DOCUMENT RESUME

ED 279 699 TM 870 086

AUTHOR	Spencer, Bruce D.
TITLE	Efficient Methods for Sampling Out-of-School Seventeen-Year-Olds in the National Assessment of Educational Progress.
PUB DATE	11 Sep 86
NOTE	20p.; One of 46 papers commissioned by the Study Group on the National Assessment of Student Achievement and cited in Appendix B to their final report "The Nation's Report Card" (TM 870 049). For other papers in this group, see TM 870 050-094.
PUB TYPE	Viewpoints (120)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Adolescents; Cohort Analysis; *Dropouts; Educational Assessment; Educational Testing; *Error of Measurement; *High School Graduates; High Schools; Longitudinal Studies; National Surveys; Research Methodology; *Sample Size; *Sampling; *Testing Problems; Testing Programs; Test Interpretation
IDENTIFIERS	*National Assessment of Educational Progress

ABSTRACT

The National Assessment of Educational Progress (NAEP) currently tests seventeen-year-old students enrolled in public and private secondary schools, but it does not test "out-of-school" seventeen-year-olds who have either graduated or dropped out. Estimating that one of five seventeen-year-olds is out of school, the interpretability of NAEP findings is weakened by omitting those persons from assessment. This paper considers precision and sample size, as well as sampling strategies, for assessing out-of-school seventeen-year-olds. Four types of statistics (comparisons of average scores for states, average scores for population subgroups, comparisons across states of proportions of seventeen-year-olds performing at a given level, and proportions of subpopulations performing at a given level) are considered and illustrated with tables to indicate sample sizes necessary to attain alternative target levels of precision in the statistics. Two kinds of sampling strategies are described, one based on sampling households and other places of residence and the other on sampling at younger ages for recontact at age seventeeen. It is concluded that sampling out-of-school seventeen-year-olds will be best accomplished through cooperation with other ongoing survey efforts, such as the Current Population Survey of the Bureau of the Census and the National Educational Longitudinal Studies. (LMO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED279699

Efficient Methods for Sampling Out-Of-School Seventeen-Year-Olds in the National Assessment of Educational Progress

Bruce D. Spencer

Northwestern University and NORC

Paper commissioned by

THE STUDY GROUP ON THE NATIONAL ASSESSMENT OF STUDENT ACHIEVEMENT

1986

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
B. D. Spencer
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

EFFICIENT METHODS FOR SAMPLING
OUT-OF-SCHOOL SEVENTEEN-YEAR-OLDS
IN THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Bruce D. Spencer
Northwestern University and NORC[1]

September 11, 1986

1. Introduction

The National Assessment of Educational Progress (NAEP) currently tests seventeen year-old students enrolled in public and private secondary schools but it does not test "out-of-school" seventeen year-olds.[2] "Out-of-school" refers to students not enrolled in elementary or secondary schools. The out-of-school seventeen-year-olds include both those seventeen year-olds who have dropped out of school as well as those who graduated from high school. As we estimate that 1 out of every 5 seventeen-year-olds is "out-of-school"[3], the omission of these persons from NAEP weakens the interpretability of NAEP findings. For

[1]The author is Associate Professor, Departments of Statistics and Human Development and Social Policy and Center for Urban Affairs and Policy Research, Northwestern University, and Director, Methodology Research Center, NORC.

[2]Here a seventeen-year-old is a person who had attained age 17 by October 1 of whatever academic year or calendar year is under consideration.

[3]This figure is derived as follows. The 1980 census counted a total of 77,000 seventeen-year-olds (as of 4/1/80) in institutions or group quarters (U.S. Bureau of Census, 1980 Census of Population, PC80-2-4D, Tables 2 and 44). The civilian noninstitutional population of seventeen-year-olds (as of 10/1/80) numbered 4,097,000, of which 3,393,000 were "in school" (U.S. Bureau of Census, Current Population Reports P-20, No. 400, Table 41). Ignoring the minor differences in the definition of seventeen-year-old, and conservatively estimating that the group quarters and institutional seventeen-year-olds were not "in school", we estimate that about 3,393/4,174 or 81% were "in school". We assume that this proportion holds constant over time.

example, an improvement in the drop-out rate for students accompanied by no change in the performance levels of all seventeen-year-olds would cause a decline in NAEP scores for seventeen-year-olds.[4] Here NAEP findings would appear to indicate a deterioration in the educational system, when in fact there was an improvement--fewer students dropped out, and performance levels of students were unchanged. Or, for another example, a finding that 90% of the tested seventeen-year-olds could read at a certain level could well mean that fewer than 80% of all seventeen-year-olds could read at that level.[5]

Generally, we are interested in the performance of all seventeen-year-olds, not the selected majority who are secondary school students. Of course, attention may focus on all seventeen-year-olds who live in a particular state or who belong to a particular subgroup. When one considers assessing the performance of all seventeen-year-olds, it is natural to consider assessing the performance of other out-of-school populations as well. For example, what is the literacy level of adults aged 21-25 and is literacy improving over time? What knowledge (as opposed to attitudes) does the voting-age population possess of nuclear energy, the strategic strength of the nation, the proportion of the federal budget that goes to education or other social services or defense? The viability of a democracy must depend in part on the knowledge of the voters, on an educated electorate. Assessment of adult

---

[4]This decline arises because those students who in past years would have dropped out of school but now remain in school tend to score lower than the population of seventeen-year-olds as a whole.

[5]Theoretically, as few as 72% could read at that level in the (implausible) event that none of the out-of-school seventeen-year-olds could read at that level.

("out-of-school") populations is vital to ensuring that the voting population is sufficiently educated to carry out its responsibilities, yet distressingly little such assessment is conducted. While noting the importance of assessing persons at older ages, we restrict consideration in this paper to (the out-of-school) seventeen-year-olds.

2. Considerations of Precision and Sample Size

Statistics produced by any data collection program are subject to various kinds of errors, and the goal of efficient survey design is to strike a favorable balance between the level of error and the cost of gathering the data and producing the statistics. NAEP utilizes sampling, and an important source of error is sampling variability. To measure sampling variability we will use the notion of a standard error, which may be simply interpreted as the typical size of the difference between a statistic calculated from a particular sample and what the statistic would have been if the entire population had been surveyed. Typically, the chances are about two out of three that the difference (or error) is smaller in size than one standard error, and they are about nineteen out of twenty (95%) that the difference (or error) is smaller than two standard errors. The coefficient of variation (c.v.) expresses the standard deviation as a percentage of the number we are trying to estimate. For example, if we estimate the proportion of seventeen-year-olds who can read at a certain level to be 40% and the standard error is 8%, then the c.v. is estimated to be 20% (= .08/.40).

One can reduce the standard error by (1) increasing the sample size, (2) by keeping the sample size fixed and increasing the efficiency of the sample design (i.e., changing the way the sample is selected so that the precision is increased), or (3) by a combination of changing the sample size and changing the efficiency. Increasing the efficiency of a sample design may increase the cost, and for that reason the most cost-effective strategy may not be to use the most efficient design. To facilitate comparisons of alternative sample designs we will use the notion of effective sample size, which we will denote by n_{eff}. Those readers familiar with the concept of design effect or deff will note that the number of seventeen-year-olds in the sample, say n, is equal to the product of the design effect and the effective sample size, or $n_{eff} = n/deff$.

How large a sample of seventeen-year-olds is needed depends of course on what statistics are of interest and how accurate they need to be. We will consider four types of statistics here and we will indicate the sample sizes necessary to attain alternative target levels of precision in the statistics. The statistics include comparisons of average scores for states, average scores for population subgroups, comparisons across states of proportions of seventeen-year-olds performing at a given level, and proportions of subpopulations performing at a given level. The scores will be assumed to be standard scores, i.e., scores expressed in standard units, with standard deviations equal to 1.

The standard error of a comparison of the average standard score in one state with the average standard score in another state is

ERIC

approximately equal to the square-root of the sum of the reciprocals of the effective sample sizes in the two states. Table 1 shows what sample sizes are needed to yield various standard errors when the same size sample is taken in eac' ' cate.

TABLE 1

Standard Error of Comparison of Average Standard Scores in Two States as a Function of Sample Sizes

Effective Sample Size in Each State

Standard Error										
1%	2%	2.5%	3%	4%	5%	6%	7%	8%	9%	10%
20,000	5,000	2,263	2,222	1,250	820	556	408	312	247	200

Thus, an effective sample size of 312 in each state yields comparisons of state average standard scores with standard errors of 0.08. Table 1 may also be used to analyze the sample sizes necessary for comparing subgroups across states. For example, to compare the average standard score for males in state A with females in state B with a standard error of 10%, it is sufficient to have an effective sample size of 200 males from state A and 200 females from state B.

Table 2 shows the sample sizes needed to attain various standard errors, not of comparisons across states, but simply of the average standard score within a state or other subpopulation.

TABLE 2

Standard Error of Average Standard Score
for a Subpopulation as a Function of Sample Size

Effective Sample Size in Subpopulation

Standard Error

1%	2%	2.5%	3%	4%	5%	6%	7%	8%	9%	10%
10,000	2,500	1,600	1,111	625	400	278	204	156	124	100

Sometimes we are interested in comparisons across states of the proportions of students who can perform at certain levels (e.g., answer a given item correctly). The precision of such a comparison depends on the actual proportions in each state as well as on the sample sizes. Table 3 shows how the standard error and coefficient of variation of such a comparison vary with the underlying proportions and the effective sample sizes. For example, if the underlying proportions are each near 25% and an effective sample size of 400 is taken from each state, the comparison of the estimated proportions will have a standard error of about 0.03 and a coefficient of variation of about 12%.

ERIC

TABLE 3

Standard Errors and Coefficients of Variation of Comparisons of Proportions for Two States as Functions of Sample Sizes

Standard Errors (C.V.'s) of Comparisons

Effective Sample Size in Each State	Actual Proportions in Each State				
	.10	.25	.50	.75	.90
100	.042 (42%)	.061 (24%)	.071 (14%)	.061 (8%)	.042 (5%)
124	.038 (38%)	.055 (22%)	.063 (13%)	.055 (7%)	.038 (4%)
156	.034 (34%)	.049 (20%)	.057 (11%)	.049 (7%)	.034 (4%)
204	.030 (30%)	.043 (17%)	.049 (10%)	.043 (6%)	.030 (3%)
278	.025 (25%)	.037 (15%)	.042 (8%)	.037 (5%)	.025 (3%)
400	.021 (21%)	.031 (12%)	.035 (7%)	.031 (4%)	.021 (2%)
625	.017 (17%)	.025 (10%)	.028 (6%)	.025 (3%)	.017 (2%)
1,111	.013 (13%)	.018 (7%)	.021 (4%)	.018 (2%)	.013 (1%)
1,600	.011 (11%)	.015 (6%)	.018 (4%)	.015 (2%)	.011 (1%)
2,500	.008 (8%)	.012 (5%)	.014 (3%)	.012 (2%)	.008 (1%)
4,444	.006 (6%)	.009 (4%)	.011 ([illegible]%)	.009 (1%)	.006 (1%)
10,000	.004 (4%)	.006 (2%)	.007 (1%)	.006 (1%)	.004 (Z)

Z - less than 0.5%

The next table shows the sample sizes needed to attain various levels of precision for an estimated proportion for a subpopulation by itself (i.e., not as part of comparison).

TABLE 4
Standard Errors and Coefficients of Variation of Proportions as Functions of Sample Sizes

Standard Errors (C.V.'s) of Proportions

Effective Sample Size in Sub-Population	Actual Proportion in Subpopulation .10	.25	.50	.75	.90
100	.030 (30%)	.043 (17%)	.050 (10%)	.043 (6%)	.095 (3%)
124	.027 (27%)	.039 (16%)	.045 (9%)	.039 (5%)	.085 (3%)
156	.024 (24%)	.035 (14%)	.040 (8%)	.035 (5%)	.077 (3%)
204	.021 (21%)	.030 (12%)	.035 (7%)	.030 (4%)	.066 (2%)
278	.018 (18%)	.026 (10%)	.030 (6%)	.026 (3%)	.018 (2%)
400	.015 (15%)	.022 (9%)	.025 (5%)	.022 (3%)	.015 (2%)
625	.012 (12%)	.017 (7%)	.020 (4%)	.017 (2%)	.012 (1%)
1,111	.009 (9%)	.013 (5%)	.015 (3%)	.013 (2%)	.009 (1%)
1,600	.008 (8%)	.011 (4%)	.012 (2%)	.011 (1%)	.008 (1%)
2,500	.006 (6%)	.009 (3%)	.010 (2%)	.009 (1%)	.006 (1%)
4,444	.004 (4%)	.006 (3%)	.008 (2%)	.006 (1%)	.004 (Z)
10,000	.003 (3%)	.004 (2%)	.005 (1%)	.004 (1%)	.003 (Z)

Z - less than or equal to 0.5%

Tables 1 and 3 show that accurate state-level comparisons will not generally be possible unless the overall sample size is quite large. For example, to compare proportions in two states so that the standard error of the estimated difference is sure to be less than about .02 requires an effective sample size of more than 1,111 per state (from Table 3, line 15, column 3). If such accuracy were needed for all states, the total effective sample size would need to exceed 50,000 seventeen-year-olds.

Reasonable levels of precision in statistics for larger subgroups are attainable at more modest sample sizes. To estimate the average standard score for a subpopulation with a standard error of 0.025 requires an effective sample size of 1600 for the subpopulation. If the subgroup is as large or larger than about one-sixteenth of the cohort of seventeen-year-olds (e.g., the subpopulation of black males is about this large) and the sampling rate for that subgroup is the same as for the cohort as a whole, then a total effective sample size of 25,600 is needed.

The remainder of this paper will assume that the sample will have an effective size of about 26,000. This assumption implies that the actual number of students responding to the NAEP questionnaires must well exceed 26,000, due to the inevitable complexity of the sampling design to be used. The complexities include clustering of students in schools, unequal selection probabilities for students in different schools, unequal sampling rates for in-school and out-of-school seventeen-year-olds, and other factors. For the present purposes, we may approximate the relationship between the needed number of students to be assessed, n, and the effective sample size, n_{eff}, by the formula

$$n = d_{clus} \cdot d_{disp} \cdot n_{eff}$$

where d_{clus} represents the design effect due to clustering of students in schools and d_{disp} represents the design effect due to disproportionate sampling of in-school and out-of-school seventeen-year-olds. Past experience suggests that d_{clus} will be about 1.5 or so. We may approximate d_{disp} by the formula

$$d_{disp} = .04/f + .64/(1-f)$$

with f = the fraction of the sample allocated to the out-of-school seventeen-year-olds. Determination of the optimal allocation of the sample to in-school versus out-of-school seventeen-year-olds is beyond the scope of this paper.

Table 5 shows some alternative sample allocations yielding effective sample sizes of 26,000.

TABLE 5

Illustrative Sample Allocation and Effective Sample Sizes
(see text for assumptions)

in-school sample respondents	out-of-school sample respondents	total sample respondents	effective sample size
41,909	3,592	45,501	26,000
35,036	5,005	40,041	26,000
32,500	6,500	39,000	26,000

Note that the last allocation in Table 5--with a total of 39,000 respondents--is not necessarily less costly than the others. The cost-effectiveness of the three allocations depends on the relative costs of assessing in-school versus out-of-school seventeen-year-olds.

Recently NAEP has assessed on the order of 30,000 seventeen-year-olds in a given year. In light of this fact, it is plausible that the desired number of out-of-school seventeen-year-old participants in NAEP should be around 5,000 or 6,000.

3. Sampling Strategies

The key to efficiently sampling the out-of-school seventeen-year-olds is to be able to identify them reliably and inexpensively. We will consider two kinds of sampling strategies. One strategy is based on sampling households (and other places of residence). The other strategy is based on sampling seventeen-year-olds at younger ages when they are almost all in school.

First we consider household surveys. The U.S. population resides in approximately 85 million households.[6] The Bureau of the Census reports that the households in 1980 contained about 4.1 million seventeen-year-olds; due to declining birth rates in the 1960's, in 1990 the households will contain only about 3.5 million seventeen-year-olds. Making the simple but reasonable approximation that each household contains no more than one seventeen-year-old, we may estimate that somewhat fewer than 1 in 25 households will contain a seventeen-year-old. As only 1 out of 5 seventeen-year-olds is out-of-school, somewhat fewer than 1 in 125 households will contain an out-of-school seventeen-year-old.

These figures--1 in 25 households containing a seventeen-year-old and 1 in 125 households containing an out-of-school seventeen-year-old-

[6]The current number of households is slightly less, but in a few years the number will be about 85 million, or perhaps even 90 million.

- imply that a large number of households will need to be screened to yield a sample of the desired size.

A sample of 100,000 households can be expected to yield 4,000 seventeen-year-olds. Even if areas with high proportions of out-of-school seventeen-year-olds could somehow be efficiently[7] identified and oversampled, it is doubtful that more than 1,000 - 2,000 out-of-school seventeen-year-olds could be sampled in this way. The screening costs alone for a 100,000 household survey would likely exceed 4 million dollars. Thus, conducting a special purpose household survey to sample out-of-school seventeen-year-olds will be very expensive and not very fruitful.

An attractive alternative is to capitalize on existing large household surveys which could perform the screening at minimal cost. The Current Population Survey (CPS), fielded by the Bureau of the Census, is a monthly survey of some 60,000 households. Each successive month some households are retained from previous months and other, new, households are selected, so that over the course of a year some 200,000 households are surveyed. By spreading out the screening for the out-of-school seventeen-year-olds for a twelve month period, it should be possible to sample some 2,000 out-of-school seventeen-year-olds. Not all of the out-of-school seventeen-year-olds will be residents of households, but it is possible to unambiguously link each non-household-resident seventeen-year-old to a residence; for example, a college student living in a dormitory would be eligible for selection

[7]It is necessary that all of the out-of-school seventeen-year-olds have approximately equal selection probabilities, or else the effective sample size will decrease.

if the student's parents' STET (or step parent's, or legal guardian's, etc.) household were sampled.[8] In this way, a sample of approximately 2,000 out-of-school seventeen-year-olds will be attainable.

It is clear that 2,000 is too few out-of-school seventeen-year-olds to support an effective overall sample size of 26,000 unless some 115,000 in-school seventeen-year-olds are sampled. Either this sample of 2,000 needs to be augmented, or we will have to live with a smaller effective sample size. For example, a combination of 2,000 out-of-school plus 33,000 in-school seventeen-year-olds yields an effective sample size of almost 17,000.

We now turn to consider a second sampling strategy, which may be used either by itself or in combination with the CPS-based sample described above. The basis for this second strategy is the fact that almost all children in the U.S. aged 14 and below are enrolled in school at ages below 15. Thus, one way to obtain a sample of seventeen-year-olds is to sample thirteen-year-olds in a school-based sample, keep track of the sampled students, and then recontact them four years later when they are aged 17 and include them in NAEP.

To better describe the details of the strategy, we will first consider its hypothetical application to NELS:88, the National Educational Longitudinal Study of the eighth grade class of 1988. In 1988 (the base year) NELS:88 will sample approximately 30,000 eighth grade students from about 1000 public and private schools containing eighth

---

[8] An application of this technique is described in M.R. Frankel, H.A. McWilliams, and B.D. Spencer, National Longitudinal Survey of Labor Force Behavior, Youth Survey (NLS), Technical Sampling Report, NORC, August 1983.

graders. As 68% of the students in the eighth grade class of Fall 1981 were age 13[9], we may expect the NELS:88 base-year sample to yield more than 20,000 thirteen-year-olds, representing 65% - 70% of the cohort. NELS:88 will follow these students (or a subsample of them) and interview them at age 15 in 1990 and at age 17 in 1992. The seventeen-year-olds in 1992 would need to respond to the NAEP questionnaire as well. In this way, all but between 30% and 35% of the seventeen-year-old cohort may be sampled.

To see how to sample the remaining 30% to 35% of the seventeen-year-olds, consider the breakdown of seventeen-year-olds by status at age 13:[10]

enrolled in grade 8	68%
enrolled in grade 7	18%
enrolled in grades 1-6	3%
enrolled in grades 9-12	8%
in U.S. but not in school	1%
not in U.S.	2%
all seventeen-year-olds by status four years earlier	100%

Nearly all of the 18% of the thirteen-year-olds who are enrolled in grade 7 attend schools eligible for inclusion in the NELS:88 sample, and they could easily be screened and included in the NELS:88 sample at slight extra cost beyond the currently planned NELS:88 surveying. Some portion of the 3% of the thirteen-year-olds enrolled in grades 1-6 and some slight fraction of the 8% enrolled in grades 9-12 also attend

---

[9]Enrollment figures based on U.S. Bureau of Census, Current Population Reports, Series P-20 No. 400, Table 15; data for persons not in U.S. based on U.S. Bureau of the Census, 1980 Census of Population, PC 80-2-2A, Table 1.

[10]Same as above.

schools eligible for inclusion in the NELS:88 sample, hence they too are easily sampled by NELS:88. However, the remainder of the thirteen-year-olds enrolled in grades 1-6 and 9-12 do not attend schools that NELS:88 will sample, and so an extra sample of schools would be necessary to sample these students. Assuming these students comprise about 9% of the thirteen-year-olds, an extra 70-80 schools would suffice to yield samples of these students with similar sampling weights and clustering characteristics to the main NELS:88 sample of thirteen-year-olds; of course, alternative sampling allocations could also be used. Thus, by slightly expanding NELS:88 97% of the seventeen cohort may be sampled.

The remaining 3% of the cohort is comprised of the 2% who immigrate after age 13 and the 1% who are not enrolled in school at age 13. Those immigrants who enter the school system after age 13 may be screened for by NELS:88 or NAEP during later school surveying. Those immigrants who do not enter the school system are problematic to sample. Similar comments apply to the 1% not enrolled in school at age 13; if they enter the school system later they may be screened for in school surveying operations, and otherwise they are problematic. I do not have good suggestions at this time for these difficult-to-sample groups; perhaps there is some comfort in the fact that they comprise perhaps 2% or less of the cohort. On the other hand, the CPS is prone to possibly even larger under-coverage of the out-of-school seventeen-year-olds.

This second strategy, of linking the seventeen-year-old sampling to NELS:88 (or other longitudinal survey of younger students) has the potential of yielding substantial numbers of out-of-school seventeen-year-olds. For example, the sampling strategy for NELS:88 could yield 30,000 seventeen-year-olds, of which 6,000 would be expected to be out-of-school. (The other 24,000 would not receive NAEP questionnaires, except through chance overlap of the NELS:88 and NAEP school samples.)

An alternative to linking the sample of out-of-school seventeen-year-olds to a longitudinal survey is to link it to a cross-sectional survey, such as NAEP samples of schools containing fifteen-year-olds. In particular, NAEP could select a sample of fifteen-year-olds from the sampled schools and follow up the students two years later when they are aged seventeen.

To see how this strategy would work, consider the breakdown of seventeen-year-olds by status at age 15[11]:

enrolled in grade 10	62%
enrolled in grade 9	22%
enrolled in grades 7-8	3%
enrolled in grades 1-6	Z
enrolled in grades 11-12	8%
in U.S. but not in school	3%
not in U.S.	2%
all seventeen-year-olds by status two year earlier	100%

Z - less than 0.5%

Assuming that NAEP will sample at least approximately 325 schools containing eleventh grade students and 475 schools containing eighth grade students, there should be no particular difficulty in obtaining a

[11]Enrollment figures based on U.S. Bureau of the Census, Current Population Reports, Series P-20 No. 400, Table 15; data for persons not in U.S. based on U.S. Bureau of the Census, 1980 Census of Population, PC80-2-2A, Table 1.

sample of more than 30,000 fifteen-year-olds enrolled in grades 7-10. A very small proportion of fifteen-year-olds are enrolled in grades 1-6, and these could easily be screened for as part of the NAEP sampling of schools containing fourth grades. The remaining 5% of the cohort is comprised of the 2% who immigrate after age 15 and the 3% who are not in-school at age 15. These individuals will be somewhat difficult to handle, but we may attempt to sample them in the manner described earlier for the NELS:88-linked sample.

The difficulty of sampling the immigrants and the out-of-school fifteen-year-olds is more problematic for the NAEP-linked strategy than for the NELS:88-linked strategy, because the numbers of these students are larger (5% of the cohort versus 3%) for the former than the latter. In principle, one could apply the NAEP-linked strategy to fourteen-year-olds rather than to fifteen-year-olds, but then the seventeen-year-old sample would have to occur three years after NAEP, rather than two years after.

If one can wait until 1992 for the out-of-school seventeen-year-old sample then the NELS:88-linked strategy will be more cost-effective than NAEP-linked sampling of thirteen-year-olds because NELS:88 will be following the majority of these students anyway, so data collection costs will be lower than for the NAEP-linked strategy. The NAEP-linked strategy has the potential to yield 6,000 out-of-school seventeen-year-olds before 1992, so if a sample of out-of-school seventeen-year-olds in 1989, 1990, or 1991 is needed then a NAEP-linked strategy is recommended.

Conclusions

Sampling out-of-school seventeen-year-olds will be best accomplished through cooperation with other ongoing survey efforts. The Current Population Survey has the potential capability to yield a sample of up to approximately 2,000 out-of-school seventeen-year-olds. NELS:88, suitably expanded, has the capability of producing a sample of 6,000 out-of-school seventeen-year-olds in 1992. By screening for fourteen-year-olds or fifteen-year-olds as part of ongoing NAEP surveys it is possible to yield a sample of 6,000 out-of-school seventeen-year-olds prior to 1992, perhaps as early as 1989. The needed sample size depends on the uses to be made of the data, but it is probably closer to 6,000 than to 2,000.

The NAEP-linked strategy will be more expensive than the NELS:88-linked strategy, so if 1992 is a satisfactory time for the assessment of out-of-school seventeen-year-olds, the NELS:88-linked strategy is preferable. If the assessment is needed sooner, the NAEP-linked strategy is preferable.

ERIC