IMPROVING THE COVERAGE OF PRIVATE ELEMENTARY-SECONDARY SCHOOLS

Betty J. Jackson, Nancy R. Johnson, Richard L. Frazier
Betty J. Jackson, Bureau of the Census, Washington, DC 20233

KEY WORDS: Data Collection, Evaluation, Traditional, Kindergarten-terminal

I. PURPOSE OF THIS PRESENTATION

This paper discusses the "traditional" universe of private elementary and secondary schools developed by the Census Bureau for the National Center for Education Statistics or NCES. This universe was initially developed in 1987, and subsequently updated five (5) times, with the sixth update currently in progress. Results of earlier updates have been previously reported, so this presentation focuses on the most recent updates, in 1995. Key results of the updates, including an analysis of the sources of added schools, their characteristics, and the impact of the adds on the universe will be discussed. Additionally, development of the Kindergarten-Terminal or K-terminal frame and results of the capture/recapture analysis will be discussed.

II. BACKGROUND

As background, it is useful to provide definitions that pertain to this paper.

A. K-TERMINAL

A K-Terminal school contains an educational program primarily for 5-year-old children who will enter first grade in the upcoming school year. This includes transitional kindergartens and/or first grades if these children are expected to enter first grade upon completing these programs. Some of these K-Terminal programs may contain nursery or preschool age children. The 1995 PSS estimated approximately 7,300 private K-Terminal schools in the nation.

B. PRIVATE SCHOOL UNIVERSE

It is useful to review the definition of the private school universe and how it is used. The private school universe is defined as including all schools that provide educational services for at least one of grades 1-12, have one or more teachers, are not administered by a public agency, and are not operated in a private home.

The private school universe is used in two major data collection efforts:

1) First, all of the schools on this universe are included in the Private School Survey or PSS. PSS is a census of private elementary and secondary schools conducted bi-annually for NCES beginning with the 1989-90 school year. PSS has a two-fold purpose.

a) First, it generates bi-annual data on the total number of private schools, along with the number of students, teachers, and graduates at these schools.

b) Second, the results it generates are used to build an accurate and complete list of private schools for NCES to use for other private school surveys.

The 1995 PSS estimated that there are 27,686 private elementary-secondary schools in the nation.

2) The second major data collection effort using this universe is the Schools and Staffing Survey or SASS. SASS selects a sample of approximately 3,500 private schools from the private school universe.

It is also useful to discuss the methodology for compiling and updating the K-Terminal universe and the traditional private school universe.

C. TRADITIONAL PRIVATE SCHOOL UNIVERSE

The traditional private school universe consists of two coverage improvement operations -- List Frame updating and an Area Search Frame. List Frame updating is a national coverage improvement operation designed to locate private elementary and secondary schools not already on the existing private school universe. The updating operation uses lists from private school associations, the 50 states and Washington, D.C., and private vendors. Area Search Frame updating is a coverage improvement operation consisting of an independent search in a nationally representative sample of counties. This operation is used to locate private schools still missing from the private school universe resulting after list frame updating.

As mentioned earlier, the private school universe was initially developed in 1987 with Quality Education Data Incorporated (QED) providing us with a list of private elementary and secondary schools. List Frame updating

was the first step in improving the coverage of this universe. For this update, 22 of the largest private school associations in the country were contacted and their lists of schools were requested. These lists were matched to the QED list and eligible non-matched schools were added to the universe.

The next step in improving the coverage of the first Private School Universe was area frame updating. For this update, a national sample of 75 PSUs was selected and field representatives were instructed to use up to ten (10) different sources such as the Yellow Pages to create an independent list of all private elementary and secondary schools in these sample areas. These independent lists were matched to the universe resulting from the list frame updating within each of the sample PSUs. The in-scope schools that did not match were weighted up to represent the schools that were missing from the updated list frame.

Since the initial development and updating of the private school universe in 1987, the universe has been updated every two years. In 1989, the List Frame updating was done using only 12 association lists due to budget constraints. We picked the lists based on the following criteria:

a)   not too large
b)   had a significant difference in the total number of schools reported between 1987 and 1989.

The Area Frame Updating in 1989 was done in a sample of 120 PSUs.

For updates in 1991, 1993, 1995, and 1997 we used many more lists for the list frame updating. These updates included lists from as many as 44 private school associations, the 50 states and Washington, D.C., QED and Josten's Education Data. For the area frame, we continued to use sets of 120 PSUs as we did in 1989.

D. K-TERMINAL PRIVATE SCHOOL UNIVERSE

In 1993-94, we began to collect information on K-Terminal school programs and build a K-Terminal frame. As lists were collected from the 50 states and Washington, D.C., and Associations for the 1993 list frame updating and from the sample PSUs for the 1993 area frame updating, we began to identify and separate those programs that indicated that they contained at most a kindergarten or were primarily for 5-year-old children.

In 1995, the K-Terminal updating again consisted of a list frame updating and an area frame updating operation. In addition to what was done in 1993, more of an effort was made to contact states or other alternative private organizations to specifically ask for a list of their private kindergartens. This was added to the 1995 operation to evaluate alternative sources for lists of kindergartens and to improve coverage of schools containing kindergartens.

The results of hte 1995 updating operations will now be presented. This analysis is done separately for the list frame updating, the area frame updating, and the K-Terminal operations. The 1995 results from both the list frame and area frame are contrasted with the 1993 results. Note that unless otherwise stated the results for 1995 are similar to those for 1993.

III.   LIST FRAME UPDATING ANALYSIS

In 1995, we added about 2,400 in-scope schools to the universe during the "traditional"list frame updating. The corresponding 1993 number was 2,300.

In terms of the **sources** of the adds, 62% came from the state lists and 38% came from the association lists.

Overall, the state lists were most effective with a total of about 1,500 adds. As might be expected, the list from the state of California provided the largest number of adds (about 25% of the total of the state list adds). The next nine (9) states provided another 43% of the adds, such that the top 10 states accounted for about 2/3 of the state adds in 1995.

The list from Arkansas was the most effective list since about 16% of the schools on the list were in-scope adds. The next three most effective state lists had effectiveness rates above 12%. They were Tennessee, Montana, and Georgia.

The four least effective state lists were from Kansas, Iowa, North Dakota, and Oklahoma with no adds.

Association Lists were also effective, adding about 900 schools. We are not able to do any further detailed analysis of adds from association lists because the information was lost.

The **characteristics** of the adds will now be presented. Regarding the **religious orientation** of the added schools:

●   57% of schools were Other Religious schools

●   39% of schools were Nonsectarian schools

●   4% of schools were Catholic schools

In terms of the **grade level** of the added schools:

- 49% of schools were Elementary schools

- 40 % of schools were Combined schools

- 11% of schools were Secondary schools

In terms of the **size** of the added schools, we see that schools added from the list frame were predominantly small schools contributing 69%. The next largest schools contributed 18%, and the larger schools contributed at most 7%.

In terms of the **percent of minority students** at the added schools, we see that schools with the lowest minority percentage contributed the most to the list frame with 35%. Schools with the next highest minority percentage contributed 23%, and schools with the higher minority percentage contributed 32%.

In terms of the **school type** of the added schools, more than half (58%) are **regular elementary/secondary** schools. Each of the other school types contribute at most 18%.

Looking at the **impact** on the universe estimates, we find that, overall list frame adds represented:

- 8% of schools on the universe

- 3% of students on the universe

- 4% of teachers on the universe

- 1% of graduates on the universe

These percentages were close to what they were in 1993 with the exception of graduates where the impact on the universe was 3% in 1993.

The impact varied considerably for the **religious orientations** and showed that the list frame updating had a substantial impact on improving coverage of Nonsectarian and Other Religious schools and very little impact on Catholic schools.

- Nonsectarian schools led the way with 15% impact

- Other Religious schools followed with 10% impact

- Catholic schools had a minimal 1% impact

The impact for the **school grade levels** showed less variation and indicated that the list frame updating had an impact on improving the coverage for all 3 grade levels.

- Combined schools led the way with a 12% impact

- Secondary schools followed with a 9% impact

- Elementary schools were next with a 7% impact

The impact varied considerably for the different **sized** schools. An inverse relationship exists between the size of school and the size of this impact. The smallest schools had a 19% impact and the largest schools had a 1% impact.

The impact varied only slightly for schools with different **percent of minority students**.

In terms of the impact of **schools of different types**, we see that Voc. Tech., Montessori, and Alternative schools had at least a 24% impact whereas the other 4 had at most a 16% impact. This is somewhat different than what it was in 1993.

IV.   AREA FRAME UPDATING ANALYSIS

In 1995, we identified a weighted estimate of 2,386 in-scope area frame schools during the updating. The corresponding 1993 number was 2,026.

The **characteristics** of the adds will now be presented. Regarding the **religious orientation** of the added schools:

- 62% of schools were Other Religious schools

- 35% of schools were Nonsectarian schools

- 3% of schools were Catholic schools

In terms of the **grade level** of the added schools:

- 47% of schools were Elementary schools

- 49 % of schools were Combined schools

- 4% of schools were Secondary schools.

In terms of the **size** of the added schools, we see that schools added from the area frame were predominantly small schools contributing 77%. The next largest schools contributed 12%, and the larger schools contributed at most 5%.

In terms of the **percent of minority students** at the added schools, we see that schools with the lowest minority percentage contributed the most to the area frame with 37%. Schools with the next highest minority percentage contributed 29%, and schools with the higher minority percentage contributed 20%.

In terms of the **school type** of the added schools, we see that 8 out of 10 schools are either **regular elementary/secondary or Alternative schools**. Each of the other types contribute at most 7% each.

The characteristics of the area frame adds were somewhat similar to those of the list frame adds for Religious Orientation, Grade Level, Enrollment, Percentage of Minority Students, and Type of School.

Looking at the **impact** of these adds on the universe estimates, we find that, overall, area frame adds represented:

- 8% of schools on the universe

- 3% of students on the universe

- 4% of teachers on the universe

- 1% of graduates on the universe

The impact varied considerably for the **religious orientations** and showed that area frame updating had a substantial impact on improving coverage of Nonsectarian and Other Religious schools and very little impact on Catholic schools.

- Nonsectarian schools led the way with 13% impact

- Other Religious schools followed with 11% impact

- Catholic schools had a minimal 1% impact

The impact for the **school grade levels** showed less variation and indicated that area frame updating had an impact on improving the coverage for all 3 grade levels.

- Combined schools led the way with a 14% impact

- Secondary schools followed with a 7% impact

- Elementary schools were next with a 4% impact

The impact varied considerably for the different **sized** schools. An inverse relationship exists between the size of school and the size of this impact.

The impact of schools with different **percent of minority students** varied only slightly.

In terms of the impact of **schools of different types**, we see that **Alternative and ECC/Daycare schools** have a combined 60% impact whereas the others have at most a 16% impact each. This is somewhat different than what it was in 1993.

The characteristics of the area frame adds were somewhat similar to those of the list frame adds for Religious Orientation, Enrollment, and Percentage of Minority Students.

V.   K-TERMINAL UPDATING ANALYSIS

Regarding the **religious orientation** of the added schools:

- 74% of schools were Other Religious schools

- 25% of schools were Nonsectarian schools

- 1% of schools were Catholic schools

In terms of the **grade level** of the added schools:

- 30% were Kindergarten only

- 70% were Kindergarten and less

Looking at the **impact** of these K-Terminal adds on the universe estimates, we find that, overall, these adds represented:

- 41% of schools on the K-Terminal universe

- 34% of students on the K-terminal universe

- 32% of teachers on the K-Terminal universe

The impact varied somewhat for the **religious orientations** and showed that the K-Terminal updating had a substantial impact on improving coverage of all 3 religious orientations.

- Nonsectarian schools led the way with 43% impact

- Other Religious schools followed with 37% impact

- Catholic schools had a smaller although significant impact at 25%

The impact for the **school grade levels** showed some variation as well and again indicated that the K-Terminal updating had an impact on improving the coverage for both grade levels.

- Kindergarten only schools led the way with a 58% impact

- Kindergarten and less schools followed with a 36% impact

VI.   CAPTURE/RECAPTURE ANALYSIS

In this section, the capture/recapture methodology and how it was used to estimate the number of schools on the 1995 PSS universe is discussed.[1] We will compare the capture-recapture estimate of the number of schools to the final weighted PSS estimate (traditional estimate) of the number of schools to estimate the coverage of private schools on the 1995 PSS universe.

The capture-recapture estimate is based on the following assumptions:

1.   The list frame and area frame are independent of one another.

2.   There are no out-of-scope records on either frame.

3.   There are no duplicate school records.

    AND

4.   The probability of observation of a school from a frame has the same expected value for all units.

This can be likened to estimating the number of fish in a pond. There is some unknown quantity (x) of fish. Draw a sample of ten and tag them. The probability of a tagged fish from this first sample is: $P(t) = 10/x$. Throw the

tagged fish back into the pond and draw another sample of ten fish. This time there are 2 tagged fish and 8 untagged fish. Since P(t) is the probability of being tagged in the first capture, 10P(t) should equal the expected number of tagged fish in the recapture. Thus, $10P(t) = 100/x = 2$ and solving for x, we estimate that there are 50 fish in the pond.

In the original list frame, 25,300 schools were "captured" and "tagged". Thus, the probability of inclusion in the list frame can be expressed as $P(t) = 25,300/x$ where x is the population of private schools in the United States.

In the subsequent 2nd sample (area frame), 22,247 schools were "captured", of which 19,861 were "recaptured" or "already tagged". The "recaptured" schools were identified during the area search frame matching operation. Any area search frame school that matched to the list frame can be said to have been "recaptured".

So, $22,247P(t) = 22,247*(25,300/x) = 19,861$. Solving for x reveals a capture-recapture estimate of private schools equal to 28,339. As noted earlier, the traditional estimate of schools is 27,686.

Thus, when comparing the traditional PSS estimate of schools to the capture-recapture estimate of schools, we estimate that the coverage of schools on the 1995 PSS universe to be 97.7%.

It's likely that the private school coverage has been overestimated based on the violation of assumption 1 (that the two frames are independent of one another) which was violated during the area frame operation.

Based on data presented in this paper assumption 4 is also violated to a certain extent. Violation of this assumption tends to underestimate the under coverage. Concerns about the validity of our coverage estimate due to violation of assumption 4 (that the probability of a school from a frame has the same expected value for all units) can be alleviated by poststratification. This is important particularly if steps are taken to address the violation of assumption 1. Poststratification involves computing a capture-recapture estimate for each one of a set of cells, with cells chosen to be correlated with the likelihood of being captured by a particular frame.

---

[1] A discussion of the model and assumptions as it applies to decennial census data can be found in Wolter, K.M. (1986): Some Coverage Error Models for Census Data. Journal of the American Statistical Association, 81,. 338-346.

Private Schools Survey-Capture/Recapture Estimates by Sets of Poststratification Cells

Sets of Poststratification Cells	Estimate
Typology	28,693
Grade Level	28,446
Religious Orientation	28,636
Religious Orientation Within Grade	28,637
TOTAL	28,339

The above table shows that the capture/recapture estimate from each of the four sets of poststratification cells is fairly close to the total capture/recapture estimate. Thus, the poststratification cells that give us the highest capture/recapture estimate will be used. Using the highest estimate would make it least likely that assumption 4 would be violated. Thus, the estimate used is given by the typology cells (28,693).

Thus, when comparing the traditional PSS estimate of schools to the capture-recapture estimate of schools using typology, we estimate that the coverage of schools on the 1995 PSS universe is 96.5%.

VII. CONCLUSIONS

1. First, the "traditional" list frame updating continues to be effective in improving the coverage of private schools - as it added about 9% to the universe.

2. Secondly, since the 1995 area frame estimated that we're still missing 8% of this universe, we need to continue our efforts in this updating to achieve a more complete universe of **all** private schools.

3. Thirdly, coverage improvement operations are especially needed for improving the coverage of small schools, Other Religious and Nonsectarian schools, and non-regular types of schools.

VIII. ADDITIONAL ANALYSIS

The following additional analysis will be done.

We will be factoring in the costs of these operations to do a cost-benefit analysis. These results can be used to develop future updating strategies for different budget scenarios - such as a tight or reduced budget.

We will look at additional K-terminal analysis.

We will also be analyzing the results of the 1997 PSS updating when they become available.