

## **Appendix B:**

### **Methodology and Technical Notes**

This documentation was taken from the NCES report indicated below. The report can be found at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007361>.

Nolle, K.L., Guerino, P., and Dinkes, R. (2007). *Crime, Violence, Discipline, and Safety in U.S. Public Schools: Findings From the School Survey on Crime and Safety: 2005-06* (NCES 2007-361). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

This page is intentionally left blank

## Methodology and Technical Notes

The School Survey on Crime and Safety (SSOCS) is managed by the National Center for Education Statistics (NCES) on behalf of the U.S. Department of Education (ED). SSOCS collects extensive crime and safety data from principals and school administrators of U.S. public schools. Data from this collection can be used to examine the relationship between school characteristics and violent and serious violent crimes in primary schools, middle schools, high schools, and combined schools. In addition, data from SSOCS can be used to assess what crime prevention programs, practices, and policies are used by schools. SSOCS has been conducted in school years 1999–2000, 2003–04, and 2005–06. A fourth collection is planned for school year 2007–08.

SSOCS was developed by the National Center for Education Statistics (NCES) and is funded by the Office of Safe and Drug-Free Schools of the U.S. Department of Education. The 2005–06 SSOCS (SSOCS:2006) was conducted by the U.S. Census Bureau. Data collection began on March 17, 2006, when questionnaire packets were mailed to 3,565 public schools, and continued through May 26, 2006. A total of 2,724 public schools submitted usable questionnaires: 726 primary schools, 956 middle schools, 954 high schools, and 88 combined schools.

### Sample Design

The sampling frame for SSOCS:2006 was constructed from the 2003–04 NCES Common Core of Data (CCD) Public Elementary/Secondary School Universe data file. The CCD is an annual survey system of all public K–12 schools and school districts. Certain types of schools are excluded from the SSOCS:2006 sampling frame, including special education schools, vocational schools, alternative schools (e.g., adult continuing education schools and remedial schools), newly closed schools, home schools, ungraded schools, schools with high grades of kindergarten or lower, overseas Department of Defense schools, schools sponsored by the Bureau of Indian Affairs, and schools in the U.S. outlying areas<sup>1</sup> and Puerto Rico. Public charter schools are not excluded. The use of the CCD as a sampling frame in SSOCS:2006 deviates from the 1999–2000 SSOCS (SSOCS:2000) and the 2003–04 SSOCS (SSOCS:2004), which both utilized a modified version of the Schools and Staffing Survey (SASS) sampling frame. This deviation was necessary because SSOCS:2006 occurred between SASS collections.

The objectives of the SSOCS sampling design were twofold: to obtain overall cross-sectional and subgroup estimates of important indicators of school crime and safety and to yield precise estimates of change in these indicators between 1999–2000, 2003–04, and 2005–06. To attain these objectives, a stratified sample of 3,565 regular public schools was drawn for SSOCS:2006 using the same general sampling design as in the previous survey administrations for stratification variables, number of strata, method of sample allocation, and sorting of variables before selection.<sup>2</sup> As in the 2003–04 SSOCS, there was no attempt to minimize overlap between the SSOCS:2006 sample and samples for other NCES surveys.

---

<sup>1</sup>The U.S. outlying areas include the following: America Samoa, Guam, the Commonwealth of the Northern Mariana Islands, and the U.S. Virgin Islands.

<sup>2</sup>Adopting the same basic design for all survey administrations increases the precision of the estimates of change.

The initial goal of SSOCS:2006 was to collect data from at least 2,550 schools, taking nonresponse into account. One possible method of allocating schools to the different sampling strata would have been to allocate them proportionally to the U.S. public school population. However, while the majority of U.S. public schools are primary schools, the majority of school violence is reported in middle and high schools. Proportional allocation would, therefore, have yielded an inefficient sample design because the sample composition would have included more primary schools (in which crime is an infrequent event) than middle or high schools (in which crime is a relatively more frequent event). As a result, a larger proportion of the target sample of 2,550 schools was allocated to middle and high schools. The target sample was allocated to the four instructional levels as follows: 640 primary schools, 895 middle schools, 915 high schools, and 100 combined schools. Schools in the 1999–2000 SSOCS (SSOCS:2000) and SSOCS:2004 were allocated to instructional levels in a similar manner.

The same variables and categories used to create strata in SSOCS:2000 and SSOCS:2004 were used to create strata in SSOCS:2006. The population of schools was stratified (grouped) into four instructional levels,<sup>3</sup> four types of locale settings,<sup>4</sup> and four enrollment size categories.<sup>5</sup> These variables were chosen because they have been shown to be associated with school crime (Miller 2004). The sample of schools in each instructional level was allocated to each of the 16 cells formed by the cross-classification of the four categories of enrollment size and four types of locale. In order to obtain a reasonable sample size of lower enrollment schools while giving a higher probability of selection to higher enrollment schools, the sample was allocated to each subgroup in proportion to the sum of the square roots of the total student enrollment in each school in that stratum.

The effective sample size within each stratum was then inflated to account for nonresponse. Once the final sample sizes were determined for each of the 64 strata, the subgroups were sorted by region and percent minority enrollment, and an initial sample of 3,565 schools was selected. Sorting by these variables before selection has the same effect as stratification with proportional allocation of schools to the strata. For more information on the sample design, see chapter 2 of the *School Survey on Crime and Safety: 2005–06 Data File User's Manual* (Bauer et al. 2007).

## **Data Collection**

SSOCS:2006 was conducted as a mail survey with telephone follow-up. Four months before the onset of data collection, NCES began working with the school districts of sample schools that required prior approval to participate in the survey. On March 10, 2006, advance letters were sent to school administrators of sample schools that included the date of the first questionnaire mailing and a toll-free number to call with any questions. On March 17, 2006, questionnaires were sent via FedEx directly to the principals of the sample schools, with a cover letter describing the importance of the survey and a promotional SSOCS pen. See appendix D for a copy of the questionnaire.

---

<sup>3</sup>The four instructional levels are primary, middle, high, and combined.

<sup>4</sup>The four types of locales are city, urban fringe, town, and rural.

<sup>5</sup>The four categories of enrollment size are less than 300 students, 300–499 students, 500–999 students, and 1,000 students or more.

Upon distribution of the SSOCS questionnaire to schools, letters were mailed to chief state school officers and superintendents to inform them that schools within their states and districts, respectively, had been selected for SSOCS:2006. The letters included information about the survey and were accompanied by a promotional SSOCS pen, an informational copy of the questionnaire, and the SSOCS brochure. The letters were not designed to request permission from these officials to participate in the survey, but rather as a vehicle to enhance participation.

During the 2 weeks following the first questionnaire mailing, a screener telephone operation was conducted to verify that sample schools had received the questionnaire and were, in fact, eligible to participate. One week after the screener ended, a reminder telephone operation began, which was conducted in two 1-week phases. The primary objective of the reminder operation was to follow up with the principal or school contact to determine the status of the questionnaire; however, during the 2nd week, the interviewer could complete the SSOCS interview over the phone at the respondent's request. Data collection ended on May 26, 2006.

Returned questionnaires were examined for quality and completeness using both manual and computerized edits. If a questionnaire did not meet predetermined levels of completeness, the respondent was contacted again to resolve issues related to the missing data, irrespective of whether the items missing data were considered "critical."<sup>6</sup> The criteria used to determine completeness are detailed in section 3.1 of the *School Survey on Crime and Safety: 2005–06 Data File User's Manual* (Bauer et al. 2007). If a satisfactory resolution could not be reached, imputation was used to resolve data quality issues for questionnaires in which at least 60 percent of all items, 80 percent of critical items, 60 percent of item 16, and 60 percent of item 22 had been completed. Questionnaires that did not meet the imputation criteria were considered incomplete and are excluded from the analyses in this report.

## **Weighting**

Sample weights allow inferences to be made about the population from which the sample units were drawn. Because of the complex nature of the SSOCS:2006 sample design, weights are necessary to obtain population-based estimates, to minimize bias arising from differences between responding and nonresponding schools, and to calibrate the data to known population characteristics in a way that reduces sampling error. The procedures used to create the SSOCS sampling weights are described below.

An initial (base) weight was first determined within each stratum by calculating the ratio of the number of schools available in the sampling frame to the number of schools selected. In order to reduce the potential of bias from nonresponse, weighting classes were determined by using a statistical algorithm similar to CHAID (chi-square automatic interaction detector) to partition the sample so that schools within a weighting class were homogenous with respect to the probability of responding. The predictor variables for the analysis were school instructional level, region, enrollment size, percent minority enrollment, student-to-teacher ratio, percentage of students eligible for free or reduced-price lunch, and number of full-time-equivalent teachers. When the number of responding schools in a class was small, that weighting class was combined with

---

<sup>6</sup> The critical items in SSOCS:2006 included all subitems associated with items 7, 8, 14, 15, 16, 17, 20, 22, 23, 24, 25, 26, 28, 31, 32, and 33.

another class to avoid the possibility of large weights. After combining the necessary classes, the base weights were adjusted so that the weighted distribution of the responding schools resembled the initial distribution of the total sample.

The nonresponse-adjusted base sampling weights were then calibrated to agree with known population counts obtained from the sampling frame to reduce bias in the estimates due to undercoverage. The calibration process, a form of poststratification, separates the sample into a number of classes (poststrata), defined by a cross-classification of variables. The known population counts may be available for the individual cells of the cross-classification or only for certain margins of it. In the latter situation, the calibration proceeds iteratively, one margin at a time, and is often called “raking.” Poststratification works well when the noncovered population is similar to the covered population in each poststratum. Thus, to be effective, the variables that define the poststrata must be correlated with the outcome of interest (school crime, in this report). They must also be well measured in the survey, and the control totals must be available for the population as a whole. As in SSOCS:2004, these requirements were satisfied in SSOCS:2006 by the two margins set up for the raking ratio adjustment of the weights: (1) instructional level and school enrollment size and (2) instructional level and locale. All three variables—instructional level, school enrollment size, and locale—have been shown to be correlated with school crime (Miller 2004).

### **Imputation Procedures**

Files containing missing data can be problematic because, depending on how the missing data are treated, analysis of incomplete datasets may cause different users to arrive at different conclusions. Another problem with missing data is that certain groups of respondents may be more likely than others to skip survey items, creating bias in the survey estimates. Completed SSOCS:2006 surveys contained some level of item nonresponse after the conclusion of the data retrieval phase, and imputation procedures were used to create values for all questionnaire items with missing information.

The imputation methods utilized in SSOCS:2006 were tailored to the nature of the survey item. Four methods were used: aggregate proportions, best match, logical, and clerical. These methods are described in detail in section 4.4 of the *School Survey on Crime and Safety: 2005–06 Data File User’s Manual* (Bauer et al. 2007).

### **Unit Response Rates**

A unit response rate is, at its most basic level, the ratio of surveys completed by eligible respondents to the total count of eligible respondents. In some surveys, this calculation can be rather complicated because it is difficult to distinguish between eligible and ineligible units. For school surveys, however, the Department of Education updates its list of known schools on a fairly regular basis, so estimating eligibility among nonrespondents is relatively straightforward.

Unit response rates can be unweighted or weighted and are traditionally reported because they reflect the potential effects of nonsampling error and indicate whether portions of the population are underrepresented due to nonresponse. In order to calculate any of these measures, it is first necessary to know the disposition (outcome) of each sampled case. Table B-1 shows the

dispositions of the 3,565 cases initially selected for participation in SSOCS:2006, as well as the unweighted and weighted unit response rates by selected school characteristics.<sup>7</sup> The overall weighted unit response rate was 81 percent.

### **Analysis of Unit Nonresponse Bias**

The existence of nonresponding schools has the potential to introduce bias into survey estimates, depending on the magnitude of the nonresponse and whether differences exist between responding and nonresponding schools in characteristics related to the estimates of interest. A unit-level nonresponse bias analysis was conducted to evaluate the extent of this bias in SSOCS:2006. Responding and nonresponding schools were compared across the characteristics available for both groups: instructional level, enrollment size, type of locale, percent minority enrollment, region, number of full-time-equivalent teachers, student-to-teacher ratio, and percentage of students eligible for free or reduced-price lunch. This analysis indicated that there were no measurable differences between the responding schools and the full sample of schools, leading to the conclusion that nonresponse bias is not an issue. For more information on the analysis of unit nonresponse, please see section 3.6 of the *School Survey on Crime and Safety: 2005–06 Data File User’s Manual* (Bauer et al. 2007).

### **Item Response Rates**

Just as principals sometimes chose not to respond to the SSOCS:2006 survey request, they occasionally chose not to answer all of the survey items. Unweighted item response rates are calculated by dividing the number of sample schools responding to an item by the number of schools asked to respond to the item. Respondents sometimes gave answers to items in earlier portions of the SSOCS:2006 questionnaire that were inconsistent with answers in latter portions of the questionnaire. All inconsistent answers were blanked and imputed to maintain consistency. However, these cases were considered respondents when calculating the item-level response rates. Weighted item-level response rates in SSOCS:2006 were generally high, ranging from 66.3 percent to 100.0 percent. Of the 231 subitems on the SSOCS questionnaire, most (205) had response rates greater than 95 percent, 13 had response rates between 85 percent and 95 percent, and 13 had response rates less than 85 percent. The 13 subitems with response rates less than 85 percent are listed below:

- C0234—Number of part-time security guards
- C0236—Numbers of full-time school resource officers
- C0238—Numbers of part-time school resource officers
- C0242—Number of part-time sworn law enforcement officers
- C0326—Number of attacks with a weapon
- C0330—Number of attacks without a weapon
- C0406—School allows outside suspension with no services available
- C0408—School used outside suspension with no services available
- C0542—Number of paid part-time special education teachers

---

<sup>7</sup>While it is reasonable to assume that the ineligible rate among nonrespondents is not zero, a zero ineligibility rate was assumed when calculating the unweighted and weighted response rates. This is the most conservative approach.

Table B-1. Response status, unweighted and weighted unit response rates, by selected school characteristics, SSOCS:2006

School characteristic	Initial sample	Completed survey <sup>1</sup>	Non-respondents <sup>2</sup>	Ineligible <sup>3</sup>	Unweighted response rate (percent) <sup>4</sup>	Weighted response rate (percent) <sup>5</sup>
Total	3,565	2,724	789	52	77.5	81.3
Level						
Primary	909	726	168	15	81.2	83.0
Middle	1,254	956	275	23	77.7	79.9
High school	1,272	954	313	5	75.3	78.8
Combined	130	88	33	9	72.7	75.7
Enrollment size						
Less than 300	469	372	76	21	83.0	83.2
300–499	631	516	103	12	83.4	84.7
500–999	1,324	1,030	280	14	78.6	79.9
1,000 or more	1,141	806	330	5	71.0	72.5
Urbanicity						
City	1,014	697	295	22	70.3	75.4
Urban fringe	1,369	1,046	310	13	77.1	80.3
Town	332	281	48	3	85.4	86.7
Rural	850	700	136	14	83.7	85.5
Percent minority enrollment						
Less than 5 percent/missing	551	470	75	6	86.2	89.5
5 to 20 percent	949	766	175	10	81.6	82.8
20 to 50 percent	898	678	210	10	76.4	79.3
50 percent or more	1,167	810	331	26	71.0	76.7
Region						
Northeast	679	495	177	7	73.7	78.0
Midwest	899	705	172	22	80.4	83.2
South	821	647	164	10	79.8	82.5
West	1,166	877	276	13	76.1	80.9

<sup>1</sup>In SSOCS:2006, a minimum of 60 percent of the 231 total subitems on the questionnaire were required to be answered for the survey to be considered complete. Of the 231 subitems, this includes a minimum of 80 percent of the 103 critical subitems, 60 percent of item 16 subitems (18 out of 28 total), and 60 percent of item 22 subitems (19 out of 30 total).

<sup>2</sup>Nonrespondents include schools whose districts denied permission to NCES and those eligible schools that either did not respond or responded but did not answer the minimum number of items required for the survey to be considered complete. In total, there were 40 schools whose districts denied permission to NCES, 345 schools that did not send back a questionnaire, and another 404 that were other noninterviews including refusals, undeliverables, and the partially completed questionnaires that did not qualify as an interview.

<sup>3</sup>Ineligible schools include those that had closed, merged with another school at a new location, changed from a regular public school to an alternative school, or are not a school, and those in which district permission to survey schools was denied to NCES.

<sup>4</sup>The unweighted response rate is calculated as a ratio: completed cases / (total sample - known ineligible).

<sup>5</sup>The weighted response rate is calculated by applying the base sampling rates to a ratio: completed cases / (total sample - known ineligible).

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2005–06 School Survey on Crime and Safety (SSOCS), 2006.

- C0546—Number of paid part-time special education aides
- C0550—Number of paid part-time regular classroom teachers
- C0554—Number of paid part-time regular classroom aides/paraprofessionals
- C0558—Number of paid part-time counselors

### **Analysis of Item Nonresponse Bias**

An item-level bias analysis was performed to determine the extent to which, for items with response rates less than 85 percent, nonresponding schools differed from responding schools. This analysis was done because differences between the schools that did and did not respond to an item can lead to bias in estimates.

The magnitude of item nonresponse bias is determined both by the level of item response and by the differences between item respondents and item nonrespondents on a survey item. Because the values of the survey items are not known for item nonrespondents, the distributions of eight sampling frame variables<sup>8</sup> were compared between the nonrespondents and respondents for the 13 subitems with response rates of less than 85 percent.

Among the items examined, ten (C0234, C0236, C0238, C0242, C0326, C0330, C0542, C0546, C0554, and C0558) were identified as having negligible nonresponse bias. The other three items (C0406, C0408 and C0550) had significant differences in their distributions across most of the key variables examined. The distributions between respondents and nonrespondents for survey items associated with C0406, C0408 and C0550 were then examined. Based on these analyses, it was determined that the increased potential for bias in these items was not enough to warrant their exclusion from the data file. More detailed information on the analysis of item nonresponse, including the specific comparisons that were significant in the tests outlined above, is available in section 3.7 and appendix L of the *School Survey on Crime and Safety: 2005–06 Data File User’s Manual* (Bauer et al. 2007).

### **Sampling Variability**

Estimates derived from a probability sample are subject to sampling error because only a small fraction of the target population is surveyed. In surveys with complex sampling designs, such as SSOCs:2006, estimates of standard errors that assume simple random sampling typically underestimate the variability in the point estimates. The standard errors in this report were produced using the jackknife replication method. The standard errors for a range of survey estimates can be computed by hand or by using a statistical package such as SAS, Stata, or SUDAAN. For guidance on how to produce survey estimates and their related standard errors using SSOCs data, please see sections 2.6 and 2.7 of the *School Survey on Crime and Safety: 2005–06 Data File User’s Manual* (Bauer et al. 2007).

---

<sup>8</sup>The eight 2003–04 CCD frame variables used in this analysis are instructional level, school enrollment size, locale, percent minority enrollment, region, number of full-time equivalent teachers, student-to-teacher ratio, and percentage of students eligible for free or reduced-price lunch.

## Statistical Tests

The tests of significance used in this analysis are based on Student's  $t$  statistic. The  $t$  statistic between estimates from various subgroups presented in the tables can be computed by using the following formula:

$$t = \frac{x_1 - x_2}{\sqrt{SE_1^2 + SE_2^2}}$$

where  $x_1$  and  $x_2$  are the estimates to be compared (e.g., the means of sample members in two groups) and  $SE_1$  and  $SE_2$  are their corresponding standard errors.

Due to the large sample size, many differences (no matter how substantively minor) are statistically significant. Thus, only differences of 5 percentage points or more between groups are mentioned in the findings.

## References

Bauer, L., Tang, S., Derby, T.L., Parmer, R.J., and Swaim, N.L. (2007). *School Survey on Crime and Safety: 2005–06 Data File User's Manual* (NCES 2007-358). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Miller, A.K. (2004). *Violence in U.S. Public Schools: 2000 School Survey on Crime and Safety* (NCES 2004-314 REVISED). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.