DOCUMENT RESUME

ED 320 936 TM 015 179

AUTHOR Kolstad, Andrew

TITLE The Impact of Clustering in the Sample Design of the 1987 High School Transcript Study on Estimates of Sampling Variability.

PUB DATE Apr 90

NOTE 16p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).

PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Cluster Analysis; *Educational Assessment; Error of Measurement; Estimation (Mathematics); High Schools; High School Students; Methods Research; *National Surveys; Research Methodology; *Sampling; Testing Problems

IDENTIFIERS Hierarchical Linear Modeling; *High School Transcript Study 1987; *National Assessment of Educational Progress

ABSTRACT

Use of clustering methods with data from the 1987 High School Transcript (HST) component of the 1986 National Assessment of Educational Progress (NAEP) is discussed in an attempt to persuade data analysts and researchers that they should be interested in problems traditionally turned over to samplers. Nearly all surveys used by the National Center for Education Statistics incorporate a stratified, multi-stage cluster sample design due to the substantial economies involved in the cost of data collection. Although clustering is a simple idea for reducing costs, the assumption of independent and identically distributed errors is violated by the clustering of students within schools. Weighting alone does not correct for the lack of independence of the observations when computing estimates of standard errors. No explicit mathematical solution can generally apply to the problem of clustering. The sample design and measures of course-taking used in the 1986 NAEP are described. Data for 21,446 students in the 1987 HST Study are included. Results, including comparisons with 1982 High School and Beyond results, indicate that the combined impact of large cluster sizes and large intra-cluster correlations is so large that one should not try to estimate standard errors without adopting a replication method. Secondly, since large within-cluster correlations exist, it is likely that further work using a hierarchical linear modeling approach is likely to be fruitful. Four data tables are included. (TJH)

The Impact of Clustering in the Sample Design of the 1987 High School Transcript Study on Estimates of Sampling Variability

By

Andrew Kolstad
National Center for Education Statistics
555 New Jersey Avenue, NW
Washington, DC 20208

Paper prepared for presentation to the annual meetings of the American Educational Research Association
Boston, Massachusetts, April 18, 1990

INTRODUCTION

Over the last ten years, high school transcript studies have become an ongoing statistical program at NCES. The idea behind our data collection program is simply to collect school records on the courses students have taken when they graduate from high school, in order to report on changes over a period of years in what successive cohorts of students study. Last year at AERA, I reported results on the changes in national patterns of course-taking over the period from 1982 to 1987. The 1982 results came from the transcript component of the High School and Beyond (HS&B) survey, while the 1987 results came from a transcript component added on to the 1986 11th grade sample of the National Assessment of Educational Progress (NAEP). This year my report is more methodological than substantive in nature.

Three more such transcript studies are currently in the planning stages: Both in 1990 and in 1992 the national (though not the state) NAEP surveys will include a transcript collection component for the 12th graders. The National Education Longitudinal Study of 1988 (NELS:88) will collect data on its eighth grade cohort after this group graduates from high school in 1992.

Clustered designs in NCES surveys. Nearly all NCES surveys, and certainly all our large-scale surveys, incorporate a stratified, multi-stage cluster sample design, because of the substantial economies involved in the cost of data collection. NCES is not an innovator in this practice; our surveys follow standard statistical procedures and have done so for many years.

Clustering is a simple idea for reducing costs. Its benefit derives from the fact that it is far cheaper to visit one school and collect fifty transcripts than it is to visit fifty schools and collect one transcript each. The economies are so great that NCES will never conduct a simple random survey. However, the clustered design creates some drawbacks in the resulting data, the principal one being that students from the same school are subject to common influences that make them more alike than those from different schools. The lack of complete independence and the homogeneity of population clusters increases the true variance of sample estimators. The principal statistical problem is that the standard errors of means and proportions are underestimated when normal procedures are followed.

The assumption of independent and identically distributed errors is violated by the clustering of students within schools. The fifty students in the one school are more alike in term of their course-taking patterns than the fifty students from fifty schools. With the high school as a cluster, this extra degree of homogeneity happens not only because the students have some impact upon one another, but more importantly, because they are subject to the common influences, such as community resources and school educational policies. Schools are rich or poor, and the courses their students take are influenced by the resources of the school. The courses schools offer or fail to offer influence the courses students can take. The graduation requirements that schools set influence the courses students take. Many such factors cooperate to produce similarities in the course-taking patterns among students from the same school.

The problems of sample design and cost efficiency have traditionally been the concern of the sampling statistician, whose expertise is devoted to measurement and reduction of sampling error. While AERA is a community of many interests, there are few sampling statisticians here. Much more common at AERA are modellers, who are concerned with the social and educational processes by which learning comes about.

Analysts typically develop their models in an adaptive process as they progress. When, through trial and error, or through hypothesis testing, they see systematic differences among subgroups, they try to incorporate the explanatory factors in their models. The process of analysis is an evolutionary one in which the final parameters to be estimated are based on the population structure that the analysts tease from the data. An important part of the analysis is determining when relationships among variables, or differences among groups, are large enough to be significant.

The sample design problem arises for analysts in that the criterion for the significance of group differences, or of model parameters, is usually the standard error of these differences or parameters. Unless the survey data were collected using a simple random sample (which is rarely the case) the estimates of these standard errors are usually too small, and confidence intervals too narrow, leading the analyst to think that groups are more different than they really are. Sampling statisticians have known about this underestimation of standard errors for many years, and have developed some methods for understanding and dealing with the problem.

Objective. My objective in this paper is to persuade data analysts and researchers that they should be interested in problems that have traditionally been turned over to samplers. Recently, some new methodological developments, particularly those that come under the heading of hierarchical linear models, have solved old problems and produced a convergence of concerns between sampling statisticians interesting in estimating and reducing survey errors and survey analysts interested in exploring and explaining relationships among educational phenomena (Skinner, Holt, and Smith (eds), 1989). The statistical problems of clustering in educational surveys are important for modellers to consider, because we are interested in educational processes, and in the school influences that bring about similarity among students. The findings I want to report today should illustrate quantitatively how important the problem of clustering is in the 1987 High School Transcript Study to report.

COMPLEX SAMPLE DESIGNS

What makes sample designs complex is not stratification. Stratifying the sample to ensure that one gets a sufficient number of cases of one kind or another is not much of a problem from a statistical point of view, in that the impact of stratification is mathematically tractable and already incorporated in most statistical packages. The proper procedure is to use a weight proportional to the probability of selection in computing a mean or a proportion.

What makes sample designs complex is clustering. Weighting alone does not correct for the lack of independence of the observations when computing estimates of standard errors. Choosing a group of people from a cluster is economical, but the people chosen are more similar to one another than would more widely scattered individuals in many, often unknown ways. The impact of clustering is mathematically intractable (hence the term "complex"), and is rarely incorporated in the common statistical packages.

Adjusting for the impact of clustering. There is no explicit mathematical solution that can generally apply to the problem of clustering. With the advent of cheap computing, sampling statisticians came up with several methods to find the appropriate standard errors. These methods generally follow a similar approach, and that is to compute repeated estimates of the parameters based on different subsets of the sample, and then use the observed variability among these estimates to derive a measure of the precision of the estimates. These methods have been labeled "replication," because the subsamples are called replicates. Some variations on the replication approach are the methods of balanced repeated replications, jackknife replications, and bootstrap methods. They are all quite expensive, because instead of computing a statistic once, the same statistic needs to be computed over and over again, until one has enough estimates to determine its variability. While the replication approach successfully produces accurate measures of precision, it provides no information on the factors that bring about within-cluster similarities.

The modeling approach to this problem tries instead to look for cluster-level factors that bring about the similarity among sample cases. This approach has been labeled "hierarchical linear modeling". It is also fairly expensive, since many statistics need to be computed for each cluster.

However, like all linear models, it typically happens that not all the cluster-level factors have been measured and incorporated, and there may remain unexplained similarities among sample cases from within the same cluster. It may take a combination of both methods to deal adequately with both controlling for and understanding the nature of the similarity among students within schools.

DATA SOURCE

In 1987, NCES sponsored a survey of course-taking by high school students. The initial idea was to take advantage of the information on schools and students already available from the 11th grade/age 17 sample for the 1986 National Assessment of Educational Progress. This section of the paper describes how the 1987 High School Transcript Study sample was designed, how the transcripts were collected and standardized, and how the weights for population estimates were obtained. The school and student samples were selected through standard survey procedures that provide known probabilities of selection, so that the findings from the survey can be considered nationally representative.

Sample design. The 1986 NAEP sample design was a stratified, multistage probability sample of schools, with students randomly selected within schools. Counties were the first stage, secondary schools the second stage, assessment sessions the third stage, and students the fourth stage.

a. Selection of Primary Sampling Units. The PSU sample design was a stratified sample with one PSU selected per stratum, with probability proportional to county population (One-third of the PSUs were so large that they were selected with certainty). A total of 94 primary sampling units (PSUs) were included in the sample. The stratification variables were region, metropolitan status, and percent minority. The values of these 1980 Census variables that determined the PSU selections were not put into the public data file. A full description of the 1986 NAEP sample design is contained in Burke, et al. (1987).

b. School sample. From a frame listing all public comprehensive and private high schools in the selected PSUs, a school sample was chosen with probability proportional to student enrollment in 11th grade (though for cost efficiency this probability was lowered for small schools and raised for high-minority schools). Schools that refused to participate were replaced by substitutes from the same PSU. Of the 479 secondary schools selected, 433 schools (90 percent) participated by supplying copies of transcripts and related information during the fall and winter of 1987. The measures of size and minority status that determined the school selections were placed in the public data file.

c. Student sample. On the whole, all students in the participating schools were listed, and students were randomly selected from the lists with a uniform probability (with the exception of handicapped students, all of whom were selected). The sample of students for the 1987 High School Transcript Study consisted of a total of 35,180 students in the participating high schools, distributed as follows: (1) 43.5 percent were nonhandicapped students who had been sampled for the 1986 NAEP survey; (2) 36.8 percent were newly sampled nonhandicapped students (newly sampled to replace the students who had participated in the 1986 NAEP, but whose identities were lost);

and (3) 19.7 percent were handicapped students, specifically oversampled as part of the Transcript Study. Transcripts were obtained for 34,140 students, or 97.0 percent of those in participating high schools. A full description of the additional factors involved in the 1987 Transcript Study sample design for schools and students is contained in Thorne, et al. (1989).

d. Comparability with 1982 High School and Beyond study. Some students were excluded from the sample in order to make valid curriculum comparisons between the 1987 data and the 1982 HS&B data. The samples of students for the present tables exclude nongraduates and those students who had participated in a special education program during high school from both the High School and Beyond study and the 1987 High School Transcript Study. In addition, some transcripts with missing or incomplete data were not usable, resulting in an actual sample size of 22,372 nonhandicapped graduates with complete records. The tables presented below focus on nonhandicapped high school graduates.

e. Average cluster sizes. The degree of clustering in the 1987 High School Transcript Study is larger than most NCES surveys, about 54 students per school, and the impact of clustering on the estimates is also expected to be larger than most. For comparison, the 1982 HS&B transcript study obtained about 12,000 transcripts from about 1,000 high schools, for an average cluster size of 12. Even though it is based on the same schools as the 1987 Transcript Study, the 1986 NAEP study has a much smaller average cluster size, because the spiral design of the instrument (that is, systematically different survey forms for students in the same school) gives the same survey form to only a subset of the 32,000 students in the age 17/grade 11 sample. For any given NAEP item, only 2,700 students respond, resulting in an average cluster size of 6. For a given NAEP scale, about 8,000 students respond, resulting in an average cluster size of 18.

Measures of course-taking. There are two principal dimensions along which course-taking can be measured: content, measured in terms of a classification of subject matter, and quantity, measured in terms of the credits earned.

a. Coding of courses on transcripts. In order to make possible the statistical summarization of a vast diversity of course content in the Nation's schools, the 1987 High School Transcript Study standardized the courses that were listed on the transcripts by classifying each course into a six-digit code, based on course content and level according to the Classification of Secondary School Courses (CSSC), containing approximately 1,800 course codes. The CSSC is detailed enough that it can distinguish an on-grade-level 10th grade English course from a below-grade-level 10th grade English course.

The CSSC was developed for use in the 1982 High School and Beyond Transcript Study. For the later 1987 High School Transcript Study, the CSSC had to be adapted to expand the vocational education course codes and to identify more accurately remedial courses and functional courses for special education students (who were largely absent from the 1982 HS&B transcript study). Unlike the 1982 HS&B Transcript Study, some additional course information was coded for each student, including the identification of courses as remedial, regular, or advanced, as offered in a different location, or as designed for handicapped students. Course catalogs and other information from participating schools were used to determine the content and level of courses. For each course on each student's transcript, information on grades earned and credit received was also standardized and transcribed.

b. Carnegie units of course credit. The standard unit of course credit is the Carnegie unit, defined as the equivalent of one course meeting five times a week for one class period throughout the school year. A one-semester course is half a Carnegie unit. If a student were to take five full-year courses a day throughout the four years of high school, the student would graduate with twenty accumulated credits. The courses and their credits were aggregated into major subject groupings for reporting purposes. The detailed listings of which courses were included in the major groups shown in the tables are available on request from NCES.

Weighting. Student transcript data were weighted for the purpose of making national population estimates of course taking. In the 1987 High School Transcript Study, the final weight attached to an individual student record reflected two major aspects of the sample design and the population being surveyed. The first component, the base weight, was used to expand sample results to represent the total population and reflected the probability of selection in the sample (estimated as the product of the probability of selection of the primary sampling unit and of the school and student within the primary sampling unit). The second component resulted from the adjustment of the base weight to account for nonresponse within the sample. Chapter 6 of the Technical Report provides details on the six factors that compose the final student weights (Thorne, et al, 1989).

Replicate weights. For the 1987 High School Transcript Study, the sampling statistician for the NCES data collection contractor (Westat) prepared a set of 36 replicate weights attached to each student record, as described in the Data File User's Manual (Thorne, et al., 1989). The 36 replicate weights differ from the final student weight in that the remaining member of a pair member is given the additional weight of the missing member when its pair is dropped out for a given replicate. Jackknife variance estimation for the 1987 High School Transcript Study is performed using these weights by repeating the estimation procedure 37 times, once using the original full set of sample weights, and once each for the set of 36 replicate weights. The variability among replicate estimates is then used to derive an approximately unbiased estimate of sampling variance.

RESULTS

The principal substantive findings of the study, for which comparisons were made with the 1982 HS&B study, have already been reported at AERA (Kolstad, 1989), and more detailed findings are starting to appear (Tuma and Gifford, 1990).

The first results relate to estimates of the average number of Carnegie credits earned in all subjects over four years, and in 21 separate subject fields, including academic, vocational, and other miscellaneous subjects. My purpose in reporting the average numbers of credits earned is not to look closely at the meaning of the patterns of course-taking, though there may be some interesting findings here. In this context, the averages are just arbitrary estimates whose precision can be examined.

Table 1 contains weighted estimates of the average number credits earned in various subjects, for the entire sample, for young men, for young women, and for several racial and ethnic groups: whites, blacks, Hispanics, Asians, and others. I also computed unweighted estimates, but except where the average number of credits is so small that rounding error is a problem, the absolute size of the bias in the unweighted estimates is only one or two percentage points, so I did not report the unweighted means.

Even though the bias due to not using the weights in this survey is small, there is no good reason not to use them. The procedures for obtaining weighted estimates are readily available in common statistical packages, and the marginal cost of computing weighted estimates is negligible.

The next set of results relate to estimates of the precision of the weighted means reported in the first table. That is, Tables 2 and 3 present two different estimates of the standard errors of the mean number of Carnegie credits earned in all subjects over four years, and in 21 separate subject fields, including academic subjects, vocational subjects, and other miscellaneous fields.

The first estimate in Table 2 is a jackknife estimate, using replicated subsamples, of the standard errors of average number credits earned in the same subjects, for the entire sample, for young men, and for young women. The second estimate is an ordinary unreplicated estimate of standard errors, for the same subjects and population groups. Weights were used for each type of estimate. The two procedures produced quite different sets of estimates: the simple standard errors, even though weighted, are smaller than the jackknifed standard errors by a average factor of 4.6 for the total sample, 3.5 for the young men, and 3.6 for the young women. This is a substantial difference, where the simple standard errors would be quite misleading.

Table 3 contains the same estimates for several racial and ethnic groups: whites, blacks, Hispanics, Asians, and others. Again, the simple estimates are quite misleading: even though weighted, the simple standard errors are smaller than the jackknifed standard errors by a factor of 4.5 for whites, 2.8 for blacks, 2.5 for Hispanics, 2.4 for Asians, and 1.3 for other racial or ethnic groups. This is not a matter of a few percentage points, but a few hundred percentage points. The simple standard errors would be quite misleading.

Compared to the problem of estimating means, the situation is reversed: The bias due to not using an appropriate method to deal with the effect of clustering on standard errors is enormous in this survey, yet the marginal cost of computing weighted estimates is large. These estimates were prepared on a mainframe computer, and the jackknifed results took more than eight times the amount of time to execute than did the simple results. Furthermore, the procedures for obtaining estimates of standard errors that adequately deal with a clustered design are not readily available in common statistical packages. The major statistical packages--such as SAS, BMDP, and SPSS-X--do not have fully-supported procedures for properly estimating sampling variances (for a survey of what is available, see Lee, Forthofer, and Lorimer, 1989). The jackknifed estimates in Tables 2 and 3 were prepared using WESVAR, a user-supported SAS procedure written at Westat, Inc.

Comparisons like these are so well known to sampling statisticians that they have quantified the efficiency of a sample design by estimating the "design effect." The design effect is the ratio of the actual variance of a clustered sample of neighbors to the variance of a simple random sample of people in which proximity was not used as a sample selection criterion and which is composed of the same number of elements. In other words, the design effect in a given comparison is the square of the ratio of the standard errors of the two estimates. In overall terms, the design effect is a function of two quantities --the intercluster correlation and the average cluster size-- according to the following relationship:

$$\text{design effect} = 1 + \text{roh}\ (\text{cluster size} - 1),$$

where roh is the intra-cluster correlation (a measure of similarity among students within schools). The cluster size plays a role because the larger the cluster size, the larger the proportion of the sample that each within-cluster correlation can affect.

Because the average cluster size and the average design effect are known, it is possible to derive an estimate of the intra-cluster correlation, by substituting known quantities into the above equation. The following table summarizes the sample sizes, the design effects, and the average cluster sizes, and shows the implied within-cluster correlations for the major demographic groups in the 1987 High School Transcript Study:

	Total	Male	Female	White	Black	Hispa	Asian	Other
Sample size	21446	10245	11201	15008	2869	2440	824	305
Average cluster size	53.36	26.25	28.33	38.61	11.86	10.34	5.18	3.11
Average ratio of jackknifed standard errors to ordinary	4.61	3.51	3.62	4.47	2.82	2.54	2.39	1.32
Average design effect (squared ratios)	24.90	14.63	15.27	22.56	8.84	7.90	7.11	2.13
Intra-cluster correlation	0.46	0.54	0.52	0.57	0.72	0.74	1.46	0.54

These relationships hold in general, but the relationship breaks down for the Asian subsample, where the sample is so skewed that the overall relationship doesn't hold. About two-thirds of the high schools had no Asians in the sample, while a quarter of the remainder had only one Asian. Only a quarter of the remainder had more than ten, yet eight schools had 25 or more Asians, and three had more than 40. The distribution is so skewed that the average is not a good measure of the distribution of cluster sizes. The general relationship implies a nonsensical intra-cluster correlation for Asians of 1.46, indicating that the relationship does not hold in this case.

The average design effects shown in this table are much higher than usually seen in most surveys. Design effects on the order of 2 or 3 are much more common. The design effects are very high in this case because the outcome measure is something upon which schools have a great impact: course-taking patterns are strongly influenced by school policies with respect to course offerings and graduation requirements, as well as by the socioeconomic resources of the community. The cluster sizes are also larger than in most household surveys and most of our student assessment surveys.

Results from the High School and Beyond study can help to put these results in context. The design effect reported for several achievement test measures was 5.2, with a cluster size of 29 students per high school (Tourangeau, et al., 1983). These figures imply an intra-cluster correlation of .15 for students from the same high school. That the estimate of intra-cluster correlation is much smaller in the High School and Beyond study is understandable, since the distribution of test scores is much less subject to school policy decisions than is course-taking.

DISCUSSION

The exercise of comparing standard errors produced by ordinary and replicated methods leads to several conclusions. First, the combined impact of the large cluster sizes and large intra-cluster correlations is so large that one should not try to estimate the standard errors from this survey without adopting one of the replication methods. Ordinary standard errors will simply be too far off the mark to allow more approximate methods to be acceptable.

Second, because large within-cluster correlations exist, it is likely that further work using a modeling approach is likely to be fruitful. While the replication approach used above doesn't tell an analyst much about the sources of the similarity, it does show that there is a fairly large amount of systematic similarity within schools. Further analysis could investigate the characteristics of schools and communities bring about the similarities among their students. In the 1987 High School Transcript Study, a number of variables are available describing of school characteristics and policies, such as: graduation requirements (overall and in English, mathematics, science, computer science, social studies, and foreign languages), socioeconomic indicators for the school population (number of dropouts, participants in the Federal school lunch program, participants in English as a second language programs, and participants in special education for the handicapped programs), school size, region, and degree of urbanism.

The hierarchical linear model takes the following approach: the average gender and racial/ethnic differences in course taking shown in Table 1 are modeled within clusters, but allowed to differ randomly from school to school. The across-school gender and racial/ethnic differences are modeled as a function of school characteristics, although some part of these differences remain unexplained and randomly different. The within-school gender and racial/ethnic differences are often based on sample sizes too small to estimate, so rather than estimating the differences individually, the HLM approach estimates parameters of the distribution (the mean and standard deviation) of the differences. The convergence of the sample design and the modeling approach occurs when enough data is available to incorporate sample selection probabilities as well as a variety of cluster characteristics. Both are iterative, and much more computationally intensive than the simple but inadequate approach (Pfefferman and Smith, 1985).

REFERENCES

J. Burke, J. Braden, M. Hansen, J. Lago, and B. Teppi.g. Final Report: National Assessment of Educational Progress--17th Year: Sampling and Weighting Procedures. Rockville: Westat, Inc. ober, 1987.

Andrew Kolstad. "Changes in High School Course Work from 1982 to 1987: evidence from two national surveys." AERA meetings, San Francisco, March 1989.

Eun Sul Lee, Ronald N. Forthofer, and Ronald J. Larimor. "Analyzing Complex Survey Data." Sage University Paper series on Quantitative Applications in the Social Sciences, 07-071. Beverly Hills: Sage Publications, 1989.

D. Pfeffermann and L. LaVange. "Regression models for stratified multi-stage cluster samples." Chapter 12 in C.J. Skinner, D. Holt, and T.M.F. Smith (eds.), Analysis of Complex Surveys, New York: John Wiley, 1989.

D. Pfeffermann and T.M.F. Smith. "Regression models for grouped populations in cross-section surveys." International Statistical Review. 53: 37-59. 1985.

Keith Rust. "Variance estimation for complex estimators in sample surveys." Journal of Official Statistics 1: 381-397, 1985.

J. Thorne, K. Rust, J. Burke, R. Marshall, N. Caldwell, D. Sickles, P. Ha, and B. Hayward. Technical Report: High School Transcript Study, 1987 National Center for Education Statistics, CS 89-447.

J. Thorne, K. Rust, J. Burke, R. Marshall, N. Caldwell, D. Sickles, P. Ha, and B. Hayward. 1987 High School Transcript Study: Data File User's Manual. National Center for Education Statistics, February 1989.

R. Tourangeau, M. McWilliams, C. Jones, M. Frankel, and F. O'Brien. High School and Beyond First Follow-up (1982) Sample Design Report. National Center for Education Statistics, 1983.

J. Tuma, A. Gifford, "Changes in high school course work among non-college-bound graduates, 1969-1987." AERA meetings, Boston, April 1990.

Table 1.--Average number of credits earned in selected major subject fields, total and by gender and race/ethnicity: 1987.

Subject Field	Total	Gender		Race/ethnicity				
		Male	Female	White	Black	Hispanic	Asian	Other
All subjects	23.01	22.88	23.13	23.06	22.54	22.87	24.51	23.18
English	4.03	4.01	4.05	3.99	4.14	4.23	4.31	4.20
History	1.90	1.92	1.88	1.88	1.88	1.78	1.97	1.99
Social studies other than history	1.43	1.39	1.47	1.42	1.43	1.45	1.67	1.26
Mathematics	2.97	3.03	2.92	2.98	2.90	2.77	3.72	2.96
Computer science, programming, and data processing	0.43	0.47	0.40	0.45	0.35	0.36	0.57	0.35
Science	2.59	2.66	2.53	2.64	2.39	2.33	3.17	2.51
Foreign languages	1.46	1.29	1.63	1.50	1.12	1.27	2.17	0.92
Non-occupationally specific vocational education	1.64	1.61	1.67	1.66	1.83	1.64	1.01	1.90
Occupationally specific vocational education								
General introductory vocational education	0.34	0.31	0.37	0.33	0.44	0.30	0.20	0.42
Agriculture	0.17	0.28	0.06	0.20	0.09	0.06	0.01	0.21
Business	0.68	0.34	1.01	0.69	0.74	0.70	0.44	0.64
Marketing and distribution	0.10	0.07	0.12	0.10	0.11	0.11	0.08	0.06
Health	0.05	0.03	0.07	0.04	0.09	0.05	0.03	0.05
Occupational home economics	0.10	0.05	0.15	0.09	0.19	0.09	0.05	0.10
Trade and industry	0.56	0.96	0.18	0.57	0.50	0.62	0.25	0.72
Technical	0.01	0.02	0.01	0.01	0.02	0.00	0.01	0.01
Visual and performing arts	1.43	1.24	1.60	1.48	1.20	1.35	1.12	1.51
Physical education, sports, and health	1.97	2.13	1.81	1.94	2.01	2.40	2.57	2.12
Other personal and social	0.77	0.65	0.89	0.76	0.69	0.95	0.95	0.97
Religion and theology	0.25	0.27	0.23	0.27	0.12	0.13	0.12	0.11
All courses other than above	0.12	0.15	0.09	0.08	0.30	0.29	0.07	0.17

NOTE: Credits measured in Carnegie units (a unit is defined as a class that meets for one hour five days per week throughout an academic year).

SOURCES: U.S. Department of Education, National Center for Education Statistics, special tabulation prepared under contract by Westat, Inc., Rockville, Maryland, from the 1987 High School Transcript Study.

Table 2.--Standard error of average number of credits earned by high school graduates in selected major subject fields, total and by gender: 1987.

Subject Field	Jackknifed standard errors			Ordinary standard errors			Ratio of jackknifed to ordinary		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
All subjects	0.1553	0.1642	0.1555	' 0'98	0.0297	0.0263	7.84	5.53	5.91
English	0.0181	0.0182	0.0208	0.0058	0.0085	0.0078	3.12	2.14	2.67
History	0.0225	0.0282	0.0214	0.0048	0.0071	0.0064	4.69	3.97	3.34
Social studies other than history	0.0481	0.0517	0.0489	0.0056	0.0079	0.0078	8.59	6.54	6.27
Mathematics	0.0301	0.0345	0.0292	0.0068	0.0102	0.0091	4.43	3.38	3.21
Computer science, programming, and data processing	0.0177	0.0209	0.0168	0.0043	0.0065	0.0055	4.12	3.22	3.05
Science	0.0461	0.0467	0.0484	0.0073	0.0111	0.0096	6.32	4.21	5.04
Foreign languages	0.0510	0.0520	0.0557	0.0092	0.0127	0.0130	5.54	4.09	4.28
Non-occupationally specif vocational education	0.0512	0.0518	0.0573	0.0097	0.0144	0.0131	5.28	3.60	4.37
Occupationally specific:									
General introductory vocational education	0.0189	0.0192	0.0203	0.0051	0.0073	0.0071	3.71	2.63	2.86
Agriculture	0.0274	0.0462	0.0119	0.0053	0.0099	0.0041	5.17	4.67	2.90
Business	0.0295	0.0193	0.0439	0.0083	0.0068	0.0138	3.55	2.84	3.18
Marketing	0.0093	0.0071	0.0153	0.0035	0.0046	0.0052	2.66	1.54	2.94
Health	0.0052	0.0048	0.0084	0.0027	0.0025	0.0046	1.93	1.92	1.83
Occupational home economics	0.0080	0.0057	0.0125	0.0035	0.0037	0.0057	2.29	1.54	2.19
Trade and industry	0.0350	0.0606	0.0181	0.0092	0.0169	0.0069	3.80	3.59	2.62
Technical	0.0026	0.0040	0.0015	0.0010	0.0017	0.0011	2.60	2.35	1.36
Visual and performing arts	.0429	0.0446	0.0474	0.0111	0.0153	0.0158	3.86	2.92	3.00
Physical education, sports and health	.0654	0.0696	0.0658	0.0083	0.0128	0.0106	7.88	5.44	6.21
Personal/social	0.0304	0.0230	0.0398	0.0063	0.0079	0.0097	4.83	2.91	4.10
Religion/theology	0.0421	0.0580	0.0530	0.0059	0.0085	0.0083	7.14	6.82	6.39
All courses other than above	0.0091	0.0107	0.0098	0.0044	0.0073	0.0052	2.07	1.47	1.88

NOTE: Credits measured in Carnegie units (a unit is defined as a class that meets for one hour five days per week throughout an academic year).

SOURCE: U.S. Department of Education, National Center for Education Statistics, special tabulation prepared under contract by Westat, Inc., Rockville, Maryland, from the 1987 High School Transcript Study.

Table 3.--Standard error of mean number of credits earned by high school graduates in selected major subject fields, by race/ethnicity: 1987 and 1982.

Subject Field	Jackknifed standard errors					Ordinary standard errors					Ratio of jackknifed to ordinary				
	White	Black	Hispa	Asian	Other	White	Black	Hispa	Asian	Other	White	Black	Hispa	Asian	Other
English	.0225	.0486	.0548	.0575	.0326	.0061	.0175	.0221	.0401	.0480	3.70	2.78	2.48	1.43	0.68
History	.0290	.0607	.0396	.0684	.0374	.0057	.0147	.0123	.0218	.0370	5.09	4.13	3.22	3.14	1.01
Social studies other than history	.0471	.05 1	.0995	.1926	.0558	.0067	.0161	.0150	.0337	.0371	7.03	3.17	6.63	5.72	1.50
Mathematics	.0374	.0551	.0471	.0935	.1049	.0082	.0181	.0183	.0339	.0484	4.56	3.04	2.57	2.76	2.17
Computer science, programming, and data processing	.0219	.0219	.0203	.0408	.0401	.0051	.0116	.0118	.0234	.0335	4.29	1.89	1.72	1.74	1.20
Science	.0575	.0726	.0619	.1087	.0747	.0089	.0184	.0175	.0439	.0533	6.46	3.95	3.54	2.48	1.40
Foreign languages	.0635	.0743	.0588	.1194	.1723	.0113	.0219	.0237	.0458	.0633	5.62	3.39	2.48	2.61	2.72
Non-occupationally specific voc education	.0664	.0498	.0448	.1104	.0696	.0119	.0258	.0257	.0407	.0790	5.58	1.93	1.74	2.71	0.88
Occupationally specific:															
General intro voc education	.0223	.0503	.0397	.0438	.0809	.0060	.0163	.0150	.0190	.0379	3.72	3.09	2.65	3.36	2.13
Agriculture	.0381	.0160	.0217	.0068	.0217	.0071	.0086	.0083	.0080	.0406	5.37	1.86	2.61	0.85	0.53
Business	.0361	.0478	.0320	.0625	.0524	.0099	.0239	.0245	.0305	.0693	3.65	2.00	1.31	2.05	0.76
Marketing	.0129	.0113	.0213	.0463	.0288	.0042	.0101	.0109	.0118	.0176	3.07	1.12	1.95	3.97	1.64
Health	.0056	.0240	.0118	.0094	.0055	.0031	.0103	.0060	.0124	.0199	1.81	2.33	1.97	0.76	0.28
Occupational home economics	.0088	.0312	.0096	.0239	.0443	.0037	.0143	.0096	.0092	.0219	2.38	2.18	1.00	2.60	2.02
Trade and industry	.0424	.0529	.0612	.0385	.0483	.0112	.0261	.0271	.0245	.0798	3.79	2.03	2.26	1.57	0.61
Technical	.0023	.0109	.0024	.0028	.0101	.0010	.0039	.0036	.0024	.0084	2.30	2.79	0.67	1.17	1.20
Visual and performing arts	.0562	.0560	.0561	.0843	.1111	.0138	.0263	.0299	.0436	.0889	4.07	2.13	1.88	1.93	1.25
Physical educatn sports, and healt	.0779	.1124	.0896	.1612	.0667	.0099	.0227	.0248	.0400	.0673	7.87	4.95	3.61	4.03	0.99
Personal/social	.0375	.0578	.0647	.0910	.0576	.0073	.0169	.0217	.0346	.0592	5.14	3.42	2.98	2.63	0.97
Religion/theology	.0467	.0495	.0445	.0406	.0642	.0077	.0118	.0131	.0228	.0285	6.0[illegible]	4.19	3.40	1.78	2.25
All courses other than above	.0098	.0518	.0494	.0132	.0476	.0042	.0177	.0183	.0161	.0328	2.33	2.93	2.70	0.82	1.45

NOTE: Credits measured in Carnegie units (a unit is defined as a class that meets for one hour five days per week throughout an academic year).

SOURCE: U.S. Department of Education, National Center for Education Statistics, special tabulation prepared under contract by Westat, Inc., Rockville, Maryland, from the 1987 High School Transcript Study.