Institute of Education Sciences
National Center for Education Statistics

NATIONAL INSTITUTE OF STATISTICAL SCIENCES
TECHNICAL REPORT

CONSTRUCTING FULL SAMPLE AND REPLICATE WEIGHTS FOR NAEP TEACHER DATA

TABLE OF CONTENTS

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

NAEP FULL SAMPLE AND REPLICATE WEIGHTS FROM TEACHERS

EXECUTIVE SUMMARY

Data weights are constructed in order for statistical analyses of data to correctly represent results presented on a national scale, to accurately reflect the composition of the national population and to provide estimated standard errors for all reported statistics. The goal of this study was to explore the feasibility and utility of constructing full sample and replicate weights for the set of teachers whose data is collected by the National Assessment of Educational Progress (NAEP).

Two sources of data for fourth grade mathematics teachers were compared with respect to national averages for selected teacher characteristics: NAEP (using the reconstructed weights) and the 2010–11 Schools and Staffing Survey (SASS).[1] The selected characteristics were both compared marginally and jointly, using estimated standard errors calculated employing NAEP replicate weights and SASS replicate weights.

There are two principal findings:

1. Using NAEP school weights as teacher weights, and with a straightforward, national calibration of the NAEP weights to the SASS weights, the two sets of national estimates for five teacher characteristics common to both the NAEP and SASS datasets are essentially indistinguishable. In other words, the procedure to create teacher weights from these two sources works.
2. In general, estimated NAEP standard errors are smaller than estimated SASS standard errors. We believe that this is largely the result of the larger sample size in NAEP.

The implications for the National Center for Education Statistics (NCES) will require careful investigation. If, as these results suggest, NAEP has the potential to be superior to SASS as a mechanism for collecting data about teachers, then the NAEP teacher questionnaire will require re-design, because NAEP collects only limited information about the teachers themselves. Further discussion appears in the full text of this report.

The potential for using NAEP to explore relationships between teacher characteristics and student performance is addressed only tangentially in this study. How such exploration *might* be done is illustrated in the full text using the NAEP weights as constructed here. However, it must be stressed that the results presented only address the question of whether teachers with different characteristics teach student populations with different characteristics.

[1] Such characteristics must, of course, be present in both data sets; see §2.3.

NAEP FULL SAMPLE AND REPLICATE WEIGHTS FOR TEACHER DATA

I. INTRODUCTION

The sample design for NAEP is built for sampling schools so that accurate inferences can be drawn about performance and attributes at the school level. NAEP does collect valuable teacher data as well. However, correct inferences about teacher characteristics require the use of different sampling weights, i.e., weights equivalent to those based on a design for drawing a sample of teachers, not schools. The objective of the research presented here is to find a way to construct full sample and replicate weights for use with NAEP teacher data, and then to evaluate the performance of these re-constructed weights. The approach taken is to start from the NAEP school weights and, using an auxiliary source for calibration, to revise the school weights to create teacher weights. Evaluation is based on the comparative performance of estimates (and standard errors) produced using the reconstructed NAEP weights and estimates (and standard errors) produced by the auxiliary source.

II. DATA AND TEACHER CHARACTERISTICS

This study focused on fourth grade mathematics teachers in 2010–11 (for SASS) and 2011 (for NAEP). Because most fourth grade teachers teach both mathematics and reading, there is no reason to suspect that the results would be different for reading. Whether essentially the same sampled populations of teachers arise for both NAEP and SASS is not completely clear; see §4 for further discussion. As described in §1.2, a crucial assumption is that (in effect), if a school is sampled for fourth grade NAEP, then all fourth grade teachers in that school are sampled. This assumption may not be valid for the eighth and twelfth grades, and in any event would need to be examined more closely.

2.1 Datasets

The data employed in this study were provided to the National Institute of Statistical Sciences (NISS) by NCES under an amendment to NISS' existing data license. The data employed in this study were derived from two datasets:

NAEP: The dataset M42NT1AT contains the 2011 fourth grade mathematics student and teacher data. This file is student-indexed, and contains 214,205 records for students from 8,505 schools. There are entries for 30,117 teachers. However, 3,278 of these contain no data, leaving 26,389 teachers.[1]

[1] If the absence of data reflects teacher nonresponse, then the nonresponse rate is 10.9%. According to NCES' Statistical Standards, a nonresponse bias analysis is not required. In any event, such an analysis is not possible, since NAEP does not collect frame information for teachers. Were NAEP to become a primary data collection vehicle for teachers, this point would merit more detailed consideration.

Because of special sampling procedures associated with them, teachers from Department of Defense and Bureau of Indian Affairs schools were dropped, leaving a final dataset size of 26,320.

The NAEP data files contain full sample and 62 sets of replicate weights for schools.

SASS: SASS files contain both full sample weights and 88 sets of replicate weights for teachers.

The SASS data contains, as separate files pubteall and priteall, data for 37,497 public school teachers and 4,523 private school teachers. These files were concatenated, and then, using the variable **T0075**, reduced to a file containing only teachers who taught fourth grade students, of which there are 4,368.

Based on guidance from Kathryn Chandler of NCES, the SASS data were filtered further by requiring that EITHER

- **T0098** ≥ 1, corresponding to teachers in self-contained classroom OR
- Any of (**T0110** = 1 AND **T0120** = 04), . . ., (**T0119** = 1 AND **T0129** = 04), corresponding to departmentalized teachers and elementary specialists.

The resultant dataset contains 1,956 teachers.

Table 1 shows counts of teachers and sums of full sample weights for the two datasets, broken by school control (public or private). The NAEP weights have been adjusted for teacher nonresponse. In theory, the sums of weights, which are national estimates of numbers of fourth grade mathematics teachers, should agree, which clearly they do not. How we dealt with this issue is described in §1.2.

During the study, we considered other filters for the SASS data:

1. The dataset of size 4,368 was selected on the basis of only **T0075** = 1. The sums of weights are 584,808.7699 for public schools and 103,343.0468 for private schools.[2]
2. Keith Rust of Westat proposed the filter **T0075** = 1 AND **T0090** = 102, which produces weights sums of 191,846.5268 for public schools and 34,273.8863 for private schools.

The latter matches NAEP weights better than the NCES-provided filter, but we employed the latter because of the expert knowledge on which it is based. See further discussion in §4.

2.2 Construction of the NAEP Teacher Weights

Reflecting previous AIR/NISS formulations of the problem, discussions with NCES personnel, and the "Westat Memo,"[3] we adopted the "School Weight Approach." That is, we assume that all teachers in each relevant grade in each NAEP-sampled school are in the teacher sample, and initially, assign each teacher a weight equal to the school full sample weight. The alternative, "Student Weight Approach" discussed in §4 may be necessary for the eighth and twelfth grades.

Because of the mismatch between totals of NAEP weights and SASS weights shown in Table 1, we calibrated the NAEP full sample and replicate weights to have the same totals for public and private schools as the SASS weights. The rationale for treating SASS rather than NAEP as "correct" is that SASS is designed to be

[2] An earlier version of this report, dated June 17, 2014, employed this filter.

[3] By this we mean the memo dated July 8, 2013 from Keith Rust of Westat to William Ward of NCES, with the subject "Analysis of Teacher Data in NAEP."

nationally representative. We emphasize, however, that SASS is not—we believe—designed to be nationally representative of fourth grade mathematics teachers. The calibration factors are:

For private schools: 39,821.5980 / 22,924.8401 = 1.7371;

For public schools: 293,825.9769 / 179,654.9696 = 1.6355.

These factors were applied to NAEP full sample and replicate weights.

There is sufficient information in the datasets to have done the calibration at the state level. Indeed, in an earlier version of the study, that is what we did. Because in that case there was no material difference between the two cases, we employed the simpler approach here.

2.3 Teacher Variables

There is limited overlap between the sets of teacher variables in NAEP and SASS. We employed five variables, which effectively are the entire overlap. These are:

Race, recoded to have values Black, White and Other. (Original values were American Indian/Alaskan native, Asian, black, Hawaiian native/Pacific Islander, white and multiple races.)

Hispanic ethnicity, with values Yes and No.

Highest degree, with original values 1 (associate degree or less), 2 (Bachelors degree), 3 (Masters degree), 4 (advanced graduate study) and 5 (doctorate or professional degree), and recoded to have the values ≤ Bach[elors degree] and > Bach[elors degree].

Certification, a categorical variable originally with the values 1 (regular), 2 (except for probationary period), 3 (additional coursework needed), 4 (certification program in order to continue teaching) and 5 (none), and recoded to have the values Regular and Other.

Years of teaching experience, an integer-valued variable, recoded to ranges of 1–5, 6–10, 11–20, 21–30 and 31+.[4]

There are fourteen categories when each variable is considered separately, and 120 for the variable that results from crossing them. i.e., the full contingency table.

Table 2 shows the exact variables NAEP and SASS variables we have employed. Even though questionnaire wording is not identical between NAEP and SASS, mapping of concordant responses is completely straightforward.

III. RESULTS

Figure 1 shows the distributions of the calibrated NAEP weights (n = 26,320) and SASS weights (n = 1,986).

Statistical comparisons between national count estimates derived from NAEP and national count estimates derived from SASS are made of the basis of Z-statistics of the form

[4] In particular, the recoding removes a possible inconsistency between how NAEP and SASS treat teachers who are in their first year.

$$Z = \frac{|\mathrm{EST}_{\mathrm{NAEP}} - \mathrm{EST}_{\mathrm{SASS}}|}{\sqrt{\widehat{\mathrm{SE}}^2_{\mathrm{NAEP}} + \widehat{\mathrm{SE}}^2_{\mathrm{SASS}}}}, \quad (1)$$

where

- $\mathrm{EST}_{\mathrm{NAEP}}$ is the NAEP estimate, for example, of the number of white teachers in the U.S., calculated using the NAEP data and the calibrated NAEP full sample weights.
- $\widehat{\mathrm{SE}}_{\mathrm{NAEP}}$ is the estimated standard error of $\mathrm{EST}_{\mathrm{NAEP}}$, calculated using the 62 sets of NAEP replicate weights.
- $\mathrm{EST}_{\mathrm{SASS}}$ is the SASS estimate, calculated using the SASS data and the SASS full sample weights.
- $\mathrm{SE}_{\mathrm{SASS}}$ is the estimated standard error of $\mathrm{EST}_{\mathrm{SASS}}$, calculated using the 88 sets of SASS replicate weights.

Under the assumption that NAEP and SASS are independent, which we believe is plausible, and the null hypothesis that there is no difference between distributions of the NAEP data and the SASS data, Z as defined in (1) has approximately the distribution of $|Z^*|$, where Z^* is normally distributed with mean 0 and variance 1. In Appendix A, we discuss an alternative approach using T-statistics. As elucidated there, we believe that this approach is not appropriate.

3.1 Five Variables Individually

Table 3 contains the results of testing whether the Z-statistic defined by (1) is non-zero for each category of each of the five variables defined in S2.3. The total number of tests is 14.

The column headings in Table 3 are nearly self-explanatory. From left to right, they are: the variable; the category; the NAEP-estimated count and NAEP-estimated standard error; the SASS-estimated count and SASS-estimated standard error; the Z-statistic and the associated p-value; and whether the p-value is significant at the level .05, using the false discovery rate (FDR) method (Benjamini and Hochberg, 1995).

Even without multiplicity adjustment[5], at the .05 level, none of the 14 tests is significant. The same is true with the FDR adjustment.

For additional insight, Table 5 contains the same estimates as in Table 3, but of proportions rather than counts. Tests of significance have been omitted because they are the same as for counts. The final column in Table 5 is the relative error

$$\frac{|\mathrm{EST}_{\mathrm{NAEP}} - \mathrm{EST}_{\mathrm{SASS}}|}{½[\mathrm{EST}_{\mathrm{NAEP}} + \mathrm{EST}_{\mathrm{SASS}}]}. \quad (2)$$

We conclude that the NAEP estimates and SASS estimates are, for practical purposes, statistically indistinguishable.

[5] As recommended by the NCES statistical standards.

3.2 Five Variables Jointly

Table 4 contains the same information as Table 3, but for the 120-category variable created by fully crossing the reduced versions of the five variables, constructed as described in §1.3. Only 71 of the 120 categories appear in this table; for the others only one of NAEP and SASS contained data, or else NAEP estimated standard errors were unstable because of zero counts in complementary VPSUs. Even without adjustment for multiplicity, no differences are significant!

Figure 2 shows the results of regressing the full set of 85 NAEP-estimated counts on the same SASS-estimated counts. The fit is nearly perfect: $r^2 = .9974$, the root mean square error (RMSE) is 3140, there is no visible structure to the residuals, and the slope estimate of .9909 is nearly equal to one. However, there is no clear explanation for the estimated intercept of 140.8, although for categories with large counts, this is ignorable. Forcing the intercept to be zero produces a slope estimate of 1.006 with a standard error of 0.005, and an immaterially increased root mean square error of 3147.

The message remains clear: the NAEP-generated estimates and SASS-generated estimates do not differ statistically.

3.3 Comparison of Standard Errors

Table 6 contains the ratios of the NAEP-estimated standard errors to the SASS-estimated standard errors, for all 85 cases appearing in Tables 3 and 4. These ratios range from 0.109 to 44, with a median value of 3.766. This latter value is approximately equal to the reduction in standard errors that would be expected on the basis of sample size alone, which is 3.640 = $\sqrt{26320/1986}$. More complete distributional information is contained in Figure 3.

IV. ANALYSES USING THE NAEP DATA

In this section, we illustrate that the NAEP teacher weights we have constructed can be used to perform analyses involving items collected by NAEP but not by SASS.

4.1 NAEP-Specific Variables

Table 7 is a straightforward application of the NAEP weights to estimation of counts and standard errors for four variables present only in NAEP:

- Class size;
- Hours per week of mathematics instruction;
- Access to computers;
- Availability of resources.

In the mathematics section of the 2011 NAEP Teacher Questionnaire National Center for Education Statistics (2011), these correspond to questions 1, 3, 13 and 15, respectively. There is nothing especially notable about the results.

4.2 Relating Teacher and Student Characteristics

We begin with the same disclaimer that appears in Executive Summary. The material here is descriptive, and meant only to exemplify how availability of teacher weights for NAEP might enable exploration of relationships between teacher characteristics—the five variables from §1.3—and student performance. The comparisons that appear below address the question of *whether teachers with differing characteristics teach student populations with differing characteristics*. They are not analyses of the performance of individual students. Consequently, it is appropriate to weight them using the NAEP full sample teacher weights, because the unit of analysis is teachers. No causality should be inferred from these analyses. Indeed, in several cases, two oppositely directed causal relationships are equally plausible *a priori*.

The student performance variable, at the teacher level, is mean, over all pupils associated with each teacher, of the NAEP variable MPSTM1. The latter, student-level, variable is the mean of the posterior distribution from which the plausible values associated with the sub-score "numbers and operations" are drawn. For more refined analyses, it would make sense to use plausible values themselves, which also account for measurement error.

We interpret the mean, over all students taught by a teacher, of these posterior means, as a descriptor of the population of students taught by that teachers.

The analyses were performed using the "Fit Y by X" functionality of SAS® JMP®. The same analyses can be carried out for the four other posterior means—MPSTM2 (measurement), MPSTM3 (geometry), MPSTM4 (data analysis and probability) and MPSTM5 (algebra).

Figures 4–8 contain the results. In each, there are:

- Box plots and quantiles of the student performance measure for each category of the variable;
- Plots of the associated cumulative distribution functions (CDFs);
- A nonparametric assessment of the statistical significance of the differences between or among the distributions. For binary variables (Hispanic ethnicity, highest degree and certification), a Kolmogorov-Smirnov test is employed. For variables with more than two categories, all pairwise comparisons were made using a Wilcoxon test.

There is no need to discuss each set of results. Illustratively, for race (Figure 4) and as measured by MPSTM1, white teachers have, on the average, higher-performing students than teachers "of all other races," who in turn have higher-performing students than students taught by black teachers. All differences are highly significant. The other analyses, which correspond to Figures 5–8, can be interpreted similarly.

V. DISCUSSION

The preceding sections show that, as long as the calibration step in §1.2 is performed, construction of full sample and replicate weights for teachers sampled by NAEP is both simple and effective, at least for the fourth grade. For variables common to NAEP and SASS, standard errors calculated using NAEP are smaller—principally, we believe, because of the larger sample size.

These findings suggest that for the purpose of collecting information about teachers, NCES may wish to consider use of NAEP as a supplement to, or even replacement for, SASS. To be sure, such a decision cannot

be taken lightly. Some factors and issues that NCES may wish and/or need to consider if such a path were pursued are discussed next.

1. Without question, the most important issue is the inability to "match" the NAEP and SASS weights, which necessitates calibration of the former. It is not completely clear how close the match "should" be, but it seems—at least without better understanding of the reasons, that it should be better than in Table 1. Were NAEP considered as a replacement for SASS for collection of data about teachers and if calibration were still necessary, it would have to be based on an alternative data source.

   Whether the assumption that if a school is selected for fourth grade NAEP mathematics, then all fourth grade teachers in that school are selected, may also affect this issue. (In our analyses, this assumption was operationalized by assigning full sample and replicate *school* weights to teachers.) However, as Table 1 indicates, weighted NAEP counts are lower than SASS counts, which would not arise from including too many NAEP teachers. Table 1 also indicates that issue is more severe in private schools than in public schools, which those with specialized knowledge may be able to explicate.

2. It is not clear that the "School Weight Approach" is applicable for eighth and twelfth graders, where teachers are subject-specific. Nor, however, is it obvious that it is necessary to use the "Student Weight Approach" of calculating explicitly the probability that a teacher is sampled as the probability that one or more of the students he or she teachers is sampled, using the student weights as inverse probabilities of selection.[6] The reason "it is not clear" is the weight calibration step requires only that using school weights as initial teacher weights be relatively—not absolutely—correct. What is clear is that for other grades the target populations of teachers will need to be defined very carefully.

   We stress that this issue may also present even for fourth graders. While coverage for mathematics and reading teachers may be acceptable, NAEP does not sample art or music teachers who do not teach reading and mathematics.

3. Collection of many SASS data elements by means of NAEP may be inefficient or lead to problems with nonresponse. Other than the five variables used throughout this study, current NAEP teacher data pertain mainly to classroom practices, teacher roles and professional development. Collecting data such as salaries via NAEP, at least using teacher-completed questionnaires, may result in substantial measurement error and item nonresponse. The 2011 grade 4 nonresponse rate of 10.9% may not be a problem, but there may be problems with more, and more sensitive, items. Collecting some teacher data directly from schools may be an effective or efficient alternative.

4. The material in §3.2 notwithstanding, the ability of NAEP to support principled analyses of relationships between teacher characteristics and student performance is untested. NCES may wish to convene a Technical Expert Panel (TEP) to address this question.

[6] This process, which the Westat memo terms an "elaborate calculation," is in fact completely straightforward, However, it does entail assumptions regarding how students are assigned to teachers, and also requires knowledge of how many teachers, by grade and subject, there are in each NAEP-sampled school. This latter information does not seem to be routinely collected by NAEP, and so an alternative source, such as the CCD or PSS, would be necessary.

APPENDICES

Appendix A: References

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A New and Powerful Approach to Multiple Testing. *J. Royal Statist. Soc. Series B*, 57:1289–1300.

National Center for Education Statistics (2011). NAEP Reading and Mathematics Teacher Questionnaire: 2011–Grade 4. Available on-line at https://nces.ed.gov/nationsreportcard/pdf/bgq/teacher/BQ11_Teacher_RM_G04.pdf.

Appendix B: T–Statistic Analysis

It is natural to ask whether the Z-statistic-based approach employed in §2 is the proper one, given the point of using replicate weights is to estimate standard errors for means. We believe that the Z-statistic approach is in fact the more appropriate, for reasons discussed momentarily.

For completeness, Tables 9 and 10 contains the same information presented in Tables 3 and 4, respectively, but with significance tested using a standard independent-two-sample T statistic, with replicate weight-based estimated standard errors substituted for sample variances. The calculated degrees of freedom assume a "sample size" of 62 for NAEP and 88 for SASS. As is clear from these two tables, there is only the single randomization associated with the replicate weights, leading to the same value of the T-statistic for all tests, and therefore to the same—and not significant—value of the T-statistic.

The salient point, we believe, is that T-statistics are suited for comparing population means when population variances are available. Standard errors estimated using replicate weights are not estimates of population standard deviations; rather they are estimates of the standard deviations of population means. Therefore, we feel that the Z-statistic-based analysis in §2 is not only more appropriate, but also more informative.

In any event, however, the message is exactly as in §2: NAEP-based and SASS-based do not differ significantly.

Appendix C: Figures and Tables

B.1 Tables for §2

Table 1: Teacher counts and sums of full sample weights for the NAEP and SASS datasets used in this study.

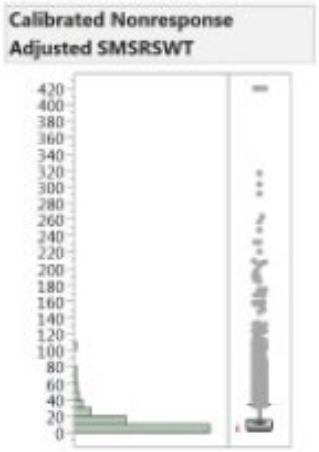
	NAEP		SASS	
Control	N(Teachers)	Σ(Weights)	N(Teachers)	Σ(Weights)
Public	25,586	179,654.9696	1,533	293,825.9769
Private	734	22,924.8401	423	39,821.5980
TOTAL	1,956	202,579.8097	26,320	333,647.5749

Table 2: NAEP and SASS variables used in the study.

Variable	NAEP Data Variable	SASS Variable
Race	TE21201–TE21205	T0528–T0532
Hispanic ethnicity	TA21101–TE21101	T0527
Certification	T096501	T0250
Highest degree	T056301	T0160
Years of teaching experience	T077101	TOTYREXP

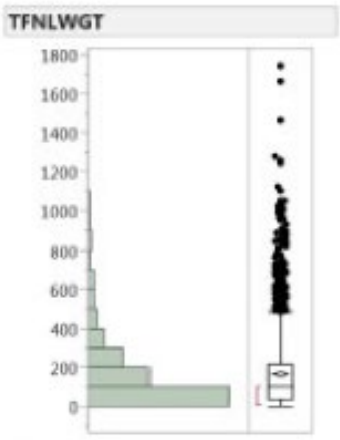
B.2 Figures and Tables for §3

Figure 1: Distributions of NAEP (left) and SASS (right) full sample weights.



Quantiles		
100.0%	maximum	419.859
99.5%		108.28
97.5%		57.1066
90.0%		27.4188
75.0%	quartile	14.9585
50.0%	median	7.40973
25.0%	quartile	3.1254
10.0%		1.94036
2.5%		1.80462
0.5%		1.78689
0.0%	minimum	1.75154

Summary Statistics	
Mean	12.676579
Std Dev	17.54179
Std Err Mean	0.1081262
Upper 95% Mea	12.888513
Lower 95% Mean	12.464646
N	26320



Quantiles		
100.0%	maximum	1745.42
99.5%		1053.06
97.5%		740.208
90.0%		400.745
75.0%	quartile	216.157
50.0%	median	98.5998
25.0%	quartile	37.9208
10.0%		18.5852
2.5%		8.76072
0.5%		4.4523
0.0%	minimum	2.7774

Summary Statistics	
Mean	167.99979
Std Dev	197.8791
Std Err Mean	4.4402795
Upper 95% Mea	176.70788
Lower 95% Mean	159.29169
N	1986

Table 3: Results of statistical comparisons of NAEP and SASS estimates for the five variables of §1.3 marginally. See text for discussion.

	Variable	Category	NAEP Count	NAEP Standard Error	SASS Count	SASS Standard Error	Z Statistic	p-Value	FDR Significant? (.05)
1	Race Recoded	White	290320.859	5180.0379538	294187.775	22570.0309	0.1669880933	0.8673794129	No
2	Race Recoded	Black	23053.6438	1536.4734478	23118.9751	11776.7676	0.005500854	0.9956109756	No
3	Race Recoded	Other	20273.0659	2101.4366387	16340.8244	7640.72714	0.4962169594	0.6197413468	No
4	Hispanic Recoded	No	308195.169	5704.0669101	309622.181	21670.47	0.0636814182	0.9492239009	No
5	Hispanic Recoded	Yes	25452.4003	2610.6319405	24025.3944	8560.67882	0.1594438649	0.8733191817	No
6	Highest Degree Recorded	> Bach	174914.859	4232.9351923	189792.039	19057.2283	0.7620853163	0.446009086	No
7	Highest Degree Recorded	< = Bach	158732.71	4464.7610943	143855.536	18639.1444	0.7762102164	0.4376248748	No
8	Certification Recoded	Regular	300009.321	4869.6206863	302755.519	23739.0243	0.1133231289	0.9097743812	No
9	Certification Recoded	Other	33638.2475	3126.2380785	30892.0563	10919.1494	0.2417875441	0.808944791	No
10	Years Teaching Certification Recoded	1 – 5	76640.3508	3929.9244087	70880.8918	14034.1303	0.3951875813	0.692704457	No
11	Years Teaching Certification Recoded	6 – 10	76795.5978	2869.9899725	83098.1219	15544.2863	0.3987169402	0.6901017843	No
12	Years Teaching Certification Recoded	11 – 20	107775.084	4513.2051088	103018.275	16985.1533	0.2706647446	0.7866488947	No
13	Years Teaching Certification Recoded	21 – 30	49731.61	3019.2409886	51720.5774	10099.4749	0.1886865309	0.8503385087	No
14	Years Teaching Certification Recoded	31 +	22704.9266	1671.3183287	24929.7085	7565.42035	0.2871489794	0.7739982504	No

Table 4: NAEP- and SASS-estimated counts and errors for (71 categories) of the 120-category variable obtained by crossing the five variables defined in §2.3.

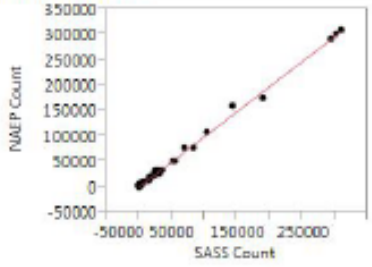
	Category	NAEP Count	NAEP StandardError	SASS Count	SASS StandardError	Z Statistic	p-Value	FDR Significant? (.05)
1	Black_No_Regular_>Bach_31+	837.981957	236.12604734	2569.48403	4673.6406	0.370010566	0.7113746177	No
2	Black_No_Other_<=Bach_1-5	999.561393	430.54844347	95.970286	15308.2902	0.0586583365	0.9532242448	No
3	White_No_Other_<=Bach_11-20	3164.73809	923.44579309	856.474108	15538.4267	0.1484807734	0.8819633584	No
4	Black_No_Regular_<=Bach_1-5	2273.10727	540.47806582	2727.59531	6205.80478	0.0729597757	0.9418381265	No
5	White_No_Other_>Bach_1-5	3026.338	863.20475884	2817.29292	18611.9379	0.0112197133	0.991048152	No
6	White_No_Regular_>Bach_1-5	15744.2755	1505.0837253	18296.8792	7387.5513	0.3385725999	0.7349317268	No
7	White_No_Regular_>Bach_6-10	31619.0684	2262.0557222	35316.3379	9320.0249	0.3855093936	0.6998600482	No
8	White_No_Regular_>Bach_11-20	51594.4395	2672.6133274	56221.1016	12073.9522	0.3741373956	0.7083020681	No
9	White_No_Regular_<=Bach_6-10	24833.5152	2417.0487719	25142.5009	7291.75455	0.0402224847	0.9679157519	No
10	Black_No_Regular_<=Bach_11-20	1793.20034	1351.0883261	1281.16791	3103.01443	0.1512921034	0.8797452982	No
11	White_No_Regular_<=Bach_1-5	32066.8474	1947.0755096	25621.464	7496.5593	0.8321682039	0.4053140085	No
12	White_No_Other_<=Bach_1-5	9103.33663	1809.303545	7573.17477	14861.9916	0.1022075504	0.9185019348	No
13	White_Yes_Other_<=Bach_1-5	957.994988	501.61173459	829.34098	20489.3846	0.0062771758	0.9949915713	No
14	Black_No_Regular_<=Bach_6-10	1418.24286	1682.6093331	2418.42671	3333.75937	0.2678357974	0.7888257147	No
15	White_No_Regular_<=Bach_31+	6124.27378	994.31205291	4682.01804	2546.36847	0.5275988424	0.5977777973	No
16	White_No_Regular_>Bach_21-30	25348.0862	2516.6030424	31503.0774	8030.59065	0.7313717893	0.4645520918	No
17	White_No_Regular_>Bach_31+	12297.5496	1439.0650885	16079.9926	6228.80642	0.5916647853	0.5540750819	No
18	White_No_Regular_<=Bach_11-20	31799.7377	3057.0406122	25190.6663	8074.46093	0.7654885657	0.4439806726	No
19	Other_No_Regular_>Bach_11-20	1888.98967	943.28045042	1078.88622	2982.46508	0.2590419038	0.7958029091	No
20	Black_No_Regular_>Bach_11-20	5263.9794	1887.6732304	358.219329	2766.36906	1.4648232108	0.1429691575	No
21	White_Yes_Regular_<=Bach_1-5	4394.65161	1638.9456209	3505.227	5019.34929	0.1684467273	0.8662318502	No
22	White_No_Regular_<=Bach_21-30	17296.2428	2294.4951228	15734.5343	6352.9519	0.2312064381	0.8171544286	No
23	Other_No_Other_>Bach_1-5	168.787915	744.71206631	889.831216	5864.19478	0.1219434375	0.9029438142	No
24	Black_No_Regular_>Bach_1-5	1649.52534	1353.7317171	584.200025	4593.62541	0.2224551659	0.8239595713	No
25	Black_No_Regular_>Bach_21-30	1459.05848	838.90320047	1797.42672	4955.6295	0.0673217713	0.9463255452	No
26	Other_No_Regular_>Bach_21-30	664.078416	1205.9187862	853.583259	7174.38776	0.0260486637	0.9792185235	No
27	Black_No_Regular_<=Bach_31+	409.170853	1011.7547599	1987.47823	3810.41015	0.4003371873	0.6889081815	No
28	Black_No_Regular_>Bach_6-10	3388.2106	1129.3012565	4901.19629	6179.31436	0.1771770142	0.8593693501	No
29	Black_No_Regular_<=Bach_21-30	780.205988	407.78809615	653.104476	2829.92397	0.04445424	0.9645423271	No
30	Other_No_Regular_>Bach_6-10	1923.94038	1693.3421098	1100.53264	1447.16086	0.3696583153	0.7116370946	No
31	White_No_Other_<=Bach_6-10	2941.90721	1403.0175891	3721.0565	12092.2494	0.0640044008	0.9489667226	No
32	Other_No_Regular_<=Bach_31+	123.480832	728.3778543	21.467043	3339.19795	0.0300201486	0.9760509842	No
33	White_Yes_Regular_>Bach_1-5	1268.7484	1470.3522698	433.643484	2004.60019	0.3359187405	0.7369321446	No
34	Other_No_Regular_>Bach_1-5	917.448567	1320.4788213	1964.43409	3343.28911	0.2912651746	0.77084852	No
35	White_No_Other_<=Bach_21-30	641.588841	453.86023905	612.625743	20171.1714	0.0014359026	0.998854635	No
36	Other_Yes_Regular_>Bach_11-20	786.928152	806.584658	531.596994	4085.56778	0.0613125081	0.9511302095	No
37	White_No_Other_>Bach_6-10	1364.22197	592.0077702	750.622384	12799.2372	0.0478891255	0.961804606	No
38	Black_No_Other_>Bach_6-10	302.936469	1173.4651179	480.474925	24500.0272	0.0072381619	0.9942248328	No
39	Black_No_Other_<=Bach_11-20	219.72408	1515.8259802	646.13348	20171.5166	0.0210797486	0.9831820896	No
40	Black_No_Other_>Bach_11-20	227.441522	2956.7401141	1974.90169	18615.4198	0.0927095124	0.9261343398	No
41	Black_No_Other_>Bach_1-5	384.629521	1831.7608139	206.530035	19281.3759	0.0091954634	0.9926631852	No
42	White_Yes_Other_>Bach_6-10	130.997409	3952.8066212	312.364597	18328.0135	0.0091933564	0.9926848662	No
43	Other_No_Regular_<=Bach_11-20	1586.64473	2302.3969889	2444.48636	3759.2883	0.1945960678	0.8457061817	No
44	White_Yes_Regular_>Bach_6-10	2102.85857	1990.877332	3041.18566	4569.95251	0.1880377034	0.8508470953	No
45	Other_No_Regular_<=Bach_1-5	1338.14158	2437.7214205	304.307282	1180.85997	0.3816753953	0.7027021526	No
46	White_Yes_Regular_>Bach_11-20	2647.93157	2645.4207247	4329.68244	5024.84576	0.2961520953	0.7671139311	No
47	Other_No_Other_<=Bach_1-5	1140.33261	2201.195287	941.430432	19361.9484	0.0103122439	0.9917721656	No
48	White_No_Other_>Bach_11-20	2244.69034	1711.6647529	4032.03459	8988.93365	0.1953285755	0.8451357279	No
49	Other_Yes_Regular_>Bach_6-10	509.467606	3896.1575612	405.486687	2204.75919	0.0234088483	0.981324147	No
50	White_No_Other_>Bach_21-30	916.543662	1897.9358422	357.119954	17376.526	0.0320038847	0.9744689529	No
51	White_Yes_Regular_<=Bach_6-10	2860.44027	2913.1672916	2286.98073	2790.92491	0.1421449838	0.8869654869	No
52	White_Yes_Regular_<=Bach_11-20	2300.23151	3737.5260892	2407.21763	3534.14795	0.0207987967	0.9834061577	No
53	Other_No_Regular_<=Bach_6-10	1008.50982	5093.6340064	653.37445	2382.01981	0.0631566284	0.9496417813	No
54	White_Yes_Regular_>Bach_21-30	611.17987	6503.3047961	943.1661	3878.73249	0.0438440231	0.9650287354	No
55	White_Yes_Regular_<=Bach_31+	248.785428	3772.7626682	126.825254	19610.0311	0.0061072755	0.9951271295	No
56	White_Yes_Other_>Bach_21-30	75.964465	1819.6767394	34.118864	13974.6653	0.0029693231	0.9976308264	No
57	White_Yes_Other_>Bach_11-20	181.177979	6703.2701471	122.358257	16200.4997	0.0022141525	0.9982333633	No
58	Black_No_Other_<=Bach_21-30	58.525267	3597.3026706	122.235169	27279.6676	0.0023153932	0.9981525852	No
59	Other_Yes_Regular_<=Bach_1-5	727.120555	5431.9941743	1509.3183	8740.51873	0.0760084643	0.9394123641	No
60	Other_Yes_Other_>Bach_1-5	137.284674	4299.5825823	175.207782	19599.8795	0.0018899249	0.998492059	No
61	Other_Yes_Regular_<=Bach_6-10	645.666276	4615.9356911	100.463826	502.433637	0.1174195609	0.9065275839	No
62	Black_Yes_Regular_<=Bach_11-20	40.842655	3417.224441	479.238931	8619.57502	0.0472809052	0.9622895856	No
63	Other_No_Regular_<=Bach_21-30	227.197632	10784.840554	66.704647	1280.89596	0.01477349	0.9882096698	No
64	Black_Yes_Regular_<=Bach_1-5	22.820859	9754.350663	38.855203	13615.338	0.0009573386	0.9992361544	No
65	Other_No_Other_>Bach_6-10	21.00747	16299.185488	1748.39221	16142.8531	0.0752992927	0.9399765842	No
66	Other_Yes_Regular_>Bach_1-5	600.864725	10746.551151	712.710405	16681.5482	0.0056364036	0.9955028264	No
67	Other_Yes_Regular_<=Bach_11-20	788.465177	13112.959462	673.527317	4023.63233	0.0083796006	0.9933141243	No
68	Other_Yes_Other_<=Bach_11-20	131.977219	17470.254797	380.628754	24927.6075	0.0081685707	0.9934824961	No
69	White_Yes_Other_<=Bach_6-10	593.320921	13841.598938	958.014251	20004.481	0.0149917272	0.9880387804	No
70	Black_No_Other_<=Bach_31+	42.966539	15021.819073	253.822377	17334.0225	0.009197033	0.9926619028	No
71	Black_Yes_Regular_>Bach_21-30	21.796465	40697.338476	76.944391	4434.39051	0.0013471015	0.9989251688	No

Figure 2: Results of regressing 85 NAEP-estimated counts on the corresponding SASS-estimated counts. See text for discussion.

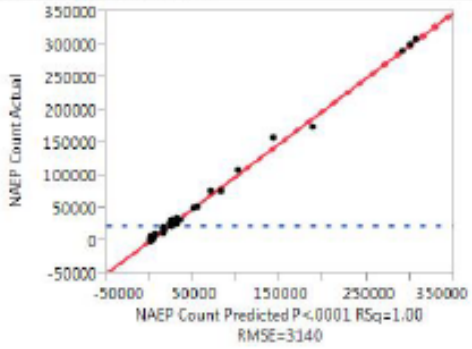
Response NAEP Count

Whole Model

Regression Plot



Actual by Predicted Plot



Summary of Fit

RSquare	0.997404
RSquare Adj	0.997373
Root Mean Square Error	3139.98
Mean of Response	23478.23
Observations (or Sum Wgts)	85

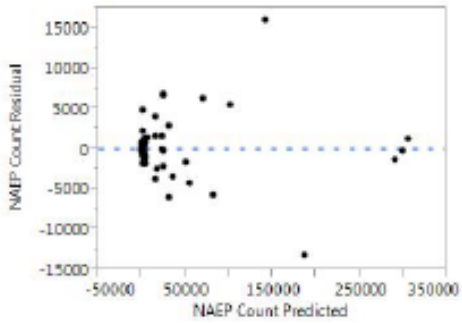
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	3.1443e+11	3.144e+11	31891.17
Error	83	818336140	9859471.7	Prob > F
C. Total	84	3.1525e+11		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	140.80202	364.79	0.39	0.7005
SASS Count	0.9909067	0.005549	178.58	<.0001*

Residual by Predicted Plot



SASS Count

Leverage Plot

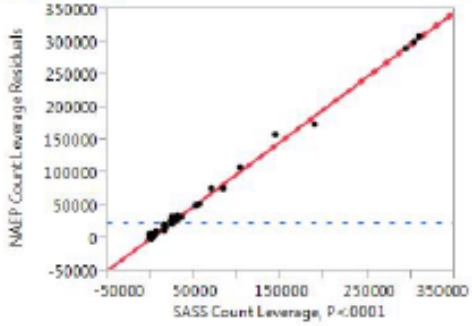


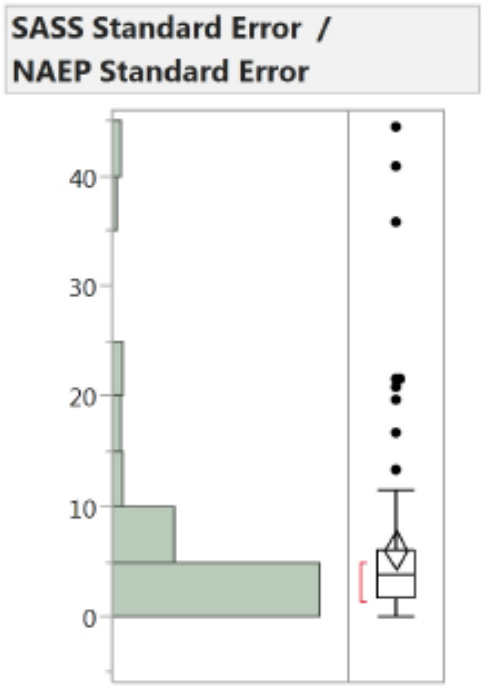
Table 5: NAEP- and SASS-estimated marginal proportions for the five variables of §2.3. See text for discussion.

	Variable	Category	NAEP Proportion	SASS Proportion	Relative Error
1	Race Recoded	White	0.8701422888	0.8817320956	0.0132313217
2	Race Recoded	Black	0.0690957943	0.0692916024	0.002829854
3	Race Recoded	Other	0.0607619169	0.048976302	0.2147950819
4	Hispanic Recoded	No	0.9237147139	0.9279917008	0.0046195087
5	Hispanic Recoded	Yes	0.0762852861	0.0720082992	0.0576826962
6	Highest Degree Recorded	> Bach	0.5242503629	0.5688398573	0.0815842894
7	Highest Degree Recorded	< = Bach	0.4757496371	0.4311601427	0.098332812
8	Certification Recoded	Regular	0.8991803008	0.907411117	0.0091119842
9	Certification Recoded	Other	0.1008196992	0.092588883	0.0851132467
10	Years Teaching Certification Recoded	1 – 5	0.2297045083	0.2124424007	0.0780831314
11	Years Teaching Certification Recoded	6 – 10	0.2301698108	0.2490595711	0.078833899
12	Years Teaching Certification Recoded	11 – 20	0.3230207372	0.3087637467	0.0451324489
13	Years Teaching Certification Recoded	21 – 30	0.1490543156	0.1550155952	0.0392099273
14	Years Teaching Certification Recoded	31 +	0.0680506281	0.0747186863	0.0934102433

Table 6: Comparison of standard errors of NAEP- and SASS-estimated counts for the five variables of §2.3, both marginally and crossed. See text for discussion.

	Variable	Category	SASS Standard Error / NAEP Standard Error
1	Race Recoded	White	4.3571368895
2	Race Recoded	Black	7.8640898813
3	Race Recoded	Other	3.6159560874
4	Hispanic Recoded	No	3.7981261932
5	Hispanic Recoded	Yes	3.2781996127
6	Highest Degree Recoded	>Bach	4.5021308894
7	Highest Degree Recoded	<=Bach	4.1747237938
8	Certification Recoded	Regular	4.8749224811
9	Certification Recoded	Other	5.4927440412
10	Years Teaching Recoded	1-5	3.571094203
11	Years Teaching Recoded	6-10	5.4161449621
12	Years Teaching Recoded	11-20	3.7634348345
13	Years Teaching Recoded	21-30	3.3450370868
14	Years Teaching Recoded	31+	4.529618403
15	Five-Variable Cross	Black_No_Regular_>Bach_31+	18.782980454
16	Five-Variable Cross	Black_No_Other_<=Bach_1-5	35.784361511
17	Five-Variable Cross	White_No_Other_<=Bach_11-20	18.884813578
18	Five-Variable Cross	Black_No_Regular_<=Bach_1-5	11.487069914
19	Five-Variable Cross	White_No_Other_>Bach_1-5	21.161403019
20	Five-Variable Cross	White_No_Regular_>Bach_1-5	4.908269899315
21	Five-Variable Cross	White_No_Regular_>Bach_6-10	4.1281570799
22	Five-Variable Cross	White_No_Regular_>Bach_11-20	4.5176577237
23	Five-Variable Cross	White_No_Regular_<=Bach_6-10	3.0168007488
24	Five-Variable Cross	Black_No_Regular_<=Bach_11-20	2.2986777321
25	Five-Variable Cross	White_No_Regular_<=Bach_1-5	3.8581636259
26	Five-Variable Cross	White_No_Other_<=Bach_1-5	8.2139743653
27	Five-Variable Cross	White_Yes_Other_<=Bach_1-5	48.865671242
28	Five-Variable Cross	Black_No_Regular_<=Bach_6-10	1.9812832706
29	Five-Variable Cross	White_No_Regular_<=Bach_31+	2.5688834258
30	Five-Variable Cross	White_No_Regular_>Bach_21-30	3.2900438458
31	Five-Variable Cross	White_No_Regular_>Bach_31+	4.3283703855
32	Five-Variable Cross	White_No_Regular_<=Bach_11-20	2.6412871744
33	Five-Variable Cross	Other_No_Regular_>Bach_11-20	3.1605810112
34	Five-Variable Cross	Black_No_Regular_>Bach_11-20	1.4654914901
35	Five-Variable Cross	White_Yes_Regular_<=Bach_1-5	3.8625003149
36	Five-Variable Cross	White_No_Regular_<=Bach_21-30	2.7687798653
37	Five-Variable Cross	Other_No_Other_>Bach_1-5	7.8744457867
38	Five-Variable Cross	Black_No_Regular_>Bach_1-5	3.2933855936
39	Five-Variable Cross	Black_No_Regular_>Bach_21-30	5.9072721337
40	Five-Variable Cross	Other_No_Regular_>Bach_21-30	5.3485125408
41	Five-Variable Cross	Black_No_Regular_<=Bach_31+	3.7681400747
42	Five-Variable Cross	Black_No_Regular_>Bach_6-10	5.471885814
43	Five-Variable Cross	Black_No_Regular_<=Bach_21-30	6.9396924805
44	Five-Variable Cross	Other_No_Regular_>Bach_6-10	0.8546181257
45	Five-Variable Cross	White_No_Other_<=Bach_6-10	0.61874898
46	Five-Variable Cross	Other_No_Regular_<=Bach_31+	4.5589726364
47	Five-Variable Cross	White_Yes_Regular_>Bach_1-5	1.3633468851
48	Five-Variable Cross	Other_No_Regular_>Bach_1-5	2.5358803345
49	Five-Variable Cross	White_No_Other_<=Bach_21-30	44.443574898
50	Five-Variable Cross	Other_Yes_Regular_>Bach_11-20	5.8659440048
51	Five-Variable Cross	White_No_Other_>Bach_6-10	21.620868837
52	Five-Variable Cross	Black_No_Other_>Bach_6-10	20.87936804
53	Five-Variable Cross	Black_No_Other_<=Bach_11-20	13.387277261
54	Five-Variable Cross	Black_No_Other_>Bach_11-20	6.2959269762
55	Five-Variable Cross	Black_No_Other_>Bach_1-5	18.526342788
56	Five-Variable Cross	White_Yes_Other_>Bach_6-10	4.8896832945
57	Five-Variable Cross	Other_No_Regular_<=Bach_11-20	1.8527715518
58	Five-Variable Cross	White_Yes_Regular_<=Bach_6-10	2.2954465545
59	Five-Variable Cross	Other_No_Regular_<=Bach_1-5	0.6864313684
60	Five-Variable Cross	White_Yes_Regular_>Bach_11-20	1.8994905173
61	Five-Variable Cross	Other_No_Other_<=Bach_1-5	8.7052468685
62	Five-Variable Cross	White_No_Other_>Bach_11-20	5.2515737312
63	Five-Variable Cross	Other_Yes_Regular_>Bach_6-10	0.5727902861
64	Five-Variable Cross	White_No_Other_>Bach_21-30	9.1554865941
65	Five-Variable Cross	White_Yes_Regular_<=Bach_6-10	0.5580375805
66	Five-Variable Cross	White_Yes_Regular_<=Bach_11-20	0.9459488247
67	Five-Variable Cross	Other_No_Regular_<=Bach_6-10	0.6676661578
68	Five-Variable Cross	White_Yes_Regular_>Bach_21-30	0.9964431782
69	Five-Variable Cross	White_Yes_Regular_<=Bach_31+	5.1907907823
70	Five-Variable Cross	White_Yes_Other_>Bach_21-30	7.6787515822
71	Five-Variable Cross	White_Yes_Other_>Bach_11-20	2.4168854388
72	Five-Variable Cross	Black_No_Other_<=Bach_21-30	7.5837889918
73	Five-Variable Cross	Other_Yes_Regular_<=Bach_1-5	1.8090810212
74	Five-Variable Cross	Other_Yes_Other_>Bach_1-5	4.5585940279
75	Five-Variable Cross	Other_Yes_Regular_<=Bach_6-10	0.6088417625
76	Five-Variable Cross	Black_Yes_Regular_<=Bach_11-20	2.5221906629
77	Five-Variable Cross	Other_No_Regular_<=Bach_21-30	0.6187681871
78	Five-Variable Cross	Black_Yes_Regular_<=Bach_1-5	1.2958220758
79	Five-Variable Cross	Other_No_Other_>Bach_6-10	0.9904885578
80	Five-Variable Cross	Other_Yes_Regular_>Bach_1-5	1.5522899265
81	Five-Variable Cross	Other_Yes_Regular_<=Bach_11-20	0.9086435536
82	Five-Variable Cross	Other_Yes_Other_<=Bach_11-20	1.4768599838
83	Five-Variable Cross	White_Yes_Other_<=Bach_6-10	1.6453169508
84	Five-Variable Cross	Black_No_Other_<=Bach_31+	1.1539029883
85	Five-Variable Cross	Black_Yes_Regular_>Bach_21-30	0.1089602091

Figure 3: Distribution of the ratio of SASS-estimated standard errors of counts to SASS-estimated standard errors of counts for the five variables of §1.3, both marginally and crossed. See text for discussion.



Quantiles

100.0%	maximum	44.4436
99.5%		44.4436
97.5%		40.0833
90.0%		14.7063
75.0%	quartile	6.12262
50.0%	median	3.76614
25.0%	quartile	1.76611
10.0%		0.75135
2.5%		0.11043
0.5%		0.10885
0.0%	minimum	0.10885

Summary Statistics

Mean	6.1192162
Std Dev	8.1170449
Std Err Mean	0.8804171
Upper 95% Mea	7.8700222
Lower 95% Mean	4.3684102
N	85

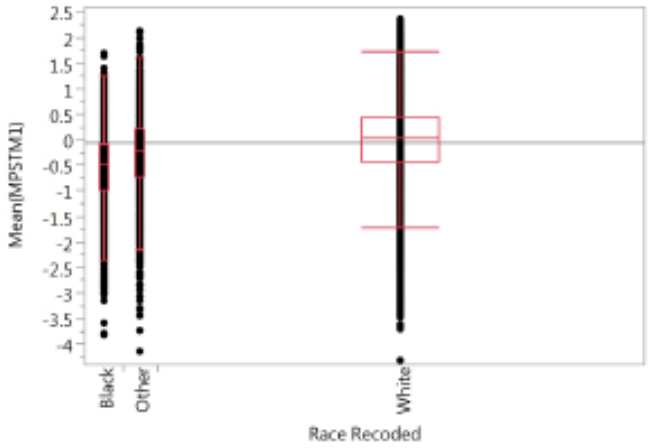
B.3 Figures for §3

Table 7: Estimated counts and standard errors for four NAEP-specific variables.

	Variable	Category	Count	Standard Error	CoeffOfVar
1	Q1 Class Size	15 or fewer	60286.3989	3594.7679013	0.007573
2	Q1 Class Size	16-18	36388.8313	2746.5817652	0.009586
3	Q1 Class Size	19-20	40378.8116	3096.2796045	0.009738
4	Q1 Class Size	21-25	117499.562	3821.3496432	0.00413
5	Q1 Class Size	26 or more	79093.965	3897.895921	0.006259
6	Q3 Hours/week	Less than 3 hours	13226.5844	1903.5629783	0.018278
7	Q3 Hours/week	3-4.9 hours	36521.4971	2976.2646773	0.01035
8	Q3 Hours/week	5-6.9 hours	193380.612	4771.0461345	0.003133
9	Q3 Hours/week	7 hours or more	90518.8757	3719.8552329	0.005219
10	Q13 Computers	Available	299792.517	5390.9097744	0.002284
11	Q13 Computers	I have access	30703.002	2714.1309709	0.011227
12	Q13 Computers	I have no access	3152.05028	1370.8423457	0.055233
13	Q15 Resources	Don't have resources	5558.84794	1001.968124	0.022891
14	Q15 Resources	Have some resources	64227.3667	2683.1981672	0.005306
15	Q15 Resources	Have most resources	178344.375	4653.8200391	0.003314
16	Q15 Resources	Have all resources	85516.9788	4577.4774855	0.006798

Figure 4: Analysis of the relationship of teacher race and the mean value, over students, of the posterior mean **MPSTM1**.

Oneway Analysis of Mean(MPSTM1) By Race Recoded



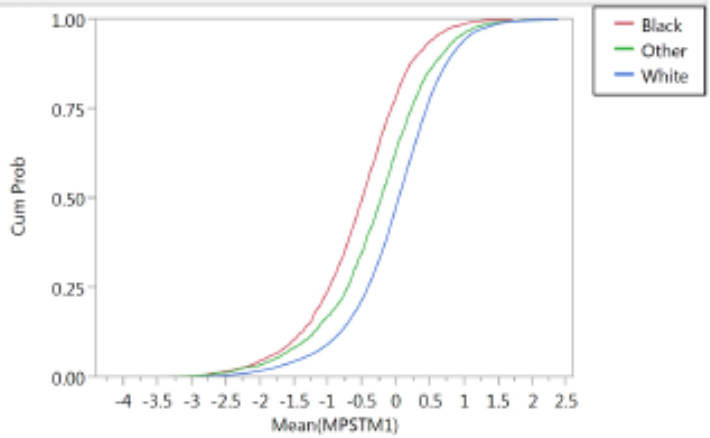
Weight Calibrated Nonresponse Adjusted SMSRSWT

Missing Rows 153

Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Black	-3.8109	-1.53665	-0.99298	-0.50505	-0.0799	0.296444	1.70384
Other	-4.1404	-1.3746	-0.72839	-0.23002	0.218594	0.65622	2.13875
White	-4.3151	-0.9639	-0.42413	0.02866	0.436646	0.810869	2.3746

CDF Plot



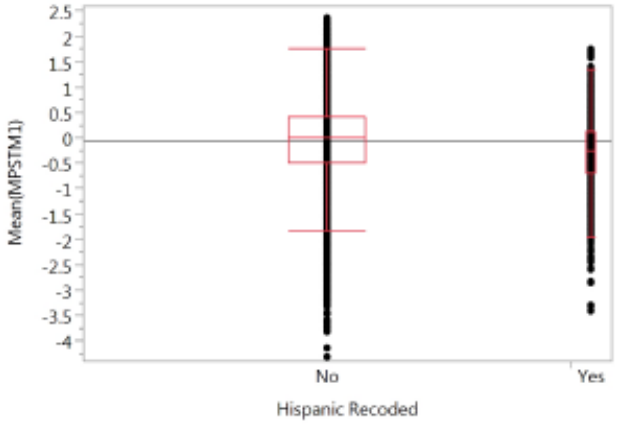
Nonparametric Comparisons For Each Pair Using Wilcoxon Method

q*	Alpha
1.95996	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
White	Black	5133.768	158.7981	32.32890	<.0001*	0.5366862	0.5061143	0.5673486
White	Other	2476.766	175.0387	14.14982	<.0001*	0.2605650	0.2248500	0.2962500
Other	Black	411.705	36.1583	11.38617	<.0001*	0.2758638	0.2290889	0.3224500

Figure 5: Analysis of the relationship of teacher Hispanic ethnicity and the mean value, over students, of the posterior mean **MPSTM1**.

Oneway Analysis of Mean(MPSTM1) By Hispanic Recoded



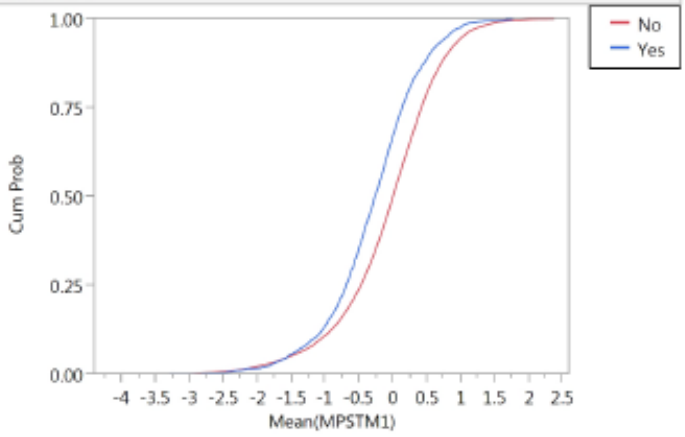
Weight Calibrated Nonresponse Adjusted SMSRSWT

Missing Rows 526

Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
No	-4.3151	-1.06328	-0.48106	-0.00807	0.413607	0.789292	2.3746
Yes	-3.4085	-1.17938	-0.69863	-0.27194	0.132243	0.527519	1.7627

CDF Plot



Kolmogorov Smirnov Two-Sample Test

Level	Count	EDF at Maximum	Deviation from Mean at Maximum
No	24000	0.432	-1.863
Yes	1794	0.605	6.815
Total	25794	0.444	

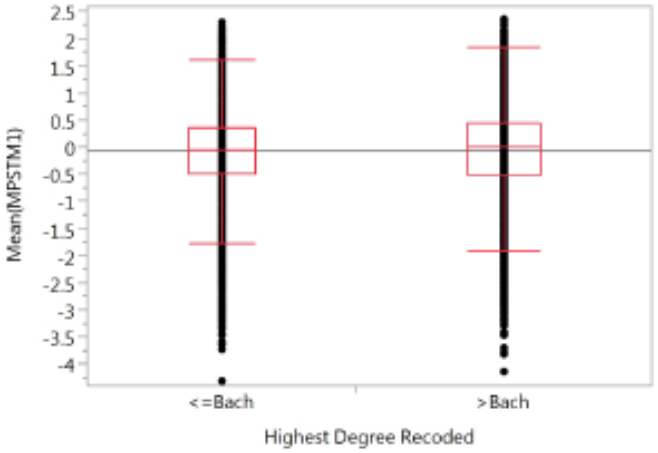
Maximum deviation occurred at observation 4679, value of Mean (MPSTM1) at maximum = -0.114025.

Kolmogorov-Smirnov Asymptotic Test

KS	KSa	D=max|F1-F2|	Prob > D	D+=max(F1-F2)	Prob > D+	D-=max(F2-F1)	Prob > D-
0.0439926	7.0654389	0.1729345	<.0001*	0.0071201	0.8443	0.1729345	<.0001*

Figure 6: Analysis of the relationship of teacher highest degree and the mean value, over students, of the posterior mean MPSTM1.

Oneway Analysis of Mean(MPSTM1) By Highest Degree Recoded



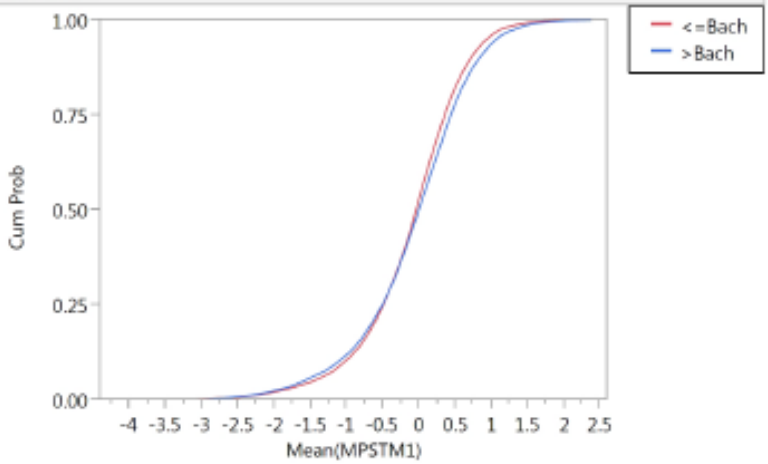
Weight Calibrated Nonresponse Adjusted SMSRSWT

Missing Rows 262

Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
<=Bach	-4.3151	-1.0286	-0.49675	-0.04577	0.35068	0.712575	2.309
>Bach	-4.1404	-1.11722	-0.5103	-0.00721	0.4366	0.827215	2.3746

CDF Plot



Kolmogorov Smirnov Two-Sample Test

Level	Count	EDF at Maximum	Deviation from Mean at Maximum
<=Bach	12091	0.725	2.910
>Bach	13967	0.676	-2.707
Total	26058	0.699	

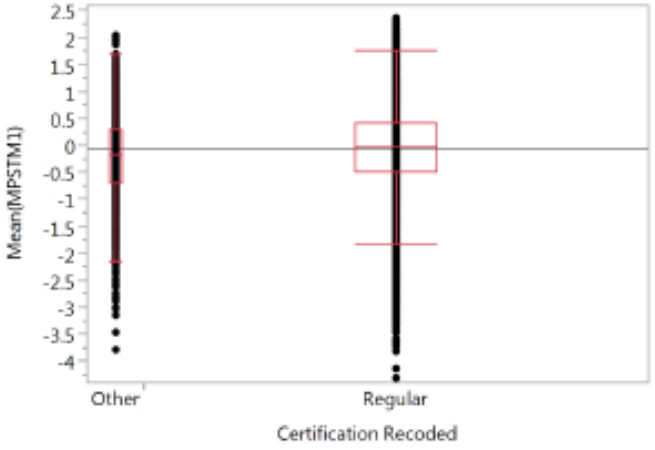
Maximum deviation occurred at observation 17776, value of Mean (MPSTM1) at maximum = 0.3046875.

Kolmogorov-Smirnov Asymptotic Test

KS	KSa	D=max|F1-F2|	Prob > D	D+=max(F1-F2)	Prob > D+	D-=max(F2-F1)	Prob > D-
0.0246216	3.9745407	0.0493713	<.0001*	0.0493713	<.0001*	0.0153907	0.0464*

Figure 7: Analysis of the relationship of teacher certification and the mean value, over students, of the posterior mean MPSTM1.

Oneway Analysis of Mean(MPSTM1) By Certification Recoded



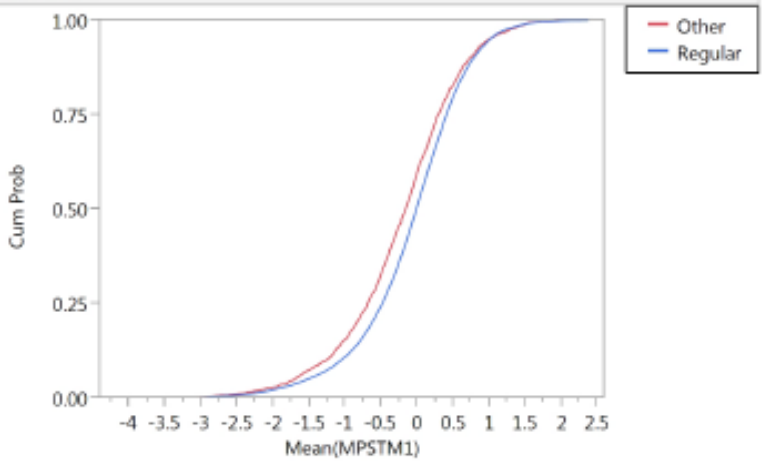
Weight Calibrated Nonresponse Adjusted SMSRSWT

Missing Rows 441

Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Other	-3.7878	-1.26409	-0.69226	-0.16252	0.290038	0.728635	2.0396
Regular	-4.3151	-1.05696	-0.48808	-0.01714	0.404056	0.77921	2.3746

CDF Plot



Kolmogorov Smirnov Two-Sample Test

Level	Count	EDF at Maximum	Deviation from Mean at Maximum
Other	1804	0.450	4.000
Regular	24075	0.348	-1.095
Total	25879	0.355	

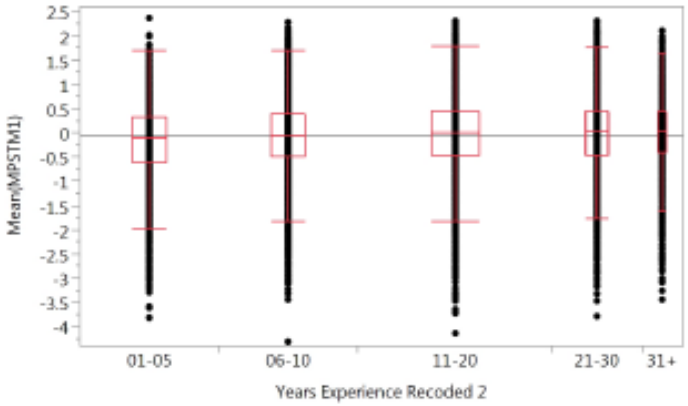
Maximum deviation occurred at observation 13718, value of Mean (MPSTM1) at maximum = -0.274025.

Kolmogorov-Smirnov Asymptotic Test

KS	KSa	D=max|F1-F2|	Prob > D	D+=max(F1-F2)	Prob > D+	D-=max(F2-F1)	Prob > D-
0.0257784	4.1469619	0.1012284	<.0001*	0.1012284	<.0001*	0.0064666	0.8690

Figure 8: Analysis of the relationship of teacher experience and the mean value, over students, of the posterior mean MPSTM1.

Oneway Analysis of Mean(MPSTM1) By Years Experience Recoded 2



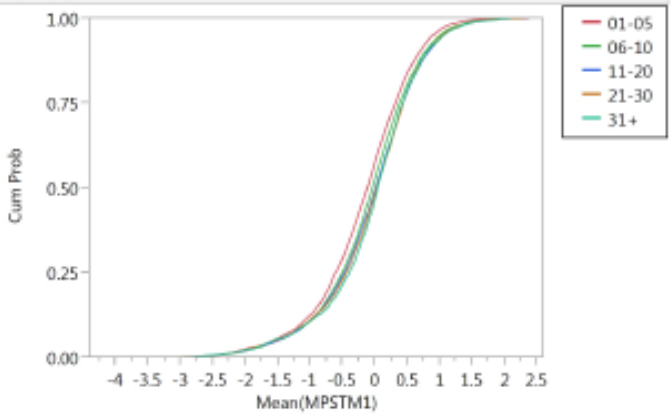
Weight Calibrated Nonresponse Adjusted SMSRSWT

Missing Rows 282

Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
01-05	-3.8109	-1.14176	-0.6069	-0.1192	0.314425	0.683991	2.3746
06-10	-4.3151	-1.05299	-0.50546	-0.04537	0.379267	0.764345	2.2935
11-20	-4.1404	-1.05507	-0.47229	0.008174	0.43375	0.823321	2.309
21-30	-3.7878	-1.07072	-0.45051	0.026514	0.433883	0.799663	2.2948
31+	-3.4468	-1.05412	-0.40005	0.038793	0.428222	0.794062	2.0938

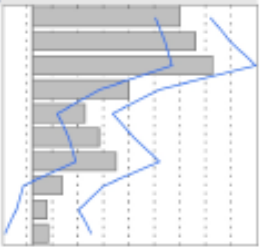
CDF Plot



Nonparametric Comparisons For Each Pair Using Wilcoxon Method

q*	Alpha
1.95996	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
11-20	01-05	734.9340	69.72889	10.53988	<.0001*	0.1282125	0.104450	0.1519500
21-30	01-05	575.8881	58.68409	9.81336	<.0001*	0.1409917	0.112957	0.1690238
31+	01-05	502.8257	59.40902	8.46379	<.0001*	0.1552500	0.119564	0.1910500
06-10	01-05	401.3391	62.62866	6.40823	<.0001*	0.0832132	0.057830	0.1086917
11-20	06-10	259.9979	69.69387	3.73057	0.0002*	0.0451333	0.021413	0.0688206
21-30	06-10	238.7653	58.59233	4.07503	<.0001*	0.0582111	0.030262	0.0860909
31+	06-10	236.1899	59.25526	3.98597	<.0001*	0.0724600	0.036955	0.1081154
31+	11-20	112.2577	74.83253	1.50012	0.1336	0.0265650	-0.008138	0.0613600
21-30	11-20	63.8314	68.24121	0.93538	0.3496	0.0125587	-0.013890	0.0390500
31+	21-30	33.8216	48.10455	0.70309	0.4820	0.0132800	-0.023861	0.0505750



B.4 Tables for Appendix A

Figure 9: Analog of Table 3 constructed using a T-statistic for independent samples.

	Variable	Category	NAEP Count	NAEP StandardError	SASS Count	SASS StandardError	T Statistic	Degrees of Freedom	p-Value for T Statistic
1	Race Recoded	White	290320.859	657.865478	294187.775	22570.0309	951.42359996	87.209767207	0.3200771218
2	Race Recoded	Black	23053.6438	195.132323	23118.9751	11776.7676	53.319271299	87.06779697	0.3200816199
3	Race Recoded	Other	20273.0659	266.88272	16340.8244	7640.72714	2371.1517443	87.301196278	0.3200742327
4	Hispanic Recoded	No	308195.169	724.417222	309622.181	21670.47	318.98920315	87.275887941	0.3200750318
5	Hispanic Recoded	Yes	25452.4003	331.550588	24025.3944	8560.67882	694.40057919	87.370274265	0.3200720539
6	Highest Degree Recoded	>Bach	174914.859	537.583307	189792.039	19057.2283	4475.5605876	87.196474852	0.3200775423
7	Highest Degree Recoded	<=Bach	158732.71	567.025226	143855.536	18639.1444	4244.6551297	87.228491678	0.3200765296
8	Certification Recoded	Regular	300009.321	618.442471	302755.519	23739.0243	718.48739351	87.167580457	0.3200784569
9	Certification Recoded	Other	33638.2475	397.032633	30892.0563	10919.1494	1117.2443847	87.32639171	0.3200734376
10	Years Teaching Recoded	1-5	76640.3508	499.100899	70880.8918	14034.1303	1866.2592439	87.312232245	0.3200738844
11	Years Teaching Recoded	6-10	76795.5978	364.489091	83098.1219	15544.2863	2787.974994	87.135767279	0.3200794646
12	Years Teaching Recoded	11-20	107775.084	573.177622	103018.275	16985.1533	1342.8600483	87.28143792	0.3200748658
13	Years Teaching Recoded	21-30	49731.61	383.443989	51720.5774	10099.4749	837.74242525	87.355839792	0.3200725089
14	Years Teaching Recoded	31+	22704.9266	212.25764	24929.7085	7565.42035	1679.1174563	87.194355383	0.3200776094

Figure 10: Analog of Table 4 constructed using a *T*-statistic for independent samples.

	Variable	Category	NAEP Count	NAEP StandardError	SASS Count	SASS StandardError	T Statistic	Degrees of Freedom	p-Value for T Statistic
1	Five-Variable Cross	Black_No_Regular_>Bach_31+	837.981957	29.988018	2569.48403	4673.6406	6611.3758975	87.010167643	0.32008145
2	Five-Variable Cross	Black_No_Other_<=Bach_1-5	999.561393	54.679707	95.970296	15398.2902	2153.366328	87.003114196	0.32008367 42
3	Five-Variable Cross	White_No_Other_<=Bach_11-20	3164.75809	117.277753	856.474108	15518.4267	3017.1542568	87.01410485	0.32008332 49
4	Five-Variable Cross	Black_No_Regular_<=Bach_1-5	3273.10727	68.640783	2727.59511	6205.80478	971.89925898	87.030212898	0.32008281 32
5	Five-Variable Cross	White_No_Other_>Bach_1-5	3026.338	109.627114	2817.29292	18611.9379	288.42486663	87.008568177	0.32008350 08
6	Five-Variable Cross	White_No_Regular_>Bach_1-5	15744.2755	191.145824	18296.8792	7187.5513	2133.511655	87.165302844	0.3200785 29
7	Five-Variable Cross	White_No_Regular_>Bach_6-10	11619.0084	287.281364	35316.3179	9120.0249	2071.9348843	87.234582189	0.32007633 71
8	Five-Variable Cross	White_No_Regular_>Bach_11-20	51594.4395	339.422232	56221.1016	12073.9522	2197.5708248	87.195126963	0.3200775 85
9	Five-Variable Cross	White_No_Regular_<=Bach_6-10	24833.5152	306.965501	25142.5009	7291.75455	162.33030154	87.437440001	0.32006990 87
10	Five-Variable Cross	Black_No_Regular_<=Bach_11-20	1793.20034	171.588389	1281.16791	3103.01443	478.1618445	87.754457682	0.32005999 87
11	Five-Variable Cross	White_No_Regular_<=Bach_1-5	32066.8474	247.278837	25621.464	7496.5593	4189.5521458	87.268624569	0.32007526 13
12	Five-Variable Cross	White_No_Other_<=Bach_1-5	9103.33663	229.78178	7573.17477	14861.3916	1064.1017784	87.05903655	0.32008185 98
13	Five-Variable Cross	White_Yes_Other_<=Bach_1-5	957.994988	63.707294	829.34098	20489.3846	270.45947356	87.002387588	0.32008369 72
14	Five-Variable Cross	Black_No_Regular_<=Bach_6-10	1418.24286	213.691599	2418.42671	3333.75937	753.05505086	88.013410676	0.32005193 23
15	Five-Variable Cross	White_No_Regular_<=Bach_31+	8124.27378	126.288297	4682.01804	2546.36847	1815.2185462	87.609329391	0.32006461 55
16	Five-Variable Cross	White_No_Regular_>Bach_21-30	25348.0862	319.608906	31503.0774	8010.59065	3106.2860609	87.39099548	0.3200714 01
17	Five-Variable Cross	White_No_Regular_>Bach_31+	12297.5496	182.761449	16079.9926	6228.80642	3306.246807	87.212562041	0.32007703 34
18	Five-Variable Cross	White_No_Regular_<=Bach_11-20	31799.7377	188.244546	25190.6663	8074.46093	2750.8644946	87.570575853	0.32006575 55
19	Five-Variable Cross	Other_No_Regular_>Bach_11-20	1888.98967	119.796737	1078.68622	2982.46508	1070.684984	87.398257741	0.32007117 23
20	Five-Variable Cross	Black_No_Regular_>Bach_11-20	5263.9794	239.73474	358.219329	2766.36906	3299.8978113	88.850222115	0.32002618 61
21	Five-Variable Cross	White_Yes_Regular_<=Bach_1-5	4194.65161	208.14681	3505.227	5019.34829	685.84608912	87.424478067	0.32007014 67
22	Five-Variable Cross	White_No_Regular_<=Bach_21-30	17296.2428	291.401172	15734.5343	6352.9519	864.10598711	87.519265758	0.32006736 62
23	Five-Variable Cross	Other_No_Other_>Bach_1-5	168.787915	94.578527	889.631216	5864.19478	1171.8035824	87.064235379	0.3200817 33
24	Five-Variable Cross	Black_No_Regular_>Bach_1-5	1649.52534	171.9241	584.200025	4593.62541	980.34817299	87.345792918	0.32007282 57
25	Five-Variable Cross	Black_No_Regular_>Bach_21-30	1459.05848	106.540813	1797.42672	4955.6296	495.29572171	87.114133621	0.32008015 02
26	Five-Variable Cross	Other_No_Regular_>Bach_21-30	664.078416	153.151839	853.583259	7174.38776	196.1055212	87.112526509	0.32008020 12
27	Five-Variable Cross	Black_No_Regular_<=Bach_31+	409.170853	128.492983	1987.47823	3810.41015	1945.1530807	87.280740184	0.32007487 86
28	Five-Variable Cross	Black_No_Regular_>Bach_6-10	3388.2306	143.421403	4501.19629	6179.31436	1228.2377961	87.133019849	0.32007965 16
29	Five-Variable Cross	Black_No_Regular_<=Bach_21-30	780.205988	51.78914	653.104476	2829.92397	354.30573037	87.082703513	0.32008114 69
30	Five-Variable Cross	Other_No_Regular_>Bach_6-10	1923.98038	215.054663	1100.53264	1447.16086	617.80967614	92.409853048	0.3199218629
31	Five-Variable Cross	White_No_Other_<=Bach_6-10	2941.90721	178.183412	3721.0565	12092.2494	693.66692753	87.053620641	0.32008206 99
32	Five-Variable Cross	Other_No_Regular_<=Bach_31+	123.488832	92.50408	21.467043	3319.19795	171.42790696	87.191775402	0.3200778 01
33	Five-Variable Cross	White_Yes_Regular_>Bach_1-5	1268.7484	186.734925	433.643484	2004.60019	719.43976323	89.136979743	0.32001747 43
34	Five-Variable Cross	Other_No_Regular_>Bach_1-5	917.448567	167.700724	1964.43409	3343.28911	999.33517336	87.620902893	0.32006417 76
35	Five-Variable Cross	White_No_Other_<=Bach_21-30	641.588841	57.640108	612.625343	20171.1714	65.131773129	87.002016645	0.3200837 09
36	Five-Variable Cross	Other_Yes_Regular_>Bach_11-20	786.928152	102.433814	531.596994	4085.56778	388.9953336	87.155216897	0.32007884 84
37	Five-Variable Cross	White_No_Other_>Bach_6-10	1364.22197	75.185062	750.622384	12799.2372	1177.0918263	87.008521786	0.32008355 23
38	Five-Variable Cross	Black_No_Other_>Bach_6-10	302.936469	149.030219	480.474625	24500.0272	184.82969514	87.009137995	0.32008348 27
39	Five-Variable Cross	Black_No_Other_<=Bach_11-20	219.72408	192.510092	646.13348	20171.5166	350.55685989	87.022493558	0.32008305 84
40	Five-Variable Cross	Black_No_Other_>Bach_11-20	227.441522	375.50637	1974.90169	18615.4198	750.18102903	87.100479048	0.32008058 31
41	Five-Variable Cross	Black_No_Other_>Bach_1-5	184.629521	212.633856	206.530035	19281.3759	122.14384183	87.035493136	0.3200826 31
42	Five-Variable Cross	White_Yes_Other_>Bach_6-10	130.997409	502.006981	312.364597	19128.0135	58.399630976	87.166569739	0.32007848 89
43	Five-Variable Cross	Other_No_Regular_<=Bach_11-20	1586.64473	202.40471	2444.48636	3759.2883	473.70912399	88.491270706	0.32003717 05
44	Five-Variable Cross	White_Yes_Regular_>Bach_6-10	2102.85857	252.841674	3040.18566	4569.95251	597.26003833	87.755266402	0.32005997 35
45	Five-Variable Cross	Other_No_Regular_<=Bach_1-5	1338.14158	309.59093	304.307282	1180.85997	540.4758865	103.3998294	0.3195445 87
46	Five-Variable Cross	White_Yes_Regular_>Bach_11-20	2647.93157	135.968768	4329.68244	5024.84576	808.73437306	88.102503467	0.3200491 68
47	Five-Variable Cross	Other_No_Other_<=Bach_1-5	1140.33261	279.552081	941.430432	19161.9484	114.09883067	87.052560361	0.32008210 35
48	Five-Variable Cross	White_No_Other_>Bach_11-20	2244.69034	217.381641	4032.03459	8988.93365	1316.5271365	87.144408363	0.32007919 08
49	Five-Variable Cross	Other_Yes_Regular_>Bach_6-10	509.467606	489.7525	405.486687	2204.75919	54.401310631	98.92005161	0.3197504395
50	Five-Variable Cross	White_No_Other_>Bach_21-30	916.543662	241.038093	357.119954	17376.526	370.99239003	87.047518242	0.32008226 36
51	Five-Variable Cross	White_Yes_Regular_<=Bach_6-10	2860.48027	369.972616	2286.98073	2790.92491	250.82353227	91.313037703	0.3199933143
52	Five-Variable Cross	White_Yes_Regular_<=Bach_11-20	2300.23151	474.666288	2407.21763	3514.14795	36.502613937	91.426551265	0.3199498 71
53	Five-Variable Cross	Other_No_Regular_<=Bach_6-10	1008.50982	646.89228	653.37445	2382.01981	88.995802995	104.53396819	0.3196197201
54	Five-Variable Cross	White_Yes_Regular_>Bach_21-30	611.17987	825.895135	943.1661	3878.73249	65.174169529	97.978818525	0.3197738186
55	Five-Variable Cross	White_Yes_Regular_<=Bach_31+	248.785428	479.141338	126.825254	19610.0311	41.129081118	87.14741174	0.32007909 56
56	Five-Variable Cross	White_Yes_Other_>Bach_21-30	75.964465	231.099177	34.118864	13974.6653	28.9544995	87.067533332	0.32008162 83
57	Five-Variable Cross	White_Yes_Other_>Bach_11-20	161.177979	851.31616	122.358257	16200.4997	7.3892381494	87.681385269	0.32006228 35
58	Five-Variable Cross	Black_No_Other_<=Bach_21-30	58.525767	456.832496	122.235189	27279.6676	22.50749741	87.069253084	0.32008157 37
59	Five-Variable Cross	Other_Yes_Regular_<=Bach_1-5	727.120555	689.86395	1509.3183	8740.51873	183.7094316	88.535412843	0.32003581 49
60	Five-Variable Cross	Other_Yes_Other_>Bach_1-5	137.284674	546.047534	175.207782	19599.8795	11.232317726	87.1916424	0.32007769 52
61	Five-Variable Cross	Other_Yes_Regular_<=Bach_6-10	645.666276	586.224419	100.463826	502.433637	150.81662362	118.26722883	0.3195214 29
62	Five-Variable Cross	Black_Yes_Regular_<=Bach_11-20	40.842655	433.987938	479.239931	8619.57502	163.35001585	87.625579737	0.3200640 31
63	Five-Variable Cross	Other_No_Regular_<=Bach_21-30	227.197652	1369.67612	66.704647	1280.89596	19.037553376	125.8548644	0.3192296052
64	Five-Variable Cross	Black_Yes_Regular_<=Bach_1-5	22.820859	1238.80377	38.855203	13615.338	2.0988090628	89.038884969	0.3287199459
65	Five-Variable Cross	Other_No_Other_>Bach_6-10	21.00747	2069.99863	1748.39221	16142.8531	135.35573077	91.037541964	0.3199611118
66	Five-Variable Cross	Other_Yes_Regular_>Bach_1-5	600.864725	1364.81336	712.710405	16681.5482	13.288012497	88.64059717	0.32003231 45
67	Five-Variable Cross	Other_Yes_Regular_<=Bach_11-20	788.665177	1665.34752	673.527337	4023.63233	11.395290442	123.99549368	0.3192580158
68	Five-Variable Cross	Other_Yes_Other_<=Bach_11-20	131.977219	2218.72458	380.628754	24927.6075	18.177678329	88.951486478	0.3200231032
69	Five-Variable Cross	White_Yes_Other_<=Bach_6-10	593.320921	1757.88482	958.014251	20004.481	33.647179178	88.902288959	0.32002460 01
70	Five-Variable Cross	Black_No_Other_<=Bach_31+	42.966539	1907.77293	253.922377	17334.0225	17.9359871	89.979323352	0.3199922038
71	Five-Variable Cross	Black_Yes_Regular_>Bach_21-30	21.796465	5168.56716	76.944391	4434.39051	1.7309909198	118.35670501	0.3198783719

Appendix D: Author and Acknowledgement

Author

Alan F. Karr, Director National Institute of Statistical Sciences

Special Recognition

The author acknowledges with appreciation the numerous insights, thoughtful advice and guidance during the course of this study.

American Institutes for Research (AIR)

Markus Broer
George Bohrnstedt
Young Yee Kim
Burhan Ogut
Michael Cohen

Westat

Keith Rust

National Center for Education Statistics

Kathryn Chandler
Taslima Rahman
William Tirre