

Section B. Description of Statistical Methodology

B.1. Respondent Universe and Sample Design

The Quick Response Information System (QRIS) consists of the Fast Response Survey System (FRSS) and the Postsecondary Education Quick Information System (PEQIS). The respondent universes and sample designs for FRSS and PEQIS that are covered by this generic clearance request are described below.

Fast Response Survey System

FRSS is designed to conduct brief surveys of state education agencies, public elementary and secondary schools, private elementary and secondary schools, public school districts, and school and public libraries. In the sections that follow, the approaches that will be used to design and select samples from the various sectors of interest are described. State education agencies are not discussed here, since they would be surveyed on a 100-percent basis without sampling.

Efficient probability sampling designs are an integral part of the FRSS. For those sectors that are surveyed frequently in FRSS (e.g., school districts and public schools), a general approach to sampling is designed and modified as necessary to meet the specific goals of the study. For example, for many FRSS surveys, probability-proportionate-to-size (PPS) sampling designs are used to ensure that both categorical and quantitative variables can be estimated reliably.

For some of the less frequently surveyed sectors, it is desirable to select a sample that is tailored to the specific needs of the individual survey. This specialization will be most efficient when pertinent data are available for sample selection purposes. Examples of situations that will necessitate designing and drawing special-purpose samples include surveys that are restricted to a particular subgroup (e.g., districts with summer migrant education programs), or surveys that require concurrent fielding of different questionnaires in the same sector (e.g., the FRSS surveys on arts education in public elementary and secondary schools).

Public Elementary and Secondary Schools

Since each new survey to be conducted under the FRSS will have unique analytic requirements, it is not possible to specify the exact form of the sample design that will be appropriate and efficient for a particular study. This can only be done after the study objectives have been clearly delineated. However, past experience suggests that many design features

that have been employed successfully in prior studies will also be applicable to future FRSS studies.

First, the frame to be used to select the required samples will be constructed from the most recent NCES CCD Public School Universe File. As shown in Table 1, almost 84,000 regular public schools are included in the CCD file, including about 63,000 elementary schools and about 21,000 secondary/combined schools. The school-level variables that are available in the CCD Public School Universe File include enrollment, number of teachers, instructional level, grade span, type of locale, region, percentage of minority students, percentage of students eligible for free lunch, and others. Such information is critical for designing efficient samples for the QRIS. As in prior studies, it is anticipated that enrollment size, type of locale, and instructional level will be used to define the primary strata for sampling purposes. It should be noted, however, that it may be necessary to merge information from other data sources (e.g., SASS) to obtain relevant auxiliary variables useful for stratification or weighting purposes. Also, some variables in the CCD file frequently have missing values (e.g., the free lunch data), and thus will have to be imputed or treated as a separate category if used for stratification purposes.

Table 1. Number of schools in the 2001-02 CCD Public School Universe file, by level and enrollment size class

Enrollment	Total*	Instructional Level	
		Elementary	Secondary or combined
1. <300	22,717	16,462	6,255
2. 300-499	23,647	20,270	3,377
3. 500-999	28,690	23,126	5,564
4. 1000-1499	5,461	2,702	2,759
5. 1500+	3,327	428	2,899
Total	83,842	62,988	20,854

*Counts reflect only regular schools in the CCD file. For example, special education, vocational education, and other alternative schools are excluded.

The allocation of the total sample to the primary strata will be made in proportion to an appropriate size measure. For example, proportional

allocation is generally efficient for estimating the proportions of schools having a given characteristic (e.g., proportion of schools that have computers for instructional purposes). On the other hand, allocation in proportion to enrollment is generally more effective for estimation of aggregate statistics that are correlated with enrollment (e.g., the number of secondary school students that have direct access to the internet). Although many prior FRSS samples have been designed primarily to estimate the latter type of aggregate statistics, there has generally been an equal interest in estimating proportions. For this reason, allocation in proportion to the square root of enrollment offers a compromise solution that is often used in many FRSS surveys. Again, the specific requirements of a particular survey will dictate how the sample is allocated to strata.

Within the primary strata, schools can also be selected at varying rates depending on the goals of the study. This can be accomplished either by using a PPS systematic sampling algorithm (Hansen, Hurwitz, and Madow, 1953), or by forming appropriate size classes (strata) of equal aggregate measure of size and selecting equal numbers of schools from each size stratum. The latter procedure is often used for reasons of simplicity. It should be noted, however, that if schools are to be drawn for the purpose of selecting a sample of teachers or students, then strict PPS sampling (i.e., sampling in proportion to estimated size) will be efficient (see the discussion below on sampling teachers or students within schools).

Finally, we note that the total sample size for a typical FRSS survey of public schools has been in the range of 800-1,000 respondents. For a sample of this size, the standard error of an estimated proportion for the total sample can be expected to be in the range of 0.015 to 0.020. For a 50 percent item, a standard error of 0.020 corresponds to coefficient of variation (cv) of 4 percent. Moreover, the sample is large enough to provide reasonably reliable estimates for broad subsets of the population (e.g., one-way classifications by type of locale or size class). Table 2 illustrates the levels of precision that can be expected for a sample of 1,000 public schools. The standard errors presented in this table are given for illustrative purposes only; the actual standard errors to be obtained for any given FRSS survey may be smaller or larger than those shown.

Private Elementary and Secondary Schools

The general approach described previously for public schools will also apply to private schools. Samples of private schools will be selected from the most current NCES Private School Survey (PSS) Universe File. Note that the PSS frame consists of two parts: a list frame and an “area frame.” The latter is actually an area probability sample that represents schools not included in the list frame. Consequently, the schools in the area sample must be weighted to represent the unlisted portion of the private school universe. Table 3 summarizes the weighted distribution of private schools in the 1996-97 PSS frame by affiliation and instructional level. To select the sample of private schools, stratification by instructional level (elementary, secondary, and combined) and by type of affiliation (Catholic, other religious, and nonsectarian) may be employed. Within each primary stratum, the private school frame can be sorted by enrollment size, geographic region, or other characteristics available in the PSS file to induce additional implicit stratification. Depending on the goals of the survey, the total sample can be

Table 2. Illustrative standard errors for an estimated proportion based on a sample of 1,000 public schools, by Census region, type of locale, and size class

Subset of sample	Expected sample size	Estimated proportion equal to:		
		0.80	0.50	0.30
Total sample	1,000	0.014	0.017	0.016
Instructional level				
Elementary	600	0.018	0.022	0.021
Secondary/combined	400	0.022	0.027	0.025
Census region				
Northeast	172	0.032	0.040	0.037
Midwest	275	0.026	0.032	0.030
South	344	0.023	0.029	0.027
West	210	0.031	0.039	0.036
Type of locale				
Cities	233	0.029	0.036	0.033
Urban fringe	243	0.028	0.035	0.032
Towns	270	0.026	0.032	0.029
Rural areas	254	0.026	0.033	0.030
Enrollment size				
Small (less than 500)	330	0.022	0.028	0.025
Medium (500-999)	325	0.022	0.028	0.025
Large (1,000+)	345	0.022	0.027	0.025

Table 3. Number of schools in the 1996-97 PSS private school frame by affiliation and instructional level

Type of affiliation	Total	Instructional level		
		Elementary	Secondary	Combined
Catholic	8,003	6,581	1087	335
Other religious	11,299	5,805	548	4,946
Nonsectarian	3,426	2,188	278	960
Total	22,728	14,574	1,913	6,241

allocated to the primary strata in different ways (e.g., proportionately, in proportion to enrollment, or in proportion to the aggregate square root of the enrollment). Table 4 illustrates the levels of precision that might be expected for a sample of 1,000 private schools.

Table 4. Illustrative standard errors for an estimated proportion based on a sample of 1,000 private schools, by level of instruction and affiliation

Subset of sample	Expected sample size	Estimated proportion equal to:		
		0.80	0.50	0.30
Total sample	1,000	0.015	0.018	0.017
Level				
Elementary	488	0.020	0.025	0.023
Secondary	257	0.028	0.036	0.033
Combined	256	0.029	0.036	0.034
Affiliation				
Catholic	455	0.021	0.026	0.025
Other religious	377	0.023	0.029	0.027
Nonsectarian	169	0.034	0.043	0.040

Public School Districts (LEAs)

The sampling frame to be used to select the required samples of public school districts (LEAs) will be constructed from the most recent NCES CCD Public Elementary and Secondary Agency Universe File. As shown in Table 5, over 14,000 regular public school districts are included in this file. We anticipate that for most of the district surveys to be conducted under FRSS, stratification by size class, region, metropolitan status, poverty status, or other variables will be used to improve the precision of overall estimates, and to ensure minimum sample sizes for the analytic domains of interest. Further, we expect that a probability-proportionate-to-square-root-of-size design will be efficient for the goals of the study. However, this basic design can be modified as necessary to meet the particular objectives of a given study.

Table 5. Distribution of public school districts in the 2001-02 NCES Common Core of Data Local Education Agency Universe Survey by enrollment size class and poverty status

Enrollment size class	Number of districts [†]	Percent of children below poverty level*			
		Missing	Under 10	10 to 19.9	20+
Less than 1,000	6,886	239	1,961	2,702	1,984
1,000 to 2,499	3,429	19	1,319	1,297	794
2,500 to 9,999	3,098	6	1,340	1,035	717
10,000 to 99,999	791	0	260	322	209
100,000+	25	0	6	9	10
TOTAL	14,229	264	4,886	5,365	3,714

*Based on Title 1 poverty measures provided by NCES.

†Includes district types 1 (local school district not part of a supervisory union) and 2 (local school district component of a supervisory union). Counts exclude districts with 0 or missing enrollment as reported in the CCD universe file.

The sample size for prior FRSS district surveys has traditionally ranged from 800-900 districts (respondents). Assuming that the sample is allocated to strata in rough proportion to aggregate square root of enrollment, the expected sample sizes would be those shown in Table 6. With these sample sizes, survey-based estimates for the total sample and for selected subgroups are expected to be reasonably precise. Table 6 illustrates the levels of precision to be expected for a sample of 850 public school districts.

Table 6. Illustrative standard errors for an estimated proportion based on a sample of 850 public school districts, by size class, OE region, and metropolitan status

Subset of sample	Expected sample size	Estimated proportion equal to:		
		0.80	0.50	0.30
Total sample	850	0.017	0.022	0.020
District size class				
Less than 2,500	296	0.028	0.034	0.032
2,500 to 9,999	305	0.025	0.031	0.029
10,000 or more	248	0.028	0.035	0.032
Region				
Central	250	0.032	0.040	0.037
Northeast	165	0.039	0.049	0.045
Southeast	170	0.039	0.049	0.044
West	264	0.031	0.039	0.036
Metropolitan status				
Urban	166	0.034	0.043	0.039
Suburban	370	0.026	0.033	0.030
Rural	313	0.029	0.036	0.033

Libraries

As an important provider of educational services, libraries (both school and public libraries) have been the focus of recent research efforts. For example, public libraries have been surveyed in FRSS 66 (survey on programs for adults in public library outlets) and FRSS 47 (survey on library services for children and young adults). In addition, Westat has conducted surveys of public and private school libraries for the U.S. Department of Education as part of the Assessment of the Role of School and Public Libraries in Support of the National Education Goals.

If required for an FRSS survey, samples of public libraries will be drawn from the most recent NCES Public Library Survey (PLS) universe file. Westat is familiar with this file, having recently used it to design and select a sample for FRSS 66. As shown in Table 7, almost 17,000 public library outlets are included in the 1997 PLS file (which was used for FRSS 66), of which

9,000 are central or main libraries, 7,000 are branches of library systems, and about 800 are bookmobiles/books-by-mail services.

Table 7 Number of public libraries in the 1997 NCES Public Library Survey (PLS) file, by type of outlet and size of legal service area

Size of legal service area*	Type of outlet			Total
	Central/Main	Branch	Bookmobile/ Mail	
1-4,999	3,915	69	—	3,984
5,000-24,999	3,165	690	—	3,855
25,000-49,999	865	767	—	1,632
50,000+	998	5,621	—	6,619
NA	—	—	829	829
Total	8,943	7,147	829	16,919

*Estimated population of legal service area.

Postsecondary Education Quick Information System (PEQIS)

The PEQIS universe includes all 2-year, 4-year, and graduate-level higher education institutions (IHEs) located in the 50 states and the District of Columbia. IHEs are defined for PEQIS as those institutions eligible to award Title IV federal financial aid and that grant degrees at the associate's level or higher. A total of 4,175 institutions were eligible for inclusion in the 2002 PEQIS sampling frame, constructed from the 2000 Integrated Postsecondary Education Data System (IPEDS) Institutional Characteristics file. The PEQIS sampling frame is stratified by instructional level (4-year, 2-year), control (public, private nonprofit, private for-profit), highest level of offering (doctor's/first-professional, master's, bachelor's, associate's), and total enrollment. Within each of the strata, institutions are sorted by region (Northeast, Southeast, Central, West) and whether the institution has a relatively high minority enrollment. The sample of institutions is allocated to the strata in proportion to the aggregate square root of total enrollment. Institutions within a stratum are sampled with equal probabilities of selection.

The PEQIS panel is a nationally representative sample of institutions from the PEQIS universe. The PEQIS panel was originally selected and recruited in 1991-92, and is periodically updated to reflect changes in the

postsecondary education universe that have occurred since the original panel was selected. A modified Keyfitz approach is used to maximize overlap between the existing panel and the periodic updates. In 2002, the sample includes a total of 1,610 institutions, with an 81 percent overlap of institutions. Table 8 summarizes the 2002 PEQIS universe counts and sample sizes by level, type of control, and highest level of offering. The PEQIS panel will be updated again in 2006 to reflect changes in the postsecondary universe since 2002.

Each institution in the PEQIS panel is asked to identify a campus representative to serve as survey coordinator. The campus representative facilitates data collection by identifying the appropriate respondent for each survey and forwarding the questionnaire to that person.

Less-than-2-year and non-degree-granting institutions are not included in the PEQIS universe and panel because of the great volatility of these types of institutions. These schools, many of which are proprietary, open and close at a much faster rate than other kinds of postsecondary institutions. This means that any portion of the PEQIS panel allotted to less-than-2-year and non-degree-granting institutions would be outdated very quickly -- that is, it would no longer represent an up-to-date universe of these schools. Further, NCES does not anticipate that there will be many survey requests that include these institutions. Thus, NCES decided that when a survey was requested through PEQIS that included less-than-2-year or non-degree-granting institutions, the most recent IPEDS Institutional Characteristics file would be used to draw an up-to-date supplementary sample of these institutions to be used for that survey. This approach means that the basic PEQIS panel will remain up-to-date (i.e., will accurately reflect the current universe of sampled institutions) for a longer period of time, and the supplementary samples of less-than-2-year and non-degree-granting institutions will also be up-to-date for the specific surveys for which these supplementary samples are drawn. NCES believes that this approach is the best compromise between the efficiencies of a standing panel of postsecondary institutions and the need for any such panel to reflect the current universe of institutions.

Nonresponse weight adjustments will be used to correct for unit nonresponse in surveys. Variances will be estimated using the jackknife replication method. Estimates produced during the PEQIS panel design stage, based on characteristics of the institutions, yielded coefficients of variation (CVs) in the range of 2 to 4 percent for most national estimates, with estimates for subgroups somewhat higher.

Table 8. Distribution of higher education institutions in 2002 PEQIS universe and panel

Level	Control	Highest level of offering	Number of institutions in PEQIS frame	Number in PEQIS panel
4-year	Public	Doctorate	257	218
		Masters	273	163
		Bachelors	92	41
	Private, nonprofit	Doctorate	353	157
		Masters	684	217
		Bachelors	510	124
	Private, for profit	—	277	45
2-year	Public	—	1,075	542
	Private, nonprofit	—	142	22
	Private, for profit	—	512	81
Total			4,175	1,610

B.2. Statistical Methodology

The statistical methodology is described in detail in Section B.1.

B.3. Methods for Maximizing the Response Rate

Telephone followup for nonresponse, which will be conducted by the staff of Westat's Telephone Research Center, will begin about 3 weeks after questionnaires have been mailed to the institutions. Experienced telephone interviewers will be trained in administering the questionnaire and will be monitored by Westat supervisory personnel during all interviewing hours. The response rate for the quick response surveys with single-stage samples completed to date through FRSS ranges from 85 to 99 percent, with most surveys above 90 percent, and on PEQIS ranges from 91 to 96 percent. Similar response rates are anticipated for future FRSS and PEQIS surveys. Ratio-weighting within adjustment cells will be used to partially compensate for the expected 10 percent (or less) nonresponse to each survey.

B.4. Tests of Procedures and Methods

Following the procedures for NCES quick-response surveys (PEQIS and FRSS) established during the current QRIS generic clearance (1850-0733), a pretest with nine institutions is conducted prior to OMB review for each survey to determine what problems respondents might have in providing the requested information and to make appropriate changes to the questionnaire, if necessary. Responses and comments on the questionnaire are collected by telephone during the pretest, and the results are summarized as part of the documentation for the survey.

B.5. Reviewing Statisticians

Statistician Adam Chu of Westat (301-251-4326) was consulted about the statistical aspects of the PEQIS panel design. Adam Chu is also the statistician for FRSS samples.

QRIS surveys are sponsored by NCES. Westat is the contractor currently conducting the QRIS surveys for NCES. For each survey, Westat will mail the questionnaires; collect data by Web, mail, and telephone; edit, code, key, and verify the data; and produce tabulations and the survey report.