#435-77

INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

RESPONSE ERROR IN EARNINGS FUNCTIONS FOR NONBLACK MALES

William T. Bielby and Robert M. Hauser

UNIVERSITY OF WISCONSIN-MADISON

RESPONSE ERROR IN EARNINGS FUNCTIONS FOR NONBLACK MALES

William T. Bielby
University of California-Santa Barbara

Robert M. Hauser
University of Wisconsin-Madison

July 1977

This research was supported by NSF Grants GI-31604 and GI-44336, NICHD Grant HD-05876, NIGMS Grant GM-01526, and by institutional support from the College of Agricultural and Life Sciences. We have worked closely with David L. Featherman in the design, collection, and management of the data on which this analysis is based. Any opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the National Science Foundation or other supporting agencies.

ABSTRACT

Biases due to measurement errors in an earnings function for nonblack males are assessed by estimating unobserved variable models with data from the Income Supplement Reinterview program of the March 1973 Current Population Survey and from the remeasurement program of the 1973 Occupational Changes in a Generation-II survey. We find that reports of social origins, educational and occupational attainments, labor supply, and earnings of nonblack males are subject to primarily random response errors. Logarithmic earnings is one of the most accurately measured indicators of socioeconomic success. Further, retrospective reports of status variables are as reliable as contemporaneous reports. When measurement errors are ignored for nonblacks, the total economic return to schooling is underestimated by about 16 percent, and the effects of some background variables are underestimated by as much as 15 percent. The total effects of first and current job status are underestimated by about 20 percent when measurement errors are ignored, as are the unmediated effects of current job status. Conflicting evidence is presented on whether respondents tend to understate the consistency between their earnings and educational attainments in the Current Population Survey. If there is such a tendency, unmediated effects of education are modestly understated when response errors are ignored, and they are overstated if no such tendency exists.

INTRODUCTION

Structural equation models of the social and economic determinants of earnings have been used by social scientists of diverse perspectives. Sociologists have specified earnings functions as part of "status attainment models" in order to examine the relative impact of schooling and social origins upon socioeconomic success (Duncan, Featherman, and Duncan, 1972; Jencks, et al., 1972; Alexander, Eckland, and Griffin, 1975; Sewell and Hauser, 1975; Griffin 1976; Hauser and Daymont, 1977). Similar functions have been used by economists to construct "human capital models" of the generation of income inequality (Becker, 1964; Mincer, 1974; Blinder, 1976; Rosenzweig, 1976). Marxist sociologists and economists, while taking issue with the substantive foundations of status attainment and human capital models, have also employed structural equation models of the determinants of earnings in their empirical representations of the generation of income inequality (Bowles, 1972; Bowles and Nelson, 1974; Bowles and Gintis, 1976; Wright and Perrone, 1977). Thus, while disagreements continue over conceptualization, the magnitudes of structural coefficients, and the appropriate specification of structural models, there exists remarkable consensus that a structural equation model is an appropriate empirical representation of the determinants of economic success.

It is also generally agreed that response errors (as well as the omission of ability or other common causes of schooling and of socioeconomic achievement) may bias estimates of the socioeconomic returns to schooling and social origins. However, the size the importance of such biases have been points of controversy. Jencks, et al., drawing on the work of Siegel and Hodge (1968), conclude that "random measurement error is of relatively little importance in research of the kind described here" (1972:336). Bowles (1972:S222) asserts that "social class background is considerably more important as a determinant of both educational attainment and economic success than has been indicated in recent analogous statistical treatments by Duncan and others." Bowles argues that retrospective reports of parental statuses are much less reliable than respondents' reports of their own attainments and that the effects of origin variables are consequently underestimated. Unfortunately, attempts to assess biases due to response errors have been flawed by a lack of appropriate data, by inadquate specifications and by crude estimation procedures (see Bielby (1976:11-61) and Bielby, Hauser and Featherman (1977a) for a more detailed discussion of these issues).

In order to help resolve the controversy surrounding response bias in models of socioeconomic success, a carefully designed remeasurement program was included as part of the 1973 Occupational Changes in a Generation study (Featherman and Hauser, 1975). Recently developed statistical procedures that allow for unobservable variables in structural equation models (Jöreskog, 1970, 1973) were applied to these data and to data from the March 1973 Current Population Survey income reinterview program in order to assess the extent of response error in measures of socioeconomic variables and to assess the biases in structural coefficients resulting from those errors. Our findings for a model of occupational attainment have been reported elsewhere. For nonblack males, we found compelling evidence that response errors in reports of social origins, education, and occupational attainments are mutually uncorrelated, and we found that retrospective reports of social origins are reported about as accurately as more recent attainments (Bielby, 1976; Bielby, Hauser and Featherman, 1977a, 1977b).[1] We found that response errors resulted in underestimates of the occupational return to some social origin variables by as much as 22 percent, but, contrary to what Bowles and other had hypothesized, they also resulted in underestimates in occupational returns to schooling by as much as 15 percent. Finally, we found that when measurement errors are ignored, variation in educational attainment that is independent of social origins is overstated by about 10 percent, and variation in occupational attainments attributable to neither schooling nor social origins is overestimated by as much as 27 percent.

The research reported here extends our earlier findings to include the determinants of earnings among nonblack males. Unless the errors in reports of earnings are correlated with errors in reports of other variables, our findings should be similar to those summarized above (since we already know the relative quality of most predetermined variables in the earnings equation). Furthermore, since earned income does not appear as a predetermined variable in any of our models, random measurement error in reports of earnings cannot bias metric structural coefficients; it can only cause an overestimate of residual variance and uniform underestimates of standardized structural coefficients in the earnings equation.

1973 OCG AND CPS DATA

The 1973 OCG (Occupational Changes in a Generation) study (Featherman and Hauser, 1975) was designed to replicate the 1962 OCG study conducted by Blau and Duncan (1967). The replicate study, executed in conjunction with the Current Population Survey, represents approximately 53 million males in the civilian noninstitutional population between the ages of 20 and 65 in March 1973. Data from the 1973 study allow us to estimate and test a variety of models of response error and to assess the effects of plausible error structures on parameters of the achievement process.

The data were collected in four surveys during 1973. First, educational and labor force data were obtained from the March 1973 CPS household interviews; in about three-fourths of the cases the CPS respondent was the spouse of the designated male. Second, a subsample of households containing about 1000 CPS male respondents was selected for inclusion in the March CPS Income Supplement Reinterview survey (ISR). Beginning about one week after completion of the CPS interview, personal (and in some cases telephone) interviews were conducted with respondents in these households to obtain a second report of selected CPS labor force and income items. Third, the CPS data were supplemented in the fall of 1973 with social background and occupational career data from the mail-out, mail-back OCG questionnaire (OCGQ); in about three-fourths of the cases the OCGQ respondent was the designated male. Responses to OCGQ were obtained from this questionnaire or subsequent telephone or personal follow-ups for more than 27,000 members of the experienced civilian labor force; the overall response rate was greater than 88 percent. Fourth, a random subsample of about 1000 OCGQ respondents (600 nonblacks and 400 blacks) was selected for inclusion in the OCG remeasurement survey (OCGR). Approximately three weeks after the mail return of their OCG questionnaires, telephone (and in a few cases personal) interviews were conducted with these respondents to obtain a second report of selected items on the OCG questionnaire.

Table 1--Timing of Measurements in the 1973 CPS and OCG Surveys

Variable	March 1973 CPS household interview (CPS)	Spring 1973 supplement reinterview (ISR)	Fall 1973 OCG Questionnaire (OCGQ)	Fall 1973 OCG remeasurement interview (OCGR)
1. Father's occupational status (FO)	--	--	x_{13}	x_{14}
2. Father's educational attainment (FE)	--	--	x_{23}	x_{24}
3. Parental income (PI)	--	--	x_{33}	x_{34}
4. Educational attainment (ED)	x_{41}	--	x_{43}	x_{44}
5. Occupational status of first job after completing schooling (O1)	--	--	x_{53}	x_{54}
6. Current occupational status (March or Fall) (OC)	x_{61}	--	--	x_{64}
7. Weeks worked in 1972 (WKS)	x_{71}	x_{72}	--	--
8. Earned income in 1972 (LNEARN)	x_{81}	x_{82}	--	--
9. Experience (years since began first job after completing schooling) (EX)	--	--	x_{93}	--
10. $(\text{Experience} - 20)^2/10$ (EX2)	--	--	$x_{10,3}$	--
11. Working full-time (FT)	$x_{11,1}$	--	--	--
12. Labor union membership (UN)	--	--	$x_{12,3}$	--

Table 1 shows which variables were measured on each of the four occasions: CPS, ISR, OCGQ, and OCGR. Educational attainment (x_{41}), current (March) occupation (x_{61}), weeks worked in 1972 (x_{71}), earned income in 1972 (x_{81}), and whether the sample person usually worked full- or part-time in 1972 ($x_{11,1}$) were ascertained in the March CPS interview. A second report of weeks worked (x_{72}) and earned income (x_{82}) were obtained from the income supplement reinterview. Reports of the three social background variables, father's (or other head of household's) occupation (x_{13}) and educational attainment (x_{23}) and parental family income (x_{33}), were obtained from the fall OCG questionnaire. Also, the fall questionnaire ascertained the man's first full-time civilian job after completing schooling (x_{53}) and a second report of his educational attainment (x_{43}). (The second measurement of ED was not intended to supplant the CPS item, but rather to improve the respondent's recall of the timing of schooling and labor force entry.) Also obtained from the fall OCG questionnaire were a measure of labor force experience (x_{93}, $x_{10,3}$), and whether or not the sample person belonged to a labor union ($x_{12,3}$). Within the OCGR subsample, each of the social background variables (x_{14}, x_{24}, x_{34}), education (x_{44}) and first job (x_{54}) was measured again. We were not able to ascertain March 1973 occupation in the OCGR interviews, and instead a report of current (Fall 1973) occupation (x_{64}) was obtained. While some job mobility occurred between the spring and fall surveys, we disregard it here on the argument that occupational status changes were negligible over the six or seven month period. Consequently, our estimates of unreliability in the reporting of current occupational status include effects of short-term job mobility as well as response error.

In summary, we have: (1) two measures of education and one measure of each other variable for the full CPS-OCGQ sample; (2) two measures each of education, earnings, and weeks worked and one measure of each other variable for the ISR subsample; and (3) three measures of education, two measures of each indicator of social origins, two measures each of first and current job status, and one measure of each of the remaining variables for the OCGR subsample.

Each of the occupation reports was scaled using Duncan SEI scores for detailed 1960 Census occupation, industry, and class of worker categories (Duncan 1961). Thus, our estimates of the quality of occupation reports do not pertain to descriptions of occupations per se, but to a particular transformation of detailed job descriptions into a status metric (Featherman and Hauser, 1973). Educational attainment is coded in exact years of schooling completed, and parental income is coded as the logarithm of price adjusted dollars.[2] The measure of weeks worked is computed from seven category midpoints; exact weeks worked, available for the ISR subsample, correlates .986 with this measure. Earned income in 1972 is computed as the natural logarithm of the sum of wage and salary, self-employed farm, and self-employed nonfarm income. Experience is computed as the number of years since the year the sample person started the first full-time civilian job he held after completing schooling.[3] The quadratic experience variable is computed as $(\text{experience} - 20)^2/10$. Working full-time and labor union membership are coded as dummy variables.

SPECIFICATION OF AN EARNINGS FUNCTION

We specify the following earnings function among "true" measures for nonblack males who report at least $1000 in 1973 earnings in the CPS interview:

$$\text{LNEARN} = \alpha + \beta_1\text{EX} + \beta_2\text{EX2} + \beta_3\text{FO} + \beta_4\text{FE} + \beta_5\text{PI} + \beta_6\text{ED}$$
$$+ \beta_7\text{OI} + \beta_8\text{OC} + \beta_9\text{WKS} + \beta_{10}\text{FT} + \beta_{11}\text{UN} + u, \qquad [1]$$

where the disturbance has the usual classical properties. We restrict the sample to those reporting at least $1000 in earnings for two reasons. First, men below that cut-off usually have marginal labor force attachments and are likely to be subject to qualitatively different determinants of earnings. Second, reports of yearly earnings less than $1000 appear to be subject to disproportionately large errors of measurement (see the first two lines of appendix Table 1A).

The semi-logarithmic form is specified for both substantive and methodological reasons. Human capital theory suggests that the increase in earnings capacity due to a year of schooling is proportional to the earnings forgone during the year, and consequently the log of earnings capacity will be approximately a linear function of schooling (Mincer, 1974; Blinder, 1976).
It is also reasonable to suppose that there is a constant proportionate return in earnings to a unit increase in the resources provided by social origins, rather than a constant dollar return.

Methodologically, using the logarithm of earnings allows for a more reasonable measurement model of response errors in reports of earnings, and it minimizes the effects of "outliers" which are often attributable to coding, keypunch, and transcription errors. One would expect the amount of error variation in earnings reports of those who earn, say, $35,000 to $40,000 to be considerably larger than the error variation in the reports of those who earn, say, \$10,000 to \$15,000. An additive measurement equation in the logarithm of earnings allows for this kind of heteroscedasticity in response errors (since it implies that error variation is a constant <u>proportion</u> or earnings), while an equation in dollar earnings requires that error variation be constant across levels of earnings. (Heteroscedasticity in conditional earnings distributions suggests a similar justification for using the logarithm of earnings in the structural equation.)

The effect of "outliers" on the correlation between two measures of dollar earnings can be quite dramatic. Our data suggest that a substantial proportion of the cases with reports of very large earnings involve coding or keypunch errors. Among the ISR subsample, ten (of more than 800) respondents reported \$40,000 or more in earnings in either the CPS interview or the ISR reinterview. In three of those ten cases, the interview and reinterview reports differ by more than \$30,000 (see appendix Table 1A). When these cases are included, the correlation between two reports of dollar earnings is .765, while the correlation between the reports of logarithmic earnings is .930. When the cases with at least \$40,000 earnings are excluded, the correlation for dollar income is .976, for logarithmic income, .970. Clearly, the logarithmic measure is much less sensitive to the presence of the outliers. To further minimize their effects, we assign a value of \$50,000 to all reports greater than that amount (yielding a correlation of .875 between two reports of dollar income, .940 between reports of logarithmic income).

The structural coefficients of the earnings function in equation 1 represent the net or direct effects of each of the determinants of earnings. However, we are also interested in the total and indirect effects of the determinants, especially of education and social origins. Consequently, we shall present estimates of reduced and semi-reduced forms as well as estimates of the structural form of equation 1. In doing so, we are implicitly assuming equation 1 to be imbedded in a larger block-recursive model with the following causal ordering:

$$\begin{array}{l} EX \\ EX2 \\ FO \\ FE \\ PI \end{array} \Longrightarrow ED \Longrightarrow OI \Longrightarrow OC \Longrightarrow FT \Longrightarrow \begin{array}{c} WKS \\ FT \\ UN \end{array} \Longrightarrow LNEARN$$

The experience variables are considered exogenous because they represent cohort differences (year of entry to first job cohorts, rather than birth cohorts) as well as labor market experience <u>per se</u>.[4]

Our strategy is to specify and estimate a measurement model for the ISR and OCGR subsamples, and then to apply those estimates to the full CPS-OCGQ sample. In this way we estimate substantive parameters in the full sample which have been corrected for response error. It is instructive to compare the corrected estimates with naive estimates for the full sample, i.e., estimates assuming perfect measurement. The models estimated in this paper apply to nonblack males in the experienced civilian labor force of March, 1973 who reported at least $1000 in 1972 earned income in the March, 1973 Current Population Survey. There are 24,352 nonblack males in the full CPS-OCGQ sample, 823 in the ISR subsample, and 556 in the OCGR subsample.

SPECIFICATION OF A MEASUREMENT MODEL

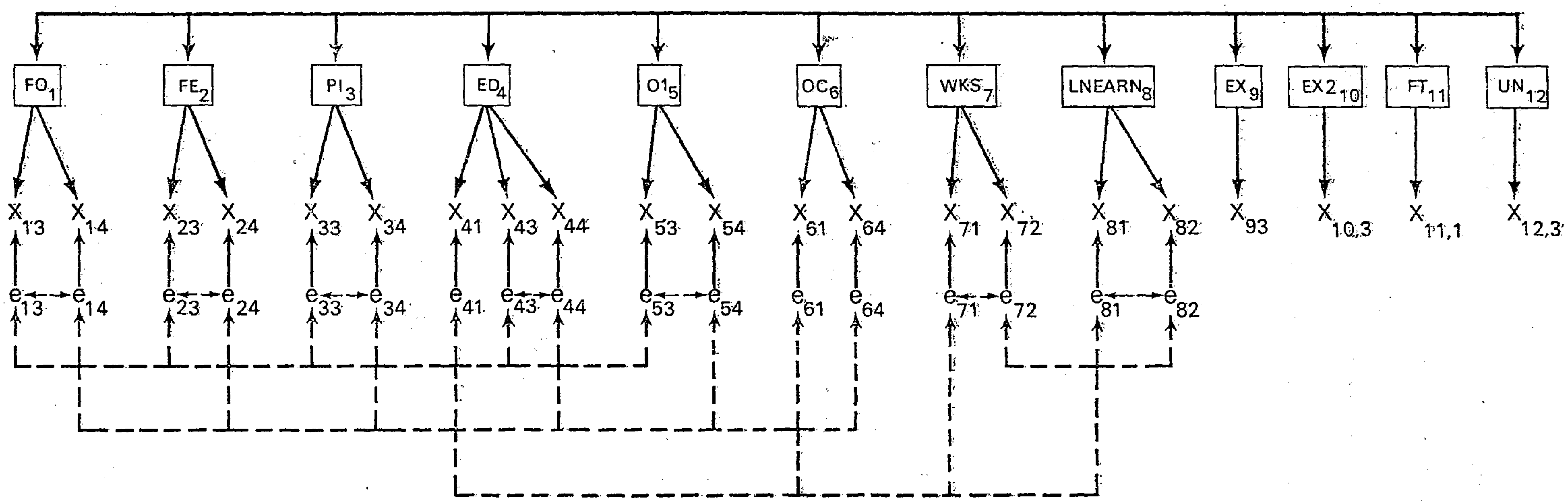
Our measurement model is presented in the path diagram of Figure 1. It shows the most general (least restricted) structure of response errors that we have estimated in the two subsamples. Ultimately, we eliminated all but two of the correlations among reporting errors. The variables enclosed in boxes (FO, FE, PI, O1, OC, WKS, LNEARN, EX, EX2, FT, UN) are unobserved "true scores";[5] the last four are assumed to be measured without error.[6] The term x_{ij} refers to the measure of the i^{th} variable obtained on the j^{th} occasion, and e_{ij} is the error component of x_{ij}.

In algebraic form, the measurement model is:

$$
\begin{aligned}
x_{13} &= \lambda_{13}FO &+ e_{13} \quad &[2]\\
x_{14} &= \lambda_{14}FT &+ e_{14} \quad &[3]\\
x_{23} &= \lambda_{23}FE &+ e_{23} \quad &[4]\\
x_{24} &= \lambda_{24}FE &+ e_{24} \quad &[5]\\
x_{33} &= \lambda_{33}PI &+ e_{33} \quad &[6]\\
x_{34} &= \lambda_{34}PI &+ e_{34} \quad &[7]\\
x_{41} &= \lambda_{41}ED &+ e_{41} \quad &[8]\\
x_{43} &= \lambda_{43}ED &+ e_{43} \quad &[9]\\
x_{44} &= \lambda_{44}ED &+ e_{44} \quad &[10]\\
x_{53} &= \lambda_{53}O1 &+ e_{53} \quad &[11]\\
x_{54} &= \lambda_{54}O1 &+ e_{54} \quad &[12]\\
x_{61} &= \lambda_{61}OC &+ e_{61} \quad &[13]\\
x_{64} &= \lambda_{64}OC &+ e_{64} \quad &[14]\\
x_{71} &= \lambda_{71}WKS &+ e_{71} \quad &[15]\\
x_{72} &= \lambda_{72}WKS &+ e_{72} \quad &[16]\\
x_{81} &= \lambda_{81}LNEARN &+ e_{81} \quad &[17]\\
x_{82} &= \lambda_{82}LNEARN &+ e_{82} \quad &[18]
\end{aligned}
$$

plus four identities: $x_{93} = EX$, $x_{10,3} = EX2$, $x_{11,1} = FT$, and $x_{12,3} = UN$.

Figure 1. Measurement Model for Variables in an Earnings Function for Nonblack Males.



Note: Variables are defined in Table 1.

To complete the model, the pattern of covariation among response errors must be specified. When multiple responses are obtained from the same individuals, three types of covariation among response errors appear particularly plausible. First, response errors in the report of a variable may covary with the respondent's true score on that variable. For example, individuals of high status may tend to understate their status while those of low status overstate their status. The implication for the measurement structure would be a non-unit slope of the population regression relating the observed measure, x_{ij}, to its corresponding true score, λ_{ij}, while maintaining the lack of correlation between the true score and e_{ij}. A second source of covariation in response error would be a tendency for respondents to overstate the consistency between different variables ascertained on a single occasion. This "within-occasion between-variable correlated error" is represented in Figure 1 by the dotted lines showing correlations among the e_{i1}, for i = 4,6,7,8 (response errors of reports obtained from the March CPS household interview), among the e_{i2}, for i = 7,8 (response errors of reports obtained from the Spring ISR reinterview), among the e_{i3}, i = 1,2,3,4,5 (response errors of reports obtained from the Fall OCGQ questionnaire), and among the e_{i4}, for 1,2,3,4, 5,6 (response errors of reports obtained from the Fall OCGR remeasurement interview). A third source of correlated response error would be contamination of the respondent's report of a given variable by his recollection of his earlier report of that variable. It seems plausible that recall contamination might occur between CPS and ISR responses obtained about a week apart and between OCGQ and OCGR obtained about three weeks apart. The former "within-variable between-occasion correlated error" is represented in Figure 1 by correlations among pairs of response errors e_{i1} and e_{i2}, for i = 7,8, and the latter by correlations between e_{i3} and e_{i4}, for i = 1,...,5. Note, however, that we assume that recall contamination does not occur between the respective Spring (CPS, ISR) and Fall (OCGQ, OCGR) reports, obtained more than five months apart and often from different persons.

We establish a metric for the true scores by fixing $\lambda_{13} = \lambda_{23} = \lambda_{33} = \lambda_{41} = \lambda_{53} = \lambda_{61} = \lambda_{71} = \lambda_{81} = 1.0$. That is, we fix the metric of the true scores to be the same as that of the observed reports which are used in models for the full CPS-OCGQ sample; the metrics of FO, FE, PI, and OI are identical to those of the corresponding OCGQ reports, and the CPS reports define the metrics for ED, OC, WKS, and LNEARN. Normalization of this kind is necessary because the metric of an unobserved variable is arbitrary, and consequently the slope coefficients with respect to indicators are identifiable only relative to each other. For example, given our normalization, a coefficient λ_{i2} greater (smaller) than unity indicates a slope of the ISR report on the corresponding true score which is steeper (flatter) than the slope of the CPS report on the true score. However, the absolute values of the two slopes are indeterminate. This normalization is imposed on all of our models. Another way of stating this normalization is that only the ratio of the slopes is identifiable. A more common normalization is to assume unit variances of true scores. However, this normalization does not allow the computation of metric coefficients relating unobservables. Error variances and reliabilities (squared true score-observed score correlations) are invariant with respect to normalization, although true score variances (and structural coefficients) do depend on which λ_{ij} are fixed to unity.

Our measurement models are all based on equations 2 to 18. In order to estimate the parameters of all 17 measurement equations, we combine estimates of different subsets of equations from the ISR and OCGR subsamples. Within subsamples, we vary the specification of the covariances among the e_{ij} and the restrictions imposed upon the λ_{ij}. We proceed by estimating and testing models in four stages, each involving a subsample and a subset of measurement equations. First, from the ISR subsample we estimate a three-variable, six-equation model for reports of education, weeks worked and earnings. Second, we briefly discuss estimates for earners in the OCGR sample of the six-variable, thirteen-equation measurement model of social origins, education, and occupational attainments that has been examined in detail in our earlier research. In the third stage, we "borrow" the OCGR subsample estimates of measurement model parameters for equations 2,4,6,11, and 13 (OCG reports of social origins and status of first job, and the CPS report of status of current job) so the measurement model in the ISR subsample can be extended to include social origins and occupational attainments. Similarly, in the fourth stage, we use the third-stage estimates of measurement model parameters for equations 15 (CPS report of weeks worked) and 17 (CPS report of earnings) to include weeks worked and earnings in a model for the OCGR subsample.

At each stage, we assess which correlations among reporting errors provide a significant improvement in fit over a model with random errors. We also look to see which λ_{ij} can be restricted to 1.0 without significantly altering the fit of the model. Our results from the four stages are then combined to provide point estimates of parameters in the full 17-equation measurement model. Finally, we use these estimates to correct the structural coefficients in equation 1 for measurement error in the full CPS-OCGQ sample.

ESTIMATION OF MEASUREMENT MODELS

Assuming the joint distribution of the reports of status variables is multivariate normal, we obtain maximum likelihood estimates of parameters of our measurement models using Jöreskog's (1970) general method for the analysis of covariance structures. The estimates have been computed from pairwise present correlations for nonblack males 20 to 65 years old in the experienced civilian labor force in March, 1973 who reported at least $1000 in 1973 earned income in the CPS interview.[7] Correlations among the 11 measures to be used in estimating models from the ISR subsample appear in Table 2, and correlations among the 15 measures to be used in estimating models from the OCGR subsample appear in Table 3. Correlations for measures available in the full CPS-OCGQ sample appear in Table 4. Corresponding means and standard deviations appear in the first two columns of the top panel of Table 6, the bottom panel of Table 6, and Table 7, respectively. There are only a few notable differences in moments between the full sample and the two subsamples. The OCGR subsample is somewhat restricted in variation on weeks worked, earnings, and education (as reported in the CPS interview). It seems likely that the OCGR subsample disproportionately includes those with accurate earnings responses, since correlations between earnings and some other variables (particularly social origins and education) are notably higher in the OCGR subsample. The correlation between CPS and OCGQ reports of education is lowest in the OCGR subsample, and examination of cross-tabulation revealed that it is

TABLE 2--Observed Correlations among Variables Measured in CPS Interview, Income Supplement Reinterview, and OCG Questionnaire: ISR Subsample of Nonblack Males in the March 1973 Experienced Civilian Labor Force Reporting Earnings of at least $1000 in the CPS (N=823)

		(1)	(2)	(3)	(4)		(5)	(6)	(7)		(8)	
Variable	Measure	x_{13}	x_{23}	x_{33}	x_{41}	x_{43}	x_{53}	x_{61}	x_{71}	x_{72}	x_{81}	x_{82}
1. FO	x_{13}	--										
2. FE	x_{23}	.547	--									
3. PI	x_{33}	.408	.405	--								
4. ED	x_{41}	.406	.438	.438	--							
	x_{43}	.389	.429	.460	.884	--						
5. OI	x_{53}	.420	.330	.332	.673	.682	--					
6. OC	x_{61}	.331	.283	.293	.584	.585	.634	--				
7. WKS	x_{71}	-.096	-.055	-.036	.037	.053	.055	.151	--			
	x_{72}	-.101	-.044	-.032	.061	.073	.054	.146	.940	--		
8. LNEARN	x_{81}	.080	.055	.127	.276	.301	.339	.429	.540	.544	--	
	x_{82}	.172	.077	.139	.293	.299	.345	.441	.550	.571	.940	--

NOTE: See Table 1 for definition of variables.

TABLE 3--Observed Correlates among Variables Measured in CPS Interview, OCG Questionnaire, and OCG Remeasurement Interview: OCGR Subsample of Nonblack Males in the March 1973 Experienced Civilian Labor Force Reporting Earnings of at least $1000 in the CPS (N=556)

Variable	Measure	(1)		(2)		(3)		(4)			(5)		(6)		(7)	(8)
		x_{13}	x_{14}	x_{23}	x_{24}	x_{33}	x_{34}	x_{41}	x_{43}	x_{44}	x_{53}	x_{54}	x_{61}	x_{64}	x_{71}	x_{81}
1. FO	x_{13}	--														
	x_{14}	.872	--													
2. FE	x_{23}	.579	.573	--												
	x_{24}	.591	.586	.937	--											
3. PI	x_{33}	.410	.415	.459	.453	--										
	x_{34}	.414	.428	.467	.463	.909	--									
4. ED	x_{41}	.408	.406	.458	.460	.467	.483	--								
	x_{42}	.424	.424	.449	.447	.412	.426	.795	--							
	x_{43}	.437	.436	.482	.492	.474	.492	.919	.835	--						
5. O1	x_{53}	.392	.411	.292	.301	.362	.351	.635	.578	.642	--					
	x_{54}	.401	.410	.324	.321	.360	.345	.636	.580	.644	.847	--				
6. OC	x_{61}	.370	.404	.309	.326	.315	.311	.577	.521	.602	.614	.617	--			
	x_{64}	.345	.382	.298	.302	.288	.296	.540	.500	.560	.578	.595	.801	--		
7. WKS	x_{71}	-.065	-.043	-.061	-.035	-.039	-.018	.071	.017	.068	.002	.031	.148	.133	--	
8. LNEARN	x_{81}	.184	.183	.132	.118	.168	.176	.363	.319	.368	.321	.346	.450	.398	.373	--

NOTE: See Table 1 for definition of variables.

TABLE 4 --Observed Correlations among Variables Measured in CPS Interview and OCG Questionnaire: Full CPS-OCGQ Sample of Nonblack Males in the March 1973 Experienced Civilian Labor Force Reporting Earning of at Least $1000 in the CPS (N=24,352)

Variable	Measure	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		x_{13}	x_{23}	x_{33}	x_{41}	x_{43}	x_{53}	x_{61}	x_{71}	x_{81}	x_{93}	$x_{10,3}$	$x_{11,1}$	$x_{12,3}$
1. FO	x_{13}	--												
2. FE	x_{23}	.531	--											
3. PI	x_{33}	.423	.434	--										
4. ED	x_{41}	.408	.468	.431	--									
	x_{43}	.388	.461	.420	.853	--								
5. OI	x_{53}	.395	.331	.309	.636	.621	--							
6. OC	x_{61}	.334	.277	.285	.572	.542	.621	--						
7. WKS	x_{71}	-.027	-.037	.005	.034	.043	.066	.140	--					
8. LNEARN	x_{81}	.107	.060	.145	.264	.271	.310	.411	.437	--				
9. EX	x_{93}	-.224	-.349	-.273	-.382	-.370	-.237	-.100	.138	.149	--			
10. EX2	$x_{10,3}$	-.014	-.028	-.087	-.148	-.150	-.102	-.105	-.114	-.240	.304	--		
11. FT	$x_{11,1}$	-.071	-.076	-.037	-.046	-.007	.023	.045	.182	.331	.110	-.142	--	
12. UN1OW	$x_{12,3}$	-.134	-.113	-.096	-.197	-.197	-.242	-.287	-.013	.042	.077	-.015	.053	--

NOTE: See Table 1 for definition of variables.

due to just a few "outliers" that happened to be selected into the OCGR subsample. The ISR subsample is somewhat restricted in variation on OCGQ reports of father's and respondent's education and exhibits slightly more variation in weeks worked than does the full CPS-OCGR sample. Correlations between the OCGQ education report and other variables are generally higher in the ISR subsample, suggesting that those giving more accurate OCGQ education reports were disproportionately included in the ISR subsample. (The error variation in the OCGQ report estimated from the ISR subsample is substantially lower than that estimated from the OCGQ subsample.) Curiously, the correlations between earnings and weeks worked are unusually high in the ISR subsample (about .54) and low in the OCGR subsample (.37). It should be stressed, however, that we have focused upon the largest discrepancies, and subsample moments are subject to considerably more sampling variability than are the full sample moments.

Goodness-of-fit tests for our various measurement models appear in Table 5. The likelihood-ratio test statistic contrasts the null hypothesis that constraints on the observed variance-covariance matrix are satisfied in the population with the alternative that the variance-covariance matrix is unrestricted. In large samples this statistic has a chi-square distribution with degrees of freedom equal to the difference between the number of variances and covariances and the number of independent parameters estimated under the hypothesized model. Moreover, when two measurement models are nested, that is, when one model can be obtained by constraining the parameters of a more general model, the difference in chi-square values provides a likelihood-ratio test of the constraints on the parameters.

Goodness-of-fit tests for six-equation, three-variable models of education, weeks worked, and earnings, estimated from the ISR subsample,

Table 5--Chi-square Goodness-of-fit Tests for Measurement Models: Nonblack Males in the Experienced Civilian Labor Force, March 1973, Who Reported at Least $1000 in Earnings in the Current Population Survey.[a]

Model	χ^2	df	p
1. Three-variable ISR model (N=823)			
A. Random measurement error--no slope restrictions	28.02	6	.000
B. Within-occasion correlated errors $\rho_{e_{41},e_{71}}$, $\rho_{e_{41},e_{81}}$, $\rho_{e_{71},e_{81}}$, $\rho_{e_{72},e_{82}}$	3.62	2	.163
C. Within-occasion correlated errors $\rho_{e_{41},e_{81}}$ and $\rho_{e_{72},e_{82}}$	5.44	4	.246
D. Within-occasion correlated error $\rho_{e_{72},e_{82}}$	14.29	5	.014
E. Within-occasion correlated errors $\rho_{e_{41},e_{81}}$ and $\rho_{e_{72},e_{82}}$ and 12 slope restrictions	5.88	6	.437
2. Six-variable OCGR model (N=556)			
Random measurement error and 5 slope restrictions	40.78	55	.923
3. Eight-variable ISR model (N=823)			
A. Random measurement error and 2 slope restrictions	48.69	23	.001
B. Within-occasion correlated errors $\rho_{e_{41},e_{81}}$ and $\rho_{e_{72},e_{82}}$ and 2 slope restrictions	23.91	21	.297
C. Within-occasion correlated error $\rho_{e_{72},e_{82}}$ and 2 slope restrictions	30.61	22	.104
4. Eight-variable OCGR model (N=556)			
A. Random measurement error and 5 slope restrictions	58.14	69	.821
B. Within-occasion correlated errors $\rho_{e_{41},e_{71}}$; $\rho_{e_{41},e_{81}}$; $\rho_{e_{61},e_{71}}$; and $\rho_{e_{61},e_{81}}$; and 5 slope restrictions	56.34	65	.769
C. Random measurement error except $\rho_{e_{41},e_{81}}$ fixed at -.14 and 5 slope restrictions	60.12	69	.768

[a]Maximum likelihood estimates were computed with the ACOVSF program described in Jöreskog, Gruvaeus and van Thillo (1970).

are presented in the top panel of Table 5. Each model is based upon equations 8, 9, 15, 16, 17 and 18. Our simplest model assumes mutually uncorrelated errors and no restrictions (other than the normalizations) on the λ_{ij}. This model corresponds to the random measurement error models of Siegel and Hodge (1968:51-52), Jencks, et al. (1973:330-336), Treiman and Hauser (1977), and the one implicitly used by other researchers applying "corrections for attenuation" (cf. Bohrnstedt, 1970). The 21 observable variances and covariances among the 6 measures provide more than enough information to identify the 15 unknown model parameters: 3 true score variances, 3 true score covariances, 6 error variances, and 3 slope coefficients (λ_{ij}). The χ^2 value of 28.02 with 6 degrees of freedom (p = .000) on line 1A suggests that the restrictions on observable moments implied by the random error model probably do not hold in the population. Model 1B introduces the three within-occasion correlations among response errors in CPS reports of education, weeks worked, and earnings, and the single correlation between response errors in ISR reports of weeks worked and earnings. Contrasting line 1B with line 1A, it is clear that adding the four within-occasion error correlations significantly improves the fit of the measurement model. We reduce χ^2 by more than 24 points while using only 4 more degrees of freedom. Furthermore, two of the four error correlations are not significantly larger than their standard errors. Model 1C introduces only the two larger error correlations, $\rho_{e_{41},e_{81}}$ and $\rho_{e_{72},e_{82}}$. Contrasting line 1C with line 1B, we see that nearly all of the improvement in fit can be attributed to these two error correlations: a negative correlation of -.17 between response errors in CPS reports of education and earnings, and a positive correlation of .33 between errors in ISR reports of weeks worked and earnings (their respective standard errors are .07 and .08). Since a tendency for respondents to understate the consistency between their education and earnings in the CPS can have a substantial effect on estimated returns to schooling, we want to be relatively certain that the correlation $\rho_{e_{41},e_{81}}$ does indeed contribute to the fit of the model. Consequently, we estimate a model where $\rho_{e_{41},e_{81}}$ is constrained to equal zero while $\rho_{e_{72},e_{82}}$ remains in the model. Contrasting lines 1D and 1C, it is clear that $\rho_{e_{41},e_{81}}$ makes a statistically significant contribution to the fit (eliminating $\rho_{e_{41},e_{81}}$ increases χ^2 by 8.85 and adds 1 degree of freedom, $p < .005$). Finally, two of the three free λ_{ij} were estimated to be nearly 1.0. However λ_{43} was estimated to be 1.08 (with a standard error of .02). Thus, as our earlier research based on the OCGR subsample has shown, the slope of the OCGQ report of education on the true score appears significantly steeper than the slope for the CPS report. Model 1E, imposing $\lambda_{72} = \lambda_{82} = 1.0$, increases the χ^2 value marginally over that of 1C, while adding two degrees of freedom. Thus, in the third stage, we combine the specification of model 1E with selected parameters estimated within the OCGR subsample.

While we have been able to test for the presence of some types of correlated error in the ISR subsample, we have not been able to test for the presence of the plausible within-variable error correlations involving weeks worked and income ($\rho_{e_{71},e_{72}}$, $\rho_{e_{81},e_{82}}$). These are not identified in any of the ISR subsample models examined here. Therefore, neither can any such correlations be detected, nor could they affect the fit of the measurement model. However, within-variable error correlations may affect estimated structural relations among true variables and we shall consider this possibility below.

Next, we briefly discuss the six-variable, thirteen-equation measurement model for social origins, education, and occupational attainments. The specification of Model 2 is identical to that of the "final" measurement model for nonblack males developed in our earlier research (Bielby, Hauser and Featherman, 1977a): mutually uncorrelated response errors and two free slope parameters, $\lambda_{43} = \lambda_{44}$ and λ_{64}. The present estimates for Model 2 differ from those resported elsewhere because males reporting less than $1000 in earnings in the CPS (22 cases) have been excluded from the subsample. The 36 parameters of Model 2 do remarkably well in representing the 91 variances and covariances among the reports (χ^2 = 40.78 with 55 degrees of freedom, p = .923). Direct tests and indirect evidence of the extent of within-occasion and within-variable correlations among reporting errors are presented in our earlier paper. We found virtually no evidence of any error correlations. Point estimates of measurement parameters based upon the 556 OCGR earners are nearly identical to those estimated previously from the 578 OCGR males in the experienced civilian labor force. Thus, in the fourth stage, we combine the specification of Model 2 with selected parameters estimated in the third stage.

In the third stage of estimation and specification we add equations 2, 4, 6, 11 and 13 for OCGQ reports of education and first job status, and the CPS report of current job status to the specification of Model 1E in order to obtain an eight-variable, eleven-equation model which can be estimated in the ISR subsample. Error variances of e_{13}, e_{23}, e_{33}, e_{53}, and e_{61} are borrowed from the second stage results (Model 2) and appear in column 3 of the top panel of Table 6. Relying on Model 2, we assume mutually uncorrelated response errors among reports of social origins, schooling, and occupational achievements. Model 3A assumes all errors to be mutually uncorrelated, constrains λ_{72} and λ_{82} to equal 1.0, but leaves λ_{43} unconstrained. The 11 reports provide 66 observed moments from which to solve for 43 parameters: 8 true score variances, 28 true score covariances, 6 error variances (two others are known), and one slope coefficient. The χ^2 value of 48.69 with 23 degrees of freedom ($p = .001$) indicates that, as expected (given the results of Model 1A), the restrictions on observable moments implied by Model 3A probably do not hold in the population. (The degrees of freedom are not strictly correct, since the five borrowed error variances are sample estimates, not known population parameters. Nevertheless, differences in degrees of freedom for nested models are correct.) Model 3B introduces the two error correlations, $\rho_{e_{41},e_{81}}$ and $\rho_{e_{72},e_{82}}$, detected in Model 1C. Contrasting line 3B with line 3A, we again see a significant improvement in fit due to the two within-occasion error correlations; χ^2 drops by nearly 25 points, using only 2 degrees of freedom, ($p < .001$). Point estimates of the error correlations are -.14 and .37 (their respective standard errors are .05 and .09). To further examine the degree to which $\rho_{e_{41},e_{81}}$ contributes to the fit of the model, we estimate Model 3C where it is constrained to equal zero while $\rho_{e_{72},e_{82}}$ remains in the model. Contrasting lines 3C with 3B, we again find that $\rho_{e_{41},e_{81}}$ makes a statistically significant contribution to the fit; eliminating $\rho_{e_{41},e_{81}}$ increases

χ^2 by 6.70 and adds one degree of freedom ($p < .01$). Thus, we accept Model 3B as the final measurement model estimated from the ISR subsample. Parameter estimates for Model 3B appear in the top panel of Table 6. They are discussed in detail below.

In the fourth stage, we add equations 15 and 17 for CPS reports of weeks worked and earnings to the specification of Model 2 in order to obtain an eight-variable, fifteen-equation model to be estimated in the OCGR subsample. Error variances of e_{71} and e_{81} are borrowed from the third stage results (Model 3B) and appear in column 3 of the bottom panel of Table 6. Relying on the results reported above, we allow just two free slope parameters, λ_{64} and $\lambda_{42} = \lambda_{43}$. In Model 4A, we specify all fifteen response error terms to be mutually uncorrelated. In that model, the 15 reports provide 120 observed moments from which to solve for 51 model parameters: 8 true score variances, 28 true score covariances, 13 error variances (two others are known), and two slope coefficients. The χ^2 value of 58.14 with 69 degrees of freedom ($p = .821$) indicates that the restrictions implied by Model 4A provide a reasonable fit to the observable moments. (As in the third stage models, the degrees of freedom are not strictly correct, since two of the "known" fixed parameters, $\sigma_{e_{71}}$ and $\sigma_{e_{81}}$, are actually estimates from the OCGR subsample). In model 4B, we add four within-occasion correlations among errors in CPS reports. Two of the four, $\rho_{e_{41},e_{71}}$ and $\rho_{e_{41},e_{81}}$, were identified and estimated in the ISR subsample, and we detected a statistically significant effect of the latter. The other two, $\rho_{e_{61},e_{71}}$ and $\rho_{e_{71},e_{81}}$, were not

Table 6--Observed Moments and Measurement Model Parameter Estimates: Nonblack Males in the Experienced Civilian labor Force, March 1973, Who Reported at Least $1000 in the Current Population Survey.

Variable	Measure	Income Supplement Reinterview (N=813)[c] (1) Mean $\hat{\mu}_{ij}$	(2) Observed Std. Dev. $\hat{\sigma}_{ij}$	(3) Std. Dev. of Error[a] $\hat{\sigma}_{e_{ij}}$	(4) Relative Slope[a] λ_{ij}	(5) Reliability Coefficient $(\hat{\sigma}^2_{T_i}/\hat{\sigma}^2_{x_{ij}})\lambda^2_{ij}$	(6) Percent of Cases with Data Present
1. FO	x_{13}	32.74	23.78	9.26	1.00	.848	95
	x_{14}	--	--	--		--	--
2. FE	x_{23}	9.055	3.895	1.14	1.00	.914	94
	x_{24}	--	--	--	--	--	--
3. PI	x_{33}	8.750	.8637	0.32	1.00	.861	90
	x_{34}	--	--	--	--	--	--
4. ED	x_{41}	12.33	3.019	1.08 (.06)	1.00	.872	100[b]
	x_{43}	12.19	3.220	1.03 (.07)	1.08 (.02)	.897	94
	x_{44}	--	--	--	--	--	--
5. OI	x_{53}	35.51	25.12	9.93	1.00	.845	88
	x_{54}	--	--	--	--	--	--
6. OC	x_{61}	43.24	25.44	9.82	1.00	.850	100[b]
	x_{64}	--	--	--	--	--	--
7. WKS	x_{71}	47.89	8.898	2.38 (.17)	1.00	.927	100[b]
	x_{72}	47.87	8.784	1.93 (.20)	--	.953	97
8. LNEARN	x_{81}	9.130	.6389	0.178 (.011)	1.00	.922	100[b]
	x_{82}	9.136	.6266	0.128 (.014)	--	.958	93

Variable	Measure	OCG Remeasurement Subsample (N=556)[d] (1) Mean $\hat{\mu}_{ij}$	(2) Observed Std. Dev. $\hat{\sigma}_{ij}$	(3) Std. Dev. of Error[a] $\hat{\sigma}_{e_{ij}}$	(4) Relative Slope[a] λ_{ij}	(5) Reliability Coefficient $(\hat{\sigma}^2_{T_i}/\hat{\sigma}^2_{x_{ij}})\lambda^2_{ij}$	(6) Percent of Cases with Data Present
1. FO	x_{13}	32.73	24.21	9.22 (.55)	1.00	.854	96
	x_{14}	33.29	23.54	7.83 (.60)	1.00	.889	95
2. FE	x_{23}	8.930	4.139	1.14 (.09)	1.00	.923	95
	x_{24}	8.904	4.093	0.91 (.10)	1.00	.951	94
3. PI	x_{33}	8.710	.9323	0.32 (.021)	1.00	.876	89
	x_{34}	8.773	.8943	0.21 (.028)	1.00	.943	90
4. ED	x_{41}	12.18	2.821	0.97 (.04)	1.00	.884	100[b]
	x_{43}	12.01	3.407	1.78 (.06)	1.07 (.02)	.717	94
	x_{44}	12.16	2.903	0.60 (.06)	1.07 (.02)	.956	96
5. OI	x_{53}	34.86	24.77	9.89 (.54)	1.00	.839	90
	x_{54}	32.31	24.28	9.30 (.55)	1.00	.856	95
6. OC	x_{61}	41.69	25.26	9.52 (.81)	1.00	.857	100[b]
	x_{64}	39.91	24.87	12.44 (.62)	0.92 (.04)	.749	100[b]
7. WKS	x_{71}	48.58	7.165	2.38	1.00	.889	100[b]
	x_{72}	--	--	--	--	--	--
8. LNEARN	x_{81}	9.150	.5753	0.178	1.00	.904	100[b]
	x_{82}	--	--	--	--	--	--

[a] Approximate standard errors of parameter estimates appear in parentheses.

[b] Missing values have been allocated for NA cases.

[c] In the ISR subsample, $\hat{\rho}_{e\ e}$ = -.14, $\hat{\rho}_{\sigma\ \sigma}$ = .37.

estimated in the ISR subsample. Contrasting lines 4B and 4A, it is clear that the four error correlations contribute virtually no improvement to the fit of the model, decreasing χ^2 by only 1.80 while using four degrees of freedom ($p > .75$). Indeed, if a decrease of that amount could be attributed to just one of the error correlations, it would not be statistically significant (i.e. $p > .10$ for $\chi^2 = 1.80$ with one degree of freedom). Thus, the fourth stage OCGR models do not replicate the finding from the ISR subsample of a significant negative correlation among reporting errors in CPS reports of education and earnings. The point estimate of $\hat{\rho}_{e_{41},e_{81}}$ in Model 4B is .085, compared to $\hat{\rho}_{e_{41},e_{81}} = -.14$ under Model 3B. Model 4C, which is identical to Model 4A except that $\rho_{e_{41},e_{81}}$ is fixed at -.14, provides a worse fit than Model 4A. However, the comparison of fourth stage OCGR models are not quite as conclusive as they may appear. The error correlations in question are affected by parameter estimates "borrowed" from the ISR subsample. Consequently, sampling error in the borrowed estimates and differential selection bias in the two subsamples may have reduced our ability to detect these particular within-variable error correlations.[8]

Model 4A is our final measurement model in the OCGR subsample, and Model 3B is our final measurement model in the ISR subsample. Together, they provide at least one point estimate of every parameter in the 17-equation measurement model (see Table 6), but they provide substantially different estimates for two of those parameters. One is the response error variation in the OCGQ report of educational attainment; $\sigma_{e_{43}}$ is estimated as 1.03 years from the ISR subsample and 1.78 years from the OCGR subsample. This difference is rather large relative to the respective standard errors of the variance components. As we noted above, the OCGR subsample contained a disproportionate number of cases with large discrepancies between CPS and OCGQ reports of educational attainment, while those with more accurate OCGQ responses seemed disproportionately represented in the ISR subsample. This appears to account for the different estimates. Furthermore, this discrepancy will not affect bias due to our assessment of response errors in the earnings function, for only the CPS report of education is used in estimating equation 1.

The discrepancy in estimates of $\rho_{e_{41},e_{81}}$ is more serious because it affects estimates of the returns to schooling. In our "listwise" estimates for the ISR subsample (excluding the 290 cases with any missing or allocated responses), we detected no such error correlation, suggesting that it might be an artifact of either the Census allocation procedure or of the "pairwise present" computing scheme.[9] Nor was the error correlation detected in (pairwise) ISR subsample estimates using dollar income rather than logarithmic income. Because of the ambiguous status of the negative correlation between errors in CPS reports of education and earnings, in assessing biases due to response errors we will first assume that the error correlation is zero and then assess the effect of an error correlation of -.14.

Within-variable, between-occasion correlations among response errors, for example, contamination of a later report of a variable by recall of an earlier report, have not been discussed in detail. As noted above, our earlier research (Bielby, Hauser and Featherman, 1977a) has produced strong evidence that no such correlations exist among errors in reports of social origins, education, and occupational attainments. Given those results, there appears to be no *prima facie* reason why such correlations should be presumed to exist among reports of either weeks worked or earnings. Nevertheless, we assess below the effect on estimates of the earnings function of a within-variable correlation between errors in reports of earnings, $\rho_{e_{81},e_{82}}$, as high as 0.5. A value of $\rho_{e_{81},e_{82}}$ greater than zero would increase estimates of $\sigma_{e_{81}}$ and $\sigma_{e_{82}}$ (and therefore lower reliabilities of earnings below levels reported in Table 6). Since it would leave the covariance between e_{41} and e_{81} unchanged, it would decrease the estimate of $\rho_{e_{41},e_{81}}$. All other estimates would remain unchanged, as would the tests of models in Table 5.

Estimates in Table 6 of measurement model parameters involving social origins, education, and occupational achievements are nearly identical to those reported in our earlier research. OCGR remeasurement interview reports have uniformly lower error variances than corresponding OCGQ questionnaire reports, and social origins appear to be measured no less accurately than socioeconomic achievements. Both ISR and OCGR subsample estimates show the regression of the CPS report of education on the true score to be flatter than the regression for the OCGQ report.[10] The CPS report of current occupational status has a steeper slope than does the Fall remeasurement report.

The Income Supplement Reinterview, conducted by experienced Census personnel, often in person, appears to be more accurate than the CPS interview. Error variation ($\hat{\sigma}_{e_{ij}}$) in the reinterview report of weeks worked is about 20 percent lower than that in the CPS report, and error variation in the reinterview report of earnings is nearly 30 percent lower than that in the CPS report. The standard deviation of errors in CPS reports of earnings is 17.8 percent; that is, the error variation is predicted to be $890 for true earnings of $5000, $1780 for true earnings of $10,000 and $4,450 for true earnings of $25,000. The standard deviation of errors in ISR reports of earnings is 12.8 percent, which implies error variations of $640, $1280, and $3200 at the respective true levels of earnings cited above. In contrast, a measurement model in dollar earnings yields estimates of $2450 and $2330 for error variation in CPS and ISR reports, respectively, regardless of the level of earnings; as noted earlier the data are not consistent with the assumption of constant error variation in dollar earnings. Moreover, the estimates based on dollar earnings suggest there is more error in earnings near the middle of the earnings distribution ($9000 to $11,000) than do the estimates based on our semi-logarithmic specification.

To summarize, response errors in reports of socioeconomic variables by nonblack males appear to be almost completely random. Respondents appear to overstate the consistency between earnings and weeks worked in the Income Supplement Reinterview, and less conclusive results suggest that they may understate the consistency between CPS reports of educational attainment and earnings. Overall, retrospective reports of social origins by nonblack males are about as accurate as reports of their own educational attainments and socioeconomic achievements. Estimates of a 17-equation measurement model obtained from the ISR and OCGR subsamples can now be applied to the full CPS-OCGQ sample in order to assess the biases in the coefficients of an earnings function that are attributable to response error.

ESTIMATION OF AN EARNINGS FUNCTION IN THE FULL CPS-OCGQ SAMPLE

Estimates from the two subsamples of parameters of measurement model equations 2, 4, 6, 8, 11, 13, 15 and 17 are used to obtain corrected full sample moments among variables entering the earnings function (equation 1). By computing the earnings function from both corrected and uncorrected full-sample moments we can assess the extent to which response errors bias estimates of the coefficients of equation 1.[11]

Estimates of measurement model parameters for social origins (equations 2, 4, and 6), and occupational attainments (equations 11 and 13) are taken from OCGR subsample Model 4A; that is, they are selected from the bottom panel of Table 6. Estimates of measurement model parameters for weeks worked and earnings (equations 15 and 17) are taken from ISR subsample Model 3B; that is, they are selected from the top panel of Table 6. Estimates of measurement model parameters for educational attainment (equation 8) are pooled from both subsamples. Since the λ_{ij} are normalized to 1.0 for all reports used to estimate the earnings function in the full sample, only estimates of the error variation ($\sigma_{e_{ij}}$) and $\rho_{e_{41},e_{81}}$ need be taken from Table 6. The estimates of error variation appear in column 3 of Table 7; $\hat{\sigma}_{e_{41}}$ is the average of the OCGR and ISR estimates weighted by the inverse of their respective standard errors. Under the specification of the measurement model, the true score variances are computed as the observed variances in the full sample less the corresponding response error variances from the subsamples (column 4 of Table 7). Except for the covariance of educational attainment and earnings, the true score covariances are set equal to the observed full sample covariances. Under the specification

Table 7--Full CPS-OCGQ Sample Observed Moments (N=24,352), Subsample Estimates of Standard Deviations of Errors, and Combined Estimates of True Standard Deviations: Nonblack Males in the Experienced Civilian Labor Force, March 1973, Who Reported Earnings of at Least $1000 in the Current Population Survey (N=24,352).

Variable		Measure	(1) Mean $\hat{\mu}_{ij}$	(2) Observed Std. Dev. $\hat{\sigma}_{x_{ij}}$	(3) Std. Dev. of Error $\hat{\sigma}_{e_{ij}}$	(4) Std. Dev. of True Score $\hat{\sigma}_{T_{ij}}$	(5) Percent of Cases with Data Present
1.	FO	x_{13}	31.02	22.90	9.22	20.96	94
2.	FE	x_{23}	8.775	4.020	1.14	3.855	94
3.	PI	x_{33}	8.683	.9210	0.32	.8636	90
4.	ED	x_{41}	12.10	3.050	1.01	2.878	100[a]
		x_{43}	11.96	3.402	--	--	94
5.	OI	x_{53}	34.00	24.68	9.89	22.61	88
6.	OC	x_{61}	41.53	25.02	9.52	23.14	100[a]
7.	WKS	x_{71}	47.93	8.416	2.38	8.072	100[a]
8.	LNEARN	x_{81}	9.093	.6534	0.178	.6287	100[a]
9.	EX	$x_{9,3}$	19.65	13.44	--	--	99
10.	EX2	$x_{10,3}$	18.07	17.81	--	--	99
11.	FT	$x_{11,1}$	0.964	0.186	--	--	100
12.	UN	$x_{12,3}$	0.305	0.460	--	--	99

$\rho_{e_{41}, e_{81}} = 0$, the true score covariance of education and earnings is also equal to the observed full sample covariances. Under the specification $\rho_{e_{41}, e_{81}} = -.14$, the true score covariance is equal to the observed full-sample covariance less the estimated covariance of the errors (where the latter is $\rho_{e_{41}, e_{81}} \sigma_{e_{41}} \sigma_{e_{81}}$). The corrected correlations obtained from the true score variances and covariances are given in Table 8.

Table 9 presents corrected estimates of structural and reduced form parameters of the earnings function under the assumption that response errors are random. Metric and standardized coefficients are presented for the structural form (line 6), and reduced forms of the earnings function (lines 1 through 5). In addition, a simple schooling and experience function is estimated (line 7) and the bivariate schooling regression is presented (line 8).

The first reduced form equation in Table 9 reveals that while each of the three social origin variables has an independent effect on earnings, parental income (PI) has by far the largest with an elasticity of nearly .12. The total effect of father's occupational status (FO) is about 2 percent for 10 SEI points, and the effect of a year of father's education (FE) is 0.76 percent. The experience variables (EX, EX2) have the largest relative impact on earnings, and the earnings-experience profile, net of social origins, peaks at about 27 years of experience.[12]

Education (ED) mediates virtually all of the effects of FO and FE (compare lines 2 and 1), but nearly half of the effect of parental income is not mediated by schooling. That is, father's occupational status

TABLE 8 --Corrected Correlations: Full CPS-OCGQ Sample of Nonblack Males in the March 1973 Experienced Civilian Labor Force Reporting Earnings of at Least $1000 in the CPS (N=24,352)[a]

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1. FO	--											
2. FE	.605	--										
3. PI	.493	.483	--									
4. ED	.472	.517	.488	--								
5. O1	.471	.377	.360	.736	--							
6. OC	.395	.312	.329	.656	.733	--						
7. WKS	-.031	-.040	.006	.038	.075	.158	--					
8. LNEARN	.121	.065	.161	***[b,c]	.352	.462	.474	--				
9. EX	-.245	-.364	-.291	-.405	-.259	-.108	.144	.155	--			
10. EX2	-.015	-.029	-.093	-.157	-.111	-.114	-.119	-.249	.304	--		
11. FT	-.078	-.079	-.039	-.049	.025	.049	.190	.344	.110	-.142	--	
12. UN	-.146	-.118	-.101	-.209	-.264	-.310	-.014	.044	.077	-.015	.053	--

[a]Correlations have been corrected with measurement model parameters estimated from subsamples of 823 (ISR) and 556 (OCGR) observations.

[b]$\rho_{ED,LNEARN} = .291$ when $\rho_{e_{41},e_{81}} = 0.00$.

[c]$\rho_{ED,LNEARN} = .305$ when $\rho_{e_{41},e_{81}} = -.14$.

TABLE 9--Corrected Estimates of an Earnings Function for 1972 Earned Income (LNEARN), Assuming Random Errors: Nonblack Males in the Experienced Civilian Labor Force, March 1973, Who Reported at Least $1000 in the Current Population Survey (N = 24,352).

Equation	Predetermined Variables											R^2	Components of Variation[a]		
	EX	EX2	FO[b]	FE	PI	ED	OI[b]	OC[b]	WKS	FT	UN		Residual σ_u	Explained $\sigma_{\hat{t}}$	Total σ_t
1.	.0161 (.344)	-.0118 (-.336)	.0277 (.092)	.0076 (.046)	.118 (.162)	--	--	--	--	--	--	.177	.570	.264	.6285
2.	.0195 (.417)	-.0109 (-.309)	.0079 (.026)	-.0069 (-.043)	.053 (.075)	.0841 (.385)	--	--	--	--	--	.266	.538	.324	
3.	.0189 (.405)	-.0108 (-.305)	-.0103 (-.034)	-.0029 (-.018)	.062 (.085)	.0407 (.186)	.0773 (.278)	--	--	--	--	.299	.526	.344	
4.	.0161 (.344)	-.0101 (-.285)	-.0151 (-.050)	-.0017 (-.010)	.056 (.077)	.0175 (.080)	.0291 (.105)	.0910 (.335)	--	--	--	.345	.509	.369	
5.	.0128 (.273)	-.0081 (-.229)	-.0078 (-.026)	-.0020 (-.012)	.050 (.069)	.0184 (.084)	.0285 (.103)	.0736 (.271)	.0274 (.352)	--	--	.459	.462	.426	
6.	.0112 (.239)	-.0065 (-.185)	-.0030 (-.010)	-.0020 (-.012)	.046 (.063)	.0216 (.099)	.0261 (.094)	.0840 (.309)	.0248 (.319)	.707 (.210)	.220 (.161)	.527	.432	.456	
7.	.0194 (.416)	-.0110 (-.311)	--	--	--	.0897 (.411)	--	--	--	--	--	.261	.540	.321	
8.	--	--	--	--	--	.0635 (.291)	--	--	--	--	--	.085	.601	.183	

NOTE: Standardized coefficients appear in parentheses. Estimates of error variances are based on subsamples of 823 (ISR) and 556 (OCGR) observations.

[a] Components are expressed as standard deviations. The additive decomposition was $\sigma_t^2 = \sigma_{\hat{t}}^2 + \sigma_u^2$.

[b] Variables expressed in the metric of Duncan SEI scores have been divided by 10 and corresponding coefficients multiplied by 10.

and father's educational attainment affect earnings by increasing or decreasing the length of schooling, but the financial status of the family of origin affects earnings both through schooling and in other ways. The total return to a year of schooling, net of experience and social origins, is more than 8 percent. Overall, schooling, experience, and social origins account for over one-fourth of the variance in earnings.

In line 3, we find that there is almost an 8 percent return to 10 SEI points of first job status (01), net of education, experience, and social origins, with 01 mediating about one-half the return to schooling, but none of the effect of parental income. There is a 9 percent return to 10 SEI points of current job status (line 4), with OC mediating more than half the return to first job status. The return to schooling net of both occupational statuses is less than 2 percent, but the elasticity of parental income remains almost .06. Schooling, experience, social origins, and occupational achievements account for over one-third of the variance in earnings.

In line 5, we find a 2.74 return to each additional week worked (WKS), the largest effect relative to those of the other variables in the earnings equation (compare standardized coefficients). Labor supply, as indicated by WKS, mediates about 20 percent of the effect of experience, probably through the decreased labor supply of older workers. WKS mediates a little less than 20 percent of the effect of current occupation.

The full equation (line 7) accounts for over half the variance in the log of earnings. Net of all other variables in the model, union members (UN) earn about 25 percent more than nonmembers ($e^{.22} = 1.25$), and those employed full-time (FT) earn about twice as much as those working part-time

($e^{.707} = 2.03$). The net return to an additional week worked is 2.5 percent in the full equation, and the return to 10 SEI points of current occupational status is over 8 percent. The return to schooling net of experience, social origins, occupational attainments, labor supply, and unionization is about 2 percent. As others have found (Sewell and Hauser, 1975; Treiman and Hauser, 1977), there is a direct effect of parental income on earnings which is not mediated by schooling, experience, occupational achievements, or labor supply variables. The elasticity of earnings with respect to parental income in the final earnings equation is .046, which is about 40 percent of its total effect.

The total effect of schooling appears to be overstated by about 7 percent when social origins are excluded from the earnings function, and social origins make a negligible contribution to the variation in earnings net of schooling and experience (compare lines 2 and 7). Finally, experience has a "suppressor effect" on the total return to schooling. Experience and schooling both enhance earnings, but the two are negatively correlated. Schooling and work are alternative uses of the time of individuals so, within a birth cohort, men with more schooling will have less work experience. Moreover, men in older cohorts, who have more experience, obtained less schooling in the aggregate than more recent cohorts. Thus the ED coefficient is almost 30 percent lower in line 8 than in line 7 of Table 9.

Table 10 presents structural and reduced forms of the same earnings function estimated from uncorrected CPS-OCGQ full sample moments. Comparing corresponding estimates in Tables 9 and 10 allows us to assess the apparent bias in the uncorrected estimates, assuming response errors are mutually uncorrelated.

TABLE 10--Uncorrected Estimates of Parameters of an Earnings Function for 1972 Earned Income (LNEARN): Nonblack Males in the Experienced Civilian Labor Force, March 1973, Who Reported at Least $1000 in the Current Population Survey (N = 24,352).

	Predetermined Variables												Components of Variation[a]		
Equation	EX	EX2	FO[b]	FE	PI	ED	OI[b]	OC[b]	WKS	FT	UN	R^2	Residual σ_u	Explained $\sigma_{\hat{t}}$	Total σ_t
1.	.0158 (.326)	-.0119 (-.324)	.0238 (.084)	.0093 (.057)	.103 (.146)	--	--	--	--	--	--	.160	.599	.261	.6534
2.	.0189 (.389)	-.0111 (-.303)	.0092 (.032)	-.0028 (-.017)	.054 (.077)	.0706 (.330)	--	--	--	--	--	.231	.573	.314	
3.	.0187 (.384)	-.0109 (-.298)	-.0024 (-.008)	-.0018 (-.011)	.054 (.077)	.0421 (.197)	.0606 (.229)	--	--	--	--	.261	.562	.334	
4.	.0165 (.340)	-.0103 (-.282)	-.0073 (-.026)	-.0016 (-.010)	.048 (.068)	.0231 (.108)	.0306 (.116)	.0715 (.274)	--	--	--	.302	.546	.359	
5.	.0133 (.274)	-.0084 (-.230)	-.0027 (-.010)	-.0014 (-.009)	.044 (.062)	.0224 (.105)	.0287 (.109)	.0597 (.229)	.0256 (.329)	--	--	.403	.505	.415	
6.	.0117 (.241)	-.0069 (-.187)	.0009 (.003)	-.0011 (-.007)	.042 (.059)	.0254 (.119)	.0283 (.107)	.0664 (.254)	.0232 (.299)	.734 (.209)	.200 (.141)	.464	.478	.445	
7.	.0185 (.380)	-.0111 (-.302)	--	--	--	.0781 (.364)	--	--	--	--	--	.225	.575	.310	
8.	--	--	--	--	--	.0566 (.264)	--	--	--	--	--	.070	.630	.172	

NOTE: Standardized coefficients appear in parentheses.

[a]Components are expressed as standard deviations. The additive decomposition is $\sigma_t^2 = \sigma_{\hat{t}}^2 + \sigma_u^2$.

[b]Variables expressed in the metric of Duncan SEI scores have been divided by 10 and corresponding coefficients multiplied by 10.

First, we note that the total variation in earnings, σ_t, is overstated by about 4 percent when response errors are ignored. Residual variation, σ_u, which includes measurement errors in reports of earnings in the uncorrected estimates in Table 10, is overestimated 5 to 9 percent in the reduced form equations (lines 1 to 5), and 11 percent in the structural equation (line 6). Thus, if we ignore response errors, we slightly overstate the total amount of earnings inequality which <u>cannot</u> be attributed to the factors represented in our earnings function. (The biases in estimates of unexplained variation were much higher, up to 27 percent, in equations for educational and occupational attainments; see Bielby, Hauser, and Featherman, 1977a). The amount of variation in earnings attributable to the variables in the earnings function, explained variation $\sigma_{\hat{t}}$, is underestimated by no more than 3 percent in the reduced-form and structural equations. In all, there is a 12 to 14 percent understatement of the proportion of variance explained (R^2) in earnings in the reduced form and structural equations.

The estimated reduced form effect of one social origin variable, father's education (FE), is actually overestimated by 22 percent using uncorrected moments (see line 1), but the reduced form effects of father's occupational status (FO) and parental income (PI) are underestimated by 14 and 12 percent, respectively. Further, when response errors are ignored the unmediated effect of parental income is underestimated by 12 to 15 percent in lines 3 through 5 and by 8 percent in the structural equation (line 6).

While the total effect of education (ED) is underestimated by 16 percent (a return of about 7 percent per year in line 2 of Table 10 and nearly 8.5 percent in line 2 of Table 9), the effects of education net of current occupation (OC), weeks worked (WKS), full-time employment status (FT), and union membership (UN) are overstated considerably in the uncorrected equations. The unmediated effect of education is overstated by 32 percent in equation 4, 22 percent in equation 5, and 18 percent in equation 6 when we assume errors in reports of education and earnings are uncorrelated.

The total economic return to occupational status of first job after completing schooling (O1) is underestimated by 22 percent in line 3 (a return of about 6 percent for 10 SEI points in Table 10 and 7.73 percent in Table 9). However, the net effect of O1 is overstated by about 8 percent in the structural equation (line 6), and by a bit less than that in line 4. In contrast, the total and net effects of occupational status of March 1973 job (OC) are all underestimated by about 20 percent when measurement errors are ignored.

The total and net effects of weeks worked in 1972 (WKS) are understated slightly in the uncorrected estimates (about 7 percent in line 5 and 6 percent in line 6). The effect of full-time employment status is slightly overstated in the structural equation (by about 4 percent), and the effect of union membership is understated by 9 percent (despite the fact that both are assumed to be measured perfectly).

The coefficients of the experience variables, EX and EX2, are barely affected by measurement errors. The largest bias is about 6 percent in the coefficient of EX2 in the structural equation (line 6). The peak of the earnings-experience profile differs at most by about one year in the corrected and uncorrected equations.

The specification bias in the total effect of schooling (ED), when social origin variables are ignored (line 7 versus line 3) actually appears larger when the comparison is made from uncorrected (Table 10) instead of corrected (Table 9) coefficients. The overstatement is about 7 percent in Table 9 and 11 percent in Table 10. This finding, contrary to what Bowles (1972) and others have implied, occurs because social origins are measured about as reliably as educational attainment.

Our findings about bias under the specification of random measurement error may be summarized as follows. The total effects (coefficient of variation in the first equation in which it appears) of two of three social origin variables, FO and PI, and of educational attainment are modestly underestimated (by 12 to 16 percent) when response errors are ignored. The total effects of occupational attainments, O1 and OC, are underestimated by an even greater amount, over 20 percent, and the net returns to current occupational status are also underestimated by about 20 percent. Biases in effects of labor supply variables, WKS and FT, and of union membership are considerably smaller. The return to parental income that is unmediated by other variables is consistently underestimated by 8 to 15 percent. In contrast, the unmediated effects of schooling are overestimated by as much as 32 percent when we ignore measurement error. Variation in earnings not attributable to factors represented in the earnings function is overstated by 5 percent (in the reduced form) and by 11 percent (in the structural form). Overall, the proportion of variance explained in earnings is underestimated by no more than 13 percent when random response errors are ignored.

These comparisons have been based on the assumption that there is no negative correlation between response errors of CPS reports of education and earnings. Table 11 presents corrected estimates under the assumption that $\rho_{e_{41}, e_{81}} = -.14$. Most of these coefficients are quite close to the corrected estimates of Table 9, obtained under the assumption that $\rho_{e_{41}, e_{81}} = 0$. The estimated components of variation in earnings, σ_u, σ_t, $\sigma_{\hat{t}}$, change little under the alternative measurement model, and, of course, the estimates of equation 1 are identical, since schooling does not appear in it.

The total effects of each independent variable are identical under the two specifications, but some of the unmediated net effects change drastically. This is particularly true of the net effects of schooling (ED), unmediated by occupational attainments, labor supply, and/or union membership. Under the alternative specification, the uncorrected return to schooling appears to be <u>underestimated</u> by 15, 14, 20, and 19 percent in equations 3 through 6, respectively, instead of substantially over-estimated, as our earlier comparisons show. The results are not surprising, since the negative error correlation implies a larger true covariation between education and earnings than does the random error model.

The returns to parental income (PI) in equations 2 through 6 are somewhat lower in the corrected estimates under the alternative measurement model, eliminating much of the apparent bias shown in the earlier comparisons. In addition, the unmediated returns to status of first job (O1) in lines 4 through 6 are considerably lower under the alternative measurement model, implying that the uncorrected estimates substantially overstate the unmediated net returns to career origins.

Thus, the significant implication of a negative association between errors in CPS reports of schooling and earnings is an increase in the net unmediated return to schooling in semi-reduced form and structural equations at the expense of returns to parental income and status of

TABLE 11--Corrected Estimates of Parameter of an Earnings Function for 1972 Earned Income (LNEARN), Assuming $\rho_{e_{41}, e_{81}} = -.14$: Nonblack Males in the Experienced Civilian Labor Force, March 1973, Who Reported at Least $1000 in the Current Population Survey (N = 24,352).

	Predetermined Variables												Components of Variation[a]		
Equation	EX	EX2	FO[b]	FE	PI	ED	OI[b]	OC[b]	WKS	FT	UN	R^2	Residual σ_u	Explained $\sigma_{\hat{t}}$	Total σ_t
1.	.0161	-.0118	.0277	.0076	.118	---	---	---	---	---	---	.177	.570	.264	.6285
	(.344)	(-.336)	(.092)	(.046)	(.162)										
2.	.0197	-.0109	.0067	-.0078	.050	.0892	---	---	---	---	---	.277	.534	.331	
	(.422)	(-.308)	(.022)	(-.048)	(.068)	(.409)									
3.	.0192	-.0107	-.0100	-.0041	.057	.0494	.0709	---	---	---	---	.305	.524	.347	
	(.411)	(-.304)	(-.033)	(-.025)	(.078)	(.226)	(.255)								
4.	.0164	-.0100	-.0146	-.0030	.052	.0270	.0242	.0880	---	---	---	.348	.507	.371	
	(.351)	(-.284)	(-.049)	(-.018)	(.071)	(.123)	(.087)	(.324)							
5.	.0131	-.0081	-.0073	-.0033	.046	.0279	.0236	.0706	.0274	---	---	.462	.461	.427	
	(.280)	(-.228)	(-.024)	(-.020)	(.063)	(.128)	(.085)	(.260)	(.352)						
6.	.0115	-.0065	-.0025	-.0033	.042	.0312	.0211	.0809	.0248	.713	.218	.530	.431	.456	
	(.246)	(-.184)	(-.008)	(-.020)	(.057)	(.143)	(.076)	(.298)	(.319)	(.212)	(.160)				
7.	.0198	-.0110	---	---	---	.0933	---	---	---	---	---	.273	.536	.328	
	(.422)	(-.310)				(.427)									
8.	---	---	---	---	---	.0666	---	---	---	---	---	.093	.599	.192	
						(.305)									

NOTE: Standardized coefficients appear in parentheses. Estimates of error variances are based on subsamples of 823 (ISR) and 556 (OCGR) observations.

[a]Components are expressed as standard deviations. The additive decomposition was $\sigma_t^2 = \sigma_{\hat{t}}^2 + \sigma_u^2$.

[b]Variables expressed in the metric of Duncan SEI scores have been divided by 10 and corresponding coefficients multiplied by 10.

first job. However, the total effects of each variable are essentially unchanged. While the ambiguous status of the error correlation confounds our interpretation of the way in which schooling confers advantages in the labor market, we are relatively certain that the total true return to an additional year of schooling in our earnings function is on the order of 8.5 to 9 percent, and consequently uncorrected estimates understate the total return to an additional year by 16 to 20 percent.[13]

To this point we have assumed that the response errors in CPS and ISR earnings reports are uncorrelated. As noted above, our earlier research found no evidence of within-variable error correlation for reports of social origins, educational attainment, or occupational statuses, and it seems reasonable to assume that errors in earnings reports are also uncorrelated. Nevertheless, the correlation between the two reports of earnings is surprisingly high, .94, and it is possible that some contamination in response errors of earnings occurs across measurement occasions. Consequently, we briefly assess the effect of a modest error correlation, $\rho_{e_{81},e_{82}} = .25$, and a substantial error correlation, $\rho_{e_{81},e_{82}} = .50$, upon estimation of the earnings function. The latter value was assumed by Bowles (1972).

First, it should be noted that the only effects of a positive value of $\rho_{e_{81},e_{82}}$ on the other measurement model parameters are: (1) the error variations of reports of earnings, $\sigma_{e_{81}}$ and $\sigma_{e_{82}}$, increase; (2) the true score variation in earnings decreases; and (3) the error correlations, $\rho_{e_{41},e_{81}}$ and $\rho_{e_{72},e_{82}}$, decrease (without changing the corresponding error covariances). Consequently, the implications for the corrected estimates of the substantive model are: (1) the unexplained component of variation, σ_u, is decreased as error variation increases; (2) corrected metric coefficients in the earnings function are unchanged, since true score covariances and true score variances of predetermined variables are unchanged; and (3) standardized coefficients in the earnings function increase in proportion to the ratio of the original and modified standard deviations of true earnings.[14]

The relationship between $\rho_{e_{81},e_{82}}$ and estimates of the error variation in CPS and ISR reports of logarithmic earnings is as follows:

	$\rho_{e_{81},e_{82}}$		
	.00	.25	.50
$\sigma_{e_{81}}$	.178	.199	.235
$\sigma_{e_{82}}$	.128	.155	.199

For the CPS report, error variation increases by just 6 percent of the level of true earnings ($e^{.235 - .178} = 1.06$) as the within-variable error correlation goes from 0.0 to 0.5. The effect is more pronounced for error variation in the reinterview report, but this has no effect on our estimates of the earnings function.

Table 12 shows the effects of shifts in $\rho_{e_{81},e_{82}}$ on estimates of components of variation in the earnings function, assuming no correlations among errors across variables (the results are quite similar if we assume $\rho_{e_{41},e_{81}} = -.14$). For each equation (rows) and each measure of variation

Table 12--Proportion of Variance Explained, Unexplained Variation, and Total Variation for Uncorrected Estimates and Corrected Estimates Assuming no Within-Occasion Correlations Among Errors and Varying Within-Variable Correlation Among Errors in Reports of Earnings: Nonblack Males in the Experienced Civilian Labor Force, March 1973, Who Reported at Least $1000 in Earnings in the Current Population Survey (N=24,352).

Equation	R^2				σ_u				σ_t			
	(1) Uncor-rected	(2) $\rho_{e_{81}e_{82}}=.00$	(3) $\rho_{e_{81}e_{82}}=.25$	(4) $\rho_{e_{81}e_{82}}=.50$	(1) Uncor-rected	(2) $\rho_{e_{81}e_{82}}=.00$	(3) $\rho_{e_{81}e_{82}}=.25$	(4) $\rho_{e_{81}e_{82}}=.50$	(1) Uncor-rected	(2) $\rho_{e_{81}e_{82}}=.00$	(3) $\rho_{e_{81}e_{82}}=.25$	(4) $\rho_{e_{81}e_{82}}=.50$
1	.160	.177	.180	.187	.599	.570	.563	.550	.6534	.6285	.6223	.6097
2	.231	.266	.271	.282	.573	.538	.531	.516				
3	.261	.299	.306	.318	.562	.526	.519	.503				
4	.302	.345	.352	.366	.546	.509	.501	.485				
5	.403	.459	.469	.488	.505	.462	.454	.436				
6	.464	.527	.537	.559	.478	.432	.423	.405				

(panels) the effect of varying $\rho_{e_{81},e_{82}}$ (columns) is small. Under the most extreme assumption, $\rho_{e_{81},e_{82}} = .50$, uncorrected estimates would overstate total variation in logarithmic earnings, σ_t, by 7 percent, rather than by 4 percent as stated above (compare columns 1 and 2 and columns 1 and 4 in the third panel of Table 12). In the reduced form equation of line 1, uncorrected estimates would understate R^2 by 14 percent rather than by 10 percent if $\rho_{e_{81},e_{82}}$ were 0.50 rather than 0.0. Similarly, unexplained variation in the reduced form equation would be overestimated by 9 percent instead of 5 percent. In the structural equation (line 6), the bias in R^2 due to ignoring measurement error would be 17 percent instead of 12 percent if $\rho_{e_{81},e_{82}}$ were 0.50 instead of zero. Residual variation in the structural equation would be overestimated by 18 percent instead of by 11 percent. Differences of similar magnitude appear in lines 2 through 5. All standardized coefficients would be increased by a factor of .6285/.6223 = 1.001 if $\rho_{e_{81},e_{82}} = .25$ and by a factor of .6285/.6097 = 1.031 if $\rho_{e_{81},e_{82}} = .50$. In summary, within-variable correlation among errors in reports of earnings can affect only a subset of parameters of our earnings function (components of variation and standardized coefficients), and the effects are quite small. Even a substantial within-variable correlation between errors in reports of earnings does not change our overall interpretation of the effects of response error on the parameters of earnings functions.

CONCLUSION

Several sociologists and economists have noted possible biases in effects of social background and schooling when intergenerational models of occupational and economic success are based on retrospective survey reports of status variables. The prevailing view has been that effects of social background are biased downward by errors in retrospective reports; consequently, effects of schooling are biased upward, at least relative to those of social background. But research on these biases has been inconclusive because appropriate data and statistical models have not been available. Using data from the Income Supplement Reinterview program of the March 1973 Current Population Survey and the remeasurement program of the 1973 Occupational Changes in a Generation survey, we have overcome some of these shortcomings by estimating and testing comprehensive structural models which incorporate both random and nonrandom response errors.

Our earlier research presented evidence that the reports of social origins, education, and occupational attainments by nonblack males are subject to only random response error. The results presented here suggest that Current Population Survey reports of weeks worked and earnings are probably also subject to random errors, although conflicting evidence was presented on whether there is a slight tendency for CPS respondents to understate the consistency between their earnings and education. We also detected a tendency for respondents to overstate the consistency between their 1972 earnings and the number of weeks worked in 1972 in their Income Supplement Reinterview responses.

Ignoring response error, we underestimate the effects of two social origin variables, father's occupational status and parental income, by 14 and 12 percent, respectively. Contrary to some previous expectations, we also underestimate the total effect of schooling by an even greater amount, 16 percent, and we detect even larger understatements, more than 20 percent, in total returns to occupational attainments when measurement errors are ignored. The unmediated effects of parental income and March 1973 occupational status are also moderately understated in uncorrected estimates. We are less certain about the effects of response errors on estimates of the net returns to schooling, unmediated by occupational attainments, labor supply, and/or union membership. If response errors are indeed random, then the unmediated effects of schooling are overestimated by as much as 32 percent when those errors are ignored. However, if CPS respondents do indeed understate the consistency between their schooling and earnings, then the unmediated effects of schooling are understated by as much as 20 percent.

Overall, we were rather surprised at the relative accuracy with which earnings are reported in the Current Population Survey. Indeed, logarithmic earnings appears to be one of the most accurately measured statuses in social surveys of the type examined here. Consequently, response errors bias estimates of coefficients of an earnings function by no more than they bias estimates for equations representing determinants of educational and occupational attainments. Components of variation in earnings are considerably less affected by response error than are components of variation in schooling and occupational attainments.

Our results failed to confirm the hypothesis of Bowles (1972) and others that because of errors of measurement, estimates of earnings functions substantially overstate the influence of schooling vis à vis social origins. The results presented here do <u>not</u> directly address whether response errors bias the extent to which the total and net effects of schooling <u>transmit</u>

inequality in social origins. Our earlier research (Bielby, Hauser, and Featherman, 1977a, 1977b) suggests that the degree to which ignoring measurement error leads to understatement of the transmission of social inequality by schooling is neither trivial nor overwhelming.

In closing, we note that measurement and scaling procedures appear to have substantial effects on the quality of data even where survey methods are highly standardized. Our experiments with scaling earnings show that simple transformations of a measure can greatly affect its quality. The procedures employed to collect data on the four measurement occasions appear systematically related to the reliability of the data. Personal interviews appear to be more accurate than telephone interviews which in turn appear to be more accurate than self-administered questionnaires. Our earlier research (Bielby, 1976; Bielby, Hauser, and Featherman, 1977b) indicated that at least one attribute of the respondent, race, affects considerably the reliability of data obtained from surveys. Finally, our alternate measurement model specification, incorporating the single within-occasion error correlation, demonstrated that if such "response effects" do indeed exist, they can have substantial effects on estimates of some coefficients in structural equation models. Sound research design demands that all of these factors be given careful attention. We hope that our results, based on data collected from one carefully designed social survey, will aid in the design and analysis of other surveys.

APPENDIX

Cross-Tabulation of Reinterview Report of Earnings by Original Report of Earnings

Table 1a--Earned Income--Percentage Distribution of Reinterview Response by CPS Response for Nonblack Males Aged 20-65 in the Experienced Civilian Labor Force, March 1973[a]

CPS Report--Earned Income	Reinterview Report--Earned Income 0	1	2	3	4	5	6	7	8	9	10	11	12	NA	Total	CPS Report Percentage	Number of Cases
0. None	48.5	0.0	29.1	22.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.6	5
1. $1-999	7.3	75.3	0.0	0.0	5.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.5	100.0	1.8	15
2. $1000-1999	0.0	0.0	82.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.3	100.0	2.5	21
3. $2000-3999	0.0	0.0	0.0	81.1	9.4	2.2	0.0	0.0	0.0	0.0	0.0	2.7	0.0	4.6	100.0	5.6	47
4. $4000-5999	0.0	0.0	1.4	3.0	78.2	7.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.9	100.0	8.3	72
5. $6000-9999	0.0	0.0	0.0	1.2	3.9	89.0	3.9	0.0	0.0	0.0	0.0	0.0	0.4	1.6	100.0	29.5	249
6. $10000-14999	0.0	0.0	0.0	0.0	0.0	1.8	92.2	2.0	0.0	0.0	0.0	0.0	0.0	4.0	100.0	27.6	234
7. $15000-19999	0.0	0.0	0.0	0.0	1.1	0.0	4.1	89.1	1.8	0.0	0.0	0.0	0.0	4.0	100.0	11.2	95
8. $20000-24999	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	90.4	4.9	0.0	0.0	0.0	4.6	100.0	2.4	20
9. $25000-29999	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.8	91.2	0.0	0.0	0.0	0.0	100.0	2.5	21
10. $30000-39999	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.6	0.0	0.0	83.1	0.0	0.0	7.3	100.0	1.4	12
11. $40000-49999	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.1	1
12. $50,000+	0.0	0.0	0.0	0.0	0.0	17.7	0.0	0.0	0.0	0.0	0.0	0.0	82.3	0.0	100.0	0.7	6
NA--Allocated	0.0	0.0	0.0	4.3	8.5	5.8	10.3	4.3	0.0	0.0	0.0	2.0	0.0	64.8	100.0	5.6	47
Remeasurement Report Percentages	0.4	1.3	2.4	5.5	9.0	27.9	27.7	10.9	2.6	2.4	1.1	0.4	0.7	7.6	100.0	100.0	---
Number of Cases	3	11	20	47	76	236	234	92	22	20	10	3	6	64	---	---	845

NOTE: Index of dissimilarity between row and column percentages = 3.24. Percentage non-NA cases on diagonal = 91.5.

[a]Percentages and marginal number of cases are based on weighted observations, not on true counts of sample cases. Caution should be used in making comparisons based on small numbers of observations.

Notes

[1]While our findings for black males were less conclusive, they suggested that the pattern of measurement error for blacks is substantially more complex than the pattern for nonblacks. Comparisons of black and nonblack error structures and their implications for racial differences in the occupational achievement process are presented in Bielby, Hauser, and Featherman (1977b).

[2]The OCG parental income item was: "When you were about 16 years old, what was your family's annual income?" The possible responses were: No income (or loss); \$1-\$499; \$500-\$999; \$1,000-\$1,999; \$2,000-\$2,999; \$3,000-\$3,999; \$4,000-\$4,999; \$5,000-\$5,999; \$6,000-\$6,999, \$7,000-\$7,999; \$8,000-\$8,999; \$9,000-\$9,999; \$10,000-\$14,999; \$15,000 or more. After examining plots of occupational status of first and current job and educational attainment by parental income category we determined that a logarithmic function of parental income was the appropriate functional form relating it to the achievement variables. The first two categories were collapsed and midpoints of intervals were used. A value of \$19,750 was assigned to the open-ended category on the basis of a canonical analysis with ED, O1 and OC as criterion variables. Responses to pretest probes and plots of achievement variables by parental income categories by ten-year age cohorts clearly indicated that respondents tended not to adjust their responses to current dollars. Therefore, the dollar midpoint responses were adjusted by a four-year moving average of the Consumer Price Index, with the four years weighted to reflect the uncertainty in determining exact year of birth from age in March 1973. The final scale was computed

as the natural logarithm of the price adjusted dollar category midpoints. Our scaling procedure explicitly attempted to maximize correlations between parental income and statuses of the respondent. As a consequence, intergenerational (father-son) correlations between PI and ED are larger than intragenerational (father's generation) correlations between PI and both FO and FE (see Tables 2, 3, 4, and 8).

[3]The sample person was instructed to report the year in which he actually began the job, even if he started it before completing schooling. Consequently, for some persons our definition includes labor market experience obtained before completion of schooling, and it includes military experience if a person held a full-time civilian job after completing schooling but before entering the military.

Year of first job was not reported for about 11 percent of the cases. For these cases, the year schooling was completed, if reported, was used to compute labor market experience. For the remaining 5 percent of the cases, experience was computed as age-educational attainment - 5. Labor force experience obtained before ten years of age was not counted. For less than 1 percent of the cases, no experience measure was computed, because of inconsistencies between age, educational attainment, and year schooling was completed.

[4]Experience and age are correlated .914 in the full CPS-OCGQ sample. Consequently, our results would be quite similar had we controlled year of birth (age) instead of year of entry to first job (experience). The role of age and experience in earnings functions and the appropriate

definition of experience have been the subject of considerable debate. For example, see Griffin (1977), Blinder (1976), Rosenzwerg (1976), and Rosenzwerg and Morgan (1976).

[5]The term, true score, should be interpreted cautiously. As in the classic psychometric model the true scores are defined as latent or unobservable variates which underly repeated observable measurements. For this reason our findings about the quality of specific measures do not establish their validity in global terms. For example, the true score underlying repeated self-reports of a man's occupation might differ from that underlying repeated reports by his employer. Thus "latent or unobserved variable" is synonymous with "true score" as we use the term.

[6]The two reports (OCGQ and OCGR) of union membership are identical for almost all cases, as are the two reports (CPS and ISR) of full-time employment status. The assumption that work experience is measured perfectly is less tenable. However, the quality of that measure will not be assessed here (but see Bielby, 1976:112-114). Preliminary models showed slightly larger effects for linear and quadratic age terms than for linear and quadratic experience terms. We surmise that the difference would disappear if we could adjust for the lower reliability of the experience measure, despite the different conceptual bases for the age and experience terms.

[7]The Bureau of the Census uses the "hot deck" technique to allocate nonresponses in CPS reports of education, occupation, weeks worked, and earnings, and we treat these allocations as responses. Allocated nonresponses

are assigned the observed value of the last case processed with the same attributes on several characteristics. For allocation of education and occupation, the characteristics are age, sex, and race. In addition to those three characteristics, occupation, class of worker, family relationship and weeks worked are used in allocating earnings. Age, sex, race, family relationship, and earnings are used in allocating nonresponses for weeks worked (Ono and Miller, 1969; Spiers, Coder and Ono, 1972; Spiers and Knott, 1969). Thus, allocated responses have both systematic and random components.

For non-allocated measures, we assume that the pairwise correlations accurately represent the correlations that would have been obtained were complete data available. While this is an untestable assumption, the alternatives are more problematic. Replacement with means restricts variances and would result in underestimates of error variances. Random allocation would reduce the ability to detect nonrandom response error structure, while systematic allocation would have the opposite effect. Omitting all cases that have missing data would reduce the sample size by about 40 percent and probably eliminate many of the cases with less accurate responses. Models of the achievement process are almost always estimated from pairwise-present correlations, and it is the response error structure in these analyses that we are attempting to assess.

Nevertheless, to assess the impact of both allocation and nonresponse, we performed parallel analyses to those reported here, excluding all cases with any missing data or allocated responses. In only a few instances did any of these analyses suggest patterns of response error that differ substantially from the results of the pairwise analysis. When "listwise" analysis does yield different results, it will be noted below.

[8]We noted above that there may have been a tendency for those with more accurate reports of earnings to be selected for the OCGR subsample. If so, the error variance for CPS reports of earnings borrowed from the ISR subsample estimates may be too large. For a given covariance among reporting errors, a larger error variance will reduce the correlation among the errors. If this is indeed the case, the fourth stage models may understate error correlations involving the CPS earnings report. On the other hand, when the λ_{ij} are 1.0, as they are for the reports in question, the covariance among observable reports of two variables is equal to the covariance among respective true scores plus the covariance among respective error terms, under the measurement model. That is, the covariance among observables generated by the model is not affected by the error variances. Model 3B suggests the covariance among e_{41} and e_{81} is modestly large and negative, while according to Model 4B the covariance among error terms is positive but negligible, regardless of the values of the borrowed error variances. Thus the ISR subsample and OCGR subsample estimates do seem to be at odds in their implications about the degree to which respondents understate (or overstate) the consistency between their educational attainments and earnings.

[9]The positive correlation between errors in ISR reports of weeks worked and earnings did not disappear in the listwise estimates ($\hat{\rho}_{e_{72},e_{82}}$ = .14). Since ISR reports are not used in estimating the earnings function in the full CPS-OCG sample, this error correlation does not affect our

assessment of biases due to response errors.

[10]Since the CPS report of education is more likely than the OCG report to be a proxy report (provided by the spouse of the designated male), there could be a tendency for 12 years completed to be reported in the CPS when the designated male actually completed slightly more or less than twelve years. Since the mean report is about 12 years, such a tendency would result in a negative correlation between errors and true scores. The ISR subsample shows CPS and OCGQ reports to be about equally reliable, while the OCGR subsample shows the latter to be substantially less reliable. For reasons noted above, the error variation in the OCGQ report is probably overestimated from the OCGQ subsample and underestimated from the ISR subsample.

[11]Alternatively, we could have estimated the earnings function directly from the subsamples, or pooled moments from the full sample and both subsamples in order to simultaneously estimate the measurement model and earnings function. The former procedure is certainly less efficient than the one employed here, and we do not know the relative merits and pitfalls of the latter procedure.

[12]Under our specification, if b_1 is the coefficient of EX and b_2 is the coefficient of EX2, the partial derivative of earnings with respect to experience is $b_1 + (b_2/5)(EX - 20)$.

[13]While the percentage biases in unmediated semi-reduced form and structural coefficients are often quite large and are quite sensitive to

whether or not the error correlation exists, the effects themselves are often quite small. Consequently, one should not overemphasize the substantive importance of these differences. For example, the uncorrected structural coefficient for a year of schooling is about 2.5 percent, our initial corrected estimate is about 2 percent, and our alternative corrected estimate is about 3 percent. The variation in assessments of relative biases is quite large, but only one-half of one percent return per additional year of schooling is at issue.

[14] Designating true LNEARN = t_8, then from the measurement model $x_{81} = t_8 + e_{81}$, and from the earnings function $t_8 = \hat{t}_8 + u$. So $\sigma^2_{t_8} = \sigma^2_{x_{81}} - \sigma^2_{e_{81}} = \sigma^2_{\hat{t}_8} + \sigma^2_u$, and $\sigma^2_u = \sigma^2_{x_{81}} - \sigma^2_{\hat{t}_8} - \sigma^2_{e_{81}}$. Variances $\sigma_{x_{81}}$ and $\sigma^2_{\hat{t}_8}$ are unaffected by the value of $\rho_{e_{81},e_{82}}$, while $\sigma^2_{e_{81}}$ increases with $\rho_{e_{81},e_{82}}$. Therefore σ^2_u gets smaller as $\rho_{e_{81},e_{82}}$ increases.

REFERENCES

Alexander, K.L., B.K. Eckland, and L.J. Griffin

1975 "The Wisconsin model of socioeconomic achievement: a replication." American Journ l of Sociology 81:324-342.

Becker, Gary

1964 Human Capital. New York: National Bureau of Economic Research.

Bielby, William T.

1976 Response Errors in Models of the Intergenerational Transmission of Socioeconomic Status. Unpublished Ph.D. dissertation, University of Wisconsin-Madison.

Bielby, William T., Robert M. Hauser, and David L. Featherman

1977a "Response errors of nonblack males in models of the stratification process." Journal of the American Statistical Association 72: forthcoming.

______________.

1977b "Response errors of black and nonblack males in models of status inheritance and mobility." American Journal of Sociology 83:1242-1288.

Blau, Peter M., and Duncan, Otis Dudley

1967 The American Occupational Structure. New York: Wiley.

Blinder, Alan S.

1976 "Dogmatism in human capital theory." Journal of Human Resources 11:8-22.

Bohrnstedt, George W.

1970 "Reliability and validity assessment in attitude measurement." Pages 80-99 in Gene F. Summers (ed.), Attitude Measurement. Chicago: Rand McNally.

Bowles, Samuel

1972 "Schooling and inequality from generation to generation." Journal of Political Economy 80:S219-S251.

Bowles, Samuel, and Herbert Gintis

1976 Schooling in Capitalist America. New York: Basic Books.

Bowles, Samuel, and Valerie I. Nelson

1974 "The 'Inheritance of IQ' and the intergenerational reproduction of economic inequality." Review of Economics and Statistics 56:39-51.

Duncan, Otis Dudley

1961 "A socioeconomic index for all occupations." Pp. 109-138 in Albert J. Reiss, Jr. (ed.), Occupations and Social Status. New York: Free Press.

Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan

1972 Socioeconomic Background and Achievement. New York: Seminar Press.

Featherman, David L.

1972 "Achievement orientations and socioeconomic career attainments." American Sociological Review 37:131-143.

Featherman, David L., and Robert M. Hauser

1973 "On the measurement of occupation in social surveys." Sociological Methods and Research 2:239-251.

Griffin, Larry J.

1977 "On estimating the economic value of schooling and experience: some issues in conceptualization and measurement." Sociology of Education 50: forthcoming.

Hauser, Robert M., and Thomas N. Daymont

1977 "Schooling, ability and earnings: cross-sectional findings 8 to 14 years after high school graduation." Sociology of Education 50: forthcoming.

Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson

1972 Inequality. New York: Basic Books.

Jöreskog, Karl G.

1970 "A general method for analysis of covariance structures." Biometrika 57:239-251.

Jöreskog, Karl G.

1973 "A general method for estimating a linear structural equation system." Pp. 85-112 in Arthur S. Goldberger and Otis Dudley Duncan (eds.), Structural Equation Models in the Social Sciences. New York: Seminar Press.

Joreskog, Karl G., Gunnar T. Gruvaeus, and Marielle van Thillo

1970 "ACOVS: A General Computer Program for Analysis of Covariance Structures." Research Bulletin 70-15, Educational Testing Service, Princeton, New Jersey.

Mincer, Jacob

1974 Schooling, Experience, and Earnings. New York: National Bureau of Economic Research.

Ono, Mitsuo and Herman P. Miller

1969 "Income nonresponses in the Current Population Survey." Proceedings of the Social Statistics Section, 1969, American Statistical Association, 12th Annual Edition. Pp. 277-288.

Rosenzweig, Mark R.

1976 "Nonlinear earnings functions, age, and experience: a nondogmatic reply and some additional evidence." Journal of Human Resources 11:23-27.

Rosenzweig, Mark R., and Jack Morgan

1976 "Wage discrimination: a comment." Journal of Human Resources 11:3-7.

Sewell, William H., and Robert M. Hauser

1975 Education, Occupation and Earnings: Achievement in the Early Career. New York: Academic Press.

Siegel, Paul M., and Robert W. Hodge

1968 "A causal approach to the study of measurement error." Pp. 28-59 in H.M. Blalock and A.B. Blalock (eds.), Methodology in Social Research. New York: McGraw-Hill.

Spiers, Emmett, John Coder, and Mitsuo Ono

1972 "Characteristics of income nonrespondents in the Current Population Survey." Proceedings of the Social Statistics Section, 1971, American Statistical Association, 14th Annual Edition. Pp. 369-374.

Spiers, Emmett, and Joseph J. Knott

1969 "Computer method to process missing income and work experience information in the Current Population Survey." Proceedings of the Social Statistics Section, 1969, American Statistical Association, 12th Annual Edition. Pp. 289-297.

Treiman, Donald J., and Robert M. Hauser

1977 "The intergenerational transmission of income." Pp. 271-302 in Robert M. Hauser and David L. Featherman, The Process of Stratification: Trends and Analyses. New York: Academic Press.

Wright, Erik Olin, and Luca Perrone

1977 "Marxist class categories and income inequality." *American Sociological Review* 42:32-55.