Analysis Plan for the 2007 Pilot Studies of Socioeconomic Measurement in NAEP

DRAFT FOR COMMENT ONLY

Robert M. Hauser
Rev. November 11, 2006

Table 1 describes the data that should become available from the 2007 pilot studies at each grade level.[1] The objective of the analysis will be to use these data – along with comparisons to analyses based on earlier administrations of NAEP – to construct and validate an improved measure of socioeconomic status (hereafter, SES) that may replace eligibility for the federal free and reduced lunch program as a measure of SES for NAEP reporting purposes.

Table 1. Sources of data in 2007 NAEP Pilot Study

Grade level	Sample	Micro-data	Census small area data[2]
4	ACS counties	National School Lunch Program (NSLP), Background Questionnaire (BQ), Enhanced Background Questionnaire (EBQ),[3] and test scores	Aggregate and/or imputed socioeconomic data from the 2000 Census and American Community Survey
8	ACS counties	NSLP, BQ, EBQ and test scores	Aggregate and/or imputed socioeconomic data from the 2000 Census and ACS
8	ECLS-K Student-Parent Sample	NSLP, BQ, parent survey and test scores	Aggregate and/or imputed socioeconomic data from the 2000 Census

Note that these data will be incomplete in several ways:

- In the fourth grade, parental education is not included in the BQ, and home tenure (own vs. rent) is not included in the EBQ.
- In the fourth and eighth grade ACS counties, there is no parent survey.

[1] See "NAEP SES 2007 Pilot Study" (10/06/06).

[2] This draft does not address the choice among levels of aggregation of Census data (block group, tract, or ZCTA) or the specific methods use to impute vectors of socioeconomic characteristics from such data. These may be analyzed using existing data, e.g., from ECLS-K, as well as data obtained in the 2007 pilot study. For insightful observations about some of the issues, see WESTAT memo, "SES Pilot Study Analysis Plan Overview," (11/03/06).

[3] Parent reports will be used to validate student reports of SES in the ECLS-K sample, but EBQs will not be available in that sample. Parent reports are not available in the main NAEP data collections.

- In the eighth grade ECLS-K sample, there is no EBQ.
- In the eighth grade ECLS-K sample, there is no match to ACS data.
- There will be no pilot data for 12th graders in 2007.[4]

In all, the design includes six potential sources of socioeconomic reports: NSLP (from schools), BQ and EBQ (from students), parent survey, and small area data for 2000 (Census) and 2001-2006 (ACS). Five of these sources occur jointly at grade 4 and in the 8th grade ACS county samples—all except the parent survey. However, only four sources occur jointly in the ECLS-K sample, including the parent survey data, which are presumably the most valid micro-data available. The analysis plan must work around the absence of important joint measurements.

There are two reasons to include data from Census 2000 as well as from the ACS in the design. The more important is that, because the ACS data will not be available for cases in the ECLS-K sample, we may be able to use Census 2000 data (imputed or aggregate) as a proxy for the ACS data. The second reason is that, by obtaining socioeconomic data from the ACS samples as well as the ECLS-K sample, we may be able to assess the validity of Census 2000 data as a proxy for more recent ACS data.

Table 2 displays the availability of data at the 8th grade level in another way, by schematic representation of correlations that may be observed in the pilot study data. That is, rows and column headings list the sources of available data, but one should think of them as representing the variables measured in each source. NAEP, NSLP, and BQ are

Table 2. Schematic diagram of the availability of data in 8th grade NAEP pilot study

DATA SOURCE	NAEP/NSLP/BQ	EBQ	ACS	Census 2000	PARENT
NAEP/NSLP/BQ	ACS,ECLS-K	--	--	--	--
EBQ	ACS	ACS	--	--	--
ACS	ACS	ACS	ACS	--	--
Census 2000	ACS,ECLS-K	ACS	ACS	ACS,ECLS-K	--
PARENT	ECLS-K	--	--	ECLS-K	ECLS-K

combined because those data always occur jointly. Each cell entry pertains to the availability of data in one or both of the 8th grade samples. For convenience, only the diagonal and sub-diagonal cells have been filled. Thus, "ACS,ECLS-K" in the upper left hand cell tells us that data are available from NAEP, NSLP, and BQ in both the ACS and ECLS-K samples at the eighth grade level.

Suppose that the ACS and ECLS-K samples are equivalent in representing 8th graders. Then, it would not matter if only one of the samples were available in each cell of the table. However, two features of the display stand out:

[4] The original decision to exclude the 12th grade was made in order to avoid a loss in response rates, but few ACS cases would occur in the 12th grade in any event because there are no state representative samples at that grade level.

- In only three of the 15 distinct cells do data appear for both samples: the intersections of NAEP/NSLP/BQ with Census 2000 and each other. That is, while we might like to have data from both samples in each cell, many correlations can be estimated in only one sample.
- Two cells are empty: the intersections of PARENT with EBQ and ACS. The latter may not be a problem if the Census 2000 data are good proxies for ACS data, but there is no way to substitute for the absence of joint observations of parent data with data from the EBQ.

A very serious implication of this design is that there is no *socioeconomic* criterion for the validity of EBQ measures (at the 4th grade level as well as in the 8th grade). That is, there is no direct way to tell whether use of the EBQ measures will lead to greater correspondence between a socioeconomic index based on student responses and a socioeconomic index based on parent reports (with or without supplementation by Census data). Academic achievement is the only available criterion for a direct comparison of validity between indexes with and without EBQ measures.

Goals of the Analysis

The overall goal of the analysis is to develop an improved measure of SES for NAEP reports, but there are multiple criteria for "improvement," and they are not entirely compatible. Tradeoffs will be necessary.

1. The SES measure should have construct validity. It should reflect the full range of variation in the actual social and economic circumstances of students.

2. The SES measure should be comparable across time, across major population groups, such as those defined by grade level and by race ethnicity, and across various measures of academic achievement (see "What about comparability?" below). This may be an especially problematic criterion because some items in the BQ and EBQ are not asked at the 4th grade level. Morever, the absence of data at the 12th grade level means that there is no way to assess comparability across all grade levels, except by reference to data from NAEP, NSLP, and BQ items that are available from other NAEP administrations.[5]

   (As discussed in the NAGB presentation package of March 2006 (Rev. June 15, 2006), the current measure of eligibility for the free or reduced-price school lunch program (NSLP) does not meet either of these first two criteria.)

3. The SES measure should have criterion validity. It should pertain to social and economic circumstances of students that are related to the full range of academic achievements that are assessed in NAEP. (This has been a point of controversy, both within the SES Working Group and the Standing Committee. Do we want a measure of SES that is most informative about academic achievement levels? Or

[5] This is one of the reasons for obtaining NAEP micro-data from earlier years.

one that is most plausible theoretically? Or one that is most like what parents of students would report?)

4. The SES measure should pertain directly to the circumstances of individual students, and not only to groups of students, e.g., those who occupy the same school building. (Most variation in the socioeconomic status of students, like most variation in their academic performance, occurs among students within schools, rather than between schools.)

5. The SES measure should be statistically reliable. Repeated measurements of the same variables for the same students should yield very similar values. (The pilot study will provide no information about statistical reliability of student reports.)

6. The SES measure should be valid across time, for different population groups, and for different types of academic achievement. For example, it should pertain to resources of the family of origin that are valid regardless of race-ethnicity, gender, region, or urban-rural location, and it should not depend on characteristics that are likely to become socially and economically obsolete.

7. Socioeconomic measurement should not create an undue burden on the student. That is, the SES measure should be based on facts that students know, on no more facts than it is necessary to ascertain, without intrusive questions, and with minimal student time and effort.

8. Regardless of the kinds of statistical analysis and validation studies that affect the creation of a new socioeconomic measure, the SES measure should be straightforward in construction and interpretation. It should have face validity in the eyes of lay-people.

The several data sources outlined above should provide ample candidate measures that might be used individually or in combination in a new SES measure. In various (but not all) sets of data, these include eligibility for the free and reduced-price lunch program (NSLP), parents' educational attainments (current and, for 8th graders, revised pilot versions), possessions in the home (current BQ), family composition (pilot BQ), parents' labor force status (pilot BQ), home tenure (pilot BQ for 8th graders), and home amenities (pilot BQ). The ECLS-K parent survey and Census data (Census 2000 and ACS) contain items that parallel the demographic characteristics on this list, plus some other socioeconomic characteristics, e.g., family income and occupation.[6]

By way of example, one obvious problem in establishing comparability of measures across grade levels is that there will be no student reports of parents' educational attainments or home tenure at the 4th grade level (nor have such measures been collected

[6] See "Proposed Items to Enhance the Measurement of Socioeconomic Status for the 2009 NAEP Student Background Questionnaires" (NCES, Rev. June 15, 2006) for new and revised NAEP background questionnaire items and discussions of them. See "Proposed Census Items" (June 2006) for a listing of potential items from geographically matched data from the American Community Survey.

in NAEP since 2002). Thus, it would be desirable to learn whether an improved measure can be constructed without using those variables at the 8th (or 12th) grade levels. This may be difficult, given that we know from previous research that parental education is a critical resource for and major correlate of student achievement.

Comparability problems are also raised by the use of enhanced BQ items in a new measure. To the extent that these items have not been used in NAEP previously, use of these items will make it more difficult to measure trends in SES differentials in achievement prior to the full introduction of the new measures.

Another possibility is to establish comparability across time and among population groups, e.g., grade levels, using different, but overlapping sets of items for different populations.[7] Here, one idea is to use common socioeconomic indicators as references or anchors for larger sets of items that may or may not appear (or be strictly comparable) across periods or population groups. While May (2006) develops this idea using an IRT model, the principle appears to be the same as that of using reference indicators to establish comparable metrics of latent constructs across populations in structural equation models. One possible drawback to this approach is that creation of a common metric does not imply comparable reliability across populations when the number and reliability of indicators varies across populations. Another problem is that the May's model assumes a one-dimensional factor underlies socioeconomic status, while some studies of the consequences of SES reject the concept of a latent SES factor (Hauser 1972; Hauser, Tsai, and Sewell 1983; Warren and Hauser 1997; Warren, Sheridan, and Hauser 2002).

A simpler way to establish comparability of SES across populations and periods would be to introduce multiple summary measures while overlapping the periods or populations in which pairs of measures apply. For example, one SES measure, based on continuously measured BQ items, might be used for cross-temporal comparisons from 1970 to 2002, while another SES measure, comprising some of those BQ items, but also Census items, might be used for cross-temporal comparisons from 1990 to 2010. Differentials observed in the overlapping period of coverage would provide a way of assessing the extent to which trends observed in one series can be compared with those in the other series.

Regardless of these possibilities, the following discussion focuses on the selection of variables to be used in a single, simple, composite measure of SES.

Proposed Analyses

1. **Existing NAEP micro-data files.** What is an efficient and valid composite measure of socioeconomic status based on existing BQ measures? How does the criterion validity of this measure (in relation to academic achievements) compare with NSLP? (Regressions of achievement on selected students' reports. Such analyses do not depend on the availability of pilot study data, but may be carried

[7] See May, Henry. 2006. "A Multilevel Bayesian Item Response Theory Method for Scaling Socioeconomic Status in International Studies of Education." *Journal of Educational and Behavioral Statistics* 31(1):63-79.

out with existing NAEP microdata, including data at the 12th grade level, whose acquisition has been in progress for months. Also, when these data become available, it should be possible to offer more relevant and timely examples of analyses like those reported below in respect to criterion validity.)

a. How does the validity of such a composite depend on the inclusion of items that are not available at the 4th grade level? (regression analysis)

b. How does the validity of such a composite depend on the use of simple unit weights (of standardized variables) vs. regression-based weights?

c. How do the weights of a regression-based composite vary across grade levels?

d. How do the weights of a regression-based composite vary among other population subgroups, e.g., by race-ethnicity, region, and gender?

e. How does the validity of simple weighting schemes vary across grade levels?

f. How does the validity of simple weighting schemes vary by race-ethnicity, region, and gender?

2. **ECLS-K sample at grade 8**. This sample will yield individual data from students and their parents, but no EBQ data and areal data only from Census 2000. Table 2 provides a schematic picture of the kind of data available for this sample.

a. What are the relationships among students' reports of each socioeconomic characteristic, corresponding reports by parents, and reports imputed from Census 2000 data? Are students' or Census reports biased, and in what direction(s)? How strongly are student and Census reports correlated with the (presumably more accurate) reports by parents? (Simple correlations and cross-tabulations)

b. How accurate are students', parents', and Census-based reports of each socioeconomic characteristic ascertained from at least two sources? For example, we should not assume, when parents' and students' reports differ, that only the students' reports are in error. Note that parental reports of education, occupation, and income have been obtained on several occasions in ECLS-K. (Confirmatory factor models of multiple constructs will identify the precision of each report. Such models have been widely used for the past 30 years (Mason, Hauser, Kerckhoff, Poss, and Manton 1976; Bielby and Hauser 1977; Bielby, Hauser, and Featherman 1977a; Hauser and Massagli 1982; Hauser et al. 1983; Massagli and Hauser 1983b; Hauser and Mossel 1985; Hauser and Sewell 1986; Hauser and Wong 1989; Warren and Hauser 1997; Warren et al. 2002).)

c. What is the relationship between a composite socioeconomic index created from reports by parents and a composite socioeconomic index created from reports by students? Between a composite socioeconomic index created from reports by parents and a composite socioeconomic index based on Census small area data? Between a composite socioeconomic index created from reports by students and one based on Census small area data? (Canonical correlation analyses. The canonical correlation analysis will yield weights for parent-reported and student-reported items that will maximize the correlation between the two composites. The variables included in the index based on parents' reports and the index based on students' reports need not be the same. A major problem here is that, in this sample, the student-based composite will not include EBQ items. However, in principle we can look at the relationship between Census-based and parent based indexes in the full ECLS-K sample, not just the subsample of ECLS-K that will participate in NAEP in 2007.)

d. How comparable are relationships with academic achievement between the best student-based composite and the best parent-based composite? (Regressions of academic achievement on student-based composite and parent based composite.)

e. Repeat parts *c* and *d*, but leave out student reported items that are not available from 4th graders. That is, how much validity (relative both to parent reports and to academic achievement) would be lost by establishing strict comparability between socioeconomic measures used at the 4th and 8th grades (and by implication, the 12th grade).

f. Repeat parts *b* through *e*, but add imputed data from small areas in the 2000 Census to the *student* reports. That is, ask how much could a student-based socioeconomic index be improved using imputed Census data, assessed both in relation to a parent-based index and to academic achievement.

g. Repeat parts *b* through *e*, but experiment with elimination of student BQ items. That is, ask whether it would be possible to drop some student BQ items if imputed Census data are used.

h. Consider each of the student-based, student/Census-based, and Census-based socioeconomic measures developed above. How does their criterion validity (relative to academic achievement) compare with that of NSLP? (Compare relationships of each index and NSLP eligibility with academic achievement(s). Has anything been gained by using multiple measures of SES?)

i. How does the criterion validity, relative both to parent reports and academic achievement, of indexes using student reports or student and Census 2000 data, as created in parts *c* to *h* above compare with that based on indexes with a simple weighting scheme, e.g., the average of standardized sample values? Is there any advantage to estimated vs. a prior weights?

j. Do the answers to these questions vary by race-ethnicity? By region? By gender? (Repeat above analyses for population subgroups.)

3. ACS sample at grade 8. This sample will yield EBQ as well as BQ reports from students and Census data from the ACS as well as Census 2000.

a. Comparability between the ACS and ECLS-K samples in the 8th grade. (Because of sample design differences, this may have to be done within race-ethnic groups or using weights within one or both samples. Ideally, we will find scant differences between ACS and ECLS-K samples.)

i. How do the distributions of BQ reports differ between the ACS and ECLS-K samples? (Cross-tabulations)

ii. Apply weights for SES indexes for BQ items in the ECLS-K sample to BQ items in the ACS sample. How do relationships between ECLS-K indexes and academic achievement differ from those in the ECLS-K sample? (Simple regressions and scatter plots. Use indexes based on all items available in the 8th grade BQ and indexes based only on items available in both the 4th and 8th grade BQ)

b. Comparability between areal data from the ACS and from Census 2000. (Ideally, we will find scant differences between relationships based on ACS data and those based on Census 2000 data within the ACS sample.)

i. What are the relationships between imputed values of socioeconomic characteristics obtained from ACS and from Census 2000? How do within-year relationships differ from between-year relationships, that is, is there any decay in the validity of imputations across time?

ii. Use weights for SES indexes including Census 2000 items in the ECLS-K sample to data from the ACS sample. How do relationships between ECLS-K indexes and academic achievement differ from those in the ECLS-K sample? (Simple regressions and scatter plots. Use indexes based on all items available in the 8th grade BQ and indexes based only on items available in both the 4th and 8th grade BQ)

iii. Repeat part *ii*, but use ACS data, rather than Census 2000 data. How do these findings differ from those in part *ii*?

iv. Repeat these analyses using simple unit weights for standardized index components.

c. Does the availability of EBQ data substantially improve the validity of socioeconomic reports by students, with or without supplemental Census data? Note there are two sets of EBQ data, those available at both the 4th and 8th grades and those available only at the 8th grade level. Also, there is no parent-based measure of SES to use as a standard; academic achievement is the only available criterion. For each available candidate index, look at relationships between the variables in that index and academic achievement in the total sample and within population subgroups, with and without supplementation by one or more EBQ measures. Does use of EBQ measures improve the prediction of academic achievement? Does use of EBQ measures render BQ measures unnecessary?

Note this is a multi-dimensional design. Indexes may or may not include Census data, may or may not include measures available only at the 8th grade level, and may or may not include EBQ items.

d. How does the criterion validity (relative to academic achievement) of the indexes created and analyzed above compare to that of NSLP? (Correlation analysis, graphical display)

4. ACS sample at grade 4. This sample will yield EBQ as well as BQ reports from students and Census data from the ACS as well as Census 2000.

a. Comparability between areal data from the ACS and from Census 2000. (Ideally, we will find scant differences between relationships based on ACS data and those based on Census 2000 data within the ACS sample.)

What are the relationships between imputed values of socioeconomic characteristics obtained from ACS and from Census 2000? How do within-year relationships differ from between-year relationships, that is, is there any decay in the validity of imputations across time?

b. Does the availability of EBQ data substantially improve the validity of socioeconomic reports by students, with or without supplemental Census data? As in the 8th grade ACS sample, there is no parent-based measure of SES to use as a standard; academic achievement is the only available criterion. For each available candidate index, look at relationships between the variables in that index and academic achievement in the total sample

and within population subgroups, with and without supplementation by one or more EBQ measures. Does use of EBQ measures improve the prediction of academic achievement? Does use of EBQ measures render BQ measures unnecessary?

Note this is a multi-dimensional design. Indexes may or may not include Census data and may or may not include EBQ items.

c. How does the criterion validity (relative to academic achievement) of the indexes created and analyzed above compare to that of NSLP? (Correlation analysis, graphical display)

d. How do the answers to these questions vary across subgroups defined by race-ethnicity, region, and gender?

Criterion Validation of an SES Composite

Despite the arguments in the following text, I have been persuaded, both by members of the Working Group and by the SES Standing Committee that it will be wise to avoid dependence on criterion validation, where academic achievement is the criterion. Unfortunately, because EBQ and ACS items are not available in the ECLS-K—the one sample in which parents' reports of their socioeconomic statuses are available—there is no way to avoid using academic achievement as a criterion in analyses of the 4^{th} and 8^{th} grade data from the 2007 pilot study. No other criterion is available. For that reason, I have retained this section of the draft plan of June 2006.

The focus of the proposed analyses on criterion validation of composite measures of SES has prompted a great deal of discussion, mainly focusing on the presumption that its objective is to construct an SES measure that will correlate highly with academic achievement. The presumption is incorrect. Rather, the objective of the analyses described above is to select a small number of variables that can be measured easily and accurately and that will capture most of the association between objective social and economic origins and academic achievement. That is, if we enumerate the set of all plausible socioeconomic characteristics of students, the objective is to find an economical composite of a few of those variables that will be reasonably representative of the combined effects of all of them. We do assume that the larger set of explanatory socioeconomic variables is nicely approximated by those available in the original or enhanced background questionnaires and the ACS data. The objective, then, is not to maximize explained variance, but to minimize cost without sacrificing validity.

One counter to the argument that criterion validation is in appropriate is to consider an extreme alternative: Choose an SES measure that is minimally related to academic achievement. Surely no one would propose this alternative, but to avoid it we must know, in advance, how proposed SES measures are related to academic achievement. Then, once we have an array of candidate measures and know their relationships with academic achievement, how should we choose among them? For example, how would we know

that a candidate composite SES measure is an improvement over eligibility for the National School Lunch Program? A sensible rule is to select a measure based on social and economic conditions that are known to affect academic achievement – but, to be sure, without over-fitting the data.

In the absence of the dual objectives – minimizing cost without losing validity – there could be no need for a 2007 Pilot Study. One possibility is that small group of social scientists could reach consensus on an appropriate list of constructs and indicators, and, without further ado, NAEP could proceed to measure them and write reports based on them. An uneducated guess is that this is pretty much what happened when eligibility for the National School Lunch Program (NSLP) was selected as the key measure of SES for use in NAEP. "Poverty" was widely regarded as both a major social problem and a key indicator of children's life chances. NSLP eligibility was based on the official poverty measure. Since it was already being ascertained by schools, it was both appropriate and convenient to choose NSLP eligibility as NAEP's socioeconomic indicator variable. We should ask – and learn – whether and how we can do better than that.

Another possibility is to choose a linear composite of variables available from NAEP that best predicts a composite of socioeconomic characteristics as reported by parents in ECLS-K. To be sure, it will be useful to look at relationships between candidate SES measures from NAEP and the reports of ECLS-K parents. However, it would not be wise to use this as the main criterion for a new SES measure for NAEP. First, the measures from ECLS-K parents do not exhaust the list of potential measures from NAEP and ACS. Second, the ECLS-K cases available in the 2007 Pilot Study will pertain to a population that is restricted in several ways: by grade level, school type (public), and survey attrition. Third, reports by parents, like those by students, are (differentially) subject to measurement error, thus possibly leading to inappropriate decisions about the choice of SES indicators. As long as measurement error in achievement is random, that problem will not occur when achievement is used as a validation criterion.[8]

Yet a third possibility is to extract a common factor or principal component from a set of socioeconomic measures and to choose those indicators that load most highly on the factor. This makes a good deal of sense in the context of achievement test items, where the notion is that unobserved abilities lead to observable test performance on a sample of items. As noted earlier, the factor model simply does not apply to socioeconomic measurement. That is, the measurement of socioeconomic status rightly reverses the direction of causation between indicators and constructs that is typical in psychometric measurement. Measurement decisions should be based on a model of how things work, not on irrelevant correlations. If there is such a construct as "socioeconomic status," it is not a "factor" that students carry around in their heads, but an effect of a set of social and economic conditions.

[8] Random measurement error in achievement test scores will increase the residual (unexplained) variance, but it will not affect the regression coefficients of achievement on SES nor the relative effects of SES indicators.

Richard Patz has nicely stated the argument that criterion validation of SES with academic achievement is undesirable:

> "Because the interpretations we aim to validate are focused on the relationship between achievement and socioeconomic status (SES), it is important that our measures of SES be independent of the achievement level of the measured individuals. For this reason we should not use achievement as a criterion for validating any newly constructed SES measure to be so used, and we should not select SES variables based on their statistical association with achievement, for doing so will risk contaminating (in particular, inflating) our conclusions about the strength of the relationship."

Patz also offers an apt example of conditions under which use of an achievement criterion could lead to an inappropriate explanatory construct:

> "Suppose we are interested in studying the relationship between "dietary health" and academic achievement, believing that research suggests that poor nutrition has a negative effect on the academic achievement of students. We want to create and study a "dietary health composite" that combines information from multiple variables. We start with body mass index (BMI) and find a weak statistical association with achievement. Selecting other variables using correlation with achievement as a criterion, we find that the presence of organic foods and the absence of processed animal fats (i.e., lard) have a strong association with achievement and we include them in our dietary health composite. Further research suggests that organic food is a proxy for SES (wealthier families buy more organic food), and animal fats are associated with membership in minority groups and English language learners. By using achievement to select these components of our composite, we end up with an incorrect conclusion about the association between dietary health and achievement."

Under some conditions, this example would indeed be cautionary. For example, in one prior effort to construct a socioeconomic composite for NAEP, researchers inappropriately chose two of eight components that were inherently confounded with academic achievement (Lubienski, Camburn, and Shelley II 2004). However, the NAEP SES project is in a very different situation. We have a clearly defined set of candidate measures from NAEP background questionnaires and from ACS/Census long-form data. These are all either socioeconomic characteristics of parents or characteristics of the students' homes that are indicative of the socioeconomic status of the parents, i.e., education, labor force and occupation, or income and wealth, or of family composition (number of siblings, parent absence). There are neither behavioral measures, nor measures of parent-student interactions, nor any policy variable (like the NSLP) that is intended to influence academic performance. In short, there is no chance that some mechanical variable selection procedure could yield an inappropriate set of SES components (nor has such a procedure been proposed).

One weakness of the argument that criterion validation will unreasonably maximize the observed association between an improved SES construct and academic achievement in NAEP is that it simply won't happen. There is sufficient redundancy among parental socioeconomic and family characteristics that a few measures will represent almost all of the explanatory power of a larger set of such variables.

A first illustration of this point is from a classic sociological monograph: Blau, Peter M. and Otis D. Duncan. 1967. *The American Occupational Structure*. New York: John Wiley and Sons, pp. 188-91.

The excerpt is appended to this document. It reports a series of hypothetical analysis of the effects of social background variables on the educational attainment, early occupational status, and current occupational status of a national sample of more than 20,000 American men in 1962. The message here is straightforward: For each of those outcomes, the analysis suggests that adding more social background characteristics adds very little to the explanatory power of the regression models. The limitation of this analysis is also obvious. It is purely illustrative because the additional background variables were not actually measured in Blau and Duncan's (1967) sample and, possibly, because the outcome variable is not academic achievement.

A second illustration is drawn from the Wisconsin Longitudinal Study, a nearly 50-year long study of more than 10,000 members of Wisconsin's high school graduating class of 1957.[9] For present purposes, I have regressed scores on the Henmon-Nelson Test of Mental Ability (Henmon and Nelson 1946; Henmon and Nelson 1954), administered during the junior year of high school, on a series of socioeconomic background variables, listed here in descending order of their zero-order correlations with the test scores:

> Father's educational attainment (mainly from the graduate in 1975)
> Mother's educational attainment (mainly from the graduate in 1975)
> Father's occupational status (mainly from the graduate in 1975)
> Average family income, 1957-1960 (from Wisconsin tax records)
> Number of siblings (from the graduate in 1975)
> Broken family (from the graduate in 1975)

It has been demonstrated elsewhere that these variables are highly reliable (Hauser and Xie ; Bielby, Hauser, and Featherman 1977b; Hauser and Massagli 1982; Hauser et al. 1983; Massagli and Hauser 1983a; Hauser and Mossel 1985; Hauser and Sewell 1986). However, because the participants in the study are all high school graduates and are from a single state – and because the Henmon-Nelson Test was not a test of academic achievement – the estimated relationships between SES variables and test scores in the WLS should not be taken as indicative of what we should expect to find in NAEP.[10]

[9] The present analysis is based on 8371 cases where all data are present.
[10] As mentioned above, it would be desirable to carry out some analyses of individual-level NAEP data before the 2007 Pilot Study is undertaken, and a similar exercise could be undertaken with existing NAEP background and achievement data.

Table 3 shows measures of fit (R^2) from selected regression models. The first six models simply show the correlations between each background measure and the H-N score, where the variables are listed in order from the most to least correlated with the score. The last model is a stepwise multiple regression in which the variables are entered in optimal order. Note that R^2 is only 0.059 in the first model (father's educational attainment). It rises to 0.077 in the second model (father's occupational status and father's educational attainment) and to 0.088 in the third model (father's occupational status, mother's educational attainment, and father's educational attainment). However, it increases only marginally (by 0.006) to 0.094 as three more variables are added to the equation. Thus, although six reliable measures of SES are available in the WLS data, only three are needed to represent almost all of the explanatory power of SES for the test scores.

What happens as additional regressors are added? Table 4 shows the regression coefficients in each of the steps of the regressions whose fit is shown in Table 3. Notice what happens to the standardized regression coefficients as variables are added. No variable dominates the regression, but the coefficient of each variable falls as other variables are added. That is, because the socioeconomic variables are moderately intercorrelated (but not highly collinear), they are partially redundant,[11] and each variable adds only marginally to the explanatory power of the set of variables.[12]

It is noteworthy that the relative weights of the SES measures do not matter very much. Recall that $R^2 = 0.088$ when father's educational attainment, mother's educational attainment, and father's occupational status have the (optimal) weights determined by a multiple regression analysis. For the same cases, the correlation is *unchanged* when H-N scores are regressed on a simple sum of standardized scores of the same three socioeconomic variables. Similarly, $R^2 = 0.093$ when H-N scores are regressed on father's educational attainment, mother's educational attainment, father's occupational status, and number of siblings, but the (squared) correlation is 0.091 when H-N scores are regressed on the sum of standard scores of the four variables.[13]

The modest – but highly significant – explanatory power of these models is not an artifact of the linear specification used here. For example, if we use dummy variables for all of the socioeconomic status characteristics, the explained variance is increased only to 0.097, adjusted for degrees of freedom.[14]

Thus, a very simple additive combination of the SES variables performs just about as well as an optimally weighted composite variable. This is relevant to the last of the analytic goals listed above: A simple composite of a few measures will be just about as valid as an optimally weighted composite of many variables. If we choose such a simple

[11] Excepting intact family, absolute values of correlations among the background variables range from about 0.15 to 0.51. Because few of the WLS graduates were from non-intact families, the correlations of that variable with other socioeconomic characteristics were quite small.
[12] Again, it will be desirable to confirm this pattern of findings in analyses of existing NAEP data.
[13] However, it is necessary to give number of siblings a negative sign in the linear composite.
[14] For other evidence of linearity in models of this kind, see Gasson, Haller and Sewell (1972).

Table 1. Some Models of Social Background and Test Scores: Wisconsin Longitudinal Study

Robert M. Hauser, May 20, 2006

Model Summary

Model	R	R Square	Adjusted R	Std. Error of the Estimate
1	0.242	0.059	0.059	14.360

a Predictors: (Constant), Best Measure of Head of Household Educ

Model Summary

Model	R	R Square	Adjusted R	Std. Error of the Estimate
1	0.227	0.051	0.051	14.416

a Predictors: (Constant), Best Measure of Mother's Education

Model Summary

Model	R	R Square	Adjusted R	Std. Error of the Estimate
1	0.240	0.058	0.058	14.368

a Predictors: (Constant), Best Measure of Father Occ Status in 1957

Model Summary

Model	R	R Square	Adjusted R	Std. Error of the Estimate
1	0.142	0.020	0.020	14.651

a Predictors: (Constant), Best Measure of Parental Income in 1957

Model Summary

Model	R	R Square	Adjusted R	Std. Error of the Estimate
1	0.142	0.020	0.020	14.651

a Predictors: (Constant), Ttl # of R Sis & Bro Ever Born

Model Summary

Model	R	R Square	Adjusted R	Std. Error of the Estimate
1	0.032	0.001	0.001	14.793

a Predictors: (Constant), liv w/both parent most of time in HS

Model Summary

Model	R	R Square	Adjusted R	Std. Error of the Estimate
1	0.242	0.059	0.059	14.360
2	0.277	0.077	0.077	14.247
3	0.297	0.088	0.088	14.135
4	0.305	0.093	0.092	14.126
5	0.306	0.094	0.093	14.093
6	0.307	0.094	0.094	14.089

a Predictors: (Constant), Best Measure of Head of Household Educ

b Predictors: (Constant), Best Measure of Head of Household Educ, Best Measure of Father Occ Status in 1957

c Predictors: (Constant), Best Measure of Head of Household Educ, Best Measure of Father Occ Status in 1957, Best Measure of Mother's Education

d	Predictors: (Constant), Best Measure of Head of Household Educ, Best Measure of Father Occ Status in 1957, Best Measure of Mother's Education, Ttl # of R Sis & Bro Ever Born
e	Predictors: (Constant), Best Measure of Head of Household Educ, Best Measure of Father Occ Status in 1957, Best Measure of Mother's Education, Ttl # of R Sis & Bro Ever Born, Best Measure of Parental Income in 1957
f	Predictors: (Constant), Best Measure of Head of Household Educ, Best Measure of Father Occ Status in 1957, Best Measure of Mother's Education, Ttl # of R Sis & Bro Ever Born, Best Measure of Parental Income in 1957, liv w/both parent most of time in HS

Table 4. Regression Models of Test Scores on Socioeconomic Background Variables: Wisconsin Longitudinal Study

Coefficients(a)

Model		Unstandardized B	Unstandardized Std. Error	Standardized Beta	t	Sig.
1	(Constant)	90.5	0.479		188.80	0.000
	Head of Household Educ	1.066	0.047	0.242	22.86	0.000
2	(Constant)	90.483	0.475		190.60	0.000
	Head of Household Educ	0.711	0.054	0.162	13.21	0.000
	Father Occ Status in 1957	1.01E-02	0.001	0.157	12.83	0.000
3	(Constant)	86.305	0.625		138.00	0.000
	Head of Household Educ	0.474	0.058	0.108	8.13	0.000
	Father Occ Status in 1957	9.12E-03	0.001	0.142	11.58	0.000
	Mother's Education	0.654	0.064	0.124	10.18	0.000
4	(Constant)	88.632	0.718		123.44	0.000
	Head of Household Educ	0.452	0.058	0.103	7.75	0.000
	Father Occ Status in 1957	8.50E-03	0.001	0.132	10.74	0.000
	Mother's Education	0.611	0.064	0.115	9.49	0.000
	Sis & Bro Ever Born	-0.422	0.065	-0.071	-6.55	0.000
5	(Constant)	88.567	0.718		123.36	0.000
	Head of Household Educ	0.441	0.058	0.100	7.55	0.000
	Father Occ Status in 1957	7.85E-03	0.001	0.122	9.57	0.000
	Mother's Education	0.596	0.065	0.113	9.23	0.000
	Sis & Bro Ever Born	-0.414	0.065	-0.069	-6.41	0.000
	Parental Income in 1957	8.22E-03	0.003	0.035	3.05	0.002
6	(Constant)	89.902	0.926		97.10	0.000
	Head of Household Educ	0.449	0.058	0.102	7.68	0.000
	Father Occ Status in 1957	7.79E-03	0.001	0.121	9.50	0.000
	Mother's Education	0.596	0.065	0.113	9.24	0.000
	Sis & Bro Ever Born	-0.407	0.065	-0.068	-6.30	0.000
	Parental Income in 1957	7.75E-03	0.003	0.033	2.87	0.004
	Lived with both parents	-1.281	0.561	-0.024	-2.28	0.023

a Dependent Variable: Henmon-Nelson Test of Mental Ability

measure, we need not be worried either about over-fitting the data or about loss of validity. Moreover, while the WLS illustration pertains to a set of analyses in a single population, the 2007 NAEP Pilot Study will provide data and analyses for multiple populations and achievement test scores by grade level, gender, race-ethnic group, etc. Thus, an additional criterion for simplifying the construction of an SES measure will be to choose variables that are of roughly similarly reliable and valid in these several populations.

One critic of the present plan has suggested that a criterion-validated measure of SES will in fact be "achievement," rather than SES. This is a misunderstanding. A linear (or nonlinear) composite of socioeconomic variables, however weighted, will still be no more than that. One relevant counter-example is the well-known Duncan SEI, a measure of the socioeconomic status of occupations (Duncan 1961). The Duncan SEI – along with subsequent updates and revisions of it (Stevens and Featherman 1981; Stevens and Cho 1985; Hauser and Warren 1997; Pagani, Tremblay, Vitaro, Boulerice, and McDuff 2001) – was constructed by regressing aggregate survey-based measures of the social standing of occupations (prestige) on census-based characteristics of the same occupations. The SEI was defined as the weighted combination of occupational education and income that best predicted prestige, but the SEI is a measure of occupational requirements and remuneration. It is not prestige, nor does it behave like prestige in analyses of occupational stratification (Featherman, Jones, and Hauser 1975; Hauser and Warren 1997).

There is one important sense in which the use of criterion validation will lead to a particular choice of SES components. The pilot study is scheduled for 2007, and we propose to develop an SES measure using indicators ascertained in that year. Thus, to the extent that there is differential change in the validity of SES indicators across time, the new SES measure may change in validity across time. One painful example of this is the fact that the NSLP measure is based upon a poverty standard whose validity has declined across time (National Research Council, Panel on Poverty and Family Assistance 1995). Another example is provided by the growing importance of maternal occupational standing in the life-chances of children (Kalmijn 1994). Such changes are a fact of life, and no SES measure will be right forever. Changes across time in the validity of specific indicators would be equally problematic in other data-dependent schemes for constructing an SES measure, like the two described above. One way of mitigating, if not circumventing this problem, would be to cross-validate new SES measures using items from earlier NAEP administrations and long-form Census data for small areas from 1990 or 2000.[15]

What about comparability?

The preceding text contains several references to achieving comparability in the measurement of SES across subpopulations. To that, we can add comparability with respect to different measures of academic achievement. It is not necessarily the case that

[15] This would not be a trivial undertaking. Links to Census data would have to be made from school locations, rather than student locations. Again, it would require access to micro-data from NAEP.

the socioeconomic resources in students' homes that may affect reading achievement will be the same as those that affect math achievement or science achievement.

For present purposes, a heuristic definition of comparability might be, simply, that the same set of socioeconomic characteristics have similar *relative* effects on multiple instances of academic achievement in multiple populations. The specification of similar *relative* effects leaves open the possibilities that some instances of academic achievement are more sensitive to SES than others and that academic achievement may be more highly correlated with SES in some student populations than in others. That is, drawing also on the goal of parsimony stated above, one would like to find a small set of socioeconomic variables that have similar relative weights in regressions of multiple instances of academic achievement on SES in multiple populations. Ideally, given the goal of transparency in index creation, we should also seek simplicity in the choice of relative weights.

The general idea may be expressed in equation form as follows (suppressing subscripts for individual cases):

$$Y_{jk} = \alpha_{jk} + \gamma_j \delta_k \sum_i \beta_i X_i + \varepsilon_{jk}$$

where Y_{jk} is the j^{th} instance of academic achievement in the k^{th} population, γ_j and δ_k are coefficients of proportionality pertaining to the instances of academic achievement and population, respectively, the X_i are socioeconomic variables, and the β_i are the desired relative weights. Statistical methods for such models have been developed by Hauser and Goldberger (1971), Joreskog and Goldberger (1975), and Hauser and Andrew (2006). However, in the present case, I see this scheme more as a metaphor for a desired result than a specific model to be estimated.

Reference List

Bielby, William T. and Robert M. Hauser. 1977. "Response Error in Earnings Functions for Nonblack Males." *Sociological Methods and Research* 6:241-80.

Bielby, William T., Robert M. Hauser, and David L. Featherman. 1977a. "Response Errors of Black and Nonblack Males in Models of the Intergenerational Transmission of Socioeconomic Status." *American Journal of Sociology* 82(6):1242-88.

Bielby, William T., Robert M. Hauser, and David L. Featherman. 1977b. "Response Errors of Nonblack Males in Models of the Stratification Process." *Latent Variables in Socioeconomic Models*, Eds. Arthur S. Goldberger and Dennis Aigner. Amsterdam: North Holland.

Blau, Peter M. and Otis D. Duncan. 1967. *The American Occupational Structure*. New York: John Wiley and Sons.

Duncan, Otis D. 1961. "A Socioeconomic Index for All Occupations." Pp. 109-38 in *Occupations and Social Status.* edited by Albert J. Jr. Reiss. New York: Free Press.

Featherman, David L., F. L. Jones, and Robert M. Hauser. 1975. "Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status." *Social Science Research* 4(4):329-60.

Gasson, Ruth M., Archibald O. Haller, and William H. Sewell. 1972. *Attitudes and Facilitation in the Attainment of Status*. Washington, D.C.: American Sociological Association.

Hauser, Robert M. 1972. "Disaggregating a Social-Psychological Model of Educational

Attainment." *Social Science Research* 1(2):159-88.

Hauser, Robert M. and Megan Andrew. 2006. "Another Look at the Stratification of Educational Transitions: The Logistic Response Model With Partial Proportionality Constraints." Pp. 1-26 in *Sociological Methodology 2006*, edited by Ross M. Stolzenberg. American Sociological Association and Blackwell Publishers.

Hauser, Robert M. and Arthur S. Goldberger. 1971. "The Treatment of Unobservable Variables in Path Analysis." Pp. 81-117 in *Sociological Methodology 1971*, edited by Herbert L. Costner. San Francisco: Jossey-Bass.

Hauser, Robert M. and Michael P. Massagli. 1982. "Some Models of Agreement and Disagreement in Repeated Measures of Occupation." Pp. 346-51 in *Proceedings of the Social Statistics Section* . Washington, D.C.: American Statistical Association.

Hauser, Robert M. and Peter A. Mossel. 1985. "Fraternal Resemblance in Educational Attainment and Occupational Status." *American Journal of Sociology* 91(3):650-73.

Hauser, Robert M. and William H. Sewell. 1986. "Family Effects in Simple Models of Education, Occupational Status, and Earnings: Findings From the Wisconsin and Kalamazoo Studies." *Journal of Labor Economics* 4(3, Part 2):S83-S115.

Hauser, Robert M., Shu-Ling Tsai, and William H. Sewell. 1983. "A Model of Stratification With Response Error in Social and Psychological Variables." *Sociology of Education* 56(1):20-46.

Hauser, Robert M. and John R. Warren. 1997. "Socioeconomic Indexes for Occupations:

A Review, Update, and Critique." Pp. 177-298 in *Sociological Methodology 1997*, edited by Adrian E. Raftery. Cambridge: Basil Blackwell.

Hauser, Robert M. and Raymond S.-K. Wong. 1989. "Sibling Resemblance and Inter-Sibling Effects in Educational Attainment." *Sociology of Education* 62(3):149-71.

Hauser, Seth M. and Yu Xie. "Temporal and Regional Variation in Earnings Inequality: Urban China in Transition Between 1988 and 1995." *Social Science Research* In Press, Corrected Proof.

Henmon, V. A. C. and M. J. Nelson. 1946. *Henmon-Nelson Tests of Mental Ability, High School Examination - Grades 7 to 12 - Forms A, B, and C. Teacher's Manual*. Boston: Houghton-Mifflin Company.

———. 1954. *The Henmon-Nelson Tests of Mental Ability. Manual for Administration*. Boston: Houghton-Mifflin Company.

Jöreskog, Karl G. and Arthur S. Goldberger. 1975. "Estimation of a Model With Multiple Indicators and Multiple Causes of a Single Latent Variable." *Journal of the American Statistical Association* 70:631-39.

Kalmijn, Matthijs. 1994. "Mother's Occupational Status and Children's Schooling." *American Sociological Review* 59(2):257-75.

Lubienski, Sarah T., Eric Camburn, and Mack C. Shelley II. 2004. *Reform-Oriented Mathematics Instruction, Achievement, and Equity: Examinations of Race and SES in 2000 MAin NAEP Data*. Washington, DC: National Center for Education Statistics.

Mason, William, Robert M. Hauser, Alan C. Kerckhoff, Sharon S. Poss, and Kenneth

Manton. 1976. "Models of Response Error in Student Reports of Parental Socioeconomic Characteristics." Pp. 443-519 in *Schooling and Achievement in American Society*, edited by William H. Sewell, Robert M. Hauser, and David L. Featherman. New York: Academic Press.

Massagli, Michael and Robert M. Hauser. 1983a. "Response Variability in Self and Proxy Reports of Paternal and Filial Socioeconomic Characteristics." *American Journal of Sociology* 89(2):420-31.

Massagli, Michael P. and Robert M. Hauser. 1983b. "Response Variability in Self- and Proxy Reports of Paternal and Filial Socioeconomic Characteristics." *American Journal of Sociology* 89(2):420-431.

May, Henry. 2006. "A Multilevel Bayesian Item Response Theory Method for Scaling Socioeconomic Status in International Studies of Education." *Journal of Educational and Behavioral Statistics* 31(1):63-79.

National Research Council, Panel on Poverty and Family Assistance. 1995. *Measuring Poverty: a New Approach*, Edited by Constance F. Citro and Robert T. Michael. Washington, D.C: National Academy Press.

Pagani, L., R. E. Tremblay, F. Vitaro, B. Boulerice, and P. McDuff. 2001. "Effects of Grade Retention on Academic Performance and Behavioral Development." 13(2):297-315.

Stevens, Gillian and Joo H. Cho. 1985. "Socioeconomic Indexes and the New 1980 Census Occupational Classification Scheme." *Social Science Research* 14:142--168.

Stevens, Gillian and David L. Featherman. 1981. "A Revised Socioeconomic Index of

Occupational Status." *Social Science Research* 10(4):364-95.

Warren, John R. and Robert M. Hauser. 1997. "Social Stratification Across Three Generations: New Evidence From the Wisconsin Longitudinal Study." *American Sociological Review* 62(4):561-72.

Warren, John R., Jennifer T. Sheridan, and Robert M. Hauser. 2002. "Occupational Stratification Across the Life Course: Evidence From the Wisconsin Longitudinal Study." *American Sociological Review* 67(3, June):432-55.

WILEY BOOKS BY THE SAME AUTHORS:

Peter M. Blau

Exchange and Power in Social Life, 1964

Otis Dudley Duncan and Albert J. Reiss, Jr.

Social Characteristics of Urban and Rural Communities, 1950, 1956

The American Occupational Structure

Peter M. Blau and Otis Dudley Duncan

WITH THE COLLABORATION OF ANDREA TYREE

JOHN WILEY & SONS, INC. NEW YORK · LONDON · SYDNEY

1967

from those estimated here. In the meantime we have a description of the "typical" life cycle of a cohort that is more detailed, precise, and explicit as to causal or sequential relationships than any hitherto available.

CONJECTURES AND ANTICIPATIONS

In an earlier section of this chapter we suggested that the critic might share part of the burden of proof for the proposition that our results are distorted by the omission of important variables. There is, however, evidence at hand, supplemented by judicious conjecture, to show that at least some obvious candidates for crucial omitted variables are not as formidable as might be supposed.

One kind of question has to do with the temporal relevance of our measure of father's status. The OCG questionnaire asked for father's occupation at the time the respondent was about 16 years old. Might we not suppose that father's occupation at an earlier date would have been a better choice, on the theory that occupational ambitions are developed in late childhood and early adolescence, being more or less fixed by the time a boy reaches high school age? Moreover, if the father were mobile during the respondent's youth, the sharing of the experience of mobility may have induced distinctive orientations in the respondent.

A different issue is whether we have overlooked a crucial factor in failing to procure some information about the respondent's mother. Several sociologists have recently emphasized the mother's role in the formation of achievement orientation and have called attention to her educational attainment as an indicator of her possible influence.

We shall discuss these two possibilities together because our approach in both cases is to present hypothetical calculations based on data that are largely conjectural but include a key item of information for which reasonably firm estimates are available.

Suppose the OCG survey had ascertained not only father's occupation at respondent's age 16 (variable X) but also at respondent's age 6 (variable X'). We must make two sorts of assumption. The first assumption is that X' has the same correlation with the other variables, V, U, W, and Y, as that observed for X. There is some support for this assumption. In the son's generation, as shown by the OCG data, r_{UW} is not strikingly different from r_{UY}. This suggests that in the father's generation X and X' might have similar correlations with V. As for the father-son correlations, we assume that the earlier occupation is as highly correlated with son's educational attainment and occupational achievement as is the later occupation of the father; that

TABLE 5.5. HYPOTHETICAL REGRESSION COEFFICIENTS IN STANDARD FORM (BETA COEFFICIENTS), FOR SPECIFIED COMBINATIONS OF VARIABLES, FOR MEN WITH NONFARM BACKGROUND, BASED ON PARTLY CONJECTURAL DATA

Dependent Variable[a]	Independent Variables[a]						Coeffieient of Determination (R^2)
	W	U	X'	X	V'	V	
SET 1							
U	...	...	...	.265	...	.285	.23
U	...	...	.183	.183	...	.233	.25
W	...	.450	...	.170	...	.037	.32
W	...	.434	.120	.120	...	.008	.33
Y	.279	.411	...	.103	...	-.019	.43
Y	.271	.405	.074	.074	...	-.037	.43
SET 2							
U	...	...	...	.265	...	.285	.23
U	...	...	...	.209	.196	.196	.25
W	...	.450	...	.170	...	.037	.32
W	...	.446	...	.163	.027	.027	.32
Y	.279	.411	...	.103	...	-.019	.43
Y	.279	.413	...	.107	-.014	-.014	.43

[a] V: Father's education.
V': Mother's education (conjectured).
X: Father's occ. status at respondent's age 16.
X': Father's occ. status at respondent's age 6 (conjectured).
U: Respondent's education.
W: Respondent's first job status.
Y: Respondent's occ. status in 1962.

is, that the correlations of X and X' with U, W, and Y are the same. The second assumption—and this is the crucial one—concerns the correlation of X with X'. Here we can draw on the data given earlier as well as on an OCG finding. The latter, which may be less relevant, is that for men 35 to 44 years old r_{YW} is .492. It will be recalled that there are two sources giving correlations between current occupation and occupation ten years earlier. For men 35 to 44 years old the Chicago data showed this to be .55; in the Minneapolis study it was .83. Our argument will only be weakened if we estimate $r_{XX'}$ on the low side; accordingly, we assign it the low compromise value of .60.

With these assumptions we have enough actual and hypothetical data to enter X' into a regression equation alongside X. Set 1 of Table 5.5 shows the results, in each case the previously calculated regression followed by the new hypothetical calculation in which X' is included as an independent variable. For each dependent variable the two measures of father's occupation split into equal shares the net influence formerly attributed to X alone. This particular result is without interest, as it merely reflects the assumption of equality of the respective correlations, which we assumed. The more important results—those we take to be indicative of what actual data might well

show—concern the coefficients of the other variables in the equations and the over-all change in proportion of variation determined. The most substantial change, and it is small enough, is noted with U as the dependent variable. With both occupational variables in the equation, the net influence of father's education is slightly diminished, and R^2 is two percentage points higher than with only X and V in the equation. At the other extreme, with Y as the dependent variable, we find no change in the other coefficients worth reporting and no detectable increase in R^2 due to the addition of X' to the other four variables.

Altogether, these results suggest that having much more detailed information on the father's occupational career would change very little our estimate of the relative importance of this factor as a determinant of the son's occupational achievement. The results leave open, of course, the question of the age at which the influence of father's occupation is most directly relevant to the course of the son's career, as well as the question of the particular influence a rare but extreme change in the father's career may have on that of the son.

In set 2 of Table 5.5 we have carried out the analogous exercise, considering hypothetical variable V' (mother's education) alongside measured variable V (father's education). Again we assume that their respective correlations with other variables in the system are the same. Unpublished data we have seen on educational plans and occupational aspirations of high-school youth suggest that mother's education is, at most, no more highly correlated with such variables than is father's education. Again, the crucial assumption has to do with the intercorrelation of the two key independent variables, V and V'. From the OCG data we can ascertain that there is substantial assortative mating by education in the respondent's generation. For men 45 to 54 years of age, the correlation between husband's and wife's education is .580, and for men 55 to 64 years old it is no less than .632. In 1940 Census tables on fertility we find a tabulation of education of husband by education of wife for parents of children under five years old; this correlation, computed somewhat approximately owing to broad class intervals, is .637. There should, of course, be little difference between this correlation and one computed for parents of boys 16 years old. Evidently we shall not greatly overestimate $r_{VV'}$ in setting it equal to .60.

The reader who has grasped the principle at work here will not be surprised to see in set 2 results much like those obtained in set 1. Mother's education divides with father's education the influence initially attributed to the latter, as a consequence of the assumptions

made. With U (respondent's education) as the dependent variable, inclusion of V' results in an appreciable diminution of the net influence attributed to father's occupation and a measurable increase in the proportion of variation in the dependent variable accounted for. For dependent variables W and Y, however, the additional variable contributes no additional information, since the education of neither parent has an appreciable direct effect on respondent's occupational status. It should be reiterated that these calculations do not answer the question of whether mother's or father's education exerts more influence on sons.

It is hardly conjectural to generalize from these two experiments in a certain respect. If we think of additional socioeconomic indicators applying to the respondent's family background it is fairly certain that each of them will correlate moderately highly with the two that we have measured here. We do not know for sure, but it seems rather unlikely that any of them will have a much higher simple correlation with our measures on the respondent than X or V. In this event inclusion of other family background socioeconomic variables may lead to some reinterpretation of how the effect of such variables is transmitted, or of what is their relative importance, but it will not alter greatly our over-all estimate of the importance of variables of this kind. He who thinks differently, of course, has the option of trying to support his opinion with evidence. As far as we can see there is every reason to suppose that we have not appreciably underestimated the role of the socioeconomic status of the family of orientation as an influence upon the respondent's occupational achievement.

Concerning several other omitted variables, we need not resort to conjecture but merely to anticipate a little of the content of subsequent chapters in this volume. These chapters are mainly concerned with qualitative or classificatory factors as possible influences on occupational achievement. This kind of factor is not readily introduced into the kind of causal diagram we have been working with in this chapter. We can, however, inquire whether neglect of such factors may have seriously misled us in regard to the nature of the causal relationships we have assumed. If, for example, a qualitative factor H operates as a determinant of both one (or more) of the independent variables and one (or more) of the dependent variables in our causal model, then the link between the two that we postulate is, in greater or lesser degree, spurious. In the event of this kind of spuriousness, holding the qualitative factor constant should markedly reduce, if not eliminate entirely, the apparent correlation between the two variables.

In Table 5.6 we report the amount of change in the correlation