Fraternal Resemblance in Educational Attainment and Occupational Status

Robert M. Hauser, Peter A. Mossel

Fraternal Resemblance in Educational Attainment and Occupational Status

Robert M. Hauser
University of Wisconsin—Madison

Peter A. Mossel
School of Public Health, Columbia University

This article develops simple structural equation models of the regression of occupational status on schooling in a sample of 518 Wisconsin high school graduates and their brothers. The models correct for response variability and incorporate a family variance component structure. Methodological complications follow from the facts that the sample consists of sibling pairs; that members of a cohort of high school graduates, rather than their families, are the sampling units; and that the primary respondents are informants about some of the characteristics of their brothers. The regression of occupational status on educational attainment is relatively insensitive both to response variability and to the specification of common family factors. Family membership accounts for about half the variance in schooling and more than one-third of the variance in occupational standing, but there is little evidence that failure to control family background leads to upward bias in estimates of the effect of schooling on occupational status.

Effects of family background on social and economic achievement are not well specified by the parental, familial, and contextual variables that usually appear in multivariate models of the stratification process. That

1 This research was carried out with support from the Spencer Foundation, the National Science Foundation (SES-80-10640), the National Institute of Mental Health (MH-06275), and the Graduate School of the University of Wisconsin—Madison. Computations were performed using facilities of the Center for Demography and Ecology at the University of Wisconsin—Madison, which are supported by the Center for Population Research of the National Institute for Child Health and Human Development (HD-5876). This revision was prepared while the first author was visiting professor at the University of Bergen, Norway. We thank William T. Bielby, Richard T. Campbell, Brian R. Clarridge, Robert D. Mare, Christopher S. Jencks, Michael R. Olneck, and William H. Sewell for helpful advice. The opinions expressed herein are those of the authors. Requests for reprints should be sent to Robert M. Hauser, Department of Sociology, University of Wisconsin, 1180 Observatory Drive, Madison, Wisconsin 53706.

© 1985 by The University of Chicago. All rights reserved.
0002-9602/86/9103-0007$01.50

AJS Volume 91 Number 3 (November 1985): 650–73
Fraternal Resemblance

is, explicit measures of background do not fully reflect the common influences of the family of orientation on schooling and adult achievement, so effects of schooling will be overestimated. For example, Hauser and Featherman estimate that in 1973 a little more than half of the resemblance in educational attainment between American men and their oldest brothers could be explained by measured factors of social background: father's education, father's occupational status, number of siblings, broken family, farm origin, southern birth, Spanish origin, and race (1976, pp. 116–18).

Many social and psychological factors in achievement are poorly represented by measured background variables. Siblings have a partly overlapping genetic heritage, and excepting the possibility of temporal change within the family of orientation, siblings share a set of parents and other relatives (including one another) with whom they each interact in ways that only partly reflect the social and cultural divisions in the larger society. There are also other parts of the social environment, such as neighborhood and community, that do not involve the functioning of families in a narrow sense, but whose nature and influence vary from family to family.

Sociologists and economists have long recognized the importance of measuring the effects of schooling. Its influence on such measures of success as occupational status and earnings serves on the one hand as an indicator of the role of educational institutions in fostering (or hampering) social mobility and on the other hand as an indicator of the productivity of personal and public investments in schooling. At the same time, it is well known that social and economic success may depend directly on personal characteristics and conditions of upbringing that also affect the length and quality of schooling. For these reasons, it is by no means obvious that an association of schooling with social or economic success can be interpreted in causal terms, and many studies have attempted to determine the degree to which such causal inferences are warranted.²

The effects of background, broadly conceived, on achievement can be taken into account by modeling the similarity of siblings. This has helped to motivate a number of studies of the stratification process that are based on samples of siblings, rather than of the general population, beginning with Blau and Duncan (1967, pp. 316–28) and most notably within the two major studies by Jencks and his associates (Jencks et al. 1972, 1979). Olneck estimates regressions of occupational status on years of schooling in several samples of American men, controlling for family background and ability (1979, pp. 168–69). His analysis separates effects of schooling before and after high school graduation, and he finds large biases in

² See, e.g., Griffin (1976) and Olneck (1979).
effects of elementary and secondary schooling. For example, among men with brothers in the 1962 Occupational Changes in a Generation (OCG) sample, the slope of occupational status on early schooling declined by 22%, from 2.541 to 1.980 points on the Duncan (1961) SEI when eight measured background factors were controlled; it declined by an additional 11% to 1.699 when an unmeasured family education factor was specified. Lesser biases were estimated in the effects of postsecondary schooling. For example, among brothers in the 1962 OCG study the effect of four years of college fell only from 29.7 SEI points to 27.5 points when measured background was controlled and to 25.1 points when a family factor was specified. Even smaller biases in the effects of postsecondary schooling on occupational status are estimated in Olneck’s (1976, 1977) survey of Kalamazoo brothers.

Griliches (1979) notes a potentially significant methodological twist in the use of sibling-based research designs (see also Griliches 1977). In a regression, say, of occupational status on schooling, random response variability in schooling will lead to more downward bias in the within-family estimator than in a naive regression that ignores family effects. This occurs because response variability necessarily occurs within individual responses, so a given component of unreliable variance in schooling is larger relative to within-family variance than to total variance. Thus, the biases attributable to omitted background variables and to response variability are probably opposite in effect, and it is necessary to correct both at the same time.

In the late 1960s, little was known about the sensitivity of estimated parameters of models of the stratification process to response variability. Bowles’s (1972) suggestion that retrospective proxy reports of parents’ status characteristics are especially prone to error stimulated several validation studies (also see Bowles and Nelson 1974; Bowles and Gintis 1976); these are reviewed by Hauser, Tsai, and Sewell (1983). Contrary to Bowles’s expectation, improved control of response variability has not led to massive downward revisions in estimates of the effects of schooling on occupational or economic success (Bishop 1974; Bielby, Hauser, and Featherman 1977; Bielby and Hauser 1977). Moreover, use of a sibling-based research design renders moot the question whether social background variables have been measured accurately. At the same time, Griliches’s argument makes it all the more important to correct for response variability in within-family regressions of adult success on schooling. For pioneering efforts, see Jencks et al. (1972, app. B) and Olneck (1976, pp. 166–98).
ment and occupational status for 518 male, Wisconsin high school graduates and a random sample of their brothers to develop and interpret skeletal models of the regression of occupational status on schooling that correct for response variability and incorporate a family variance component structure. We term these models “skeletal” because they do not include explicit socioeconomic background variables, mental ability and other social psychological variables, or outcomes of schooling other than occupational status. Methodological complications follow from the facts that the sample consists of sibling pairs; that primary respondents, rather than families, are the sampling units; and that primary respondents served as informants about some of the characteristics of their brothers.

Although the present analysis focuses on very simple models of sibling resemblance in educational attainment and occupational status, these models are not merely of methodological interest. First, their parameters describe the heterogeneity of status and schooling within and between families. Second, the models can be used to test hypotheses about “family bias” in the regression of occupational status on schooling; that is, they can tell us to what degree the association between schooling and occupational status is an artifact of common causation by elements of family background that are common to siblings. Third, although the models can readily be elaborated to include other variables, their exclusion does not affect the validity of the present analyses. For example, explicit measurements of social background have been left out of the analysis. Although it is interesting and important to add such variables to the model, their inclusion helps to specify the content of common family influences without affecting estimates of family bias. The models may also be elaborated to include additional causes or consequences of schooling that vary both within and between families. For example, it is desirable to measure biases in the effects of schooling that may be attributed to ability and motivation, as well as to common family factors, and it is desirable to specify the effects of schooling on earnings as well as on occupational status. These are important matters, but they are not the subject of this article.

In addition, the issues addressed here also occur in larger models and are closely paralleled in other areas of sociological research. For example, similar issues arise in analyses of neighborhood effects (Bielby 1981), husband-wife interaction (Thomson and Williams 1982), fertility (Clarridge 1983), and political identification (Jennings and Niemi 1981, chap. 4) and—quite generally—in analyses of change over time (Jöreskog and

---

4 Although our findings pertain only to brother pairs in the Wisconsin sample, Hauser (1984b) validates them using Wisconsin sister pairs and sister-brother pairs and Olneck’s (1976) Kalamazoo brother sample.
It could even be argued that models of the present form should supplant the analysis of covariance as the standard model for analyses of contextual effects (Boyd and Iversen 1979); one practical advantage is that many models are identified with only two observations per family, organization, or other unit of aggregation.

Despite the findings of Olneck (1979), we expect to find some evidence of omitted variable bias in the Wisconsin sample. For example, among 2,069 men with nonfarm origins who were working in 1964, the slope of occupational status on years of schooling declined from 8.65 SEI points per year to 8.13 points when four socioeconomic background variables were controlled and to 7.45 points when mental ability was also controlled (Sewell and Hauser 1975, pp. 72, 81). Among 1,789 men for whom high school grades, significant other's influence, and aspirations had also been ascertained, controls for ability, background, and these additional variables reduced the slope from 8.50 points to 6.12 points (Sewell and Hauser 1975, pp. 93, 98). Furthermore, with eight socioeconomic background variables controlled among 3,411 male respondents in the 1975 Wisconsin follow-up survey, the biases were 5.5% for status of the first occupation and 11.1% for status of current occupation (Sewell, Hauser, and Wolf 1980, pp. 571, 581). In a more elaborate social psychological model, Sewell et al. (1980, pp. 571, 581) estimated biases of 13.7% in the case of first full-time civilian occupation and of 32.9% in the case of current occupation.5

Following a brief description of the Wisconsin data, the second section of the article specifies a structural model with distinct regressions of occupational status on schooling for families, primary respondents, and brothers. The sampling of brothers through respondents in the Wisconsin study leads to an interesting problem of identification. After proposing a solution to the identification problem, this section of the paper compares within- and between-family structural regressions based on self-reports of educational attainment and occupational status.

The third section of the article combines the model of family resemblance with a multiple-indicator measurement model for educational attainment and occupational status. The measurement model reflects the facts that primary respondents served in some cases as informants about brothers and that in some cases the same survey items were used to obtain self-reports from primary respondents and from their brothers. The fourth section of the article compares within- and between-family struc-

---

5 Because ability and other social psychological characteristics vary both between and within families, the estimates of Sewell et al. correct in part for bias that is not attributable to shared effects of families on offspring.
Fraternal Resemblance

tural regressions that have been corrected for response variability, and it compares these with estimates that fail to compensate for response variability or for family effects. We close with a discussion of some extensions of this research.

THE WISCONSIN SIBLING DATA

The Wisconsin Longitudinal Study has followed a random sample of more than 10,000 men and women who were seniors in the state's public, private, and parochial high schools in 1957 (Sewell and Hauser 1980). Late in the senior year, detailed information was collected on the social origins, the academic ability and performance, and the educational aspirations of the students. There were successful follow-up surveys of the total sample (with approximately 90% response rates) in 1964 and 1975. Most important for the present purpose, the 1975 survey obtained a roster of the siblings of each primary respondent, including date of birth, sex, and educational attainment. For a randomly selected sibling, the survey ascertained current address and occupation. In 1977, telephone interviews were conducted with a sample of the selected siblings (aged 20–65) that had been stratified by the size of the sibship, the sex of the sibling and the primary respondent, and the birth order and educational attainment of the sibling. Of 879 brothers of male primary respondents who were selected into this supplement, telephone interviews were completed with 749 (85.2%). There is reason to believe that the achieved sample of brother pairs adequately reflects the composition of the sample of primary respondents (and their brothers) from which it was drawn (Hauser, Sewell, and Clarridge 1982, pp. 7–13). For the present analysis, we further restricted the sample to those 518 pairs of brothers aged 20–50 for whom the nine variables listed in table 1 had been ascertained. Only 19 pairs were lost because of the age restriction, but an additional 212 pairs lacked complete data. In many cases the missing data were due to school enrollment or absence from the labor force, rather than to item nonresponse.

As shown in table 1, there are two indicators of educational attainment for the primary respondent (EDEQYR, EDAT64) and two for his brother (XEDEQYR, SSBED). The first member of each pair is a self-report, and the second is a proxy report. In the case of the primary respondent, the proxy report (EDAT64) was coded from the educational history in the 1964 follow-up, and in that of the brother, the proxy report (SSBED) was given by the primary respondent in the 1975 survey. In both cases there is some slippage in time between the self- and proxy reports, and consequently some true educational mobility may appear as response variability in later models. To minimize this problem, as well as that of classify-
ing postgraduate education in years, we have followed the census practice of collapsing schooling at or beyond 17 years.

All of the occupation reports have been classified using materials from the 1970 census and coded in the Duncan SEI (Duncan 1961; Hauser and Featherman 1977, app. B). There are self-reports of the primary respondent's occupational status in 1970 (OCSX70) and in 1975 (OCSXCR). There are self-reports of the brother's occupational status in 1970 (XOCSX70) and in 1977 (XOCSXCR), and there is a proxy report (by the primary respondent) of the brother's occupational status in 1975 (OCSSIB).

As in the case of educational attainment, there is some spread in the temporal referents of these measurements, and some true status mobility may appear to be response variability. There are two reasons for our decision to treat the indicators for each brother as measures of the same occupational status construct. First, even over a period of several years, unreliability looms large relative to mobility as a component of observed change in occupational status (Bielby et al. 1977; Hauser et al. 1983). Second, our intention is not to depict the true status of the individual at an instant in time but rather a relatively stable feature of his placement in the occupational hierarchy. Thus, our concept of response variability in occupational status is inclusive of true short-run changes in status.

Table 2 reports the means and standard deviations of the nine status variables and their intercorrelations. All of the analyses discussed below are based on these data. Note that brothers have slightly less schooling than primary respondents but are more variable in schooling than respondents. There is a similar pattern in the case of occupational status. This

---

**Note.**—Occupation is scaled on Duncan's SEI.

---

**Table 1**

**Description of the Variables, Mnemonics, Source of Report, and Year of Measurement: Wisconsin Brothers (N = 518)**

<table>
<thead>
<tr>
<th>Mnemonic</th>
<th>Description</th>
<th>Source</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. EDEQYR</td>
<td>Respondent's years of schooling</td>
<td>Respondent</td>
<td>1975</td>
</tr>
<tr>
<td>2. EDAT64</td>
<td>Respondent's years of schooling</td>
<td>Parent</td>
<td>1964</td>
</tr>
<tr>
<td>3. XEDEQYR</td>
<td>Sib's years of schooling</td>
<td>Sibling</td>
<td>1977</td>
</tr>
<tr>
<td>4. SSBED</td>
<td>Sib's years of schooling</td>
<td>Respondent</td>
<td>1975</td>
</tr>
<tr>
<td>5. OCSXCR</td>
<td>Respondent's current occupation</td>
<td>Respondent</td>
<td>1975</td>
</tr>
<tr>
<td>6. OCSX70</td>
<td>Respondent's 1970 occupation</td>
<td>Respondent</td>
<td>1975</td>
</tr>
<tr>
<td>7. XOCSXCR</td>
<td>Sib's current occupation</td>
<td>Sibling</td>
<td>1977</td>
</tr>
<tr>
<td>8. OCSSIB</td>
<td>Sib's 1975 occupation</td>
<td>Respondent</td>
<td>1975</td>
</tr>
<tr>
<td>9. XOCSX70</td>
<td>Sib's 1970 occupation</td>
<td>Sibling</td>
<td>1977</td>
</tr>
</tbody>
</table>

---

6. Detailed industry and class of worker were used in some instances to refine the scale values reported by Hauser and Featherman for certain occupation lines.
Fraternal Resemblance

TABLE 2

PRODUCT-MOMENT CORRELATION COEFFICIENTS, MEANS, AND SD:
WISCONSIN BROTHERS (N = 518)

<table>
<thead>
<tr>
<th>Mnemonic</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. EDEQYR</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. EDAT64</td>
<td>.906</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. XEDEQYR</td>
<td>.404</td>
<td>.437</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. SSBED</td>
<td>.419</td>
<td>.450</td>
<td>.926</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. OCSXCR</td>
<td>.552</td>
<td>.525</td>
<td>.251</td>
<td>.252</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. OCSX70</td>
<td>.590</td>
<td>.562</td>
<td>.300</td>
<td>.295</td>
<td>.818</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. XOCSEXCR</td>
<td>.217</td>
<td>.243</td>
<td>.622</td>
<td>.568</td>
<td>.264</td>
<td>.315</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. OCSSIB</td>
<td>.217</td>
<td>.245</td>
<td>.627</td>
<td>.593</td>
<td>.265</td>
<td>.307</td>
<td>.815</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>9. XOCSX70</td>
<td>.228</td>
<td>.257</td>
<td>.628</td>
<td>.575</td>
<td>.247</td>
<td>.275</td>
<td>.819</td>
<td>.780</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Mean ............. 13.60 13.38 13.37 13.29 4.91 4.88 4.80 4.72 4.49
SD ............... 2.09 1.83 2.27 2.22 2.44 2.41 2.57 2.51 2.54

NOTE.—Correlations are based on 518 pairs of brothers for whom complete data were available. For explanation of mnemonics, see table 1. For convenience in the scaling of coefficients, values of the Duncan SEI have been divided by 10.

reflects basic differences between the populations of primary respondents and of brothers that are represented in the Wisconsin sibling data. There is a floor on the schooling of primary respondents but not on that of their brothers; none of the former obtained less than 12 years of regular schooling. Moreover, nearly all of the primary respondents were born in 1939, but the ages of their brothers varied widely over the range from 20 to 50. These cohort and age differences between the primary respondents and their brothers may also have affected the joint distributions of educational attainment and occupational status.

In modeling sibling resemblance, the usual procedure is to treat the members of a given sibling pair as unordered or indistinguishable (Jencks et al. 1972, 1979; Olneck 1977; Olneck and Bills 1980). Common family factors affect each member of the pair in the same way, and there is only one within-family regression. The analysis treats families, rather than persons, as units of analysis. For each variable and family, there are observations on each member of the fraternal pair, but it does not matter which observation is which. This greatly simplifies data analysis. For example, regardless of the pattern of common (family) causation, regressions of intrapair differences yield unbiased estimates of within-family regressions. In the present research design, with brothers sampled through a narrowly defined cohort of primary respondents, symmetry
within- and between-family regressions

The first step in our analysis was to estimate and compare the simple regressions of occupational status on schooling for men and their brothers without correcting for response variability. There was reason to find differences between the regressions for primary respondents and their brothers because there was a floor on the schooling of primary respondents and because the brothers (but not the primary respondents) varied widely in age. To provide a baseline for comparison of estimates that had been corrected for response variability or for family effects, it was desirable to estimate one or more common or pooled regressions of occupational status on schooling. In all, there were 10 zero-order regressions of a man's occupational status on his schooling, four among primary respondents and six among their brothers. Considering the heterogeneity of populations, informants, and temporal referents, these regressions were remarkably similar. We obtained the following pooled, baseline estimates of the regression of occupational status on schooling: $\hat{\beta} = 0.666 (0.057)$ among primary respondents; $\hat{\beta} = 0.679 (0.054)$ among their brothers; and $\hat{\beta} = 0.673 (0.042)$ among all siblings combined.\(^7\)

Figure 1 shows the path diagram of a simple model of sibling resemblance in educational attainment and occupational status using the notation of the LISREL model (Jöreskog and Sörbom 1978).\(^8\) The observations of educational attainment are denoted by $Y_R$ and $Y_S$, and the observations of occupational status are denoted by $Y_R$ and $Y_S$ for respondent and sibling, respectively. As shown in the central portion of the diagram, there are common family factors for educational attainment, $\xi_2$, and for occupational status, $\eta_2$, which affect the respective individual observations and are linked by the between-family regression, $\gamma_{22}$. The disturbances of the observables are the respective within-family components of educational attainment and occupational status for respondent and sibling. Thus, in the upper portion of the diagram, the within-family component of respondent's occupational status, $\eta_1$, is regressed on the within-family component of his educational attainment, $\xi_1$; in the lower portion

---

\(^7\) Each coefficient states the effect of a one-year increase in schooling on the Duncan SEI in 10-point units, and standard errors are given in parentheses. The differences among the three estimates are not statistically significant. Hauser and Mossel (in press) describe the statistical methods used to obtain these estimates, which take account of the pairing of observations across siblings and the occurrence of multiple measurements of educational attainment and of occupational status for each sibling.

\(^8\) This model was suggested to us by William T. Bielby.
FIG. 1.—A structural equation model of sibling resemblance in educational attainment and occupational status.

The path diagram in figure 1 gives the appearance that any or all parameters of the model may differ between the primary respondent and his brother. In fact we cannot make this assumption because the model is underidentified. As shown, the model has 11 parameters: three variances of $\xi$'s ($\psi$'s), three variances of disturbances in $\eta$'s ($\psi$'s), three structural regressions ($\gamma$'s), and two scale factors ($\lambda_1$ and $\lambda_2$); but there are only 10 sample moments: four variances and six covariances among the four observable indicators. In order to estimate the model, it is necessary to impose at least one restriction on the parameters. We chose to impose the two restrictions $\lambda_1 = \lambda_2 = 1$; this implies that both pairs of within-family variables are in the same metric as the family factors and so justifies comparisons of slopes among the three regressions (Bielby 1982). 9

9 We experimented with other identifying restrictions—e.g., $\psi_{11} = \psi_{33}$, which says that disturbance variances are equal in the two within-family regressions. However,
specification implies, also, that respondent-brother differences in variance are all due to within-family components of schooling and status; by construction, family effects on each sibling are the same.

The model of figure 1 differs from some other models of sibling resemblance in its expression of the within-family regressions (Olneck 1976, pp. 139–49; Olneck 1977; Corcoran and Datcher 1981, pp. 195–97). Several critics have suggested either that it would be better to express the within-family regressions in the total variates or that such a reexpression would be equivalent. That is, in the present model, the within-family regressions are written in disturbances of the factor model, and in the alternative model, the within-family regressions are written directly in the educational and occupational variables. Hauser (1984a) has compared the two models and shown that they are algebraically equivalent when the two within-family regressions are homogeneous. Otherwise, they are not equivalent, and the second model has undesirable logical implications. Moreover, when the within- and between-family regressions are nearly homogeneous in slope, as we have found in the present case, the second model exhibits symptoms of "near-underidentification."

One plausible form of causation is excluded from the model of figure 1, that is, the direct influence of one sibling on the other (Olneck 1976; Benin and Johnson 1984). All "family" influences are carried by the common family factors. Hauser and Mossel (in press) have shown how this model may be modified to incorporate unidirectional or mutual influence between siblings, but that possibility will not be pursued further here.

Table 3 shows goodness-of-fit statistics and selected parameter estimates for several versions of the model of figure 1. That model uses only one indicator of educational attainment and one of occupational status for each member of the sibling pair, and the estimates in table 3 are based on the self-reports of educational attainment and occupational status at the survey date. A priori, we take these self-reports of current statuses to be of higher quality than the others. We begin with a model that imposes equivalent scales on all of the variables, and we then test whether the

---

10 Also, see Hauser 1984b.

11 We obtained similar results in parallel analyses of the reports of the primary respondent about himself and about his brother in the 1975 survey and of the self-reports of educational attainment and of occupational status in 1970 (Hauser and Mossel, in press).
Table 3

Maximum Likelihood Estimates of Models of Sibling Resemblance in Educational Attainment and Occupational Status with Latent Family Variables: Wisconsin Brothers (N = 518)

<table>
<thead>
<tr>
<th>VARIABLES AND MODEL</th>
<th>Slope</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Respondent</td>
<td>Brother</td>
<td>Family</td>
<td>$L^2$</td>
<td>df</td>
</tr>
<tr>
<td>1. $\lambda_1 = \lambda_2 = 1$</td>
<td>.620</td>
<td>.735</td>
<td>.659</td>
<td>.73</td>
<td>1</td>
</tr>
<tr>
<td>2. Add $\gamma_{11} = \gamma_{33}$</td>
<td>.691</td>
<td>.691</td>
<td>.650</td>
<td>2.28</td>
<td>2</td>
</tr>
<tr>
<td>3. Add $\gamma_{11} = \gamma_{22} = \gamma_{33}$</td>
<td>.676</td>
<td>.676</td>
<td>.676</td>
<td>2.44</td>
<td>3</td>
</tr>
<tr>
<td>4. Add $\psi_{11} = \psi_{33}$</td>
<td>.676</td>
<td>.676</td>
<td>.676</td>
<td>2.52</td>
<td>4</td>
</tr>
<tr>
<td>5. Add $\phi_{11} = \phi_{33}$</td>
<td>.676</td>
<td>.676</td>
<td>.676</td>
<td>6.65</td>
<td>5</td>
</tr>
</tbody>
</table>

Note.—Estimates are based on self-reports of educational attainment and current occupation: EDEQYR, OCSXCR, XEDEQYR, and XOCSXCR. Standard errors are in parentheses. There is no correction for response variability.

parameters for respondents, brothers, and families are similar in other respects. The model in row 1 of table 3 yields seemingly disparate slope estimates for primary respondents, brothers, and families. Indeed, the within-family slope estimate for primary respondents is quite low, but the estimate for brothers exceeds that for families. This initial, equivocal finding on bias in the schooling-occupation relationship recurs throughout our analysis.

The within-family slope estimates for respondents and brothers do not differ significantly. The common, within-family slope estimate shown in row 2 of table 3, $\hat{\gamma}_{11} = \hat{\gamma}_{33} = 0.691$, is actually larger than the common slope that we estimated without correction for measurement error or family bias ($\hat{\beta} = 0.673$). Again, there is little evidence that the omission of common family variables significantly affects these estimates.

In multilevel analyses, it is often found that regressions across population aggregates—such as cities, states, or organizations—are steeper than corresponding individual regressions. This is (partly) the basis of the well-known literature on “ecological correlation” (Duncan, Cuzzort, and Dun-
can 1961) and “aggregation bias” (Hannan 1971). For example, the occurrence of heterogeneous within-school and between-school regressions of educational aspirations on socioeconomic status (Sewell and Armer 1966) was the source of a controversy that revolved around the question whether there were emergent “contextual” effects of schools or whether the individual-level regressions were misspecified (Hauser 1972; Boyd and Iversen 1979).

Similarly, in the present case we expected to find steeper between-family than within-family regressions of occupational status on schooling, but this proved not to be the case. In the model of row 2, the within-family slope estimate is larger than the between-family slope. Moreover, as shown in row 3 of table 3, there is virtually no deterioration in the fit of the model when all three regressions are constrained to share a common slope. This says that there is no family bias—no emergent family effect—on the relationship between educational attainment and occupational success; that relationship is just what we would expect from the differential rewards of schooling across individuals. Again, the common slope estimate is virtually the same as that estimated under the model of figure 1.

In rows 4 and 5 of table 3, two more restrictions are added to the model; neither affects the slope estimates or their standard errors. First, we specify that $\psi_{11} = \psi_{33}$; this says that the disturbances in the two within-family regressions have the same variance. Under this additional restriction, there is virtually no change in fit. Second, we specify that $\phi_{11} = \phi_{33}$; this says that the within-family variances in educational attainment are the same for primary respondents and their brothers. Congruent with our expectations about selection into the sample, the data do not meet the latter restriction. Thus, with this one exception, the data do not depart significantly from the usual assumption of symmetry between siblings. Our main finding is that of homogeneity in the regressions of occupational status on schooling, without regard either to the choice of indicators or to the specification of common family factors.

### MEASUREMENT ERROR IN THE STRUCTURAL MODEL

The path diagram in figure 2 embeds our specification of fraternal resemblance within a multiple-indicator measurement model of educational attainment and occupational status. The observables ($Y_1, \ldots, Y_5$) appear only as reflections or effects of the “true” educational and occupational constructs ($\eta_1, \ldots, \eta_5$). Initially, we resolved the indeterminacy in the metrics of the latent variables (Bielby et al. 1977) by fixing the regressions

14 The fit of the model deteriorates significantly ($L^2 = 4.13$ with 1 df).

15 Hauser and Mossel (in press) describe this measurement model in detail.
Fig. 2.—A model of sibling resemblance in educational attainment and occupational status with errors in variables and latent family factors.
of the self-reports of educational attainment on true education at 1.0 for respondents and siblings and by fixing the regressions of the self-reports of current occupational status on true status at 1.0 for respondents and siblings. This implied that the constructs were in the metrics of these indicators and that their variances were the true variances of the respective indicators. This was a convenient normalizing constraint because each of the reference indicators was a self-report and because the same methods were used to ascertain and to code these variables for respondent and sibling. Ultimately, excepting the parent's report of the primary respondent's schooling, the differences among loadings within constructs were not statistically significant, and we fixed all of the loadings at 1.0 except that of \( Y_2 \) on \( \eta_2 \).

The measurement model also includes selected covariances among response errors, which are not shown in figure 2. Initially, covariances were permitted between the errors in any pair of variables that had been ascertained on the same occasion or from the same informant. Thus, the model allowed for all possible error covariances among reports by primary respondents and among reports by their brothers but no error covariances between reports by respondents and brothers, by respondents and parents, or by brothers and parents. None of the covariances between errors in the primary respondent's reports of his own and of his brother's characteristics was statistically significant, and these were deleted from the model. In addition, the final measurement model imposes selected equality constraints on error variances and error covariances (Hauser and Mossel, in press).^16^

Table 4 reports the reliabilities of the indicators and the correlations between response errors in the measurement model. The reliabilities of the indicators of educational attainment range from 0.89 to 0.95; since slope corrections are inverse to the square root of the reliability of the regressor, the implied corrections in regressions of occupational status on schooling are small. The reliabilities of the indicators of occupational status are lower than those of educational attainment, but unreliability in occupational status has no effect on the slope estimates. Four of the five

---

^16^ One such error covariance was not identified within the model—that between errors in the respondent's reports of his current occupation (OCSXCR) and his occupation in 1970 (OCSX70). We specified that error covariance to be equal to the corresponding error covariance for brothers—between XOCSXCR and XOCSX70—which is identified.

^17^ The likelihood ratio fit statistic for the final measurement model is satisfactory: \( L' = 24.39 \) with 24 df. Instead of selecting backward, we could have started with a model without correlated error and added error covariances to it as needed, but we chose to be as generous as possible to published suggestions that correlated errors are pervasive in reports of socioeconomic variables.
TABLE 4

RELIABILITIES AND ERROR CORRELATIONS IN A MEASUREMENT MODEL OF SIBLING
RESEMBLANCE IN EDUCATIONAL ATTAINMENT AND OCCUPATIONAL STATUS

<table>
<thead>
<tr>
<th>MNEMONIC</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. EDEQYR</td>
<td>.887</td>
<td>.......</td>
<td>.......</td>
<td>.093</td>
<td>.088</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
</tr>
<tr>
<td>2. EDAT64</td>
<td>.........</td>
<td>.929</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
</tr>
<tr>
<td>3. XEDEQYR</td>
<td>.........</td>
<td>.......</td>
<td>.904</td>
<td>.......</td>
<td>.......</td>
<td>.073</td>
<td>.......</td>
<td>.073</td>
<td>.......</td>
</tr>
<tr>
<td>4. SSBED</td>
<td>.........</td>
<td>.......</td>
<td>.......</td>
<td>.948</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>- .044</td>
</tr>
<tr>
<td>5. OCSXCR</td>
<td>.........</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.746</td>
<td>.078</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
</tr>
<tr>
<td>6. OCSX70</td>
<td>.........</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.267</td>
<td>.775</td>
<td>.......</td>
<td>.......</td>
</tr>
<tr>
<td>7. XOCSXCR</td>
<td>.........</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.070</td>
</tr>
<tr>
<td>8. OCSSIB</td>
<td>.........</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
</tr>
<tr>
<td>9. XOCSX70</td>
<td>.........</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
<td>.......</td>
</tr>
</tbody>
</table>

NOTE.—Entries on the main diagonal are reliabilities. Entries below the main diagonal are correlations between errors in variables. Entries above the main diagonal are error covariances, expressed as proportions of the respective observed covariances. All of the error covariances are significantly different from zero at the .05 level.

estimates are close to 0.75, and only the reliability of OCSSIB is as large as 0.84. The lower reliabilities of the indicators of occupational status may reflect temporal spread as well as errors in reporting and processing the data. Of course, all of the unreliabilities affect the estimated correlations between status variables. Observed correlations between educational attainment and occupational status range from 0.525 to 0.590 for primary respondents and from 0.568 to 0.628 for brothers. The corrected correlation between educational attainment and occupational status is 0.653 for primary respondents and 0.689 for brothers.

Correlated errors of measurement also affect the slopes and correlations between the educational and occupational constructs. The entries below the main diagonal of table 4 are correlations between errors in the constrained measurement model. There are positive correlations of approximately 0.3 between errors in self-reports of educational attainment and occupational status. These tend to compensate for random response variability by increasing the regressions (and correlations) between observed indicators of schooling and occupational status. At the same time, there is a negative correlation of about the same size between errors in the primary respondent’s reports of his brother’s educational attainment and occupational status, and this adds to the effect of random response variability by decreasing the observed correlation between those two variables. Last, there are positive correlations of approximately 0.25 between response errors in self-reports of occupational status; these positive,
within-construct error correlations add to the effect of random response variability by decreasing the observed correlations between educational and occupational indicators. As a practical matter, none of the correlated errors has a very large effect on slope estimates in the model. The error correlations are relatively large because the response error variances are relatively small. The entries above the main diagonal of table 4 express the estimated error covariances as proportions of the respective observed covariances, and none of these is as large as 10% of an observed covariance.

CORRECTED WITHIN- AND BETWEEN-FAMILY REGRESSIONS

Table 5 shows goodness-of-fit statistics and slope estimates for several versions of the model in figure 2. As in the model of table 3, the unrestricted slope estimate for primary respondents is less than that for families, which is in turn less than that for brothers. Model 2 adds the restriction of a common slope for primary respondents and brothers, and this does not significantly affect fit. The common, within-family slope estimate ($\hat{\gamma}_1 = \hat{\gamma}_2 = 0.728$) is actually larger than the between-family slope estimate ($\hat{\gamma}_3 = 0.678$). Model 3 adds the restriction that all three slopes are homogeneous; again, there is no deterioration in fit. The common slope estimate, $\hat{\gamma}_1 = \hat{\gamma}_2 = \hat{\gamma}_3 = 0.708$, is only 1.047 times larger than the uncorrected common slope in the family model of table 3 ($\hat{\gamma} = 0.676$); it is 1.051 times larger than the common slope estimate in the naive regressions ($\hat{\beta} = 0.673$). We are left with the strong impression that neither family factors nor response error has substantial effects on our estimates of the occupational effects of schooling.

Model 4 of table 5 adds the constraint that disturbance variances are the same in the two within-family regressions, and the fit is virtually unaffected by this. However, the data are not consistent with the addition of the restriction in model 5 that true within-family variances in educational attainment and occupational status of the brothers, but table 5 pertains to models in which these two coefficients have been fixed at unity in order to identify the model and normalize slope estimates. Thus, model 1 of table 5 incorporates one more restriction than the final measurement model.

19 This finding was quite unexpected, and, at the suggestion of Christopher Jencks, Hauser (1984b) validated it in other, larger subsamples of siblings drawn from the Wisconsin Longitudinal Study and in comparisons with the Kalamazoo study. The prima facie evidence of family bias in these samples was in some cases stronger than that reported here, but again observed biases disappeared when corrections were made for response variability in schooling. All of these findings apply mainly to high school graduates, and it is important to test them again in populations that are more variable in levels of completed schooling.

18 The path diagram in figure 2 shows distinct scale factors, $\gamma_4$ and $\beta$, for the effects of the family factors on the true educational attainment and occupational status of the brothers, but table 5 pertains to models in which these two coefficients have been fixed at unity in order to identify the model and normalize slope estimates. Thus, model 1 of table 5 incorporates one more restriction than the final measurement model.
### TABLE 5

**Maximum Likelihood Estimates of Models of Sibling Resemblance in Educational Attainment and Occupational Status with Errors in Variables and Latent Family Factors: Wisconsin Brothers (N = 518)**

<table>
<thead>
<tr>
<th>Model</th>
<th>Slope</th>
<th>Respondent</th>
<th>Brother</th>
<th>Family</th>
<th>$L^2$</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\gamma_4 = \beta = 1$</td>
<td>.674</td>
<td>.756</td>
<td>.684</td>
<td>26.07</td>
<td>25</td>
<td>.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.081)</td>
<td>(.057)</td>
<td>(.062)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Add $\gamma_1 = \gamma_2$</td>
<td>.728</td>
<td>.728</td>
<td>.678</td>
<td>26.74</td>
<td>26</td>
<td>.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.047)</td>
<td>(.047)</td>
<td>(.062)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Add $\gamma_1 = \gamma_2 = \gamma_3$</td>
<td>.708</td>
<td>.708</td>
<td>.708</td>
<td>27.03</td>
<td>27</td>
<td>.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.029)</td>
<td>(.029)</td>
<td>(.029)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Add $\psi_6 = \psi_7$</td>
<td>.708</td>
<td>.708</td>
<td>.708</td>
<td>27.07</td>
<td>28</td>
<td>.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.029)</td>
<td>(.029)</td>
<td>(.029)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Add $\phi_2 = \phi_3$</td>
<td>.708</td>
<td>.708</td>
<td>.708</td>
<td>32.04</td>
<td>29</td>
<td>.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.029)</td>
<td>(.029)</td>
<td>(.029)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.**—Standard errors are in parentheses.

---

The within- and between-family variances of schooling are not

1. Statistical attainment are equal for primary respondents and their brothers ($L^2 = 4.97$ with 1 df). Model 4 is our preferred measurement and structural model, and table 6 gives additional structural parameters of that model.

Even though regressions of occupational status on schooling are homogeneous across persons and families, we hasten to add that this by no means denies the importance and visibility of families in the stratification process. For example, for primary respondents, 51.4% of the variance in schooling lies between families, and for their brothers, 42.2% of that variance lies between families. Conditional on the hypothesis that true variance in schooling is the same for respondents as for their brothers—that is, on model 5 of table 5—there is little difference between the within- and between-family variance components in schooling. The restriction that $\phi_{11} = \phi_{22} = \phi_{33}$ increases the test statistic by only $L^2 = 1.04$ with 1 df relative to the restriction that $\phi_{22} = \phi_{33}$.

Of the total variance in occupational status—whether or not it is attributable to differences in schooling—39.3% lies between families in the case of respondents, and 35.9% lies between families in the case of their brothers. Similarly, there is much less unexplained variance in occupational status between than within families; 30.3% occurs between families. If we add the restriction $\psi_{11} = \psi_{66} = \psi_{77}$ to model 4 of table 5, the test statistic increases significantly by $L^2 = 15.98$ with 1 df.

---

If we add the restriction $\psi_{11} = \psi_{66} = \psi_{77}$ to model 4 of table 5, the test statistic increases significantly by $L^2 = 15.98$ with 1 df.
very different from one another, and the slopes of occupational status on schooling are also homogeneous across families and persons; thus, the low unexplained between-family variance in occupational status implies that the correlation between occupational status and schooling will be larger between than within families. Under model 4 of table 5, the within-family correlations are 0.584 for primary respondents and 0.655 for their brothers; the between-family correlation is 0.746.

DISCUSSION

We have intended this analysis to serve two purposes. First, we hope that it may serve as a methodological template for research on the stratification process and, perhaps, in other analyses that cut across levels of aggregation. Second, we think that it yields significant findings about the influence of family background in the stratification process and about the importance of response variability in survey-based socioeconomic models. We shall comment on each of these points in turn.

We have expressed a model of sibling resemblance in the LISREL framework, thus facilitating the process of model specification, estimation, and testing. A useful innovation in this model has been our specification of distinct within- and between-family regressions. Conventionally, the latter have not been made explicit (Olneck 1976, pp. 139–49; Olneck 1977; Corcoran and Datcher 1981, pp. 195–97). We believe that the between-family slopes and, especially, their contrasts with the within-family slopes, are of real sociological importance. They show whether families enter the stratification system as relatively homogeneous, but neutral, aggregates of persons, or whether they affect returns to the attributes and resources of their members (see Chamberlain and Griliches
Furthermore, we have incorporated random (and certain types of correlated) response errors in the model by obtaining multiple measurements of schooling and occupational status.

Within this framework, we have estimated regressions of occupational status on educational attainment among primary respondents and among their brothers, with and without response variability and common family factors. Paralleling Chamberlain and Griliches's (1975, pp. 428-32) analyses of schooling and income in the Gorseline data, we find little evidence that the omission of common family variables leads to bias in our estimates of the effect of schooling on occupational status. The between-family variance in schooling is about as large as the within-family variance, and there is substantial between-family variance in occupational status as well. Nonetheless, the regression of occupational status on schooling is homogeneous within and between families in the simple models we have estimated.

This does not at all imply an absence of omitted-variable bias in the relationship between schooling and occupational status. As shown by Sewell and Hauser (1975), Sewell et al. (1980), and Hauser et al. (1983), among others, the bias is substantial, but our finding suggests that intrafamily differences in such variables as ability and motivation are its sources, rather than common family influences. The relationship between schooling and occupational success across families is just what we would expect from the differential rewards of schooling across individuals.

Moreover, this finding is insensitive to our treatment of measurement error. There is substantial unreliability (or at least temporal instability) in occupational status, and there are small positive correlations between self-reports of one's own educational attainment and occupational status. At the same time, the reliability of educational attainment is extremely high. Even after we purge the variance of schooling of its large between-family component, the regression of occupational status on schooling is not substantially affected by response variability in schooling. Although this might be taken to encourage studies in which response variability in schooling is not or cannot be specified, we think such an application of our findings is not warranted and may lead to erroneous conclusions. Our findings pertain to a well-educated population in which years of schooling have been ascertained with a good deal of care and detail. Moreover, as additional (nonfamilial) explanatory variables—like mental ability—are added to the equations for educational attainment, the threat posed by response variability takes on greater importance.

Within the LISREL framework it has been straightforward to test a variety of hypotheses about symmetry between our primary respondents and their brothers in parameters of both the measurement and structural models. In extensions of the present work, we expect the similarity in
measurement models to be extremely important; in some cases we have multiple measurements of a given construct only for primary respondents, and in other cases, only for their siblings. Within the present framework, it is possible to "borrow" an estimate of error variance that is identified in one within-family segment of the model and use it in the other within-family segment of the model (Hauser and Sewell 1984).

Another straightforward modification of the model permits us to test the factorial complexity of the latent family variates. For example, we can test the hypothesis that there is a single, unobserved family factor by setting $\psi_{11} = 0$; this yields an increase in the likelihood ratio test statistic of $L^2 = 31.30$ with 1 df in the model of row 1 in table 5. Thus, we find that a single factor model is unacceptable (cf. Hauser and Dickinson 1974; Jencks 1974); similar analytic issues will recur as we add explicit family background constructs to the model (Chamberlain and Griliches 1975, 1977; Hauser and Sewell 1984).

The present model also lends itself to elaboration in a number of ways. First, it is possible to add more variables that have been observed (possibly with error) for respondent and sibling and to specify their corresponding within- and between-family components and regressions. Perhaps the two most obvious constructs to be added in this fashion are mental ability and earnings, of which the former is an antecedent of schooling and the latter is a consequence of schooling and occupational status. In the Wisconsin survey we have multiple observations of both of these variables among male primary respondents. Moreover, we have at least one observation for respondents and for siblings on each of the variables in the Wisconsin model of status attainment (Sewell and Hauser 1980; Hauser et al. 1983). By using multiple indicators throughout the model, we shall be able to address such issues as "the endogeneity of schooling" with fewer trade-offs among specifications of errors in variables, simultaneity, and family effects (cf. Griliches 1977, 1979).

Second, it is possible to add constructs to the model that are common to primary respondents and their siblings and that have no "within-family" components. Here the most obvious variables are shared characteristics of the family or community of orientation: parents' educations, occupations, and earnings; family size, ethnicity, and religious preference; and community size and location. In most cases these variables will be specified as antecedent to other "between-family" variables. Again, these variables are subject to error, and in several instances we have multiple indicators of them in the Wisconsin data.

Third, beyond the specification of cross-sibling effects, there are other, and perhaps more interesting, elaborations of the model that exploit the multiple-group feature of the LISREL program. As noted earlier, the full Wisconsin sibling sample is based on a design that crosses sex by response
status, so primary respondents of each sex are paired with randomly selected siblings of each sex. Thus, we can increase the statistical power of our analyses by fitting models within the multiple-group framework and pooling estimates where similar populations occur in different pairings, for example, male primary respondents paired with sisters as well as with brothers. More important, within this framework it is possible to contrast parameters of the model between men and women (Hauser 1984b).

Although it is common to identify the family as a source of persistent social inequality, Griliches offers several interesting speculations about ways in which the family may reduce inequality (1979, pp. S60–S63). These, too, lead to hypotheses about intergroup contrasts in parameters of sibling resemblance. For example, he suggests that families may try to invest their resources to minimize differences in outcomes between children, and he points to lower within-family than between-family regressions of schooling on IQ as possible evidence of this. Of course, the latter may also be artifacts of attenuation, since the previous argument about bias in the schooling coefficient applies equally well here. Another possibility is that family efforts to minimize differences in outcomes will be more successful as familial resources increase, say, as indicated by parents' socioeconomic status, and as sibship size decreases. Thus, we might look for reduced within-family regressions and within-family variance in smaller and higher-status families, relative to those in larger and/or lower-status families. Given observed secular changes in socioeconomic standing and in completed family size, changes in the family may contribute to reductions of social inequality.

REFERENCES


American Journal of Sociology


