Resolution of Multiple Census Returns Using a Re-interview

FINAL REPORT

This evaluation study reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Stephanie Baumgardner

Decennial Statistical
Studies Division

USCENSUSBUREAU
Helping You Make Informed Decisions

EXECUTIVE SUMMARY

The full report for this evaluation is not available because it is census confidential.

The Primary Selection Algorithm was the computer program designed to resolve the receipt of multiple responses from addresses (Census IDs) in Census 2000. The focus of this evaluation is to examine whether or not the Primary Selection Algorithm made the best decisions on the determination of the census household given the information collected from returns for the Census ID. The analysis concentrates on Census IDs with two returns since 97 percent of all Census IDs with multiple returns have exactly two returns.

To examine the Primary Selection Algorithm, a sample of Census IDs affected by the Primary Selection Algorithm was selected. An interview was conducted at each Census ID with someone familiar with the household enumerated during Census 2000. The goal of the re-interview was to determine the residency status of each person on the census returns at the Census ID. The residency statuses obtained in the re-interview were then used to determine if the Primary Selection Algorithm made the best decisions on the determination of the census household.

What are the limitations to this study?

One set of limitations to this study includes operational problems with data collection. Unit nonresponse, allowing the use of proxy respondents, and recall bias may contribute to missing or inaccurate responses. Another set of limitations to this study includes limits to the statistical analysis. This study did not investigate vacant returns or evaluate the person matching process. There were also person records with an unresolved residency status and errors in the inclusion of some Census IDs in the sample. These factors make it difficult to fully evaluate the performance of the Primary Selection Algorithm.

How well did the Primary Selection Algorithm form Primary Selection Algorithm households comprised of two returns?

A Primary Selection Algorithm household is a set of associated persons at one Census ID. If two or more returns for the same Census ID have at least one person in common (determined by person matching), then these returns form a single Primary Selection Algorithm household. At Census IDs with two returns that form one Primary Selection Algorithm household, we found that about 82 percent of the households are formed correctly. This means that there exists at least one resident on each return (besides those matched during person matching) in the Primary Selection Algorithm household.

How often did the Primary Selection Algorithm select the correct Primary Selection Algorithm household when more than one was formed at the Census ID?

When person matching did not match people across the two returns at a Census ID, these two returns formed two separate Primary Selection Algorithm households. At Census IDs with two PSA households, the re-interview determined that there were residents in both households about 38 percent of the time, residents in just one of the households about 58 percent of the time, and no residents in either household about 4 percent of the time.

At Census IDs with two Primary Selection Algorithm households where the re-interview determined that there were residents in both of the households, person matching was performed and missed a duplicate identified during the re-interview in about 16 percent of the cases. Also of interest is how often the Primary Selection Algorithm picked the "best" Primary Selection Algorithm household to represent the Census ID since both of the households contained residents. "Best" is defined here as the Primary Selection Algorithm household with the greater net number of residents. The net number of residents is the balance after subtracting the number of non-residents from the number of residents. At Census IDs with residents in two Primary Selection Algorithm households, the "best" household or a household which was identical in terms of net residents to the other household at the Census ID was selected about 80 percent of the time.

Nearly 58 percent of Census IDs with two Primary Selection Algorithm households have at least one resident in just one of those households. At about 65 percent of these Census IDs, the Primary Selection Algorithm selected the household that contained at least one resident. Of the Census IDs with two Primary Selection Algorithm households and residents in just one of those households, the effectiveness of the Primary Selection Algorithm household selection criteria was examined. As expected, higher priority selection criteria were more effective at selecting the correct Primary Selection Algorithm household than the lower priority selection criteria. However, the selection criterion "Coverage Edit Follow-up Status" did not perform as well as expected. It selected the correct Primary Selection Algorithm household only about 69 percent of the time that it was used.

Recommendations

During Census 2000, Census IDs with two eligible returns that formed two Primary Selection Algorithm households had one vacant household and one occupied household about 51 percent of the time. This evaluation cannot address this sizeable number of cases that the Primary Selection Algorithm handled because vacant returns were not considered for the re-interview. In the future, an evaluation such as this one should set out to determine if the Primary Selection Algorithm selected the right household regardless of whether or not the Primary Selection Algorithm household is occupied. Furthermore, the questionnaire used for the re-interview should be designed to manage both occupied and vacant returns. It should also not rely on census residence

rules to determine residency status for people at the Census ID since the Primary Selection Algorithm itself cannot take those rules into account when making decisions regarding the census household.