CODEBOOK SUPPLEMENT

APPENDIX #10

The material which follows originated in <u>Determinants of Early Labor Market Success Among Young Men: Race, Ability, Quantity and Quality of Schooling</u>, Andrew I. Kohen, Ph. D. dissertation, January 1973. The discussion and results reported here, though based on the National Longitudinal Surveys of Young Men 14-24, apply equally well to the Surveys of Young Women 14-24.

APPENDIX B

CONSTRUCTION OF AN INDEX OF SOCIOECONOMIC LEVEL OF PARENTAL FAMILY

There are two fundamental reasons for construction of the index of socioeconomic level of parental family which is described below. First, the concept of "family background" employed in this study is multi-dimensional in the sense that it is theorized to represent biological, social, economic and psychological influences on a young man's ability and behavior. Second, there is no unique operational measure of this concept which is universally accepted by social scientists. Hence, an index which incorporates several commonly-used, intercorrelated measures was constructed. This approach to operationalizing "family background" is also employed in two recent major pieces of social research.[1]

Preliminary examination of the data available in the rich bank of the National Longitudinal Surveys Project revealed eight variables which were potential candidates for inclusion in an index of family back-ground.[2] Table B-1 contains descriptions of these variables along with means and standard deviations based on all data available for males who

[1]See Jerald G. Bachman, Youth in Transition, Vol. II (Ann Arbor: Institute for Social Research, University of Michigan, 1970), pp. 219-30, and John C. Flanagan, et al., Project TALENT--The identification, development and utilization of human talents (Final report to the U.S. Office of Education, Project No. 635) (Pittsburgh: University of Pittsburgh, 1964).

[2]Actually, a ninth available variable (i.e., type of area residence when the respondent was 14) was rejected on the a priori grounds that its relevance was already represented by other variables, namely, father's occupation, availability of reading material and family size.

TABLE B-1

MEANS, STANDARD DEVIATIONS AND MISSING-DATA RATES[a] FOR
EIGHT MEASURES OF FAMILY BACKGROUND: MALES
14 TO 24 YEARS OF AGE IN 1966

Variable No. Name	Unit of measurement	Mean	S.D.	Missing-data. rate[a]
1 Father's education[b]	Years	9.9	3.8	27%
2 Mother's education[b]	Years	10.3	3.2	15
3 Father's occupation[c]	Duncan index	31.4	27.8	9
4 Sibling's education[d]	Years	11.6	2.6	38
5 Availability of reading material[e]	Ordinal scale	2.9	1.4	1
6 Number of siblings[f]	Persons	3.4	2.6	1
7 Family structure[g]	Ordinal scale	3.98	0.20	0
8 Immigrant status[h]	1=immigrant 0=nonimmigrant	.03	.18	1

[a]The total number of data cases is 5,225. None of the variables was available for every case, though the seventh item was codable for 5,213 respondents.

[b]There are two possible reasons for nonavilability of these variables. The first is a noncodable response or a response of "don't know" to either the question of the highest grade attended by the parent or the question of whether that grade was completed. The second is that the interviewee may not have been asked the question because (1) the parent was deceased at the time of the interview or (2) the respondent did not live with the parent either at age 14 or at the time of the interview.

[c]The variable actually refers to the occupation of the head of the household when the respondent was 14. Thus, in a few instances it refers to the occupation of the respondent's mother, grandfather, etc. Duncan's index assigns a two-digit (2-98) status score to each three-digit occupation code in the classification scheme of the Bureau of the Census.

[d]This refers to the education of the respondent's oldest older sibling. Thus, it is automatically not available for those with no siblings and for those with only younger siblings.

[e]This is a five-valued (0-4) scale based on the respondent's answers to questions about the availability of magazines, newspapers, and a library card in the home when he was 14. The first two items were each coded 0 if unavailable and 1 if available. The third item was coded 0 if unavailable and 2 if available. The value of the scale is the sum of the three scores. The higher weight is assigned to the existence of a library card in the belief that it is indicative of a stronger family commitment to educational pursuits than is the existence of either magazines or newspapers.

[f]The sum of all brothers and sisters living (in the same or different households) at the time of the survey.

[g]This is a five-valued scale (0-4) describing the respondent's living arrangements when he was 14. The coding was as follows: 4 = with both natural parents, 3 = with one natural parent and one step-parent, 2 = with one natural parent only, 1 = other family arrangement, 0 = on his own.

[h]A respondent is defined as an immigrant if he was born outside the U.S. or Canada or if both of his parents were born outside the U.S. or Canada.

were 14 to 24 years of age in 1966. On the basis of an examination of the intercorrelations among the variables and a rough assessment of a few techniques of index construction, it was decided to employ six of the eight variables in the study. However, only five are components of the index; the other enters the analysis separately.

The five elements which comprise the index are as follows: (1) father's educational attainment, (2) mother's educational attainment, (3) occupational status of the father (or head of household) when the respondent was 14 years of age, (4) educational attainment of the respondent's oldest older sibling[3] and (5) availability of reading material in the home when the respondent was 14. The first three items are the variables most commonly used, singly or jointly, by social scientists to represent socioeconomic background. The fourth and fifth items were chosen to represent additional dimensions of the quality of the home environment available to a young man--i.e., the reinforcement of achievement attitudes and behavior.

The sixth measure of family background which is retained is total number of siblings. It was excluded from the index purely for computational convenience. That is, while the five included elements are related positively to one another, this construct is negatively correlated with each of them (Table B-2). Nonetheless, the measure is used in the study because of its documented theoretical and empirical relevance. The two remaining variables were dropped from further consideration principally because of their low and irregular correlations with the other measures (Table B-2). In addition, the family structure measure exhibits virtually no variation (Table B-1).

---

[3]See note d, Table B-1.

TABLE B-2

ZERO-ORDER CORRELATIONS AMONG EIGHT MEASURES OF FAMILY BACKGROUND[a]: MALES 14 TO 24 YEARS OF AGE IN 1966[b]

		2	3	4	5	6	7	8
Father's education	1	.648 (3563)	.574 (3647)	.441 (2353)	.483 (3772)	-.291 (3770)	.003 (3410)	-.027 (3410)
Mother's education	2		.457 (4130)	.463 (2775)	.468 (4416)	-.329 (4414)	-.007 (3410)	-.088 (2531)
Father's occupation	3			.350 (2946)	.411 (4734)	-.276 (4733)	.032 (3410)	.015 (3410)
Sibling's education	4				.435 (3208)	-.301 (3220)	.090 (2225)	-.032 (2531)
Reading material	5					-.317 (5107)	.017 (3410)	-.049 (3410)
Number of siblings	6						.042 (3410)	-.022 (3410)
Family structure	7							-.014 (3410)
Immigrant status	8							1.000

[a]See Table B-1 for description of the variables and units of measurement.

[b]Numbers in parentheses are numbers of cases upon which the coefficient is based. For correlations among items 1-6 it is the maximum number of cases for which both variables were available. Correlations involving items 7 or 8 are drawn from preliminary examinations of the data and are based on several sub-samples which were restricted by availability of information on different combinations of variables.

Two basic criteria were used in selecting the method of combining the items into a summary index. These criteria were computational ease and preservation of data cases. As noted above, the former criterion strongly suggested that the number-of-siblings variable be omitted from the index. The second criterion was applied because of the high missing-data rates for two of the variables--i.e., father's and sibling's education. Finally, in the absence of any evidence for a set of "optimal" weights, it seemed reasonable to compute the summary index by weighting each of the five items equally.[4]

Hence, the value of the index of socioeconomic level of parental family is the average value of the components, where the components are linearly transformed so that each has a mean of 10.0 and a standard deviation of approximately 3.0.[5] In the interest of preserving data cases, the index was calculated for each respondent for whom information was available on at least three of the components. The linear transformation procedure is illustrated in Table B-3 and consisted of two steps. First, each component was multiplied by a scalar weight (Column 3). Then a constant was added to the weighted value (Column 6).[6]

---

[4]A principal components analysis of the five items selected for inclusion in the index yielded loadings for the first component which were substantially equal. Using the variable numbers shown in Table B-1, these loadings were as follows: 1 = .82, 2 = .80, 3 = .72, 4 = .67, and 5 = .71. In addition to the fact that factor analytic methods provide no theoretical justification for using these weights, use of them would have increased considerably the computational difficulty of constructing the index.

[5]This technique was adapted from the one used in Bachman, Youth in Transition, II, Appendix B.

[6]Because of the specialized features of the subsample used in this study, the linear transformation weights and constants were based on initial values of the components for 1,727 males who were 18 to 24 years of age and out of school in 1966. Thus, the values in columns 1, 2, and 9 in Table B-3 do not correspond to those in Table B-1.

TABLE B-3

PROCEDURE FOR COMPUTING THE SUMMARY INDEX OF SOCIOECONOMIC LEVEL OF PARENTAL FAMILY

	Initial values[a]		Scalar weight	Values after weighting		Constant added	Transformed values		Missing-data rate (Percent)[b]
	Mean	S.D.		Mean	S.D.		Mean	S.D.	
Father's education	8.7	3.5	.9	7.8	3.1	2.2	10.0	3.1	35
Mother's education	9.4	3.3	.9	8.5	3.0	1.5	10.0	3.0	19
Father's occupation	26.7	20.0	.15	4.0	3.0	6.0	10.0	3.0	8
Sibling's education	11.2	2.8	1.1	12.3	3.1	-2.3	10.0	3.1	27
Reading material	3.1	1.2	2.4	7.4	3.0	2.6	10.0	3.0	2

[a]See Table B-1 for description of variables and units of measurement.

[b]Based on 1,727 males 18 to 24 years of age who were not enrolled in school in 1966.