Statistical Methodology

Reviewing Statisticians

Adam Chu, Senior Statistician, Westat, (301) 251-4326, was consulted about the statistical aspects of the design.

Respondent Universe

The respondent universe for the proposed FRSS survey on alternative schools and programs will include all local public school districts in the United States (50 states and the District of Columbia). School districts in the outlying U.S. territories will be excluded from the survey. As indicated in Table 1, 14,214 local public school districts (i.e., districts with a type-of-agency code of 1 or 2) are included in the 2005-06 CCD universe file. Of these, 2,503 (about 18 percent) have at least one alternative school listed in the corresponding 2005-06 CCD public school universe file. Approximately 408,000 students (less than 1 percent of the total public school enrollment) are enrolled in the alternative schools listed in the 2005-06 CCD.

Table 1. Distribution of public school districts in the 2005-06 NCES Common Core of Data (CCD) Public Elementary and Secondary Agency Universe File

Number of alternative schools in CCD*	Enrollment size class	Number of districts†	Total enrollment	Number of schools	Number of alternative schools	Enroll-ment in alternative schools
1 or more	Less than 1,000	411	219,436	1,627	582	7,734
	1,000 to 2,499	564	963,378	3,205	795	24,230
	2,500 to 9,999	918	4,812,686	10,072	1,521	99,049
	10,000 to 24,999	378	5,976,179	9,784	942	94,004
	25,000 to 99,999	208	9,004,524	13,919	1,228	127,830
	100,000+	24	4,520,805	5,803	380	54,699
	Subtotal	*2,503*	*25,497,008*	*44,410*	*5,448*	*407,546*
None	Less than 1,000	6,491	2,410,561	12,132	—	—
	1,000 to 2,499	2,771	4,479,210	10,806	—	—
	2,500 to 9,999	2,164	9,703,343	17,040	—	—
	10,000 to 24,999	229	3,314,963	5,131	—	—
	25,000 to 99,999	54	2,002,670	2,963	—	—
	100,000+	2	605,542	903	—	—
	Subtotal	*11,711*	*22,516,289*	*48,975*	—	—
Total		*14,214*	*48,013,297*	*93,385*	*5,448*	*407,546*

* 2005-06 CCD Public Elementary and Secondary School Universe File.

† Counts include district type 1 (local school district that is not part of a supervisory union and type 2 (local school district component of a supervisory union sharing a superintendent and administrative services with other school districts). All other district types are ineligible for the survey.

Statistical Methodology

Only those districts that operate alternative schools or alternative programs within "traditional" schools are eligible for the study. Based on the previous FRSS study conducted in 2001, an estimated 41 percent of all regular public school districts have either alternative schools or alternative programs. However, the information available in the CCD file about the presence of alternative schools is incomplete. For example, as summarized in Tables 1 and 2, while over 82 percent of the districts do not report any alternative schools in CCD, over 30 percent of these are expected to operate either alternative schools or programs. Moreover, among the roughly 2,500 districts that report one or more alternative schools in CCD, an estimated 14 percent are expected to be ineligible for the survey (i.e., do not operate alternative education programs). The implication of these results is that a stratified sampling design with disproportionate allocation will be required to obtain the desired number of eligible districts for analysis purposes.

For the proposed study, a stratified sample of 1,800 public school districts will be selected from the 2005-06 CCD universe file. Information from the previous FRSS survey on alternative schools and programs will be used to guide the allocation of the total sample to the four major categories of districts obtained by cross-classifying according to the presence or absence of alternative schools in the CCD file and whether or not the district serves only elementary grades. Within each of the four categories, the samples will be allocated to size strata in rough proportion to the aggregate square root of the enrollment in the stratum. Such an allocation is expected to yield relatively precise estimates of proportions (e.g., the proportion of eligible districts that operate alternative programs in community centers), as well as aggregative measures related to enrollment (e.g., the number of alternative programs or students enrolled in alternative programs). Districts in the sampling frame will be sorted by metropolitan status (urban, suburban, rural) and region (Northeast, Southeast, Central, West) to induce additional implicit stratification. Within each primary stratum, districts will be selected systematically and with equal probabilities. Assuming an overall response rate of 90 percent, the initial sample of 1,800 districts will yield 1,620 completed questionnaires, of which about 960 will be for eligible districts (i.e., districts with either alternative schools or programs). Table 3 summarizes the proposed sample allocation and the expected sample yields by primary sampling stratum.

Table 2. Distribution of public school districts in 2005-06 CCD universe file and estimated numbers of alternative schools/districts

Stratum	Number of alternative schools in district as reported in 2005-06 CCD	Enrollment size class of district	Number of districts in 1998-99 CCD frame	Estimated number of districts with alternative schools/ programs*	Estimated percent of districts with alternative schools/ programs*
1	1 or more	Less than 1,000	411	238	58%
2		1,000 to 2,499	564	475	84%
3		2,500 to 9,999	918	841	92%
4		10,000 to 24,999	378	369	98%
5		25,000 to 99,999	208	206	99%
6		100,000+	24	24	100%
7	0	Less than 1,000	6,491	1,015	16%
8		1,000 to 2,499	2,771	1,082	39%
9		2,500 to 9,999	2,164	1,331	62%
10		10,000 to 24,999	229	200	87%
11		25,000 to 99,999	54	54	100%
12		100,000+	2	2	100%
Total			14,214	5,836	41%

*Estimates based on FRSS survey of alternative schools and programs conducted in 2001.

Table 3.   Proposed sample sizes for the study

Stratum	Instructional level	Number of alternative schools in district as reported in 2005-06 CCD	Enrollment size class of district	Number of districts to be sampled	Expected number of responding districts*	Expected number of responding districts with alternative schools or programs
1	Elementary grades only	——	Less than 1,000	111	100	16
2			1,000+	13	11	6
3	Unified or secondary	1 or more	Less than 1,000	24	21	12
4			1,000 to 2,499	63	56	47
5			2,500 to 9,999	184	165	151
6			10,000 to 24,999	140	126	123
7			25,000 to 99,999	126	113	112
8			100,000+	24	22	22
9		None	Less than 1,000	274	247	39
10			1,000 to 2,499	302	272	106
11			2,500 to 9,999	429	386	237
12			10,000 to 24,999	80	72	63
13			25,000 to 99,999	30	27	27
14			100,000+	2	2	2
Total				1,800	1,620	963

*Assumes an overall response rate of 90 percent.

Expected Levels of Precision

Table 4 summarizes the approximate sample sizes and standard errors to be expected under the proposed design for selected subgroups. Since the sample sizes in Table 4 are based on preliminary tabulations of the 2005-06 CCD file, the actual sample sizes to be achieved may differ from those shown. Also, note that the sample sizes represent the expected numbers of completed questionnaires with eligible districts, and not the initial numbers of districts to be sampled. The standard errors in Table 4 have been inflated by an overall design effect of 1.5. The design effect arises primarily from the use of variable sampling fractions across the major sampling strata. In particular, the design effect reflects the fact that under the proposed stratified design, large districts will be sampled at relatively higher rates (i.e., have smaller sampling weights) than small districts. The standard errors in Table 4 can be converted to 95 percent confidence bounds by multiplying the entries by 2. For example, an estimated proportion of the order of 20 percent ($P = 0.20$) for suburban districts will be subject to a margin of error of ±4.6 percent at the 95 percent confidence level. Similarly, an estimated proportion of the order of 50 percent ($P = 0.50$) for districts in the Northeast will be subject to a margin of error of ±10.2 percent at the 95 percent confidence level.

Table 4. Expected standard error of an estimated proportion under proposed design for selected analytic
domains

Domain (subset)	Expected sample size*	Standard error† of an estimated proportion equal to ...		
		P = 0.20	P = .33	P = .50
Total sample	963	0.016	0.019	0.020
Metropolitan Status				
Urban	183	0.036	0.043	0.045
Suburban	462	0.023	0.027	0.028
Rural	318	0.027	0.032	0.034
Region				
Northeast	146	0.041	0.048	0.051
Southeast	171	0.037	0.044	0.047
Central	277	0.029	0.035	0.037
West	369	0.026	0.030	0.032
District Enrollment Class				
Under 2,500	223	0.033	0.039	0.041
2,500 to 9,999	391	0.025	0.029	0.031
10,000 to 24,999	187	0.036	0.042	0.045
25,000 or more	163	0.038	0.045	0.048

* Expected number of responding eligible districts, assuming response rate of 90 percent. The
standard errors given in this table are given for illustration. Actual standard errors may differ
from those shown.

† Assumes unequal weighting design effect of 1.5.

Estimation and Calculation of Sampling Errors

For estimation purposes, sampling weights reflecting the overall probabilities of selection and adjustments for nonresponse will be attached to each data record. To properly reflect the complex features of the sample design, standard errors of the survey-based estimates will be calculated using jackknife replication. Under the jackknife replication approach, 50 subsamples or "replicates" will be formed in a way that preserves the basic features of the full sample design. A set of estimation weights (referred to as "replicate weights") will then be constructed for each jackknife replicate. Using the full sample weights and the replicate weights, estimates of any survey statistic can be calculated for the full sample and each of the 50 jackknife replicates. The variability of the replicate estimates is used to obtain a measure of the variance (standard error) of the survey statistic. Previous surveys, using similar sample designs, have yielded relative standard errors (i.e., coefficients of variation) in the range of 2 to 10 percent for most national estimates. Similar results are expected for this survey.