpretation that the diminishing importance of social origin among the highly educated is due to population heterogeneity in unmeasured yet relevant characteristics (such as ability and interest in science). As pointed out by Mare (in press), such unobserved heterogeneities could, through selective attrition, result in suppressing the true effects of family background as educational attainment increases. In the following, I will follow the example of Althauser, Spivack, and Amsel (1975, pp. 155-158) and discuss possible consequences of selectivity in unmeasured characteristics.

The first column of Table 4 clearly demonstrates the positive effect of father's socioeconomic status on attaining 16 years of education. In addition to family resources, certain unmeasured qualities such as innate abilities, which in my opinion are more or less evenly distributed across class boundaries, also affect the attainment of higher education. It can then be inferred that among all college graduates sons of lower-class fathers are more likely to possess these desirable qualities than sons of higher-class fathers, as the former group is more selective due to their lower rate of attaining college graduation. Statistical bias resulting from such unequal distributions of unmeasured characteristics is usually called heterogeneity bias. In the current case, the continued significance of the unmeasured qualities means that sons of lower-class fathers with a college degree should have higher chances of entering the scientific profession than do sons of higher-class fathers with a similar education. That is, everything else being equal, heterogeneity effect induces a negative effect of father's occupation on entry into science among college graduates. Most of the estimated parameters reported in the second column of Table 4 agree with this expectation with the notable exception of the effects of father's occupation being a professional and father's occupation being a farmer (or a farm laborer). If the heterogeneity hypothesis holds true, it still cannot explain all the findings reported here.

The findings point to the importance of early socialization. Beyond those mediated through higher education, father's influences on son's likelihood of being a scientist are likely to take the form of cultivating interests in scientific careers at an early age. This explanation is partially supported by the notion that there is a strong effect of "inheritance" in intergenerational mobility (for a review, see Hout 1983). In the present analysis, the strong direct effects of father being a professional on a scientific career and of father being a farmer (or a farm laborer) on a career in biological science suggest that socialization plays a role in occupational choice. Sons of professional fathers may be exposed

to science and encouraged to pursue scientific careers more than sons of incumbents in other occupations. Farm origin is considered a socialization factor for biological (or agricultural) science because children growing up on farms have first-hand contacts with living things as well as with the farm production process.

The comparison of four fields of science—physical science, biological science, mathematical science, and social science—does not support the hypothesis that recruitment into a softer field of science is less universalistic than that into a harder field. Instead, it is found that all of the fields are homogeneous with respect to the composition of the social origins of scientists. The only exception to this generalization is that there is a strong inheritance-like association between farm origin and being a biological scientist. After this is taken into account, all scientists can be treated as a single group in the modelling of the origin-destination associations. This finding calls for more extensive research employing multivariate regression analysis. More explanatory variables such as parent's education should be considered to differentiate scientists from nonscientists.

APPENDIX

1960 Occupational Census Codes and Occupational Titles for Scientists

Code	Occupational Title
Physical Scientists	
021	Chemists
034	Professors and instructors, chemistry
041	Professors and instructors, geology and geophysics
045	Professors and instructors, physics
052	Professors and instructors, natural sciences, n.e.c.
134	Geologists and geophysicists
140	Physicists
145	Miscellaneous natural scientists
Biological Scientists	
031	Professors and instructors, agricultural sciences
032	Professors and instructors, biological sciences
043	Professors and instructors, medical sciences
130	Agricultural scientists
131	Biological scientists
Mathematical Scientists	
042	Professors and instructors, mathematics
051	Professors and instructors, statistics

135 Mathematicians
174 Statisticians and actuaries
Social Scientists
035 Professors and instructors, economics
050 Professors and instructors, psychology
053 Professors and instructors, social sciences, n.e.c.
172 Economists
173 Psychologists
175 Miscellaneous social scientists

ACKNOWLEDGMENTS

This research was supported by a grant from the National Science Foundation and a Spencer Fellowship from the National Academy of Education. The author is grateful to Robert Althauser, Chris Bettinger, Clifford Clogg, Warren Hagstrom, Lowell Hargens, Robert Hauser, Scott Long, Robert Mare, Arthur Sakamoto, Lawrence Wu, and Kazuo Yamaguchi for advice and comments on earlier drafts of this paper. The opinions expressed herein are those of the author.

NOTES

1. As most studies have been concerned only with male workers, social origin and social status are often represented by father's occupation and son's occupation.

2. The universalism hypothesis is about the normative behavior of scientists. Admittedly, deviant cases are not uncommon in science. But the fact that they are considered by the scientific community as "deviant" is evidence for the claim that universalism is a norm in science. For a discussion of deviant behavior in science, see Hagstrom (1965). A well-known example of deviation from universalism in science is the forced exodus of Jewish scientists from Nazi Germany, an event that strengthened the scientific enterprise in the United States during and after World War II.

3. I thank Lowell Hargens for suggesting this to me.

4. It is not clear how Parsons' definition of universalism influenced Merton's, and vice versa. Merton added a citation to Parsons (1951) when Merton's (1942) original article was reprinted in a collected volume (Merton 1973).

5. This hypothesis may account for the overrepresentation of persons of Jewish background and Asians among American scientists, a phenomenon that has been frequently reported (Cole and Cole 1973; National Science Foundation 1986). Blalock (1967, pp. 99-100) suggests that the unusual degree of competition explains the absence of discrimination against Jews in American academic science.

6. In this connection, however, the reported figures are ambiguous because a large proportion of undergraduates in social science majors, especially those from high-status families, intend to pursue their careers as lawyers and other nonscientific occupations after graduating from college.

7. For a recent review of various definitions of scientists, see Citro and Kalton (1989).

8. The reduction of sample size is mostly due to the exclusion of (1) all females, and (2) males who were engineers, technicians, and teachers of primary and secondary education.

9. The part of the PTM survey used in the analysis is itself a stratified sample with thirteen occupational groups as sampling strata. However, I will treat the sample as composed of four broad strata (physical science, biological science, mathematical science, and social science) on the assumption that observations within each of these four categories are homogeneous with respect to the analytical models.

10. The major departure from Featherman and Hauser's scheme is that their "upper nonmanual" and "lower nonmanual" are regrouped into "professionals," "managers, officials, and proprietors," and "other nonmanual." Another technical departure is that the current classification is solely based on father's occupation, since father's industry and class of worker, necessary for Featherman and Hauser's classification scheme, are not available in the PTM survey.

11. The inflow rates are appropriate because the marginal distribution of destination reflects, rather than the actual distribution in the population, the particular sampling scheme, which is disproportionately stratified by type of destination.

REFERENCES

Agresti, A. 1990. *Categorical Data Analysis*. New York: Wiley.

Althauser, R.P., S.S. Spivack, and B.M. Amsel. 1975. *The Unequal Elites*. New York: Wiley.

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Blalock, H.M., Jr. 1967. *Toward a Theory of Minority-Group Relations*. New York: Wiley.

Blau, P.M., and O.D. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Citro, C.F., and G. Kalton, eds. 1989. *Surveying the Nation's Scientists and Engineers: A Data System for the 1990s*. Washington, DC: National Academy Press.

Cole, J. and S. Cole. 1973. *Social Stratification in Science*. Chicago: University of Chicago Press.

Davis, J.A. 1964. *Great Aspirations: The Graduate School Plans of America's College Seniors*. Chicago: Aldine.

Duncan, O.D. 1961. "A Socioeconomic Index for All Occupations." Pp. 109-138 in *Occupations and Social Status*, edited by A.J. Reiss, Jr. New York: Free Press.

________. 1979. "How Destination Depends on Origin in the Occupational Mobility Table." *American Journal of Sociology* 84:793-803.

Duncan, O.D., D.L. Featherman, and B. Duncan. 1972. *Socioeconomic Background and Achievement*. New York: Academic Press.

Farewell, V.T. 1979. "Some Results on the Estimation of Logistic Models Based on Retrospective Data." *Biometrika* 66:27-32.

Featherman, D.L., and R.M. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.

Fienberg, S.E. 1980. *The Analysis of Cross-Classified Categorical Data* (2nd ed.). Cambridge, MA: MIT Press.

Galton, F. 1874. *English Men of Science: Their Nature and Nurture*. London: Macmillan.

Goodman, L.A. 1978. *Analyzing Qualitative/Categorical Data: Log-Linear Models and Latent-Structure Analysis*. Cambridge, MA: Abt Books.

________. 1984. *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge: Harvard University Press.

Gortmaker, S.L. 1979. "Poverty and Infant Mortality in the United States." *American Sociological Review* 44:280-297.

Grusky, D.B., and R.M. Hauser. 1984. "Comparative Social Mobility Revisited: Models of Convergence and Divergence in 16 Countries." *American Sociological Review* 49:19-38.

Hagstrom, W.O. 1965. *The Scientific Community*. New York: Basic Books.

Hargens, L.L. 1988. "Scholarly Consensus and Journal Rejection Rates." *American Sociological Review* 53:139-151.

Hargens, L.L., and W.O. Hagstrom. 1982. "Scientific Consensus and Academic Status Attainment Patterns." *Sociology of Education* 55:183-196.

Hauser, R.M. 1971. *Socioeconomic Background and Educational Performance*. Washington, DC: Rose Monograph Series, American Sociological Association.

———. 1978. "A Structural Model of the Mobility Table." *Social Forces* 56:919-953.

———. 1979. "Some Exploratory Methods for Modeling Mobility Tables and Other Cross-Classified Data." *Pp. 413-458 in Sociological Methodology*, edited by K.F. Schuessler. San Francisco, CA: Jossey-Bass.

Hodge, R.W., P.M. Siegel, and P.H. Rossi. 1964. "Occupational Prestige in the United States, 1925-63." *American Journal of Sociology* 70:286-302.

Hout, M. 1983. *Mobility Tables*. Beverly Hills, CA: Sage.

———. 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93:1358-1400.

Jencks, C., M. Smith, H. Acland, M.J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson. 1972. *Inequality*. New York: Basic Books.

Knapp, R.H. 1963. "Demographic Cultural and Personality Attributes of Scientists." Pp. 205-216 in *Scientific Creativity*, edited by C.W. Taylor and F.Barron. New York: Wiley.

Lipset, S.M., and R. Bendix. 1964. *Social Mobility in Industrial Society*. Berkeley, CA: University of California Press.

Long, J.S., P.D. Allison, and R. McGinnis. 1979. "Entrance into the Academic Career." *American Sociological Review* 44:816-830.

Manski, C.F., and D. McFadden. 1981. "Alternative Estimators and Sample Designs for Discrete Choice." Pp. 2-50 in *Structural Analysis of Discrete Data with Econometric Applications*, edited by C.F. Manski and D. McFadden. Cambridge, MA: MIT Press.

Mare, R.D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75:295-305.

———. In press. "Educational Stratification on Observed and Unobserved Components of Family Background." Forthcoming in *A Comparative Study of Educational Stratification in Fourteen Countries*, edited by Y. Shavit and H.P. Blossfeld. Boulder, CO: Westview Press.

Merton, R.K. 1942. "A Note on Science and Democracy." *Journal of Legal and Political Sociology* 1:115-126.

———. 1970. *Science, Technology and Society in Seventeenth-Century England*. New York: Harper & Row.

———. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*, edited by N.W. Storer. Chicago: University of Chicago Press.

National Science Foundation. 1986. *Women and Minorities in Science and Engineering*. NSF 86-301. Washington, DC: National Science Foundation.

Parsons, T. 1951. *The Social System*. New York: Free Press.

Pearson, W., Jr. 1985. *Black Scientists, White Society, and Colorless Science: A Study of Universalism in American Science*. New York: Associated Faculty Press.

Pearson, W., Jr., and J.R. Earle. 1984. "Race-Gender Variations in the Demographic Characteristics of Doctoral Scientists." *Sociological Spectrum* 4:229-248.

Prentice, R.L., and R. Pyke. 1979. "Logistic Disease Incidence Models and Case-Control Studies." *Biometrika* 6:403-411.

Raftery, A.E. 1986. "Choosing Models for Cross-Classifications (Comment on Grusky and Hauser)." *American Sociological Review* 51:145-146.

Roe, A. 1952. *The Making of a Scientist*. New York: Dodd, Mead.

Sewell, W.H., and R.M. Hauser. 1975. *Education, Occupation, and Earnings*. New York: Academic Press.

Sobel, M.E., M. Hout, and O.D. Duncan. 1985. "Exchange, Structure, and Symmetry in Occupational Mobility." *American Journal of Sociology* 91:359-372.

Storer, N.W. 1967. "The Hard Science and the Soft: Some Sociological Observations." *Bulletin of the Medical Library Association* 55:75-84.

Terman, L.M. 1954. "Scientists and Nonscientists in a Group of 800 Gifted Men." *Psychological Monographs: General and Applied* 68, No. 7 (Whole No. 378):1-44.

Treiman, D.J., and H.B.G. Ganzeboom. 1990. "Cross-National Comparative Status-Attainment Research." Pp. 105-127 in *Research in Social Stratification and Mobility*, Vol. 9, edited by A. Kalleberg. Greenwich, CT: JAI Press.

U.S. Bureau of the Census. 1969. *Characteristics of America's Engineers and Scientists: 1960 and 1962*. Technical Paper No. 21. Washington, D.C.: U.S. Government Printing Office.

Visher, S.S. 1947. *Scientists Starred: 1903-1943 in "American Men of Science."* Baltimore, MD: The Johns Hopkins Press.

Xie, Y. 1989a. "The Process of Becoming a Scientist." Doctoral dissertation, University of Wisconsin, Madison.

———. 1989b. "Structural Equation Models for Ordinal Variables: An Analysis of Occupational Destination." *Sociological Methods and Research* 17:325-352.

Xie, Y., and C.F. Manski. 1989. "The Logit Model and Response-Based Samples." *Sociological Methods and Research* 17:283-302.

Yamaguchi, K. 1983. "The Structure of Intergenerational Occupational Mobility: Generality and Specificity in Resources, Channels, and Barriers." *American Journal of Sociology* 88:718-745.

Zuckerman, H. 1977. *Scientific Elite: Nobel Laureates in the United States*. New York: The Free Press.

Zuckerman, H., and R.K. Merton. 1973. "Age, Aging, and Age Structure in Science." Pp.497-559 in *The Sociology of Science*, by R.K. Merton, edited by N.W. Storer. Chicago: University of Chicago Press.