

THE NATIONAL SCHOOL LUNCH PROGRAM DIRECT CERTIFICATION IMPROVEMENT STUDY: ANALYSIS OF UNMATCHED RECORDS (SUMMARY)

Background

Direct certification increases access to school meals for students eligible for the National School Lunch Program (NSLP) and School Breakfast Program (SBP). It also decreases burden on families and district staff by limiting the amount of paperwork they must prepare and process. States conduct direct certification by matching school enrollment data with data from the Supplemental Nutrition Assistance Program (SNAP) and other programs that confer categorical eligibility.

The purpose of this study is to gain a better understanding of the categorically eligible children who are not matched in the direct certification process and to identify potential matching process improvements that might increase the number of matched children.

Methods

Seven States were selected in the study to ensure there was a mix of high-performing, improving, and low-performing States. States participated in this study by submitting SNAP caseload data and NSLP application data from sampled districts. These data were used in two sets of analyses: descriptive analysis of SNAP records and independent match of SNAP records to NSLP applications.

In the first analysis, statewide SNAP participant data from two States were analyzed to compare characteristics of children who were directly certified and those who were not directly certified. Because all children in this sample are categorically eligible for free school meals, these comparisons identified patterns in age, name characteristics, and local area school and economic characteristics associated with more or less success with direct certification.

In the second part of the study, data were analyzed for children certified for school meal benefits by application based on categorical eligibility. These

data, drawn from randomly sampled districts in all seven participating States, represented categorically eligible students who could have been matched in direct certification but were not matched. A two-stage matching process was used to identify these categorically eligible applicants in State-level SNAP participation files. The first stage was to conduct a deterministic match, requiring exact matches on key variables such as name and date of birth. The second stage included probabilistic matching that incorporated more flexible algorithms and allowed inexact matches between data fields.

Findings

Descriptive Analysis of SNAP Records

Students' age distribution varied for SNAP participants by direct certification status, but was not consistent across the two States. The age distribution of directly certified school-age SNAP participants differed significantly from that of other school-age SNAP participants. In the State where almost half of SNAP children were directly certified, 18-year-olds were less likely to be directly certified. Only 1.3 percent of directly certified children were 18 years old compared to 9.3 percent of SNAP children not directly certified. This could be due to children who have dropped out of school and do not appear on the school enrollment lists.

In contrast, in the other State, where about 95 percent of all SNAP children were directly certified, 5-year-olds were less likely to be directly certified. This could be due to 5-year-olds not being enrolled in school yet, or the enrollment data may not have included them at the initial direct certification match.

SNAP participants who were not directly certified tended to have longer, less common names than students who were directly certified. Commonality of last names was calculated using 2000 Decennial

Census¹ list that contains all last names that appeared at least 100 times in that year's census. For example, a child in the 80th first name commonality percentile has a more common first name than 80 percent of people nationally born from 1994 to 2007. In both study States, the first and last names of directly certified school-age SNAP participants were significantly more common names than those of SNAP participants who were not directly certified. These differences were particularly large for last names. In one State, the average last name was at percentile 52.0 for directly certified SNAP participants and percentile 43.8 for other SNAP participants; in the other State, these percentiles were 53.0 and 50.5, respectively.

Directly certified school-age SNAP participants are less likely than other school-age SNAP participants to live in counties with higher private school enrollment. In both States, directly certified school-age SNAP participants lived in counties with significantly lower average private school enrollment rates than other SNAP participants. Categorically eligible students attending private school could be less likely to be matched because private schools are less likely than public schools to participate in the NSLP.

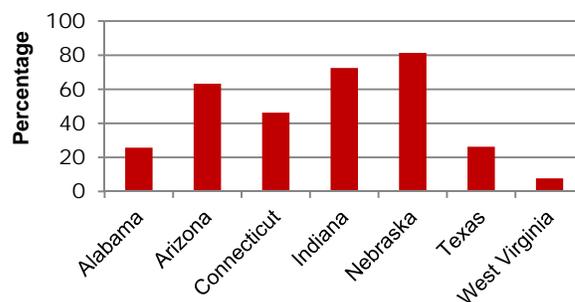
Study States exhibit differences by direct certification status in the average local economic conditions and urbanicity of school-age SNAP participants. In both States, directly certified SNAP participants lived in counties with significantly higher unemployment rates, on average, than other SNAP participants. In one State, they also lived in counties with significantly higher poverty rates. In addition, directly certified SNAP participants in both States were less likely than other SNAP participants to live in urban areas.

Independent Match of SNAP Records to NSLP Applications

The rate of categorically eligible applicants identifiable in State SNAP records through the independent match process varied widely by State. Including both deterministic and probabilistic

matches, match rates ranged from less than 8 percent to 81 percent (Figure 1).

Figure 1. Analysis Matching Rate, by State



Complete data make matching easier. For instance, Nebraska had the highest match rate of 81 percent, and no students were missing multiple elements used for the match. Conversely, West Virginia had the lowest match rate of 8 percent, and nearly 80 percent of its students were missing multiple elements.

Probabilistic matching might offer a way to overcome barriers related to data recording difficulty and data completeness. The study results indicate that probabilistic matching helps to resolve issues related to harder-to-record data items, such as longer, less common names and addresses. For instance, in some States, address is the data field that prevents a deterministic match. A primary challenge to matching using street addresses is the existence of multiple correct variations. A large portion of inexact matches on street address resulted from variations such as *Street* versus *St.* or different ways to represent apartment numbers.

For More Information

Moore, Q., Gothro, A., Conway, K., Kyler B., (2014). National School Lunch Program Direct Certification Improvement Study: Analysis of Unmatched Records. Prepared by Mathematica Policy Research. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support, Project Officers: Allison Magness and Joseph F. Robare. Available online at: www.fns.usda.gov/research-and-analysis.

¹ U.S. Census Bureau. "Genealogy Data: Frequently Occurring Surnames from Census 2000." Available at <http://www.census.gov/genealogy/www/data/2000surnames/index.html>.