

Technical Notes

**FRSS 84: Distance Education Courses for Public Elementary
and Secondary School Students: 2002–03**

Technical Notes — Distance Education Courses for Public Elementary and Secondary School Students: 2002–03

Fast Response Survey System

The Fast Response Survey System (FRSS) was established in 1975 by the National Center for Education Statistics (NCES), U.S. Department of Education. FRSS is designed to collect issue-oriented data within a relatively short time frame. FRSS collects data from state education agencies, local education agencies, public and private elementary and secondary schools, public school teachers, and public libraries. To ensure minimal burden on respondents, the surveys are generally limited to three pages of questions, with a response burden of about 30 minutes per respondent. Sample sizes are relatively small (usually about 1,000 to 1,500 respondents per survey) so that data collection can be completed quickly. Data are weighted to produce national estimates of the sampled education sector. The sample size permits limited breakouts by classification variables. However, as the number of categories within the classification variables increases, the sample size within categories decreases, which results in larger sampling errors for the breakouts by classification variables.

Sample and Response Rates

The sample for the FRSS survey on distance education courses consisted of 2,305 public school districts in the 50 states and the District of Columbia. It was selected from the 2001–02 NCES Common Core of Data (CCD) Local Education Agency Universe file, which was the most current file available at the time of selection. The sampling frame included 14,229 regular public school districts and 989 “other education agencies” with at least one charter school. For the purposes of the study, “regular” school districts included any local school district that was not a component of a supervisory union (i.e., Education Agency type 1 on the CCD), or was a local school district component of a supervisory union sharing a superintendent and administrative services with other local school districts (i.e., Education Agency type 2 on the CCD). Excluded from the sampling frame were districts in the outlying U.S. territories and regular districts with no enrollments.

The school district sampling frame was stratified by district type (regular or charter), enrollment size (less than 1,000, 1,000 to 2,499, 2,500 to 9,999, 10,000 to 99,999, and 100,000 or more), and percentage of children in the district ages 5–17 in families living below the poverty level (less than 10 percent, 10 to 19.99 percent, 20 to 29.99 percent, and 30 percent or more).¹ Districts in the frame were then sorted by metropolitan status (urban, suburban, rural) and region (Northeast, Southeast, Central, West) to induce additional implicit stratification.

All survey data were reported for the 2002–03 12-month school year when the survey was conducted. Questionnaires and cover letters for the study were mailed to the superintendent of each sampled district in November 2003. Respondents were asked to have the questionnaire completed by the

¹ Poverty estimates for school districts were based on Title I data provided to the U.S. Department of Education by the Bureau of the Census and contained in U.S. Department of Commerce, Bureau of the Census, Current Population Survey (CPS) “Small Area Income and Poverty Estimates, Title I Eligibility Database, 1999.” The No Child Left Behind Act of 2001 directs the Department of Education to distribute Title I basic and concentration grants directly to school districts on the basis of the most recent estimates of children in poverty. For income year 1999, estimates were derived for districts according to their 2001–02 boundaries based on 2000 census data and model-based estimates of poverty for all counties. For detailed information on the methodology used to create these estimates, please refer to www.census.gov/hhes/www/saie.html.

person who was most knowledgeable about the district's distance education courses. Respondents were offered the option of completing the survey via the Web or by mail. Telephone follow-up for survey nonresponse and data clarification was initiated in December 2003 and completed in April 2004. During data collection, 13 districts were found to be ineligible for the survey because they no longer existed or did not meet some other criteria for inclusion in the sample. For the eligible districts, the response rate was 94 percent (2,158 responding districts divided by the 2,292 eligible districts in the sample). The weighted response rate was 96 percent. Of the districts that completed the survey, 40 percent completed it by mail, 29 percent completed it by Web, 13 percent completed it by telephone, and 12 percent completed it by fax.

Although item nonresponse was very low, data were imputed for the 29 items with a response rate of less than 100 percent. The missing items included both numerical data, such as counts of enrollments in distance education courses, and categorical data, such as which technologies were used as primary modes of instructional delivery for distance education courses. The missing data were imputed using a "hot-deck" approach to obtain a "donor" district from which the imputed values were derived. Under the hot-deck approach, a donor district that matched selected characteristics of the district with missing data (the recipient district) was identified. The matching characteristics included district type, region, metropolitan status, district enrollment size class, and poverty concentration. Once a donor was found, it was used to derive the imputed values for the district with missing data. For categorical items, the imputed value was simply the corresponding value from the donor district. For numerical items, the imputed value was calculated by taking the donor's response for that item (e.g., number of distance education course enrollments) and dividing that number by the total number of students enrolled in the donor district. This ratio was then multiplied by the total number of students enrolled in the recipient district to provide an imputed value. All missing items for a given district were imputed from the same donor whenever possible. Imputation flags are included in the data.

Weighting Procedures and Sampling Errors

The response data were weighted to produce national estimates (see table 1). The weights were designed to adjust for the variable probabilities of selection and differential nonresponse. FRSS survey data are based on complex sample designs that require the use of weights to compensate for variable probabilities of selection, differential response rates, and possible deficiencies in the sampling frame. The reciprocal of the probability of selection, referred to as the "base weight," will produce unbiased (or consistent) estimates of population totals and ratios if there is no nonresponse in the survey. Since a stratified sample design was employed for the survey, the base weight for the i -th school in stratum h was computed as $w_{hi} = 1/f_h$ where f_h is the overall sampling rate used to select schools in stratum h .

Although the survey had a high response rate, adjustment of the base weights was necessary to compensate for the survey nonrespondents (i.e., whole questionnaire or unit nonresponse). To compensate for unit nonresponse, an adjustment factor was computed within selected weighting classes, and this factor was then used to inflate the base weights of the schools in the weighting class. The weighting classes used for this purpose were subsets of the sample defined by enrollment size class, poverty status, and type of locale. Within each subset, a nonresponse-adjustment factor was computed as the inverse of the base-weighted response rate. The factor was then applied to the base weights of the responding schools in the subset to obtain the final nonresponse-adjusted weight.

The findings in the survey report titled *Distance Education Courses for Public Elementary and Secondary School Students 2002–03* are estimates based on the sample selected and, consequently, are subject to sampling variability. The standard error is a measure of the variability of an estimate due to

sampling. It indicates the variability of a sample estimate that would be obtained from all possible samples of a given design and size. Standard errors are used as a measure of the precision expected from a particular sample. If all possible samples were surveyed under similar conditions, intervals of 1.96 standard errors below to 1.96 standard errors above a particular statistic would include the true population parameter being estimated in about 95 percent of the samples. This is a 95 percent confidence interval. For example, the estimated percentage of public school districts with students regularly enrolled in distance education courses is 36.4 percent and the standard error is 1.2 percent. The 95 percent confidence interval for the statistic extends from $[36.4 - (1.2 \times 1.96)]$ to $[36.4 + (1.2 \times 1.96)]$, or from 34.0 to 38.8 percent. The coefficient of variation (“c.v.,” also referred to as the “relative standard error”) of an estimate (y) is defined as $c.v. = (s.e. / y) \times 100$, where s.e. is the standard error of the estimate y.

Table 1. Number and percent of public school districts in the study, and the estimated number and percent in the nation, for the total sample and for districts with students regularly enrolled in distance education courses in 2002–03, by district characteristics: 2003

District characteristic	Total sample				Districts with students regularly enrolled in distance education courses in 2002–03			
	Respondents (unweighted)		National estimate (weighted)		Respondents (unweighted)		National estimate (weighted)	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
All public school districts	2,158	100	15,040	100	812	100	5,480	100
District enrollment size								
Less than 2,500.....	1,039	48	11,080	74	381	47	4,060	74
2,500 to 9,999.....	722	34	3,100	21	232	29	1,010	18
10,000 or more	394	18	820	5	199	25	410	7
Metropolitan status								
Urban.....	282	13	1,220	8	99	12	280	5
Suburban.....	1,052	49	6,150	41	337	42	1,700	31
Rural.....	824	38	7,660	51	376	46	3,500	64
Region								
Northeast.....	459	21	3,040	20	108	13	640	12
Southeast.....	355	16	1,750	12	182	22	790	14
Central.....	700	32	5,390	36	297	37	2,500	46
West	644	30	4,850	32	225	28	1,540	28
Poverty concentration								
Less than 10 percent.....	751	37	4,850	35	249	31	1,620	30
10 to 19 percent.....	776	38	5,330	38	331	41	2,220	41
20 percent or more	519	25	3,690	27	222	28	1,560	29

NOTE: For the FRSS study sample, there were 3 cases for which district enrollment size was missing and 112 cases for which poverty concentration was missing. Detail may not sum to totals because of rounding or missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Fast Response Survey System (FRSS), “Distance Education Courses for Public School Elementary and Secondary School Students: 2002–03,” FRSS 84, 2003.

Estimates of standard error were computed using a technique known as jackknife replication. As with any replication method, jackknife replication involves constructing a number of subsamples (replicates) from the full sample and computing the statistic of interest for each replicate. The mean square error of the replicate estimates around the full sample estimate provides an estimate of the variance of the statistic. To construct the replications, 50 stratified subsamples of the full sample were created and then dropped 1 at a time to define 50 jackknife replicates. A computer program (WesVar) was used to calculate the estimates of standard errors.

Nonsampling Errors, Coding, and Editing

The survey estimates are also subject to nonsampling errors that can arise because of nonobservation (nonresponse or noncoverage) errors, errors of reporting, and errors made in data collection. These errors can sometimes bias the data. Nonsampling errors may include such problems as misrecording of responses; incorrect editing, coding, and data entry; differences related to the particular time the survey was conducted; or errors in data preparation. While general sampling theory can be used to determine how to estimate the sampling variability of a statistic, nonsampling errors are not easy to measure and, for measurement purposes, usually require that an experiment be conducted as part of the data collection procedures or that data external to the study be used.

To minimize the potential for nonsampling errors, the questionnaire was pretested with respondents at institutions like those that completed the survey. During the design of the survey and the survey pretest, an effort was made to check for consistency of interpretation of questions and to eliminate ambiguous items. The questionnaire and instructions were extensively reviewed by NCES.

Manual and machine editing of the questionnaire responses were conducted to check the data for accuracy and consistency. Cases with missing or inconsistent items were recontacted by telephone. A coding source file and editing manual were produced to identify cases requiring data retrieval or clarification and prepare cases for key entry. The source file served as a data dictionary and included the data file layout, a description of each data item, and a list of valid response codes, range formats, as well as codes for nonresponse, inapplicable responses, and defined skip patterns. The coding source file was used to develop the ACCESS database for data verification while the codebook served as the main tool for coding, editing, and processing questionnaires received by mail, fax, or telephone.

Logics, ranges, and validation checks were prepared prior to data collection and included online edit checks, manual logic checks, and automated checks using SAS. Online checks were incorporated into the online data entry system while manual edits were conducted to process cases received by mail, fax, or telephone. In both cases, where electronic and hardcopy survey data were submitted, steps were taken to ensure that the method of entering the data is the same, regardless of mode. For example, to enter survey data received by mail/fax or telephone, we accessed the survey website as “respondents” and “complete” the survey using the responses on the hardcopy survey. Subjecting all survey responses to the same set of built-in logics, ranges, and validation checks helps to ensure that data entry does not produce system differences in the survey data. In addition, all hardcopy data were subject to 100 percent verification and “double-keyed” in a simulated website interface for subsequent data checks.

Definitions of Selected Analysis Variables

District Enrollment Size—This variable indicates the total number of students enrolled in the district based on data from the 2001–02 CCD. Data on this variable were missing for three districts; districts

with missing data were excluded from all analyses involving district enrollment size. The variable was collapsed into the following three categories:

Less than 2,500 students
2,500 to 9,999 students
10,000 or more students

Metropolitan Status—This variable indicates the type of community in which the district is located, as defined in the 2001–02 CCD (which uses definitions based on U.S. Census Bureau classifications). Metropolitan status is the classification of an education agency’s service area relative to a Metropolitan Statistical Area (MSA). An MSA is an area consisting of one or more contiguous counties (cities and towns in New England) that contain a core area with a large population nucleus, as well as adjacent communities having a high degree of economic and social integration with that core. An area is defined as an MSA if it is the only MSA in the immediate area and has a city of at least 50,000 population or it is an urbanized area of at least 50,000 with a total metropolitan population of at least 100,000 (75,000 in New England). The categories are described in more detail below.

Urban—Primarily serves a central city of an MSA
Suburban—Serves an MSA but not primarily its central city
Rural—Does not serve an MSA

Region—This variable classifies districts into one of the four geographic regions used by the Bureau of Economic Analysis of the U.S. Department of Commerce, the National Assessment of Educational Progress, and the National Education Association. Data were obtained from the 2001–02 CCD Local Education Agency Universe file. The geographic regions are

Northeast—Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont

Southeast—Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia

Central—Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin

West—Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming

Poverty Concentration—This variable indicates the percentage of children in the district ages 5–17 in families living below the poverty level, based on the Title I data provided to the U.S. Department of Education by the Bureau of the Census. Data on this variable were missing for 112 districts; districts with missing data were excluded from all analyses involving poverty concentration. The variable was collapsed into the following three categories:

Less than 10 percent
10 to 19 percent
20 percent or more

Data Disclosure Warning

Under law, public use data collected and distributed by the National Center for Education Statistics (NCES) may be used only for statistical purposes.

Any effort to determine the identity of any reported case by public-use data users is prohibited by law. Violations are subject to Class E felony charges of a fine up to \$250,000 and/or a prison term up to 5 years.

NCES does all it can to assure that the identity of data subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted or modified in the dataset to protect the true characteristics of individual cases. Any intentional identification or disclosure of a person or institution violates the assurances of confidentiality given to the providers of the information. Therefore, users shall:

- Use the data in this dataset for statistical purposes only.
- Make no use of the identity of any person or institution discovered inadvertently, and advise NCES of any such discovery.
- Not link this dataset with individually identifiable data from other NCES or non-NCES datasets.
- To proceed you must signify your agreement to comply with the above-stated statutorily based requirements.

Data perturbations were conducted on some background data to preclude identification of individuals and institutions.