

Appendix D

Data Sources and Definitions

I. DATA SOURCES

National Education Longitudinal Study

The National Education Longitudinal Study of 1988 (NELS:88) is a longitudinal survey that began in the spring of 1988 with a cohort of 24,599 8th graders. A two-stage stratified probability design was used to select a nationally representative sample of schools and students. In the first stage, 1,734 schools were selected, of which 1,052 (61 percent) participated—815 public and 237 private schools. From these participating schools, 26,599 students (93 percent) participated in the survey. On average, each of the participating schools was represented by 23 students. The student, who filled out a background questionnaire and completed an assessment test, is the basic unit of analysis in the NELS:88. All other components of the study—school, teacher, and parent questionnaires—are primarily intended to supplement the student data set.

NCES has been conducting NELS follow-up surveys every 2 years. This report uses data from the NELS Base Year Survey and the First Follow-up Survey, which was conducted in 1990 (when the students were generally in the 10th grade). For the First Follow-up Survey, several components were added, including a dropout study.

Schools and Staffing Survey

The Schools and Staffing Survey (SASS) is an integrated survey of schools, school districts, school administrators, and teachers designed to explore the major issues concerning the school work force and workplace. A sample of more than 9,300 public and 3,500 private schools was drawn for the SASS that the National Center for Education Statistics (NCES) conducted in the 1987–88 school year. From these schools, a sample of 56,000 public and 11,000 private school teachers

were selected. For this report, data from the 1987–88 survey were used to construct several indicators. In addition to the 1987–88 survey, NCES also conducted the SASS in the 1990–91 school year. Since SASS does not survey students, information on the students enrolled in a particular school is only available through the items directed at the schools, school administrators, and teachers concerning the characteristics of teachers and schools.

National Assessment of Educational Progress (NAEP) and NAEP High School Transcript Study

The National Assessment of Educational Progress (NAEP) is a federally funded periodic assessment of the educational achievement of students in various subject areas. The school and student samples for the 1990 NAEP were selected using a complex multistage sampling design. The four stages included 1) selection of geographically based Primary Sampling Units (PSUs); 2) selection of schools within sampling units; 3) assignment of session types (these are based on assessment subjects) to schools; and 4) selection of students for session types within schools. The resulting samples were nationally representative of schools and students. The High School Transcript Study (HSTS), which was linked to the 1990 NAEP survey, included a representative sample of high school seniors who graduated in the 1990 calendar year. The file includes high school transcripts of approximately 21,500 students from 330 schools; 16,456 of these students were from public schools. To increase the reliability of estimates, private schools and public schools with high enrollments of Hispanic or black students were oversampled.

The schools included in the sample for the 1990 HSTS included all schools with 12th grades that were selected to participate in the 1990 NAEP assessment whether they participated or not. The students chosen, when possible, were the same students who partici-

pated in the 1990 NAEP assessment. When this was not possible, such as cases where schools had not participated in the NAEP, a sample of students from those schools was drawn specifically for the HSTS.

High School and Beyond

High School and Beyond (HS&B) is a national longitudinal survey of high school sophomores and seniors conducted by NCES. HS&B is a stratified national probability sample of students. For the first stage of the HS&B, a sample of public and private high schools was selected. In the second stage, 36 seniors and 36 sophomores were selected in each school. (In schools with fewer than 36 students in either of these groups, all eligible students were included.) Certain types of schools were oversampled to make the study more useful for policy analysis: public schools with a high percentage of Hispanic students, Catholic schools with a high percentage of minority students, alternative public schools, and private schools with high achieving students. In 1980, more than 30,000 sophomores and 28,000 seniors enrolled in 1,015 public and private high schools participated in the Base Year Survey. The survey instruments used in the Base Year Survey included a sophomore questionnaire, a senior questionnaire, a school questionnaire, a teacher comment checklist, and a parent questionnaire. HS&B participants have been followed through the 1980s (1982, 1984, and 1986) and into the 1990s, with the Fourth Follow-up of the 1980 sophomores taking place in 1992.

Technical Note

The weight (panelwt4) applied to the data reported here is the weight used for cases with data from all of the HS&B surveys: the Base Year, First Follow-up, Second Follow-up and Third Follow-up. This is necessary because the samples drawn for the follow-up surveys were not the same as the original base year sample. The base year sample included 28,000 sophomores from 1,015 schools with a target number of 36 students from each school. For the First Follow-up, 30,000 people who were 1980 sophomores were selected. The Second and Third Follow-up sophomore

cohort samples were the same, totaling 15,000. The response rate for the 1984 survey was 92 percent, and for the 1986 survey, it was 91 percent.

National Longitudinal Survey of Youth

The National Longitudinal Survey of Youth (NLSY), conducted by the Bureau of Labor Statistics, is a longitudinal survey of a representative sample of the noninstitutionalized civilian and military population born between the years of January 1, 1957, through December 31, 1964. They have been interviewed on an annual basis since 1979, with the most recent interview year being 1996. Two special subsamples of the survey are 1) a supplemental sample that overrepresents the civilian Hispanic, black, and economically disadvantaged non-black/non-Hispanic youth living in the United States; and 2) a military sample representing the youth population serving in the armed forces as of September 30, 1978, and born January 1, 1957, through December 31, 1961. Although a subsample of the original military sample has been surveyed, the military sample was dropped in 1985 as a complete subsample.

In order to categorize respondents by the community type of the high school they attended, a separate file was used that included an item for urbanicity derived from the HS&B survey. This file contains data only for respondents who attended public high schools.

II. DEFINITION OF URBANICITY AND POVERTY CONCENTRATION IN SURVEYS

Although every effort was made to categorize schools by the same criteria for location and poverty concentration across all of the surveys used in this report, the way in which these two variables were defined in each of the data sources does vary. Below is a brief description of the definition of these two variables for each data source analyzed.

National Education Longitudinal Study

School urbanicity, either urban, suburban, or rural, for the Base Year Survey of NELS was based on urban-

icity assigned to schools in the frame used to create the NELS sample, and corresponds to U.S. Bureau of the Census classifications. For this survey, the Quality Education Data (QED), a universe database of public and private parochial and nonparochial schools in the United States, was used. When urbanicity data were unavailable from the QED, Bureau of the Census data were used. For the NELS 10th-grade Follow-up Survey, the same source for school location data was used to ensure comparability with the 1988 survey. Schools classified as urban are located in central cities of Metropolitan Statistical Areas (MSAs); schools classified as suburban are located within the area surrounding a central city within a county constituting the MSA; and schools classified as rural are outside of an MSA.

For poverty concentration, the 1988 base year school administrators were asked to report the *number of students* receiving free or reduced price lunches. For the analysis in this report, the number reported was divided by the total school enrollment to get the percentage of students receiving free or reduced price lunch. This resulted in school poverty concentrations that were similar in distribution to those produced by the SASS. However, for the 10th-grade follow-up, school administrators were asked to list the *percentage of students* receiving free or reduced price lunch. The resulting distribution of schools matches more closely the data on the percentage of disadvantaged students reported for the HS&B and the NLSY databases.

Schools and Staffing Survey

Respondents reported on the urbanicity of schools in the school questionnaire administered as part of the 1987–88 SASS, while in the other surveys (NELS, NAEP, HS&B, and NLSY) school urbanicity was assigned to the school based upon the address. All subsequent administrations of the SASS survey have used a locale code assigned to the school based upon its address. As of the publication of this report, these codes have been added to the 1987–88 data to enable comparisons between the different administrations of SASS. In 1987–88, the SASS respondents chose from 10 residence categories, ranging in size from a rural or farming community to a very large city of more than

500,000 people. For this report, the 10 categories have been collapsed into the same three urbanicity categories (urban, suburban, and rural) as were used in the other data sets. Three city sizes in SASS—ranging from a medium-sized city of 50,000 to 100,000 people, a large city of 100,000 to 500,000 people, or a very large city of more than 500,000 people—have been collapsed into one urban group comprised of cities of 50,000 people or more. The suburbs of each of these cities have been collapsed into one suburban category, with military bases or stations added to this group. One rural category has been created from the following three groups: a rural or farming community, a small city or town of fewer than 50,000 people that is not a suburb of a larger city, or an Indian reservation.

The self-reported urbanicity in the 1987–88 SASS results in noticeable differences in how schools are categorized when compared with data on school location from other sources. Although the number of urban schools appears to be roughly comparable between the 1987–88 SASS and other data sources, the number of schools classified as rural is higher, and consequently, the number of suburban schools is lower, when compared with Bureau of the Census location data. For example, according to 1990 Bureau of the Census data presented in appendix table 1.1, 28 percent of public school students attended urban schools, 45 percent attended suburban, and 28 percent attended rural schools. However, according to the comparable SASS enrollment data for 1987–88 using the school-reported location, 30 percent of public school enrollment is in urban schools, 23 percent in suburban, and 47 percent in rural.

Poverty concentration for the SASS survey was defined as “the percentage of students in the school receiving free or reduced price lunch.” This figure was calculated by using the number of students receiving lunches divided by the total enrollment.

National Assessment of Educational Progress (NAEP) and NAEP High School Transcript Study

The urbanicity of schools in the National Assessment of Educational Progress (NAEP) High School Transcript Study (HSTS) survey was assigned to

schools based on school address using 1980 Bureau of Census data and definitions of urban, suburban, and rural areas. Schools classified as urban are located in central cities of Metropolitan Statistical Areas (MSAs); schools classified as suburban are located within the area surrounding a central city within a county constituting the MSA; and schools classified as rural are outside of an MSA.

The definition of poverty concentration used for the NAEP survey, which was linked to the HSTS, was the percentage of students in a school receiving free or reduced price lunch which was derived from estimates by the school administrator who filled out the questionnaire.

High School and Beyond

School urbanicity was assigned to schools in the High School and Beyond (HS&B) survey based on 1980 Bureau of the Census data and definitions for urban, suburban, and rural areas. Schools classified as urban are located in central cities of Metropolitan Statistical Areas (MSAs); schools classified as suburban are located within the area surrounding a central city within a county constituting the MSA; and schools classified as rural are outside of an MSA.

School poverty concentration was derived from school administrator responses to a school questionnaire item asking first what criteria, whether federal, state, or other guidelines, were used to classify students as disadvantaged. Then the responding school official was asked to estimate the percentage of “disadvantaged”

students in the school, according to the guidelines of Chapter 1 of the Elementary and Secondary Education Act. This figure was used to classify schools into the poverty concentration categories used in this report.

National Longitudinal Survey of Youth

School urbanicity was not a variable on the public release file of the National Longitudinal Survey of Youth (NLSY) survey. However, urbanicity codes were matched to the respondent’s high school by the Center for Human Resource Research at Ohio State University using the original HS&B sampling frame. School urbanicity was assigned to schools in the HS&B survey based on 1980 Bureau of the Census data and definitions for urban, suburban, and rural areas. Schools classified as urban are located in central cities of Metropolitan Statistical Areas (MSAs); schools classified as suburban are located within the area surrounding a central city within a county constituting the MSA; and schools classified as rural are outside of an MSA. For the purpose of this report, urbanicity codes for respondents in the nonmilitary sample who attended public high schools were made available.

Poverty concentration was defined in the same manner as it was for the HS&B survey, as discussed above. School administrators were asked what percentage of the students in their school were classified as “disadvantaged” according to the guidelines of Chapter 1 of the Elementary and Secondary Education Act.