The Study of State and Local Implementation and Impact of the
Individuals with Disabilities Education Act

Volume III: Technical Appendices for the SLIIDEA Sourcebook Report

April 2006

Prepared for
Dr. Scott Brown
Office of Special Education Programs
Potomac Center Plaza
550 12th Street, SW
Room 4076
Washington, DC 20065

Prepared by
Cristofer Price
Amanda Parsad
Robert St.Pierre
Ellen Schiller

Abt Associates Inc.

Abt Associates Inc.
4550 Montgomery Avenue
Suite 800
Bethesda, MD 20814

Preface

This report, *Volume III: Technical Appendices for the SLIIDEA Sourcebook Report*, provides the complete description of the sampling design and analytic approach used in the Study of State and Local Implementation and Impact of IDEA (SLIIDEA). SLIIDEA addressed how states, districts, and schools made progress on nine issues of concern identified by Congress in the 1997 amendments to IDEA. Abt Associates and its subcontractor, Westat, conducted this study between April 2000 and September 2006.

This is Volume III of three volumes that have been prepared to conclude this seminal policy study in special education. Volume I, the *Sourcebook*, summarizes the research results related to each congressional topic. Volume II consists of tables that display data for each data collection year, and show changes, including trends, over time in responses to individual survey items. Each volume is listed below with its citation and publication date.

- Schiller, E., Fritts, J., Bobronnikov, E., Fiore, T., O'Reilly, F., & St. Pierre, R. (2006, April). *Volume I: The SLIIDEA Sourcebook Report (1999–2000, 2002–2003, 2003–2004, and 2004–2005 School Years)*. Bethesda, MD: Abt Associates Inc.
- Schiller, E., Bobronnikov, E., Fritts, J., Parsad, A., Brown-Lyons, M., Chawla, D., Simpson, L., Mahmud, F., & Marsh, M. (2006, April). *Volume II: Data Tables for the SLIIDEA Sourcebook Report*. Bethesda, MD: Abt Associates Inc.
- Price, C., Parsad, A., St. Pierre, R., & Schiller, E. (2006, April) *Volume III: Technical Appendices for the SLIIDEA Sourcebook Report*. Bethesda, MD: Abt Associates Inc.

In addition to these three volumes, a summary report from this six-year study, called *Marking Progress of IDEA Implementation* (Schiller, O'Reilly, & Fiore, 2006),[1] has been prepared to capture the distinctive and emerging findings from SLIIDEA and to discuss the implications. The study's web site, www.abt.sliidea.org, provides public access to SLIIDEA reports and data tables from each survey year (the 1999–2000, 2002–2003, 2003–2004, and 2004–2005 school years), as well as reports for each qualitative study, and the above-mentioned volumes. A public use data set and supporting documentation are available on CD-ROM from the U.S. Department of Education.

The Office of Special Education Programs, U.S. Department of Education, supported this special education policy study from April 2000 to September 2006, under Contract No. ED-00-CO-0026. Dr. Kelly Henderson served as the project officer from April 2000 to December 2004, and Dr. Scott Brown served as the project officer between January 2005 and September 2006. The Abt SLIIDEA study team greatly benefited from their ongoing guidance and leadership. The study's Technical Work Group also provided invaluable feedback and guidance, and consisted of the following experts: Dr. Alvin Crawley (Arlington Virginia Public Schools), Dr. Mary-Beth Fafard (Brown University), Dr. Margaret Goertz (University of Pennsylvania), Dr. Bev McCoun (Mount Horeb Area School District), Dr. Margaret McLaughlin (University of Maryland), Dr. James Nuttall (Michigan Department of Education), Marshall Peter (Consortium for Appropriate Dispute Resolution in Special Education [CADRE]), Dr. Beth Rous (University of Kentucky), and Dr. Terrance Scott (University of Florida).

[1] Schiller, E., O'Reilly, F., and Fiore, T. (2006, April). *Marking Progress of IDEA Implementation*. Bethesda, MD: Abt Associates Inc.

Contents

Chapter 1
Selection of Sample Districts

This appendix describes the procedures for selecting the sample of districts included in the SLIIDEA data collection. The SLIIDEA design documents produced by the American Institutes for Research (AIR) as part of an earlier study called for a sample of 1,000 school districts and 3,800 schools. AIR anticipated an 80 percent response rate for districts and schools and planned on conducting analyses based on data collected from 800 school districts. However, at the completion of the design phase of their study, AIR had recruited only 324 districts to participate (out of a sample of 1,000). At that time, the Office of Special Education Programs (OSEP) invited Abt Associates to prepare several options for increasing the sample size.[2] OSEP selected an option that involved selection of a supplemental sample of districts with the intention of having a total of 800 recruited districts. The following strategy was used to obtain 800 recruited districts:

- AIR had already recruited 324 districts. Therefore, 476 (800–324 = 476) additional districts needed to be recruited.
- To recruit 476 districts, the plan was to select a supplemental sample of 635 districts and assume that we would get a 75 percent agree-to-participate rate (635 * 0.75 = 476).

Selection of the 635 new districts involved the following considerations:

- If possible, the sampling frame for the supplemental sample of districts was to consist of those districts that had not already been selected as part of AIR's original sample. We believed that it would be more difficult and more expensive to convert districts that had already refused to participate into participating districts than it would be to entice new districts to participate.
- Districts that were participating in one of two other OSEP-sponsored studies, the Special Education Elementary Longitudinal Study (SEELS) and the Study of Personnel Needs in Special Education (SPeNSE), were excluded from the sampling frame. It was believed that the response burden would be too great if a district that was already in SEELS or SPeNSE was asked to participate in SLIIDEA.
- Certain types of districts were not appropriate for inclusion in the sample (i.e., those that did not have regular schools associated with them, such as federally-operated institutions). National Center for Education Statistics (NCES) classifications in the Common Core of Data (CCD), described in more detail below, were used to identify the type of districts to be included.

The sampling frame for the supplemental sample of districts and sampling method are discussed below. The sampling frame discussion includes a description of the Office for Civil Rights (OCR) data, a listing of the reasons why districts were not included in the sampling frame, and a description of the region and urbanicity classifications of the districts.

---

[2] Options were presented to OSEP in a June 20, 2000 document titled *Study of State and Local Implementation and Impact of the Individuals with Disabilities Education Act*: *Task 6.1.2 Alternative Sampling Plans*.

The Sampling Frame of Districts for the Supplemental Sample

The sampling frame of districts was created from several sources. The preliminary list consisted of all records in the 1997–98 CCD. For reasons described below, we omitted many of the records from the preliminary list from the final sampling frame. Decisions about which records would be omitted and which would be retained in the sampling frame were based on information contained in the 1997–98 CCD; the 1997 and 1998 OCR data sets; the list of 1,004 districts that AIR had originally selected; and lists of districts in the SEELS and SPeNSE studies.

The OCR Data

The samples of districts in each of the 1997 and 1998 OCR databases were about a third of the total number of districts in the country in each year. A district could be selected by OCR because it:

1. Was part of the probability sample
2. Had 25,000 or more students
3. Was in a state where the total number of districts in the state was 25 or fewer (DC, DE, HI, MD, NV)
4. Was under "court order/voluntary settlement"
5. Was in a state with 13 or fewer Regional Educational Service Agencies (RESAs)
6. Had been a non-respondent in two previous OCR selection cycles

Before the CCD and OCR databases were merged to create the sampling frame, the districts that were included in the OCR samples for reasons 4, 5 and 6 were removed. We refer to districts that entered the OCR sample for reasons 4, 5 or 6 as "forced-in" districts. The resulting samples included only those districts that had been selected because they were in the probability sample, were large (25,000+ students) or were in a state with 25 or fewer districts. A single data file was created from the two OCR databases (1997 and 1998). If a district had a record in both the 1997 and 1998 OCR databases, then the information from the more recent 1998 database was used.

Deletions from the Preliminary List to Create the Sampling Frame

Districts in the preliminary sampling frame list were omitted for the following reasons:

- The preliminary list included districts from places that were not part of the 50 U.S. states or the District of Columbia. For example, districts in Puerto Rico, Guam and the Virgin Islands were removed from the list.
- Education agencies that had closed were removed from the list.
- Districts that had a grade span listed in the CCD as 0000 were removed from the list. There were only two such agencies that would not have been removed for other reasons. It is believed that the children in these districts were counted in other districts that remained in the list.
- Agencies that were regional, state, federal or other (non-local) were removed from the list, based on the NCES classification in the CCD data set (see Exhibit 1.1).

- Districts with fewer than 100 students were removed from the list because it would be inefficient to recruit and mail to districts with so few students. Sufficient numbers of small districts remained in the sampling frame.
- Districts that were in SEELS or SPeNSE were removed from the list.
- Our intention was to remove all of the districts in AIR's original sample of 1,004 districts. However, AIR's original sample included all districts with more than 40,500 students. Some of these large districts had already been recruited and were included in the list of the 324 that had agreed to participate. Among the 324, there were three areas of the country that had very few of these large districts (the urban and suburban West, and the urban Northeast). We drew a sample of nine large districts from the list of districts that had already been contacted by AIR to ensure sufficient representation in these three areas. We therefore included large West urban, West suburban and Northeast urban districts that had already been selected by AIR in the sampling frame.

Finally, after all the deletions described above, we calculated the number of new districts we needed to select from each of 12 region-by-urbanicity strata. We determined that the OCR databases contained a sufficient number of districts to allow us to restrict the sampling frame to only those districts that were in those databases. The reason for this restriction was that the information in OCR would be useful for selecting a sample of schools from the sample of districts. Sampling schools from OCR districts allows us to over-sample schools that serve children with low-incidence disabilities.

Exhibit 1.1

Types of Districts Included and Excluded From Sampling Frame, by NCES Code

Included	1 = Local school district that is not a component of a supervisory union.[1]
Included	2 = Local school district component of a supervisory union sharing a superintendent and administrative services with other local school districts.
Included	3 = Supervisory union administrative center, or a county superintendent serving the same purpose.
Excluded	4 = Regional education services agency, or a county superintendent serving the same purpose.
Excluded	5 = State-operated institution charged, at least in part, with providing elementary and/or secondary instruction or services to a special need population.
Excluded	6 = Federally operated institution charged, at least in part, with providing elementary and/or secondary instruction or services to a special need population.
Excluded	7 = Other education agencies that do not fit into the first six categories.

[1] A supervisory union is defined as an educational agency where administrative services are performed for more than one school district by a common superintendent.

Region and Urbanicity Classifications

All of the sampled districts were assigned to region and urbanicity classifications, which were then used to stratify the sampling frame. Urbanicity was defined according to the metro status code variable on the CCD data set (see Exhibit 1.2). Region was defined according to the four Census regions (see Exhibit 1.3).

Exhibit 1.2

Urbanicity Classification

Classification	NCES classification of the agency's service area relative to a Metropolitan Statistical Area (CCD variable name MSC97)
Urban	1 = Primarily serves a central city of an MSA
Suburban	2 = Serves an MSA but not primarily its central city
Rural	3 = Does not serve an MSA

Exhibit 1.3

Region Classification

Classification	State Abbreviation
West	HI, WA, OR, MT, ID, WY, CA, NV, UT, CO, AZ, NM, AK
Northeast	NY, VT, ME, NH, MA, CT, RI, NJ, PA
South	TX, OK, AR, LA, KY, TN, MS, AL, WV, VA, MD, DC, DE, NC, SC, GA, FL
Midwest	ND, SD, NE, KS, MN, IA, MO, WI, IL, MI, IN, OH

Sampling Method

We selected a stratified systematic sample of 635 districts from 15 strata. The first 12 strata represented the region-by-urbanicity classifications and included districts that were not selected in AIR's original sample. Six hundred twenty-six districts were selected from these 12 strata. A sample of nine districts was selected from three additional strata that consisted of large districts that had already been selected in AIR's original sample but had not been recruited. As described earlier, this needed to be done to ensure sufficient representation of large districts in the West and Northeast.

The number of districts sampled from each of the 12 region-by-urbanicity strata was determined by allocation proportional to the number of students.3 Prior to sampling, districts were sorted by size within each stratum. Systematic sampling after sorting by size increases the likelihood of having a wide distribution of district sizes in the selected sample.

Exhibit 1.4 shows the total number of districts in the target population, sampling frame, and sample. Contributions to the total are given separately based on the original 333 districts in the AIR sample (S1-districts), and the 626 supplemental districts in the Abt Associates sample (S2-districts).

[3] A simple example of proportional allocation is as follows. Suppose you have two strata (A and B) and want to take a sample of 10 districts. Suppose there are 300 students in stratum A and 700 students in stratum B. You would sample three districts from A and seven districts from B.

Exhibit 1.4

Number of Districts in the Population, Frame and Sample

	S1-Districts	S2-Districts	Total
Target Population	1,004	12,314	13,318
Sample Frame	1,004	11,190	12,194
Selected Sample[1]	333	626	959

[1] The supplemental S2-District sample included 635 districts, but 9 of these were large districts from the West and Northeast that had originally been part of the S1-Districts sample. In this table, these nine districts were subtracted from the 635 S2-Districts and added to the 324 S1-Districts.

Chapter 2
Initial Sampling Weights for Districts

This appendix describes the calculation of initial sampling weights for districts. These weights would be used in the analyses only if we obtained responses from all 959 districts in the sample. Chapter 3 describes adjustments in the weights to deal with district non-response. Computation of initial sampling weights for districts is complex because of the way in which we arrived at our sample of 959 districts. In the design phase, the design contractor, American Institutes for Research (AIR), had selected 1,004 districts, of which they were able to recruit 324. Abt Associates subsequently selected a supplemental sample of 635 districts, nine of which were intended to fill out region-by-urbanicity cells that did not have enough large districts. The remaining 626 districts in the supplemental sample were selected from a sample of the districts in the combined 1997 and 1998 Office of Civil Rights (OCR) database. The two components of the sample are:

S1-districts: These 333 districts are in the final sample and were also in AIR's original sample. They include 324 districts that AIR recruited plus nine additional districts that Abt Associates reselected in order to fill out three region-by-urbanicity cells that did not contain enough large districts.

S2-districts: These 626 districts are in Abt Associates' supplemental sample of districts, but were never part of the AIR sample. They include the 635 districts selected in the supplemental sample, minus the nine large districts that were selected in the supplemental sample but had originally been part of AIR's sample.

The Size of the Target Population

The target population is the population of districts to which we want our sample to generalize. The target population is broader than the sampling frame because, for example, even though we excluded districts that were included in OSEP's Special Education Elementary Longitudinal Study (SEELS) and the Study of Personnel Needs in Special Education (SPeNSE) from the sampling frame for the supplemental sample, we would like our results to generalize to the whole population that includes those districts. The size of the target population is 13,318 districts and consists of the following:

- Type 1, 2 and 3 districts as classified by the National Center for Education Statistics in the Common Core of Data (i.e., excludes regional, state, federal or other non-local districts)
- Districts in the 50 states and the District of Columbia
- Districts that have schools and students associated with them (i.e., the number of schools in the districts is at least one and the grade span of students served exists)
- Districts that were not closed
- Districts with 100 or more students

The calculation of sampling weights is such that the sum of all the weights is equal to 13,318 (the number of districts in the target population). The calculation of sampling weights is made more complex because districts in SEELS and SPeNSE were excluded from the S1 and S2 sampling frames.[4] The rationale for excluding them was that we did not want to overburden those districts by selecting them for multiple studies. These districts are counted as belonging to the target population of 13,318 even though they could not be in the sample, and treated as if they were missing at random.

Exhibit 2.1 shows the size of the sampling frame and the target population. The first row shows that 11,190 districts comprised the sampling frame for the S2-districts. There were 1,124 districts that were excluded from the S2 sampling frame because they were in SEELS or SPeNSE. There were 1,004 districts in the AIR sample. The cumulative sum of those three classifications of districts is 13,318. Those are the districts we are treating as our target population. There were 1,245 districts with fewer than 100 students that were excluded from the S2 sampling frame and are *not* considered to be part of the target population.

Exhibit 2.1

Size of S2-District Sampling Frame and Target Population

District Classification	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Sample Frame for S2	11,190	76.84	11,190	76.84
SEELS/SPeNSE	1,124	7.72	12,314	84.56
In AIR sample (S1)	1,004	6.89	13,318	91.45
< 100 students[a]	1,245	8.55	14,563	100.00

[a] = Not in target population

Initial Sampling Weights for S1-Districts

AIR's original sample of 1,004 districts were treated as certainty units. That is, each district had a probability of selection equal to one. The 333 S1-districts were treated as a stratified simple random sample of the 1,004 districts. The strata are the 12 region-by-urbanicity categories. Sampling weights for S1-districts are calculated as follows. Let nh be the population number of school districts (that were in AIR's original sample) in the hth stratum. Let mh be the number of school districts in the hth stratum that are in the S1 sample. Then the weight for the jth district in the hth stratum is calculated as:

$$w_{hj} = 1 \big/ \Pr(d_{hj} \in S1) = 1 / \frac{m_h}{n_h} = \frac{n_h}{m_h}. \quad \textbf{[Equation 2.1]}$$

Where $\Pr(d_{hj} \in S1)$ is the probability that the j^{th} district in the h^{th} stratum is in sample S1. Note that

$\sum_{h=1}^{12} n_h = 1{,}004$, $\sum_{h=1}^{12} m_h = 333$, and $\sum_{h=1}^{12} \sum_{j \in S1} w_{hj} = 1{,}004$.

The sum of the weights over all the districts in sample S1 is 1,004. In other words, these 333 districts represent 1,004 districts. Over all i districts in S1, we have:

$$\sum_{i \in S1} w_i = 1{,}004.$$

[4] Only those districts in SEELS and SPeNSE that were smaller than 40,500 were excluded.

Initial Sampling Weights for S2-Districts

S2-districts were selected from a sampling frame constructed from the 1997-1998 CCD and the 1997 and 1998 OCR databases. The OCR database was such that if a district was in either of the 1997 or 1998 samples, then it was included in the OCR sample (if a district was in both samples, then the record from 1998 was used). The probability of selection of an S2-district is equal to the probability that the district was selected conditional on the district being in the combined OCR sample, multiplied by the probability that the district was in the combined OCR sample. We use this notation:

$$\Pr(d_i \in S2) = \Pr(d_i \in S2 \mid d_i \in OCR) \times \Pr(d_i \in OCR)$$ **, [Equation 2.2]**

where di is district i.

Exhibit 2.2 shows the size of the sampling frame for S2-districts. The exhibit shows that 5,145 of the 11,190 districts in the original S-2 sampling frame were in the combined OCR sample. The sample of S-2 districts was selected from this group of 5,145 because information from the OCR would be useful for selecting a sample of schools.

Exhibit 2.2

Size of Sampling Frame of S2-Districts

	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Not in OCR	6,045	54.02	6,045	54.02
In OCR	5,145	45.98	11,190	100.00

The Conditional Probability of Selection

SAS Proc Survey was used to select the sample of S2-districts from the sample of 5,145 OCR districts. The SAS procedure calculated probabilities of selection for each selected district. The SAS output data set provides, for each S2-district, the probability of selection conditional on the district having been in OCR. That is, we know this part of Equation 2.2:

$$\Pr(d_i \in S2 \mid d_i \in OCR)$$

The result of summing the inverse of the conditional probability of selection over all of the districts in S2 is 5,145. Now we turn to the problem of calculating the probability that a district was in OCR.

The Probability of Selection into OCR

We obtained the probability of selection for each district in the 1997 and 1998 OCR databases from the OCR contractor. Here, we calculate the probability that a district was in the combined OCR sample. That is, we need to calculate this part of the equation:

$$\Pr(d_i \in OCR)$$

The logic here for getting the probability that a district was selected to be in either the 1997 or 1998 (or both) OCR databases is as follows:

- Let Pr(A) be the probability that a district was selected to be in the 1997 OCR database.
- Let Pr(B) be the probability that a district was selected to be in the 1998 OCR database.
- Let Pr(A or B) be the probability that a district was in the combined OCR database.

Then, Pr(A or B) = Pr(A) + Pr(B) - Pr(A and B)

Selection in 1997 was independent of selection in 1998. Therefore, Pr(A and B) = Pr(A) * Pr(B) Now, consider a district that is in the combined OCR database. That is, the district is in A or B. Suppose that the district was in A, but not B. We know Pr(A) because it was attached to the file obtained from the OCR contractor, but we do not know Pr(B). Therefore, we can't calculate Pr(A or B) as shown above. We do know that the sampling strategies were very similar for the 1997 and 1998 samples. Therefore, we used the following method to approximate the unknown Pr(B): If we don't know Pr(B) for a district (its probability of selection in 1998), we use Pr(A) in its place (its probability of selection in 1997). Our assumption is that a district's probability of selection in 1998 will be very similar to its probability of selection in 1997. Likewise, if we know Pr(B) but don't know A, then we approximate Pr(A) by Pr(B).

How good are our estimated probabilities of selection? The sum of the inverse of the probabilities of selection into OCR over the 5,145 districts that were in OCR and in the S2 sampling frame is 10,024. We expect the sum of these weights to be close to 11,190. These two numbers are not closer for two reasons. The first is that our approximation method might be inexact. The other is that there was some non-response in the OCR sample. Not every district that was selected to be in the 1997 OCR sample responded, nor did every district in the 1998 sample respond. Our approach was to adjust the probabilities of selection into OCR so that the sum of their inverses across the 5,145 districts would be 11,190. We calculated those adjustments within the 12 region-by-urbanicity categories.

Let w_{hj}^{adj} be the new weight for the j^{th} district in the h^{th} stratum after the adjustment. And let w_{hj} be the weight before adjustment. These weights are the inverse of the approximated probability of selection into OCR that has been calculated up to this point (i.e., the weights that sum to 10,024). Let n_h be the population number of school districts in the h^{th} stratum. Let m_h be the number of school districts in the sample. Then the new adjusted weight for the j^{th} district in the h^{th} stratum is calculated as:

$$w_{hj}^{adj} = w_{hj} \times \frac{n_h}{\sum_{j=1}^{m_h} w_{hj}}. \qquad \textbf{[Equation 2.3]}$$

Note that $\sum_{h=1}^{12} n_h = 11{,}190$, $\sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj} = 10{,}024$, and $\sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj}^{adj} = 11{,}190.$

Let new, adjusted probabilities be denoted as:

$$\Pr(d_i \in OCR) = 1/w_i^{adj} ,$$

where the i subscript corresponds to all the districts in all strata.

The Unconditional Probability of Selection

We now have both parts of the right-hand side of Equation 2.2, and can therefore calculate, for each district in the S2 sample, the left-hand side:

$$\Pr(d_i \in S2).$$

As usual, the weights will be equal to the inverse of the probability of selection,

$$w_i^* = 1/\Pr(d_i \in S2). \qquad \textbf{[Equation 2.4]}$$

The weights in Equation 2.4 are starred ("*") to differentiate them from the adjusted weights described in the next section. Summing the weights over all districts in S2 gives

$$\sum\nolimits_{i \in S2} w_i^* = 11{,}163.73$$

This is close to the population size of 11,190.

Reweighting so That S2-Districts Represent the Population Including SEELS and SPeNSE

We now adjust the weights shown in Equation. 2.4 so that they sum to 12,314. This is the number of districts in the S2 sampling frame plus the districts that were excluded because they were in SEELS or SPeNSE. The rationale for this adjustment is that it is reasonable to treat the S2 sample of districts as representative of the population that includes the SEELS and SPeNSE districts. If we treat the sample as if it represents the full population including the SEELS and SPeNSE districts, then we avoid repeated explanations in the results sections of our reports that the SEELS and SPeNSE districts were excluded from the sampling frame. The weighting adjustment uses the same method as described in Equation 2.3. After adjustment, summing the new weights over all districts in S2 gives:

$$\sum_{i \in S2} w_i = 12,314 .$$

Sum of the Weights of S1- and S2-Districts

Summing over the weights of all 959 districts in the sample gives

$$\sum_{i \in S1 \cup S2} w_i = \sum_{i \in S1} w_i + \sum_{i \in S2} w_i = 1,004 + 12,314 = 13,318 .$$

These weights represent the initial sampling weights. Subsequent adjustments are necessary to adjust for unit non-response. These adjustments are described in Chapter 3.

Chapter 3
Reweighting Districts for Non-response

Chapter 2 describes how the initial sample of districts was weighted to represent the target population. Here we discuss adjustments to the weights when districts failed to respond to the survey. The original sample consisted of 959 districts. By the time of the Year 4 data collection (2004-2005 school year) some districts had changed names, merged with other districts, or split. Districts that changed names, but retained their original Common Core of Data (CCD) ID number retained the sampling weight of the originally selected district. Of the original 959 districts, three were closed and not replaced, so the Year 4 sample consisted of 956 districts (959-3=956). Two districts merged, one district split, and three districts were closed and replaced. In the first case, one district replaced two, in the second case two districts replaced one, and in the last case each district was replaced by one other district. After these changes, the number of districts in the Year 4 sample remained at 956.

Of the 956 districts surveyed, 849 (89 percent) returned completed surveys. To account for unit non-response, weighting-class adjustments were made to the sampling weights (Lohr, 1999)5. Weighting classes were formed from variables with known values for all districts in the sample: district size, region, and urbanicity. The method assumes that respondents and non-respondents are similar within weighting classes. Within weighting classes, the sampling weights of responder districts are inflated to represent the non-responder districts. This method assumes a missing at random (MAR) non-response mechanism, and that, conditional on the weighting class variables, the probability of non-response is unrelated to the survey items. Fortunately, with the high 89 percent response rate, even if this assumption is not strictly true, the potential for non-response bias is low.

Fifty-one weighting classes were formed by crossing four levels of region (Northeast, South, Midwest, and West) by three levels of urbanicity (urban, suburban, rural) by five levels of district size (where the number of students is <1,500; 1,501–3,500; 3,501–8,000; 8,001–20,000; or >20,000). Some particular combinations of region, urbanicity, and size were not found in the population, therefore there were 51 weighting classes instead of the 60 that would be expected by the crossing of four levels of region by three levels of urbanicity by five levels of size. For example, there were no Midwestern, rural districts of size greater than 20,000. Within each class, the weights of responder districts were inflated by a factor equal to the inverse of the estimated probability of response within each class. The probability of response within each class is estimated by:

$$\hat{f}_c = \frac{\text{sum of weights for respondents in class } c}{\text{sum of weights for selected sample in class } c}. \qquad \textbf{[Equation 3.1]}$$

The inflation factor is equal to $\frac{1}{\hat{f}_c}$.

Let w_{ic} = the original sampling weight of responder district *i* in weighting class *c*. And let $\tilde{w}_{ic}$ = the inflated weight of responder district *i* in weighting class *c*. Then

$$\tilde{w}_{ic} = w_{ic} \times \frac{1}{\hat{f}_c}. \qquad \textbf{[Equation 3.2]}$$

[5] Lohr, S.L, (1999). *Sampling: Design and Analysis*. Brooks/Cole Publishing Company, Pacific Grove CA.

The sum of the inflated weights, $\tilde{w}_{ic}$, over all responding districts is equal to the number of districts in the target population. Exhibit 3.1 shows the number of sampled districts and responder districts in each of 51 weighting classes. Also shown are the sums of the original sampling weights, the inflation factor, and the sums of the inflated weights for districts in each weighting class. The exhibit shows that the sum of the inflated weights is equal to the total number of districts in the target population.

Exhibit 3.1

Inflation of Weights Within Weighting Classes Defined by Region, Urbanicity, and District Size

Weighting Class	Number in Sample	Sum of Original Weights	Number of Responder Districts	Sum of Original Weights	Inflation Factor	Sum of New Weights
NE Urb 1501-3500	2	13.10	2	13.10	1.00000	13.10
NE Urb 3500-8000	11	37.16	7	23.14	1.60569	37.16
NE Urb 8001-20000	8	27.57	8	27.57	1.00000	27.57
NE Sub <=1500	39	842.30	38	814.99	1.03351	842.30
NE Sub 1501-3500	45	702.54	39	622.80	1.12803	702.54
NE Sub 3501-8000	38	368.91	34	331.58	1.11260	368.91
NE Sub 8001-20000	4	39.24	3	36.03	1.08928	39.24
NE Rur <= 1500	16	564.37	15	561.09	1.00586	564.37
NE Rur 1501-8000	5	185.62	4	182.34	1.01802	185.62
SO Urb <= 1500	5	21.96	4	18.47	1.18901	21.96
SO Urb 1501-3500	6	26.15	6	26.15	1.00000	26.15
SO Urb 3501-8000	17	48.07	16	44.83	1.07231	48.07
SO Urb 8001-20000	10	29.81	8	24.34	1.22434	29.81
SO Urb > 20000	19	61.02	17	54.54	1.11887	61.02
SO Sub <= 1500	38	457.16	36	438.03	1.04369	457.16
SO Sub 1501-3500	20	171.60	20	171.60	1.00000	171.60
SO Sub 3501-8000	37	187.90	31	159.90	1.17516	187.90
SO Sub 8001-20000	26	107.68	23	95.97	1.12208	107.68
SO Sub > 20000	13	44.65	10	34.35	1.30000	44.65
SO Rur <= 1500	43	1187.88	38	1036.84	1.14567	1187.88
SO Rur 1501-3500	34	518.12	30	437.63	1.18391	518.12
SO Rur 3501-8000	27	272.03	25	236.83	1.14866	272.03
SO Rur 8001-20000	7	35.97	6	14.86	2.42029	35.97
MW Urb <= 1500	2	25.87	2	25.87	1.00000	25.87
MW Urb 1501-3500	3	15.01	3	15.01	1.00000	15.01
MW Urb 3501-8000	13	41.32	9	26.62	1.55239	41.32
MW Urb 8001-20000	14	39.80	11	31.59	1.25979	39.80
MW Urb >20000	10	25.00	10	25.00	1.00000	25.00
MW Sub <= 1500	39	853.80	34	742.30	1.15021	853.80
MW Sub 1501-3500	56	764.48	49	649.56	1.17693	764.48
MW Sub 3501-8000	32	290.75	30	278.47	1.04410	290.75
MW Sub 8001-20000	12	72.12	10	66.27	1.08821	72.12
MW Sub > 20000	2	5.85	2	5.85	1.00000	5.85
MW Rur <= 1500	48	2296.36	43	1857.19	1.23647	2296.36
MW Rur 1501-3500	18	458.43	17	455.68	1.00605	458.43
MW Rur 3501-8000	9	91.21	9	91.21	1.00000	91.21
WE Urb <= 1500	9	82.45	8	74.71	1.10350	82.45
WE Urb 1501-3500	6	32.13	6	32.13	1.00000	32.13
WE Urb 3501-8000	7	17.70	7	17.70	1.00000	17.70
WE Urb 8001-20000	10	24.60	9	22.41	1.09744	24.60
WE Urb > 20000	8	42.13	7	36.50	1.15410	42.13
WE Sub <= 1500	37	421.58	29	338.32	1.24611	421.58
WE Sub 1501-3500	29	229.37	27	211.61	1.08393	229.37
WE Sub 3501-8000	34	194.16	30	163.50	1.18749	194.16
WE Sub 8001-20000	23	106.39	20	91.53	1.16237	106.39
WE Sub > 20000	18	63.50	16	56.98	1.11450	63.50
WE Rur <= 1500	19	888.46	16	754.18	1.17805	888.46
WE Rur 1501-3500	8	149.31	8	149.31	1.00000	149.31
WE Rur 3501-8000	10	82.34	9	80.31	1.02759	82.34
WE Rur 8001-20000	3	35.89	2	4.42	8.11719	35.89
Total	956	13318.01	849	11725.74	1.16784	13318.01

Abbreviations: Region: NE = Northeast, SO = South, MW = Midwest, WE = Wes
Urbanicity: Urb = Urban, Sub = Suburban, Rur = Rural
Size: <1,200 students; 1,500-3,500 students, 3,501-8,000 students, 8,001-20,000 students, >20,000 students

Chapter 4
Selection of Sample Schools

Sampling Schools Within Sampled Districts

This chapter describes procedures for selecting the sample of schools. Recall that the selection of districts was based on a systematic sample (based on district size) within a stratified sample (based on 12 region-by-urbanicity strata). Both types of sampling had the effect of increasing the spread of districts across all combinations of region, urbanicity, and size.

In the second stage of sampling we sampled schools from the districts selected at stage one. This was done separately for each school level (elementary, middle and high) to enable separate analyses for these subpopulations. Thus, sampled schools were clustered (or nested) within sampled districts. First, we wanted the sample of schools to be representative of all public schools in the country. Second, we wanted to select a sample of elementary, middle and high schools from each of the selected districts. Third, we wanted to over-sample schools serving large numbers of children with sensory and physical disabilities because of the relatively low prevalence of such schools. Finally, we wanted to avoid enormous variation in school selection probabilities because this could lead to large variance estimates, and wide confidence intervals.

We originally intended to sample two elementary schools, two middle schools and two high schools from each district, with an emphasis on sampling schools with large numbers of sensory and physically disabled students. When sampling two or more schools of the same type (e.g., two elementary schools with similar numbers of sensory and physically disabled students), we sampled systematically by school size. The two-stage cluster sampling plan for a given school level (e.g., elementary schools) is depicted in Exhibit 4.1.

Exhibit 4.1

Two-Stage Cluster Sampling Plan for Each School Level[a]

Stage	Sampling Unit	Stratification Factors	Systematic Factors	Oversampling Factors
1	district	region, urbanicity	district size	—
2	school		school size	sensory and physically disabled students

[a] This sampling plan was used for elementary, middle, and high schools.

The School Sampling Frame

Frame and Target

Recall that a total of 959 districts were selected in the first sampling stage. This was the sum of 333 districts selected from the AIR district sampling frame, and 626 districts selected from the Abt Associates district sampling frame. The school sampling frame comprised the 13,900 schools that were in the 959 selected districts and in the 1997–98 NCES Common Core of Data (CCD), minus 665 schools that were considered inappropriate for reasons discussed below. Thus the school sampling frame consisted of 13,235 appropriate schools from the 959 sampled districts.

The school sampling frame is the list of schools from which the sample was actually drawn, and the school target population is the list of schools that the sample is supposed to represent. Because selected schools were nested in selected districts, the school target population is larger than the school sampling frame. Ignoring minor exceptions, the school sampling frame consisted of all 13,900 CCD schools in the 959 sampled districts, while the school target population consisted of all 82,947 CCD schools in all 13,318 districts. In other words, the school target population consisted of all CCD schools in the country (defined as the 50 states and the District of Columbia).

Exclusions from the Frame

The sampling frame originally included 13,900 schools. One hundred twenty-nine schools were omitted from the sampling frame because they were closed; 150 schools were omitted because they served only kindergarten or pre-kindergarten children (they did not have any students in grades 1–12); 153 schools were omitted because they were very small (15 or fewer students); and 233 schools were omitted because they were listed in the CCD database as “grade span=0000.” For the latter group of schools, we had no information on the number of children that were served or the grade levels of children served. After these exclusions, the sampling frame consisted of 13,235 schools.

Target, Frame and Sample: Elementary, Middle and High Schools

As mentioned previously, the sample was not actually drawn from the all-schools frame of 13,235 schools, but rather from three distinct frames: schools serving elementary students (8,853 schools), schools serving middle students (3,763 schools), and schools serving high school students (2,580 schools). Exhibits 4.2, 4.3, and 4.4 show the numbers of schools in the target population, sample frame, and selected sample for these three groups of schools.

Exhibit 4.2

Elementary Schools

	S1-Schools (from 333 districts in the AIR sample)	S2-Schools (from 626 districts in the Abt Associates sample)	Total
Target Population	18,320	34,360	52,680
Sample Frame	6,306	2,547	8,853
Selected Sample	905	1,272	2,177

Exhibit 4.3

Middle Schools

	S1-Schools (from 333 districts in the AIR sample)	S2-Schools (from 626 districts in the Abt Associates sample)	Total
Target Population	7,221	17,450	24,671
Sample Frame	2,692	1,071	3,763
Selected Sample	674	884	1,558

Exhibit 4.4

High Schools

	S1-Schools (from 333 districts in the AIR sample)	S2-Schools (from 626 districts in the Abt Associates sample)	Total
Target Population	5,263	13,821	19,084
Sample Frame	1,787	793	2,580
Selected Sample	589	678	1,267

Target, Frame and Sample: All Schools Combined

Exhibit 4.5 shows the number of schools in the frame and sample for All Schools Combined. The sampling frame for All Schools Combined (13,235 schools) is smaller than the sum of the three separate elementary, middle and high school frames (8,853 + 3,763 + 2,580 = 15,196 schools) because a school serving both elementary and middle students, for example, is in both the elementary and middle frames, but appears only once in the All Schools Combined frame. Similarly, the sample for All Schools Combined (4,534 schools) is smaller than the sum of the elementary, middle and high schools samples (2,177 + 1,558 + 1,267 = 5,002 schools). This is because some schools teaching multiple levels were also sampled for multiple levels.

Exhibit 4.5

Multiple School Levels

Type of School	Sampling Frame (1)	Sample – Levels Served (2)	Sample – Levels Selected (3)
Elementary School only	7,617	1,901	1,954
Middle School only	2,096	1,026	1,157
High School only	1,855	946	1,016
Elementary and Middle School	942	257	156
Elementary and High School	0	0	6
Middle and High School	431	236	184
Elementary, Middle and High School	294	168	61
Total	13,235	4,534	4,534

Not all schools serving multiple levels were selected for all of their levels. This distinction appears in Exhibit 4.5 as sampled schools *in terms of the levels served* (column two) and sampled schools *in terms of the levels sampled* (column three). For example, if a school serving both elementary and middle students is sampled, it need not be in both the elementary and middle sample. Rather it will be in either the elementary-only sample, the middle-only sample or the elementary-and-middle sample. Suppose the school was selected for the elementary-only sample. Then in terms of Exhibit 4.5, the school would appear on the "Elementary and Middle School" row for the sampling frame (column one), and for the sample in terms of levels taught (column two), but would appear on the "Elementary School Only" row for the sample in terms of levels sampled (column three).

The fact that some schools serving multiple levels were not selected for all of their levels makes some of the schools in the lower rows of column three migrate upward to form column four. This explains the apparent incongruity whereby six schools were in the elementary-and-high-school sample, even though no such schools existed in the sampling frame. These schools actually served all three levels, but were selected in the elementary and high school samples and were not selected for the middle school sample. Hence, these schools appear in the "Elementary, Middle and High School" row of column two, and the "Elementary and High School" row of column three.

Exhibit 4.6 shows, for All Schools Combined, the total number of schools in the target, frame, sample and responder sample. Contributions to the total are given separately based on the 333 S1-districts in the AIR district sample, and the 626 S2-districts in the Abt Associates district sample.

Exhibit 4.6

All Schools

	S1-Schools (from 333 districts in the AIR sample)	S2-Schools (from 626 districts in the Abt Associates sample)	Total
Target Population	27,141	55,806	82,947
Sample Frame	9,295	3,940	13,235
Selected Sample	1,978	2,556	4,534

Sampling Method

The sample for All Schools Combined can be broadly construed as a two-stage cluster design, with stratification, systematic sampling and over-sampling as depicted in Exhibit 4.7.

Exhibit 4.7

The Two-Stage Cluster Design for All Schools

Stage	Sampling Unit	Stratification Factors	Systematic Factors	Oversampling Factors
1	district	region, urbanicity	district size	—
2	school	school level	school size	sensory and physically disabled students

The stratification by school level should not be taken literally, however, because schools serving multiple levels (13 percent of all schools) appeared in multiple sampling frames. In fact, the second stage of the sampling plan was executed separately for each school level, based on the two-stage cluster design presented in Exhibit 4.1.

Sampling School Levels

Our original intention was to select two schools from each level (two elementary, two middle, two high) from each district selected in the first stage of the sample. In fact, many districts had too few schools to make this possible, and the sampling plan evolved accordingly. The typical situation was one in which many districts had enough elementary schools but too few middle and high schools. We sampled what was possible, but this left a shortfall of schools in the sample. For example, if a district had four elementary, one middle, and one high school, we sampled two elementary, one middle and one high school, leaving a deficit of two schools. As the sample proceeded in this fashion, we needed to increase the desired sample per district to, say, five schools from each level. The final result was that we obtained more elementary schools than middle and high schools. Of the 4,534 selected schools, 2,177 were elementary schools, 1,558 were middle schools and 1,267 were high schools.

This imbalance occurred, of course, because elementary schools are far more numerous (in the frame and the population) than middle and high schools (see Chapter 3). The effect was to redistribute the sample of 5,002 schools (2,177 elementary, 1,558 middle, 1,267 high) from an intended design of equal allocation (1,668 elementary, 1,667 middle, 1,667 high, say) toward a design of proportional allocation (2,914 elementary, 1,239 middle, 849 high). The common proportion here is 0.33, being the sum of the samples (5,002) divided by the sum of the frames (15,196).

Over-sampling Sensory and Physical Disabilities

Thirty-nine percent of schools in the sampling frame served children with sensory disabilities. The percentages for physical, behavioral and cognitive disabilities were 62 percent, 70 percent and 92 percent, respectively. It seemed sensible therefore, to over-sample schools serving children with sensory and physical disabilities. Schools serving children with cognitive and behavioral disabilities were not over-sampled because these types of schools were more prevalent.

The details of the over-sampling algorithm are complicated, but the main point is that whenever more than two schools of a given level (e.g., elementary) existed in a district, schools with high numbers (10 or more students) of sensory disabled students and schools with high numbers of physically disabled students were given preference in the selection process. For example, if a district had 20 elementary schools—one with many sensory cases, one with many physical cases and 18 “other” schools—and we wanted to draw three of these schools, then the algorithm would have given us one of each school type (one with many sensory cases, one with many physical cases and one “other”).

Sampling School Size

In situations where one or more schools of a given level were to be selected from a larger group of a given type, schools were selected systematically by school size using Lahari’s circular method. This has the effect of ensuring a reasonable spread of school sizes in such cases. For example, suppose we wanted to draw a sample of three from seven similar elementary schools (e.g., all with less than 10 sensory and physical cases). We would rank the schools in order of size (1, 2, 3, 4, 5, 6, 7) pick a random start (2, say), and then choose every other school (4 and 6).

Chapter 5
Initial Sampling Weights for Schools

This appendix describes the calculation of the initial set of weights for schools. These weights are appropriate assuming all schools in the sample respond to the survey. In cases where unit non-response occurs, weights were inflated.

Recall from Chapter 1 that 959 districts were selected in the first stage of the sample. This was the sum of 333 S1-districts selected from the AIR district sampling frame of 1,004 districts, and 626 S2-districts selected from the Abt Associates district sampling frame of 11,190 districts. S1-schools correspond to the sample of S1-districts, and S2-schools correspond to the sample of S2-districts.

The 333 S1-districts included the 324 districts selected by AIR plus nine additional districts selected by Abt Associates from AIR's original sampling frame of 1,004 districts. The nine additional districts were selected in order to fill out region-by-urbanicity cells that did not have enough large districts. The 626 S2-districts included the 635 districts selected in the supplemental sample, minus the nine large districts that were selected in the supplemental sample that were originally part of AIR's sampling frame. S1-districts were weighted to represent the S1-target population of 1,004 districts, and S2-districts were weighted to represent the S2-target population of 12,314 districts. The sum of these weights is 13,318, the size of the total district target population.

Weights for S1-Schools

Conditional Weights for S1-Schools

A conditional weight is the inverse of the probability of selection of a school conditional on the school's district having been selected to be in the sample. The conditional weights were calculated in two separate SAS programs. The first program calculated weights for the schools selected from the 324 districts recruited by AIR. The second program calculated weights for schools selected from the nine additional districts that were reselected in order to fill out region-by-urbanicity cells that did not have enough large districts.

Of the 324 districts recruited by AIR, three independent samples of schools were taken: a sample of elementary schools, a sample of middle schools and a sample of high schools. A sample of 662 elementary schools was selected from a sampling frame of 4,562 elementary schools. The conditional weights for those 662 sampled elementary schools were readily obtained from the output from the SAS program used to select the sample. The sum of the sampling weights for those 662 schools is 4,562. The conditional weights for middle and high schools were obtained the same way. The numbers of schools in the sampling frames, the numbers of schools selected and the sums of their conditional weights for elementary, middle and high schools are summarized in Exhibit 5.1.

Exhibit 5.1

Schools Selected from 324 Districts Recruited by AIR

School Level	Sums of Conditional Sampling Weights: Schools in the Sampling Frame	Schools Selected
Elementary Schools	4,562	662
Middle Schools	2,001	532
High Schools	1,261	498

In addition to conducting analyses at the school level, we wished to run analyses for all schools. This required calculation of the probability that a school was selected for any of the three samples:

Pr(E or M or H) = Pr(E) + Pr(M) + Pr(H)
– Pr(E and M) – Pr(E and H) – Pr(M and H) **[Equation 5.1]**
+ Pr(E and M and H).

Because the three samples were independent, all joint probabilities were the products of the marginal probabilities. Thus, Pr(E and M) = Pr(E)*Pr(M), Pr(E and H) = Pr(E)*Pr(H), Pr(M and H) = Pr(M)*Pr(H), and Pr(E and M and H) = Pr(E)*Pr(M)*Pr(H).

Combining the sampling frames for elementary, middle and high schools in the 324 districts recruited by AIR gave 6,729 unique schools (see Exhibit 5.2). A unique school is a school selected into at least one of the elementary, middle or high school samples. In this context, a school is counted only once even if it was selected for more than one sample; for example, a school teaching elementary and middle students could be selected for both elementary and middle school samples. A total of 1,579 unique schools were selected in the elementary, middle or high school samples. The program calculated the conditional probabilities of selection, Pr(E or M or H), and the initial weights (the inverse of these selection probabilities). The calculations were summed over the 1,579 unique schools. The sum of the weights was 6,815.25, which was too high by 86.25, since the 1,579 schools in the sample are supposed to represent the 6,729 schools in the combined sampling frame. The weights were correspondingly adjusted to equal 6,729.

Exhibit 5.2

Schools Selected From 324 Districts Recruited by AIR

School Level	Sums of Conditional Sampling Weights: Schools in the Sampling Frame	Schools Selected
All Schools Combined	6,729	1,579

After adjusting the weights, we ensured that no school had a weight of less than one. Out of the 1,579, 666 represented themselves (had sampling weights of one) and 913 represented the remaining 6,063 schools (6,729 – 666 = 6,063). However, the sum of the weights for the 913 schools was 6,149.25 (86.25 too many), so the weights were reduced by a factor of 6,063/6,149.25 for these 913 schools.

The new adjusted weights for the 1,579 schools summed to 6,729. The logic for calculation of weights for schools selected from the 324 districts was applied to the calculation of weights for schools selected from the nine supplemental districts (see Exhibit 5.3). Combining results from both programs, we obtain weights for all of the schools selected from S1-districts, as shown in Exhibit 5.4.

Exhibit 5.3

Schools Selected From Nine Supplemental Districts

School Level	Sums of Conditional Sampling Weights: Schools in the Sampling Frame	Schools Selected
Elementary Schools	1,744	243
Middle Schools	691	142
High Schools	526	91
All Schools Combined	2,566	399

Exhibit 5.4

Schools Selected From All 333 S1-Districts (324 + 9)

School Level	Sums of Conditional Sampling Weights: Schools in the Sampling Frame	Schools Selected
Elementary Schools	6,306	905
Middle Schools	2,692	674
High Schools	1,787	589
All Schools Combined	9,295	1,978

Unconditional Weights for S1-Schools

An unconditional weight is the inverse of the unconditional probability of selection, which is the school's conditional probability of selection multiplied by its district's probability of selection:

$$\Pr(s_{ij} \in S1) = \Pr(s_{ij} \in S1 \mid d_i \in S1) \times \Pr(d_i \in S1) \quad \textbf{[Equation 5.2]}$$

where d_i is district *i*, and s_{ij} is school *j* in district *i*. The weight variable associated with the probability that a district was selected is "wgt5". The sum of *wgt5* for S1-districts is 1,004 because the 333 S1-districts represent the 1,004 districts in the AIR frame. The conditional probability of selection:

$$\Pr(s_{ij} \in S1 \mid d_i \in S1)$$

is the inverse of the adjusted weight for the All Schools Combined weights (*allwgt2*). These are the selection probabilities output by SAS in the case of elementary, middle and high school samples. The program calculates the unconditional probability of selection,

$$\Pr(s_{ij} \in S1)$$

via the statements: scprb1=(1/wgt5)*(1/*allwgt2*), Elemprb1=(1/wgt5)*elemSelectionProb, Midlprb1=(1/wgt5)*midlSelectionProb, and Highprb1=(1/wgt5)*highSelectionProb. The weights are the inverses of the corresponding probabilities. Exhibit 5.5 shows the target population of schools in 1,004 districts; the sample of S1-schools represents this population. A unique school is a school selected into a least one of the elementary, middle or high school samples. Exhibit 5.6 shows that the sum of weights from sampled schools was slightly higher than the number of schools in the 1,004 districts, so school weights were adjusted to sum to the target population in the 1,004 districts. The following adjustments were made within the 12 region-by-urbanicity categories.

Exhibit 5.5

Numbers of Schools in the 1,004 Districts

School Level	Number of Schools
Unique Schools	27,141
Elementary Schools	18,320
Middle Schools	7,221
High Schools	5,263

Exhibit 5.6

The Sum of the Weights From the Sampled Schools

School Level	Sum of Weights
Unique Schools	30,022
Elementary Schools	20,427
Middle Schools	8,159
High Schools	5,678

Let w_{hj}^{adj} be the new weight for the j^{th} school in the h^{th} stratum after the adjustment. And let w_{hj} be the weight before adjustment. Let n_h be the population number of schools in the hth stratum. Let m_h be the number of schools in the sample. Then the new adjusted weight for the j^{th} school in the h^{th} stratum is calculated as:

$$w_{hj}^{adj} = w_{hj} \times \frac{n_h}{\sum_{j=1}^{m_h} w_{hj}} \qquad \textbf{[Equation 5.3]}$$

For unique schools we have

$$\sum_{h=1}^{12} n_h = 27{,}141 \;,\; \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj} = 30{,}021.66 \text{, and } \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj}^{adj} = 27{,}141,$$

while for elementary schools we have

$$\sum_{h=1}^{12} n_h = 18{,}320 \;,\; \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj} = 20{,}426.96 \text{, and } \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj}^{adj} = 18{,}320.$$

Similar adjustments were made for middle and high schools. The resulting adjusted weights are shown in Exhibit 5.7.

Exhibit 5.7

The Sums of Adjusted School Weights

School Level	Sum of Adjusted Weights
Unique Schools	27,141
Elementary Schools	18,320
Middle Schools	7,221
High Schools	5,263

Weights for S2-Schools

Conditional Weights for S2-Schools

Three independent samples of schools were taken from the Abt Associates supplemental district sample: a sample of elementary schools, middle schools and high schools. A sample of 1,272 elementary schools was selected from a sampling frame of 2,547 elementary schools. The weights for those 1,272 sampled elementary schools were obtained from the SAS program used to select the sample. The sum of the sampling weights for the 1,272 schools is 2,547. The weights for middle and high schools were obtained in the same way. The numbers of schools in the sampling frames, the numbers of schools selected and the sums of their conditional weights for elementary, middle and high schools are summarized in Exhibit 5.8.

Exhibit 5.8

Schools Selected from 626 Districts in the Abt Associates Supplemental Sample

School Level	Sums of Conditional Sampling Weights: Number of Schools in the Sampling Frame	Number of Schools Selected
Elementary Schools	2,547	1,272
Middle Schools	1,071	884
High Schools	793	678

In addition to conducting analyses at the school level, we wished to run analyses for all schools combined. This requires the calculation of the probability that a school was selected for any of the three samples, namely

$$\Pr(E \text{ or } M \text{ or } H) = \Pr(E) + \Pr(M) + \Pr(H) - \Pr(E \text{ and } M) - \Pr(E \text{ and } H) - \Pr(M \text{ and } H) + \Pr(E \text{ and } M \text{ and } H). \quad \textbf{[Equation 5.4]}$$

Because the three samples were independent, all joint probabilities were simply the products of the marginal probabilities. Thus, Pr(E and M) = Pr(E)*Pr(M), Pr(E and H) = Pr(E)*Pr(H), Pr(M and H) = Pr(M)*Pr(H), and Pr(E and M and H) = Pr(E)*Pr(M)*Pr(H).

Combining the sampling frames of elementary, middle and high schools in the 626 districts in the supplemental sample gave 3,940 unique schools (see Exhibit 5.9). There were 2,556 unique schools selected in the elementary, middle or high school samples. These calculations were then summed over the 2,556 unique schools. The sum of the weights was 3,948.18, which was 8.18 too many because the 2,556 sampled schools were supposed to represent the 3,940 schools in the sampling frame. The weights were adjusted to sum to 3,940.

After adjusting the weights, it was ensured that no school had a weight of less than one. Out of the 2,556, there were 1,707 representing themselves only (had a sampling weight of one) and 849 purporting to represent the remaining 2,233 schools (3,940 – 1,707 = 2,233). However, the sum of the weights for the 849 schools was 2,241.18 (8.18 too many), so weights were reduced by a factor of 2,233/2,241.18. The new adjusted weights for the 2,556 schools summed to 3,940, as it should.

Exhibit 5.9

Schools Selected From 626 Districts in the Abt Associates Supplemental Sample

School Level	Sums of Conditional Sampling Weights: Number of Schools in the Sampling Frame	Number of Schools Selected
Elementary Schools	2,547	1,272
Middle Schools	1,071	884
High Schools	793	678
All Schools Combined	3,940	2,556

Unconditional Weights for S2-Schools

Unconditional weights for S2-schools were calculated in a fashion similar to those for S1-schools. The probability of selection was:

$$\Pr(s_{ij} \in S2) = \Pr(s_{ij} \in S2 \mid d_i \in S2) \times \Pr(d_i \in S2), \qquad \textbf{[Equation 5.5]}$$

where di is district i, and sij is school j in district i. The sum of the weight variable for S2-districts is 12,314 because the 626 S2-districts represent the 12,314 S2-districts. The conditional probability of selection, $\Pr(s_{ij} \in S2 \mid d_i \in S2)$ is the inverse of the adjusted weight for the All School Combined weights. For the elementary, middle and high school samples, this is simply the probability. Exhibit 5.10 shows the population number of schools in the 12,314 districts. The sample of S2-schools is intended to represent the population of schools in the 12,314 districts.

Exhibit 5.10

Numbers of Schools in the 12,314 Districts

School Level	Number of Schools
Unique Schools	55,806
Elementary Schools	34,360
Middle Schools	17,450
High Schools	13,821

Exhibit 5.11 shows that the sum of the weights from the sampled schools was slightly lower than the number of schools in the 12,314 districts. School weights were adjusted in order that they sum to the target population of schools in the 12,314 districts. These adjustments were calculated within the 12 region-by-urbanicity categories.

Exhibit 5.11

Sum of the Weights From the Sampled Schools

School Level	Number of Schools
Unique Schools	51,263
Elementary Schools	31,130
Middle Schools	16,047
High Schools	12,744

Using a similar notation as in the adjustment of S1-schools, we let w_{hj}^{adj} be the new weight for the j^{th} school in the h^{th} stratum after the adjustment, and let w_{hj} be the weight before adjustment. Let n_h be the population number of schools in the hth stratum. Let m_h be the number of schools in the sample. Then the new adjusted weight for the j^{th} school in the h^{th} stratum is calculated as:

$$w_{hj}^{adj} = w_{hj} \times \frac{n_h}{\sum_{j=1}^{m_h} w_{hj}} \qquad \textbf{[Equation 5.6]}$$

For unique schools we have

$$\sum_{h=1}^{12} n_h = 55{,}806 \text{ , } \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj} = 51{,}263.05 \text{, and } \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj}^{adj} = 55{,}806$$

and for elementary schools we have

$$\sum_{h=1}^{12} n_h = 34{,}360 \text{ , } \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj} = 31{,}129.75 \text{, and } \sum_{h=1}^{12}\sum_{j=1}^{m_h} w_{hj}^{adj} = 34{,}360.$$

Similar adjustments were made for middle and high schools. The adjusted weights are shown in Exhibit 5.12.

Exhibit 5.12

Sums of Adjusted School Weights

School Level	Sum of Weights
Unique Schools	55,806
Elementary Schools	34,360
Middle Schools	17,450
High Schools	13,821

Summary of Initial Weights of S1-Schools and S2-Schools

The previous sections described the calculation of adjusted conditional sampling weights for S1 and S2 schools. We refer to these adjusted conditional sampling weights as the "initial sampling weights." The initial weights would be appropriate for analysis only if there were no unit non-response. In order to account for non-response, additional adjustments to the weights for responder-schools are necessary. Exhibit 5.13 shows the sums of the unconditional weights for the S1-School and S2-Schools. The sum of the weights of the combined samples of S1 and S2 schools is equal to the size of the target population of schools.

Exhibit 5.13

S1 and S2 Unconditional Weights Sum to the Number of Schools in the Target Population

School Level	Sum of Unconditional Weights		
	S1 Schools	S2 Schools	S1 + S2 Schools
Unique Schools	27,141	55,806	82,947
Elementary Schools	18,320	34,360	52,680
Middle Schools	7,221	17,450	24,671
High Schools	5,263	13,821	19,084

Chapter 6
Procedures for Determining Sampling Weights for Year 4 (Wave 4) Responding Schools

A two-stage sampling design was used to select the SLIIDEA sample of schools. In stage 1, a sample of 959 school districts was selected. In stage 2, samples of elementary, middle, and high schools were chosen from within the sample of 959 districts. Across the three school levels, 4,534 schools were selected from the sample of 959 districts. Those 4,534 schools represented a national population of 82,947 schools. A set of initial sampling weights were calculated, indicating the number of schools in the population represented by each sampled school. The initial weights are appropriate for use in the analysis only if there were no unit non-response (i.e., if every sampled school returned a completed school questionnaire). Given a certain amount of unit non-response, adjustments must be made to the weights of the responding schools so that estimates derived from those schools will be representative of the entire target population of schools. Those adjustments are described in this chapter.

Details of the sampling and derivation of initial sampling weights are included in Chapters 1, 4, and 5. Chapter 1 describes procedures for selecting school districts, and Chapter 4 describes procedures for the second stage sampling of schools, nested within the selected sample of districts. Derivation of the initial sampling weights for schools is described in Chapter 5.

The Year 4 (Wave 4) Sample of Schools and Response Rates

By the time of the Year 4 data collection, some of the 4,534 schools in the original sample had closed or consolidated with other schools. When possible, replacement schools were chosen. Replacement schools were chosen from the same school districts as the schools they replaced. If all schools in a given district were already in the sample, no replacement school was chosen. After accounting for closings, consolidations, and replacements, the size of the Year 4 sample was 4,412 schools. In districts where school closings or consolidations occurred, the sampling weights of both replacement schools and schools in the original sample were re-calculated by taking the inverse of the probability of selection into the Year 4 sample. After an additional minor adjustment,[6] the weights of the Year 4 sample of 4,412 schools summed to the original population total of 82,947 schools.

Surveys were mailed to the 4,412 schools in the Year 4 sample and 3,526 returned a completed questionnaire. Therefore, the overall school response rate was 80 percent (3526/4412 = 0.7992). In order to account for non-response, the sampling weights of the 3,526 responding schools were adjusted to represent the full target population using a response propensity scoring methodology (Sardal, Swensson, & Wretman, 1992).[7] After completion of the adjustments, the adjusted sampling weights of the 3,526 responding schools sum to the population total of 82,947 schools.

---

[6] Prior to the adjustment, the weights of the 4,412 schools in the Year 4 sample summed to 79,510. After multiplying each school's weight by a factor of (82,947/79,510), the new "initial weights" summed to 82,947. We call these the "initial weights" (for n=4,412 schools) to differentiate them from weights that were subsequently adjusted for non-response (for n=3,526 responding schools).

[7] Sarndal, C.E., Swensson, B, Wretman, J. (1992). Model Assisted Survey Sampling. Springer, New York.

Adjusting Sampling Weights for Non-response

The school sample was designed to support inferences and estimation for elementary, middle, and high schools, and for the combined population of schools across all three levels. Therefore, there are four sets of sampling weights. Although each set of weights requires its own adjustments for non-response, the methodology is the same for all four sets of weights. To avoid repetition, the discussion below focuses on adjustments to the all-schools sampling weights. The details section below shows numbers for all four sets of weights.

A propensity score methodology was used to adjust the sampling weights of the 3,526 responding schools. A logistic regression model was fit to data from the 4,412 sampled schools where the probability of response was modeled as a function of school demographic variables. The results of this model were used to calculate the estimated probability of response (or propensity to respond) for each of the 4,412 schools in the sample. The response propensities were used to create weighting classes. Schools within a particular weighting class had similar demographic characteristics and had similar propensities to respond. Within each weighting class, the weights of responding schools were adjusted such that they would sum to the population total number of schools in the class.

Step 1: Fit Models

For the logistic regression models, the response (dependent) variable is a dummy variable that took the value "1" for responding schools and took the value "0" for non-responding schools. The explanatory (independent) variables were school demographic variables constructed from sampling frame information (the NCES Common Core of Data (CCD)). Several models were fit, where each successive model dropped one non-significant term from the previous model. The final model indicated that the probability of being a responding school was related to school size, percent of students receiving free/reduced-price lunch, percent minority, percentage of students with IEPs, whether or not the school served middle school level students, whether or not the school served high school-level students, metro status (urban, suburban, rural), and region.

Unlike the remaining steps in the re-weighting procedure, the model fitting step was only done once, using data on responding and non-responding schools from the all-schools sample (i.e., it was not done separately for the elementary, middle, and high school samples).[8] The response and explanatory variables used in the logistic regression model are summarized in Exhibit 6.1. The results from the final model are summarized in Exhibit 6.2.

[8] Separate models were fit to the school sample at each level. However, the results at each level were quite similar and the added complexity of estimating propensities from four separate models did not appear to be justified. Therefore, the estimated propensities from the all-schools sample were used throughout.

Exhibit 6.1

Variables Used in Logistic Regression Model

Response Variable		**Values**
School Returned Survey		(Yes, No)
Explanatory Variables	**Variable Name**	**Values**
School Size	Sch_Size	(0-<250, 250-750, >750)
% Lunch	Sch_Lunch	(0-<25%, 25-50%, >50%)
% Minority	Sch_Minority	(0-<25%, 25-75%, >75%)
% IEP	Sch_IEP	(0-<8%, 8-15%, 15-25%, >25%)
Has elementary students	Sch_e	(Yes, No)
Has middle students	Sch_m	(Yes, No)
Has high students	Sch_h	(Yes, No
Metro Status	Sch_Metro	(Rural, Suburban, Urban)
Region	Region	(Northeast, Midwest, South, West)

Exhibit 6.2

Summary of Final Model Results

Type III Analysis of Effects

Variable	**DF**	**Wald Chi-Square**	**Pr > Chisq**
School size	2	21.7692	<.0001
% Lunch	2	11.2591	0.0036
% Minority	2	59.8694	<.0001
% IEP	3	95.5114	<.0001
Has middle students	1	6.4811	0.0109
Has high school students	1	13.7902	0.0002
Metro status	2	6.5952	0.0370
Region	3	13.3234	0.0040

The Logistic Procedure

Analysis of Maximum Likelihood Estimates

Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	2.2800	0.2446	86.8716	<.0001
School size 1	1	-0.5557	0.1336	17.2860	<.0001
School size 2	1	-0.3921	0.0992	15.6410	<.0001
% Lunch 1	1	-0.3959	0.1198	10.9237	0.0009
% Lunch 2	1	-0.2731	0.1100	6.1658	0.0130
% Minority 1	1	1.0354	0.1378	56.4620	<.0001
% Minority 2	1	0.4614	0.1169	15.5810	<.0001
% IEP 1	1	-1.4805	0.2142	47.7711	<.0001
% IEP 2	1	-0.6454	0.2068	9.7373	0.0018
% IEP 3	1	-0.5495	0.2127	6.6768	0.0098
Middle sch 1	1	-0.2076	0.0816	6.4811	0.0109
High sch 1	1	-0.3443	0.0927	13.7902	0.0002
Metro 1	1	0.1945	0.1143	2.8929	0.0890
Metro 2	1	-0.0696	0.0994	0.4913	0.4833
Region 1	1	-0.2741	0.1174	5.4515	0.0196
Region 2	1	-0.2781	0.1143	5.9163	0.0150
Region 3	1	0.0474	0.1099	0.1861	0.6661

Step 2: Use Model Results to Calculate Response Propensities

In Step 2, parameter estimates obtained from the fitted model were used to calculate the predicted probability that a school will respond to the survey. The logistic regression model is represented as:

$$\log\left(\frac{p_i}{1-p_i}\right) = b_0 + \sum_k b_{ki}\ , \quad \textbf{[Equation 6.1]}$$

where p_i is the probability that school *i* is a responding school, and the summation is over the *k* demographic predictor variables in the final model. The predicted probabilities were obtained by solving the previous equation for p_i, and substituting the parameter estimates from the fitted model in place of the parameters. The solution for the predicted probability for school *i* is given by:

$$\hat{p}_i = \frac{\exp(\hat{b}_0 + \sum_k \hat{b}_{ki})}{1 + \exp(\hat{b}_0 + \sum_k \hat{b}_{ki})} \quad \textbf{[Equation 6.2]}$$

Each school's predicted probability of response ($\hat{p}_i$) is called its "response propensity". Schools with similar response propensities have similar demographic characteristics. In particular, they are similar on the demographic characteristics that are most related to the probability of response.

Step 3: Group Schools with Similar Response Propensities into Weighting Classes

In this step, schools with similar response propensities were grouped into weighting classes. The weights of responding schools within a class were inflated so that the responding schools within the class represent the population that both the responding and non-responding schools within the class were originally sampled to represent. Exhibit 6.3 shows the distribution of response propensities for the all-schools, elementary, middle, and high school samples. The distributions of the propensity measures were similar across the all-schools, elementary, middle and high school samples. Therefore, the same boundaries were used for defining the weighting classes for each sample.

Exhibit 6.3

Distributions of Propensity Scores

All Schools (n=4,412)		Elementary Schools (n=2,140)		Middle Schools (n=1,524)		High Schools (n=1,242)	
Quantile	Estimate	Quantile	Estimate	Quantile	Estimate	Quantile	Estimate
100% Max	0.959500	100% Max	0.959500	100% Max	0.950616	100% Max	0.934457
99%	0.931857	99%	0.933312	99%	0.919644	99%	0.908410
95%	0.906467	95%	0.912149	95%	0.901630	95%	0.890894
90%	0.889245	90%	0.899725	90%	0.885360	90%	0.877365
75% Q3	0.860936	75% Q3	0.869698	75% Q3	0.853027	75% Q3	0.848180
50% Median	0.822617	50% Median	0.826927	50% Median	0.813716	50% Median	0.801693
25% Q1	0.763328	25% Q1	0.765469	25% Q1	0.761653	25% Q1	0.734410
10%	0.674781	10%	0.689793	10%	0.680406	10%	0.627318
5%	0.610897	5%	0.627536	5%	0.601400	5%	0.550013
1%	0.486617	1%	0.529066	1%	0.473879	1%	0.439655
0% Min	0.272482	0% Min	0.356567	0% Min	0.272482	0% Min	0.272482

Note: Some schools are classified as both elementary and middle, or both middle and high school. Hence, the sum of the number of elementary, middle and high schools is greater than the "all schools" total.

Weighting classes were formed to ensure that all schools within a class fell within a narrow range of propensity scores. The boundaries for the weighting classes were determined by creating approximately equal-interval propensity score groupings. The top and the bottom of each propensity interval differed by .09 to .22 points. The resulting five classes corresponded to propensities in the ranges of 27–50, 51–60, 61–70, 71–80, and 81–96 percent probability of response. Exhibit 6.4 shows the frequency and percent of schools that fell within each of the five weighting classes.

Exhibit 6.4

Number and Percent of Schools in Each of Five Weighting Classes (Results for All 4,412 Schools in the Sample)

Weighting Class	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1)Propensities .27-.50	58	1.31	58	1.31
2)Propensities .51-.60	141	3.20	199	4.51
3)Propensities .61-.70	373	8.45	572	12.96
4)Propensities .71-.80	1173	26.59	1745	39.55
5)Propensities .81-.96	2667	60.45	4412	100.00

Step 4. Within Weighting Class, Inflate Weights of Responding Schools to Sum to Population Total

Within each weighting class, the weights of all schools (both responders and non-responders) were summed. Next, the weights of just the responding schools were summed. Then, within each weighting class, new, adjusted weights of responding schools were calculated by multiplying the initial weights by a factor equal to the ratio of the sum of the weights of all schools to the sum of the weights of the responding schools. The adjusted weight for the i^{th} school in the j^{th} weighting class is represented symbolically by:

$$w_{ij}^{adj} = w_{ij} * \frac{\sum_{k \in responders\ \&\ nonresponders} w_{kj}}{\sum_{i \in responders} w_{ij}}, \textbf{ [Equation 6.3]}$$

where w_{ij} is the initial sampling weight for the i^{th} school in the j^{th} weighting class, the summation in the numerator is over all k schools in the set of responders and non-responders within weighting class j, the summation in the denominator is over all i schools in the set of responders in weighting class j, and there are j = 1,…, 5 weighting classes. The new, adjusted sampling weights sum to the population total number of schools. This result can be written symbolically as:

$$\sum_j \sum_{i \in responders} w_{ij}^{adj} = \sum_j \sum_{k \in responders\ \&\ nonresponders} w_{kj} \quad \textbf{[Equation 6.4]}$$

Exhibit 6.5 shows that the weights of the 4,412 schools in the sample sum to the total number of schools in the target population (N=82,947). Exhibit 6.6 shows the weights summed within each of the five weighting classes. The sum shown for the j^{th} weighting class (j = 1,…,5) corresponds to the numerator in the second term of Equation 6.3, i.e., the term:

$$\sum_{k \in responders\ \&\ nonresponders} w_{kj}$$

For example, the all-schools weights of the 58 schools in the first weighting class sum to:

$$\sum_{k \in responders\ \&\ nonresponders} w_{k1} = 1020.45\,.$$

Exhibit 6.5

Size of Target Population (Sum of Initial Weights for All 4,412 Schools in the Sample)

Weight Variable	N	Sum	Minimum	Maximum
nscwgt2	4412	82947.01	1.4907206	253.9748455
nElemwgt2	2140	52680.02	1.7015548	256.1317276
nMidlwgt2	1524	24671.00	1.6353688	180.4918077
nHighwgt2	1242	19083.98	1.5188006	143.2633565

nscwgt2 = the initial weight for all-schools sample (sums to the population total number of schools)
nElemWgt2= the initial weight for Elementary school sample (sums to the population total number of elementary schools)
nMidlwgt2= the initial weight for Middle school. Sample (sums to the population total number of middle schools)
nHighWgt2= the initial weight for High school sample (sums to the population total number of high schools)

Exhibit 6.6

Size of Target Population Within Each Weighting Class (Sum of Weights for All 4,412 Schools in the Sample)

Weighting Class	N	Variable	N	Sum	Minimum	Maximum
1)Propensities .27-.50	58	nscwgt2	58	1020.45	2.1814183	84.8449376
		nElemwgt2	10	423.9272795	3.8570094	107.7126039
		nMidlwgt2	24	369.7481912	2.1830459	78.5695753
		nHighwgt2	39	553.0691783	2.4710216	53.5322514
2)Propensities .51-.60	141	nscwgt2	141	2626.98	2.3473838	164.7946335
		nElemwgt2	60	1409.36	3.4098380	173.6639858
		nMidlwgt2	50	892.2100592	2.1490056	82.1846857
		nHighwgt2	59	939.9286254	1.9958256	80.7900644
3)Propensities .61-.70	373	nscwgt2	373	8047.93	1.7863324	224.0932866
		nElemwgt2	177	5238.76	2.4723364	210.7775610
		nMidlwgt2	121	2192.77	2.2157827	126.3979353
		nHighwgt2	125	2061.78	1.5188006	143.2633565
4)Propensities .71-.80	1173	nscwgt2	1173	21820.32	1.7863324	253.9748455
		nElemwgt2	521	12999.76	1.7015548	256.1317276
		nMidlwgt2	444	7180.03	1.6353688	140.3825171
		nHighwgt2	389	6762.80	1.5188006	109.7926221
5)Propensities .81-.96	2667	nscwgt2	2667	49431.33	1.4907206	228.8958631
		nElemwgt2	1372	32608.21	1.7015548	232.1915455
		nMidlwgt2	885	14036.23	2.0855625	180.4918077
		nHighwgt2	630	8766.41	1.5188006	98.3978663

nscwgt2 = initial weight for all-schools sample (sums to population total number of schools within class)
nElemWgt2= initial weight for Elementary school sample (sums to population total number of elementary schools within class)
nMidlwgt2= initial weight for Middle school sample (sums to population total number of middle schools within class)
nHighWgt2= initial weight for High school sample (sums to population total number of high schools within class)

The sum of the initial weights of the 3,526 responding schools is shown in Exhibit 6.7. We call them "initial weights" because they are the sampling weights prior to adjustment for non-response. The weights of the 3,526 responding schools sum to a number that is smaller than the size of the target population. Exhibit 6.8 shows the weights summed within each of the five weighting classes. The sum shown for the j^{th} weighting class (j = 1,…,5) corresponds to the denominator in the second term of Equation 6.3, i.e., the term:

$$\sum_{i \in responders} w_{ij} \; .$$

For example, the initial all-schools weights of the 20 responding schools in the first weighting class sum to:

$$\sum_{i \in responders} w_{i1} = 462.3130931$$

Exhibit 6.7

Initial (Unadjusted) Weights of Responding Schools (Sum of Initial Weights for All 3,526 Responding Schools)

Variable	N	Sum	Minimum	Maximum
nscwgt2	3526	65421.32	1.4907206	253.9748455
nElemwgt2	1711	41378.80	1.7015548	256.1317276
nMidlwgt2	1130	17745.89	1.6353688	180.4918077
nHighwgt2	944	14579.61	1.5188006	143.2633565

nscwgt2 = the initial weight for all-schools sample
nElemWgt2 = the initial weight for Elementary school sample
nMidlwgt2 = the initial weight for Middle school sample
nHighWgt2 = the initial weight for High school sample

Exhibit 6.8

Initial (Unadjusted) Weights of Responding Schools by Weighting Class (Sum of Initial Weights for All 3,526 Responding Schools)

Weighting Class		N	Variable	N	Sum	Minimum	Maximum
1)Propensities	.27-.50	20	nscwgt2	20	462.3130931	2.3738576	84.8449376
			nElemwgt2	3	180.2665102	3.8570094	107.7126039
			nMidlwgt2	5	141.6080041	6.5329269	70.3208589
			nHighwgt2	12	213.3778241	3.0483010	46.9607668
2)Propensities	.51-.60	73	nscwgt2	73	1381.40	2.3473838	164.7946335
			nElemwgt2	39	876.0170243	3.4098380	173.6639858
			nMidlwgt2	22	374.1515511	2.7301068	78.5695753
			nHighwgt2	18	270.6306243	1.9958256	50.6644522
3)Propensities	.61-.70	261	nscwgt2	261	5502.67	2.0760592	224.0932866
			nElemwgt2	128	3784.57	2.5049443	210.7775610
			nMidlwgt2	71	1222.88	2.2531520	126.3979353
			nHighwgt2	79	1352.29	1.9958256	143.2633565
4)Propensities	.71-.80	889	nscwgt2	889	15929.59	1.7863324	253.9748455
			nElemwgt2	384	9214.34	1.7015548	256.1317276
			nMidlwgt2	323	5034.78	1.6353688	105.2943000
			nHighwgt2	286	5134.48	1.5188006	109.7926221
5)Propensities	.81-.96	2283	nscwgt2	2283	42145.35	1.4907206	228.8958631
			nElemwgt2	1157	27323.60	1.7015548	232.1915455
			nMidlwgt2	709	10972.47	2.0855625	180.4918077
			nHighwgt2	549	7608.83	1.5188006	98.3978663

nscwgt2 = the initial weight for all-schools sample
nElemWgt2= the initial weight for Elementary school sample
nMidlwgt2= the initial weight for Middle school sample
nHighWgt2= the initial weight for High school sample

The inflation factors for each of the five weighting classes are shown in Exhibit 6.9. The inflation factors correspond to the second term of Equation 6.3, i.e., the term:

$$\frac{\sum_{k \in responders\ \&\ nonresponders} w_{kj}}{\sum_{i \in responders} w_{ij}}.$$

For example, the inflation factor for the all-school weights for the first weighting class is:

$$\frac{1020.45}{462.3130931} = 2.20727$$

Exhibit 6.9

Inflation Factors Within Weighting Classes

Weighting Class	All Schools	Elementary Schools	Middle Schools	High Schools
1)Propensities .27-.50	2.20727	2.35167	2.61107	2.59197
2)Propensities .51-.60	1.90168	1.60883	2.38462	3.47311
3)Propensities .61-.70	1.46255	1.38424	1.79312	1.52466
4)Propensities .71-.80	1.36980	1.41082	1.42609	1.31713
5)Propensities .81-.96	1.17288	1.19341	1.27922	1.15214

Exhibits 6.10 and 6.11 show the sums of the non-response adjusted weights for the 3,526 responding schools, overall and by weighting class. The adjusted weights sum to the size of the target population. The numbers shown in Exhibits 6.10 and 6.11 correspond to the left-hand side of Equation 6.4:

$$\sum_{j} \sum_{i \in responders} w_{ij}^{adj}$$

For example, the non-response adjusted all-school weights of the 3,526 responder schools sum to

$$\sum_{j} \sum_{i \in responders} w_{ij}^{adj} = 82{,}947.01,$$

and, for example, the non-response adjusted all-school weights of the 20 responding schools in the bottom weighting class sum to

$$\sum_{i \in responders} w_{i1}^{adj} = 1020.45.$$

Exhibit 6.10

Sum of Nonresponse Adjusted Weights for 3,526 Responding Schools

Variable	N	Sum	Minimum	Maximum
nscwgt3	3526	82947.01	1.7484326	347.8942271
nElemwgt3	1711	52680.02	2.0306495	361.3553425
nMidlwgt3	1130	24670.99	2.3321768	230.8891731
nHighwgt3	944	19083.99	1.7498654	218.4276473

nscwgt3 = the non-response adjusted weight for all-schools sample
nElemWgt3= the non-response adjusted weight for Elementary school sample
nMidlwgt3= the non-response adjusted weight for Middle school sample
nHighWgt3= the non-response adjusted weight for High school sample

Exhibit 6.11

Sum of Nonresponse Adjusted Weights for 3,526 Responding Schools by Weighting Class

Weighting Class		N	Variable	N	Sum	Minimum	Maximum
1)Propensities	.27-.50	20	nscwgt3	20	1020.45	5.2397457	187.2757183
			nElemwgt3	3	423.9272794	9.0704118	253.3044607
			nMidlwgt3	5	369.7481912	17.0579192	183.6125759
			nHighwgt3	12	553.0691783	7.9011085	121.7209559
2)Propensities	.51-.60	73	nscwgt3	73	2626.99	4.4639716	313.3865689
			nElemwgt3	39	1409.36	5.4858401	279.3953408
			nMidlwgt3	22	892.2100591	6.5102730	187.3587461
			nHighwgt3	18	939.9286253	6.9317124	175.9629718
3)Propensities	.61-.70	261	nscwgt3	261	8047.93	3.0363403	327.7476360
			nElemwgt3	128	5238.76	3.4674486	291.7671111
			nMidlwgt3	71	2192.77	4.0401708	226.6466052
			nHighwgt3	79	2061.78	3.0429519	218.4276473
4)Propensities	.71-.80	889	nscwgt3	889	21820.32	2.4469144	347.8942271
			nElemwgt3	384	12999.77	2.4005847	361.3553425
			nMidlwgt3	323	7180.03	2.3321768	150.1587424
			nHighwgt3	286	6762.80	2.0004644	144.6116344
5)Propensities	.81-.96	2283	nscwgt3	2283	49431.33	1.7484326	268.4667927
			nElemwgt3	1157	32608.20	2.0306495	277.0993088
			nMidlwgt3	709	14036.23	2.6678984	230.8891731
			nHighwgt3	549	8766.41	1.7498654	113.3677633

nscwgt3 = the non-response adjusted weight for all-schools sample
nElemWgt3= the non-response adjusted weight for Elementary school sample
nMidlwgt3= the non-response adjusted weight for Middle school sample
nHighWgt3= the non-response adjusted weight for High school sample

Chapter 7
Procedures for the Calculation of Unconditional Percents

This appendix describes the analytic procedures used to calculate unconditional percents for survey follow-up questions. The unconditional calculations from the School Questionnaire and the District Questionnaire were comparable, and the steps presented below for the district data are the same as those used for school data. Many items on the state,[9] district, and school surveys are in the form of a gateway question with a follow-up question. An example of a gateway question is, "Has your district publicly reported test scores of students with IEPs participating in state- or district-wide assessments?" A follow-up is, "If yes, how does your district report the data for students with IEPs?" In *Volume I: The SLIIDEA Sourcebook Report (1999–2000, 2002–2003, 2003–2004, and 2004–2005 School Years)* and *Volume II: Data Tables for the Draft SLIIDEA Sourcebook Report*, we report the percentage of all districts responding to the survey question. The calculations for the follow-up questions are not conditional on how the districts responded to the gateway. The results reported unconditionally to the example question above are presented in Exhibit 7.1.

Exhibit 7.1

Percentage of Districts Publicly Reporting Test Scores of Students With IEPs on State- or District- Wide Assessments (2002–2003 School Year)

	For Students Participating in State- or District-Wide Assessments: Percent (SE)
Reported scores of students w/IEPs	**78.3** (2.3)
Aggregated only	31.2 (2.7)
Both aggregated and disaggregated	43.6 (2.9)
Disaggregated only	3.4 (1.1)
Scores not reported	**21.7** (2.3)

Note: The data correspond to those presented in Exhibit 6.6, page 108, of the *Final 2nd Interim Report (2002-2003 School Year).*

In this example, 78 percent of districts reported that they publicly reported test scores of students with IEPs participating in state- or district-wide assessments. Forty-four percent of districts indicated that the test scores of students with IEPs participating in state- or district-wide assessments were reported both aggregated with other students and disaggregated, 31 percent reported that the scores were aggregated with other students only and 3 percent indicated that the scores were reported only in a disaggregated fashion" (from the Final 2nd Interim Report [2002–2003 School Year], Exhibit 6.6,

[9] For data from the State Questionnaire, e*stimated* numbers and percentages were calculated using the number of states that answered the gateway question as the denominator. If a state answered "yes" to the gateway question, but left the follow-up question blank, these responses were proportionally distributed among the follow-up response categories; if the gateway question was left blank, it was left as missing for the follow-up question. The percent of states that answered "no" to the gateway question was added as a new category when presenting results for the follow-up question; since this percentage is not inflated, it will always agree with the percentage presented in the gateway question.

page 108). In this example, the percent of districts reporting test scores aggregated with other students, disaggregated or both are calculated for all districts responding to the survey question.

Analytic Process for Unconditional Calculations

To calculate the unconditional percentages, both the gateway question and follow up questions need to be included in the calculations. We outline the process for each one below, using an example item from the district survey - Q10A of the 2002–2003 District Questionnaire.

Percentage Calculation for the Gateway Question

The original coding of the gateway question (Q10A of the District Questionnaire) was 1=”yes”, 2=”no”, 7=’’refused”, 8=”don’t know”, 9=”not ascertained”. The analysis variable was a recoded version of Q10A, named nQ10A and created as follows:

nQ10a=Q10a;

if Q10a in (7,8,9) then nQ10a=.;

The percent “yes” on the gateway question was calculated as follows:

10371.27 / 13249.51 = 0.7828 (see Exhibit 7.2).

Likewise, the percent “no” was calculated as follows:

2878.238 / 13249.51 = 0.2172 (see Exhibit 7.2).

Exhibit 7.2

Weighted Frequencies for nQ10a

Q10A	nQ10AA	Frequency	Cumulative Frequency
1	1	10371.27	10371.27
2	2	2878.238	13249.51
9	.	68.50393	13318.01

Unconditional Percentage Calculations for the Follow-Up Question

The unconditional percents were calculated by inflating the weights of the respondents to the follow-up question by an appropriate inflation factor. SAS was used to create the unconditional variables and inflated weights (see Exhibit 7.3). Sudaan, a statistical software package, was used with the inflated weights to calculate the unconditional percents and corresponding standard errors. An example follows. To calculate the appropriate inflation factor, we used the following formula:

Inflation factor = (Sum of weights where gateway question = Yes) divided by (Sum of weights where gateway question = Yes AND follow-up question was not missing)

Using the numbers from the first table in Exhibit 7.4, the inflation factor was calculated as follows:

(5291.734 + 3793.022 + 417.432 + 869.0828)/(5291.734 + 3793.022 + 417.432) = 1.0915

Numerator for inflation factor was calculated as follows:

(5291.734 + 3793.022 + 417.432 + 869.0828) = 10371.2708

Denominator for inflation factor was calculated as follows:

(5291.734 + 3793.022 + 417.432) = 9502.18

Exhibit 7.3

SAS Code to Create Variables, Inflate Weights

```
data temp2; set temp;
nQ10a=Q10a;
if Q10a in (7,8,9) then nQ10a=.;
ucQ10aa=Q10aa;
if nQ10a=.                          then ucQ10aa=.;
if nQ10a=1 and Q10aa in (.,7,8,9)   then ucQ10aa=.;
if ucQ10aa=3                        then ucQ10aa=1;
if ucQ10aa=4                        then ucQ10aa=2;
if ucQ10aa=5                        then ucQ10aa=3;
if nQ10a=2                          then ucQ10aa=4;

inflatewt=Yr2DWgt;
if ucQ10aa in (1,2,3) then inflatewt=Yr2DWgt * 1.0915;
run;
```

Exhibit 7.4 shows the distribution of the unconditional variable using the original weights and the inflated weights. Exhibit 7.5 shows how we arrived at the unconditional estimates for “both aggregated and disaggregated,” “aggregated only,” and “disaggregated only” in our example follow-up question. Exhibit 7.6 shows the estimates and standard errors obtained from Sudaan when using the inflated weights.

Exhibit 7.4

Distribution of the Unconditional Variable Created in Exhibit 7.3

```
Proc freq data=temp2;
Tables ucQ10aa*Q10a*Q10A/list missing:
Weight Yr2DWgt: *original weights;
run:
```

The FREQ Procedure

ucQ10aa	Q10AA	nQ10A	Q10A	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	.	.	9	68.50393	0.51	68.50393	0.51
.	9	1	1	869.0828	6.53	937.5867	7.04
1	3	1	1	5291.734	39.73	6229.32	46.77
2	4	1	1	3793.022	28.48	10022.34	75.25
3	5	1	1	417.432	3.13	10439.77	78.39
4	.	2	2	2878.238	21.61	13318.01	100.00

```
proc freq data=temp2;
tables ucQ10aa*Q10aa*nQ10a*Q10a/list missing;
weight inflatewt; * Inflated Weights;
run;
```

ucQ10aa	Q10AA	nQ10A	Q10A	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	.	.	9	68.50393	0.48	68.50393	0.48
.	9	1	1	869.0828	6.13	937.5867	6.61
1	3	1	1	5775.927	40.71	6713.514	47.32
2	4	1	1	4140.084	29.18	10853.6	76.50
3	5	1	1	455.627	3.21	11309.23	79.71
4	.	2	2	2878.238	20.29	14187.46	100.00

Exhibit 7.5

Unconditional Percents Created Using Variables From Exhibit 7.3

```
proc freq data=temp2;
tables ucQ10aa/list missing;
weight inflatewt;
where ucQ10aa in (1,2,3,4);
format ucQ10aa freqz.;
run;
```

ucQ10aa	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Both aggregated and disaggregated	5775.927	43.59	5775.927	43.59
Aggregated only	4140.084	31.25	9916.011	74.84
Disaggregated only	455.627	3.44	10371.64	78.28
No public reporting	2878.238	21.72	13249.88	100.00

Note that these percents are exactly the same as the ones reported in Exhibit 6.6 (page 108) in the *Final 2nd Interim Report*.

Exhibit 7.6

Percents and Standard Errors From Sudaan

Q10A. Reporting of test scores of IEP students	Both		Aggregated Only		Disaggregated Only		No Response		Missing	Total
	%	SE	%	SE	%	SE	%	SE	55	782
	43.6	.29	31.2	2.7	3.4	1.1	21.7	2.4		

Note that these percents are exactly the same as the ones reported in Exhibit 6.6 (page 108) in the *Final 2nd Interim Report.*

Treatment of Missing Values in the Unconditional Analytic Process

There are two ways in which the data can be said to be "missing:" 1) a 'Yes' on the gateway question but missing on the follow-up question, and 2) missing responses on the gateway and on the follow-up question. The districts that answered 'Yes' on the gateway question but were missing on the follow-up were distributed in the same proportion as the districts that answered 'Yes' on the gateway question and responded on the follow-up question. The districts that were missing a response on the gateway item were distributed in a similar manner.

In our example gateway and follow-up question (Q10A from the District Questionnaire) 869.0828 districts responded 'yes' to the gateway question but were missing the follow-up question, and 68.50393 were missing the gateway question (see Exhibit 7.4). Of the 869.0828 districts missing on the follow-up, 483.9891 were distributed as reporting test data 'Both' aggregated with other students and disaggregated:

869.0828 * (5291.734 / (5291.734 + 3793.022 + 417.432)) = 483.9891

And of the 68.50393 districts missing on the gateway, 29.8622 were distributed as 'Both,' as follows:

68.5093 * (5291.734 + 483.9890562)[10]/ 13249.51 = 29.8622

So, the unconditional number of districts reporting 'Both' was calculated as:

5291.734 + 483.9890562 + 29.86221588 = 5805.5853

And the unconditional proportion of 'Both' was calculated as: 5805.5853 / 13318.01 = 0.4359.

[10] All available decimal places are used in calculations. Answers are rounded to four decimal places.

Chapter 8
District and School Change Analysis

This appendix explains how we tested the null hypothesis that the proportion reported from one wave of data is not significantly different than the proportion reported on the same (or similar) item from another wave of data. The method will be explained using an example in which the proportion of "yes" responses on Wave 4 district item 12a "public reporting of test scores of students with IEPs on state- or district-wide assessments " is compared to the proportion of "yes" responses on a similar item from the Wave 2 district item 10a.

Test of Difference in Proportions Between Two Waves of Data

A specification of the hypothesis test and the method for calculating the p-value is shown below. This method uses data from all districts that responded to either Q10a in Wave 2 or Q12a in Wave 4.

Let p_1 denote the estimated proportion from the wave 2 sample and let p_2 denote the proportion from the wave 4 sample.

$$H_o : p_2 - p_1 = 0 \quad vs \quad H_a : p_2 - p_1 \neq 0$$

Let the *z* be the standardized deviate, calculated as: $z = \dfrac{(p_2 - p_1) - 0}{SE(p_2 - p_1)}$ **[Equation 8.1]**

Compare *z* to the quantiles of a standard normal distribution, N (0,1), to find the two-sided probability of obtaining a deviate with absolute value that is as large or larger than *z*. If the absolute value of *z* is greater than 1.96, the null hypothesis will be rejected at the p<0.05 level. The challenging part of conducting this test is calculating the variance of the difference of the wave 4 and wave 2 proportions. A method for calculating the variance (and standard error) of the difference follows (Kish, 1965).[11]

Let p_1 denote the estimated proportion from the first sample of size n_1. Let p_2 denote the proportion from the second sample of size n_2. Let m denote the amount of overlap between the two samples. We are interested in testing the difference between the two sample proportions. We can write the estimated variance of the difference between the two sample proportions as:

$$v(p_1 - p_2) = v(p_1) + v(p_2) - 2m\,\mathrm{cov}(p_1, p_2)$$

$v(p_1)$ is the estimated variance of the first proportion based on a sample of n_1 units and $v(p_2)$ is the estimated variance of the second proportion based on n_2 units. Under simple random sampling, the estimated variance of the difference in two sample proportions becomes

$$v(p_1 - p_2) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} - 2m\frac{p_{12} - p_1 p_2}{n_1 n_2} \quad \textbf{[Equation 8.2]}$$

where p_{12} is the proportion of having the attribute in both the samples based on a sample of m units. For estimating the variance under the design, we can first estimate the variance under simple random

11 Kish, L. (1965). Survey Sampling. Wiley and Sons, NY.

sampling using the formula given above but with weighted proportions. Then, we multiply the variance by the design effect. The square root of the variance gives the standard error of the difference in the two proportions, which can be used in a statistical test recognizing that we have overlapping samples and the samples are not independent.

An alternative to this is to get the variance under the design from SUDAAN for the first proportion, the variance for the second proportion and to estimate the covariance we write the covariance terms as follows.

$$\text{cov}(p_1, p_2) = r_{p_1 p_2} \sqrt{v(p_1) v(p_2)}$$

For the variance of the two proportions we again use the values from SUDAAN and estimate the correlation based on the overlap using the weighted values.

Note that the term p_{12} in the algorithm is a proportion calculated from the overlapping sample. This should be a weighted proportion. But three sets of weights could be applied to these data: (1) Wave 2 sampling weights (adjusted for Wave 2 non-response), (2) Wave 4 sampling weights (adjusted for Wave 4 non-response), and (3) the average of Wave 2 and Wave 4 weights. We used the latter weights to calculate p_{12}. Similarly, the algorithm calls for multiplying the SRS variance estimate by the design effect. Again, we used an average design effect obtained from the two waves of data.

Exhibit 8.1 shows the wording of the Wave 2 and Wave 4 items, their estimates, and standard errors, and their difference, the standard error of the difference, and the p-value for a two-sided test of the null hypothesis that the difference is zero. There is insufficient evidence to say that the proportion of districts that publicly reported test scores of students with IEPs on state- or district-wide assessments has changed from the 2002–2003 school year to the 2004–2005 school year.

Exhibit 8.1

Results From Method 1

Item / Wave	Survey Items	Estimate % Yes	SE	p-value
12a / Wave 4	During the past 12 months, has your district publicly reported test scores of students with IEPs tested with and without accommodation on state- or district-wide assessments	82.4	2.1	
10a / Wave 2	During the past 12 months, has your district publicly reported test scores of students with IEPs on state- or district-wide assessments	78.3	2.3	
	Difference (Wave 4 – Wave 2)	4.1	3.0	0.17

Comparing Proportions Across Three Waves of Data

Let p_1 denote the estimated proportion from the Wave 1 sample, let p_2 denote the estimated proportion from the Wave 2 sample, and let p_3 denote the estimated proportion from the Wave 4 sample. We conduct the following three hypothesis tests using the method described above.

$$H_o : p_3 - p_1 = 0 \quad vs \quad H_a : p_3 - p_1 \neq 0$$

$$H_o : p_3 - p_2 = 0 \quad vs \quad H_a : p_3 - p_2 \neq 0$$

$$H_o : p_2 - p_1 = 0 \quad vs \quad H_a : p_2 - p_1 \neq 0$$

Comparing Means Instead of Proportions

In some cases we tested whether the difference between means from two waves of data is equal to zero. The text below shows a method for calculating the variance (and standard error) of the difference (Kish, 1965).[12]

Let $\bar{x}_1$ denote the estimated mean from the first sample of size n_1. Let $\bar{x}_2$ denote the mean from the second sample of size n_2. We are interested in testing the difference between the two sample means. We can write the estimated variance of the difference between the two sample means as:

$$v(\bar{x}_1 - \bar{x}_2) = v(\bar{x}_1) + v(\bar{x}_2) - 2\operatorname{cov}(\bar{x}_1, \bar{x}_2)$$

Under simple random sampling, the variance of the difference can be written as

$$v(\bar{x}_1 - \bar{x}_2) = v(\bar{x}_1) + v(\bar{x}_2) - \frac{2\, r_{x_1 x_2}\, m \sqrt{v(\bar{x}_1) v(\bar{x}_2)}}{\sqrt{n_1 n_2}}$$ **[Equation 8.3]** ,

where $v(\bar{x}_1)$ is the estimated variance of the first mean based on a sample of n_1 units, $v(\bar{x}_2)$ is the estimated variance of the second mean based on n_2 units and m is the amount of overlap between the two samples. The correlation ($r_{x_1 x_2}$) is estimated based on the overlap. We can get the variance under the design from SUDAAN for the first mean and for the second mean. The square root of the variance gives the standard error of the difference in the two means, which can be used in a statistical test recognizing that we have overlapping samples and they are not independent.

A specification of the hypothesis test and the method for calculating the p-value from the test is shown below.

$$H_o : \bar{x}_2 - \bar{x}_1 = 0 \quad vs \quad H_a : \bar{x}_2 - \bar{x}_1 \neq 0$$

The test statistic is:

$$t = \frac{(\bar{x}_2 - \bar{x}_1) - 0}{SE(\bar{x}_2 - \bar{x}_1)}$$ **[Equation 8.4]**

If the observed value of *t* as calculated above is greater than the critical value from the t-distribution with n –2 degrees of freedom and a=0.05, the null hypothesis will be rejected at the p<0.05 level.

[12] Kish, L. (1965). Survey Sampling. Wiley and Sons, NY.