

INTEGRATED COVERAGE MEASUREMENT PROCESSING EVALUATIONS

Danny R. Childers, U. S. Bureau of the Census
DSSD, Room 3772-3, Washington, DC 20233*

KEY WORDS: Coverage Measurement, Census Evaluations

1. Introduction

The 1995 Integrated Coverage Measurement (ICM) was designed to collect data from one Computer Assisted Personal Interview (CAPI) instrument to produce two estimates of the population using CensusPlus and Dual System Estimation (DSE) models. There are two processes that are required for producing the two population estimators, the residence status coding operation and the matching and follow up operations for Dual System Estimation.

The CAPI interview for ICM collected a roster of people who lived at each housing unit on census day, the ICM roster. The ICM roster was compared to the people enumerated in the census at this housing unit, the census roster. Questions were asked of all people to establish residence status of the housing unit on Census Day, according to census residence rules. Each person was classified as a resident or a nonresident of the housing unit. If residence status was not established, the person was assigned a residence status of unresolved. The residence status was assigned by computer within the CAPI instrument and selected cases were reviewed clerically during the clerical residence status coding operation.

The residence status coding for each person was needed for both CensusPlus and DSE. Residence status coding identified the resolved roster for CensusPlus. The resolved roster was the roster of people who should have been counted in the census in each housing unit for CensusPlus. Residence status coding also identified the people to be included in the P-sample for DSE. The people identified as Census Day residents of the housing units in the sample blocks were in the ICM population sample, the P-sample. The census enumeration sample or the E-sample was the people enumerated in the census in the sample blocks. The P-sample is matched to the E-sample to estimate missed and erroneously enumerated people for the dual system estimator.

The objective of the residence status coding evaluation was to identify errors in the clerical coding operation that assigned residence status. Errors in applying the census residence rules result in errors of inclusion and exclusion in the resolved roster for CensusPlus and the P-sample for DSE.

For the production DSE, the P-sample and the E-sample were computer matched with a clerical review.

People identified as possible matches, all E-sample nonmatches, P-sample whole household nonmatches, and P-sample people with unresolved residence status were sent for a follow up interview. The results of the follow up were processed to obtain a final match code for the P-sample people and a final enumeration status for the E-sample people.

For the DSE evaluation, the P-sample and the E-sample were rematched in a sample of clusters. People were identified for a follow-up interview. The final codes for the production and the evaluation were compared to identify errors in the P-sample match codes and the E-sample enumeration status. The objective of the DSE evaluation was to identify errors in the DSE process. These errors were either matching errors, respondent errors, or interviewer errors.

2. Residence Status Coding

2.1 The Production

The ICM CAPI instrument collected an ICM roster of names and other information for people who should have been counted in the housing unit on census day. The people in the ICM roster were compared to the people in the census roster. For some housing units, more than one census form was returned. The people captured on the first census form for a census ID received by the processing office by May 22, 1995 were loaded in the CAPI instrument as the census roster.

If the people on the ICM roster were not matched to the people on the census roster, there were probes in the instrument about the reasons the people were not included in the census. If the people on the census roster did not match to the people in the ICM roster, there were probes about the reason the census person was not on the ICM roster. If the people in the ICM roster matched to the people in the census roster, there were probes about other residences they might have had in the past year to determine usual residence. Answers to these probes and information about other residences were collected within the CAPI instrument to assist in the residence status coding phase of the ICM. The residence status coding established, by applying census residence rules, where a person should have been counted on census day.

The residence status was first assigned by computer within the CAPI instrument. The households with unresolved residence status and interviews with notes were clerically reviewed by the Analyst Staff in Jeffersonville

and a code of resident or nonresident of the housing unit on census day was determined. If insufficient information was collected within the ICM instrument to determine the residence status, it remained unresolved.

The numbers and percentages of people classified as resident, nonresident, and unresolved for the three 1995 Census Test sites after clerical review are in Table 1. (Res=Resident, Non=Nonresident, and U=Unresolved in the table.) The numbers of people are unweighted. These numbers represent all people collected on the ICM roster and census people interviewed in the CAPI instrument. Table 1 does not include census people loaded into the CAPI instrument for whom no questions were asked because the case was a noninterview.

	Oakland		Paterson		Louisiana	
	People	%	People	%	People	%
Res	19,764	89.5	19,104	87.8	9,184	91.0
Non	1,170	5.3	1,784	8.2	570	5.6
U	1,152	5.2	881	4.0	342	3.4
Total	22,086	100.0	21,769	100.0	10,096	100

2.2 The Evaluation

For the evaluation, a group of headquarters personnel from the decennial ICM area reviewed the codes assigned during the production. On a continuous basis, errors in assigning the residence status were discussed with the Analyst Staff in Jeffersonville as a quality assurance process. In this way, the quality assurance also improved the entire process.

The quality assurance sample was selected on a flow basis after the clerical residence status coding was completed for a household. The sample was ten percent of the households where at least one person was clerically reviewed in the production. The people flagged for quality assurance within these households were assigned an evaluation residence status code. The percentage of actual people reviewed for the quality assurance in this evaluation was 4.5 percent of all people for Oakland, 5.7 percent for Paterson, and 2.9 percent for Louisiana.

Tables 2, 3, and 4 contain a cross classification of residence status for the production and the evaluation for the three test sites. The numbers for production are the numbers without the quality assurance codes and the numbers for the evaluation are the numbers with the quality assurance codes.

The numbers off the diagonal are the changes or errors identified by the quality assurance coders. The numbers on the diagonal are the number of people with no error in the

production residence status coding. The percentage of people where the quality assurance code was the same as the production code was 97.9 percent in Oakland, 98.6 percent in Paterson, and 98.6 percent in Louisiana. These numbers indicate consistency between production and quality assurance.

Even though there were a few errors identified during the quality assurance, the errors net to a small overall error. For example, in Oakland there were 3 people coded as residents who were changed to nonresidents and 1 person coded as a nonresident who was changed to a resident. In Paterson there was one resident changed to nonresident and two nonresidents changed to residents.

There were cases with a residence status of unresolved during production that were changed to resolved resident or nonresident during quality assurance. There was a tendency to code as unresolved when there existed a possibility of obtaining more information that could have helped to resolve the case with absolute certainty. Some of these cases could be resolved with the existing information and were corrected during the quality assurance. These cases were discussed with the Analyst Staff and these mistakes were not repeated, thus increasing the quality of the residence status coding.

Production	Evaluation			
	Resident	Non-resident	Un-resolved	Total
Resident	726	3	3	732 (73.9%)
Nonresident	1	149	0	150 (15.2%)
Unresolved	10	4	94	108 (10.9%)
Total	737 (74.4%)	156 (15.8%)	97 (9.8%)	990

Production	Evaluation			
	Resident	Non-resident	Un-resolved	Total
Resident	917	1	7	925 (74.2%)
Nonresident	2	231	1	234 (18.8%)
Unresolved	5	2	80	87 (7.0)
Total	924 (74.1%)	234 (18.8%)	88 (7.1%)	1,246

Production	Evaluation			Total
	Resident	Non-resident	Un-resolved	
Resident	213	0	1	214 (73.8%)
Nonresident	0	52	0	52 (17.9%)
Unresolved	2	1	21	24 (8.3%)
Total	215 (74.1%)	53 (18.3%)	22 (7.6%)	290

2.3 Conclusions/Recommendations

The percentage of people where the quality assurance code was the same as the production code was 97.9 percent in Oakland, 98.6 percent in Paterson, and 98.6 percent in Louisiana. These numbers indicate consistency between production and quality assurance.

Even though there were some errors in the residence status coding operation, the errors net to a small overall error. Therefore, the results of this evaluation indicate that the Analysts performed high quality work in the residence status coding phase of the ICM.

The percentage of people coded unresolved during the residence status coding was 5.2 percent in Oakland, 4.0 percent in Paterson, and 3.4 percent in Louisiana. This evaluation indicated that the Analysts were accurately coding the cases with an unresolved residence status. They were unresolved because there was insufficient information collected within the CAPI instrument to code the person as a resolved resident or nonresident.

The CAPI instrument in the 1996 ICM test is being designed to collect the information necessary to code residence status and, therefore, reduce the unresolved rate. This instrument is also being designed to do more of the residence status coding by computer, reducing the number of cases requiring a clerical review to code residence status.

3.0 Dual System Estimation

3.1 The Production

The P-sample was constructed from the CAPI interview. People collected in the Independent Roster who were determined from the interview to be residents of the housing unit on Census Day were defined as the P-sample. The Independent Roster contained the people who were included in the ICM roster before the census roster was revealed to the interviewer. People in the ICM roster who were collected after the census roster was seen were not independent of the census and are not in the P-sample.

The E-sample was defined as the people enumerated in the final census. The Census Roster used in the CAPI instrument could be different from the E-sample, because of later census processing activities and any census enumerations occurring after the census roster was extracted for the CAPI instrument.

For the production DSE, the P-sample and the E-sample were computer matched. The P-sample nonmatches, the E-sample nonmatches, and the possible matches were clerically reviewed. Duplicates within the block cluster were identified clerically for both the P-sample and the E-sample.

There were 9424 P-sample housing units in Oakland. The numbers and percentages interviewed, noninterviewed, and vacant are in Table 5. The percentage of noninterviews was 12.6 percent of the total P-sample housing units and 14.0 percent of occupied housing units. All noninterviews are assumed to be occupied housing units.

Of the noninterviews, 570 (47.8 percent of all noninterviews) were in households where the CAPI interviewer made changes to the entire ICM roster after viewing the census roster. Since the P-sample was no longer independent of the census, these housing units were noninterviews in the P-sample. The noninterviews also included 453 households with last resort data (38.0 percent of all noninterviews).

Interview Outcome	Housing Units	Percent
Interviewed	7,319	77.7
Noninterview	1,192	12.6
Vacant	913	9.7
Total	9,424	100.0

Tables 6 and 7 contain the unweighted numbers for the P-sample match codes and the E-sample enumeration status codes for all of Oakland. A P-sample person was matched when the person was found in the census within the block cluster and the matched E-sample person was correctly enumerated. A P-sample person was not matched when the person was confirmed to have lived at the housing unit on census day and was not found in the census within the block cluster. A P-sample person was coded as unresolved when a follow up interview could not determine the person's census day residence.

No matching was attempted for P-sample and E-sample people with insufficient information for matching and follow up. A complete name and two or more characteristics were defined as the criteria for sufficient information. Both P-sample and E-sample people must have sufficient information to code them as a match or nonmatch with confidence. In addition, there must have been sufficient information to complete a follow up

interview. P-sample people with insufficient information for matching and follow up were unresolved. E-sample people with insufficient information for matching and follow up were subtracted from the census counts with the erroneous enumerations in the dual system estimator.

In order to reduce the follow up workload, half of the clusters were not sent to the field for a follow up interview. P-sample and E-sample people sampled out and not sent for a follow up interview were included with the unresolved cases for imputation. People were removed from the P-sample when they were determined to be duplicated, fictitious, or not really residents of the housing unit on census day.

A follow up interview was conducted for the census nonmatches. The interview identified the people as correctly enumerated or erroneously enumerated in the census. Census nonmatches were classified as correct or erroneous enumerations based on whether they were residents or nonresidents of the block cluster on Census Day, according to census residence rules. Erroneous enumerations also include the census duplicates, fictitious census people, and people in housing units that were included within the block cluster in error (i.e., should have been counted in another census block).

matched to the census, not matched to the census, unresolved match status, and removed from the P-sample.

Match Codes	Production		Evaluation	
	Number	Percent	Number	Percent
Matched	1,541	75.7	1,529	75.1
Not Match	376	18.5	370	18.2
Unresolved	99	4.9	113	5.6
Removed	19	0.9	23	1.1
Total	2,035	100.0	2,035	100.0

Table 9 contains a cross classification of the P-sample match codes for the production and the evaluation. The numbers on the diagonal indicate agreement between the production and evaluation final match codes. Differences between the two final match codes are off-diagonal. The number of P-sample people with agreement between the production and the evaluation match codes divided by the total number of P-sample people was defined as the percentage with consistency between the production and the evaluation. The percentage with consistency between the production and the evaluation P-sample match codes was 96.3 percent. For cases that were resolved as matched or not matched, 99.5 percent were consistent.

There are differences between the production and evaluation final match codes, but the differences net to a small number of errors. For example, there were 17 people who were not matched during the production matching that were coded unresolved during the evaluation and 15 people who were coded unresolved during production matching that were coded not matched during the evaluation.

Comparing only the cases in production that were resolved as matched or not matched was an indication of the quality of the DSE matching. A two by two table of matched and not matched for production and evaluation identified 5 people matched during the production that were coded as not matched in the evaluation and 3 people not matched during the production that were coded as matched during the evaluation. This results in a small net error and thus indicated consistency between the production and evaluation matching.

Match Code	Number	Percent
Matched	14,104	76.4
Not Matched	1,717	9.3
Unresolved	710	3.8
Sampled Out	1,688	9.1
Removed from P-sample	247	1.3
Total	18,466	99.9

Enumeration Status	Number	Percent
Correctly Enumerated	16,684	75.8
Erroneously Enumerated	858	3.9
Insufficient Information	1,465	6.6
Unresolved	657	3.0
Sampled Out	2,353	10.7
Total	22,017	100.0

Production Match Code	Evaluation			
	Matched	Not Matched	Un-resolved	Removed
Matched	1,522	5	14	0
Not Matched	3	345	17	11
Unresolved	4	15	80	0
Removed	0	5	2	12

3.2 The Evaluation

3.2.1 P-Sample Analysis

A comparison of the P-sample results for the production and the evaluation are in Table 8. The numbers are unweighted people in the P-sample who were coded

3.2.2 E-sample Analysis

A comparison of the E-sample results for production and evaluation are in Table 10. The table contains the numbers and percentages of E-sample people coded correctly enumerated, erroneously enumerated, insufficient information for matching and follow up, and unresolved in production and evaluation.

Enumeration Code	Production		Evaluation	
	Number	Percent	Number	Percent
Correct	2,089	80.3	2,077	79.9
Erroneous	173	6.6	111	4.3
Insufficient	158	6.1	157	6.0
Unresolved	181	7.0	256	9.8
Total	2,601	100.0	2,601	100.0

Table 11 contains a cross classification of the E-sample codes for production and evaluation. The numbers on the diagonal indicate agreement between the production and the evaluation enumeration status codes. The sum of the E-sample people with codes that agree divided by the total number of E-sample people was defined as the percentage with consistency between the production and the evaluation. The percentage with consistency was 91.0 percent for the E-sample. Ignoring the unresolved cases and cases with insufficient information for matching and follow-up, the percentage with consistency is 98.4 percent, indicating consistency in the resolved cases for the E-sample.

The errors identified by the evaluation for the E-sample were a combination of clerical coding errors, respondent errors, and interviewer errors in the DSE follow up interview. It was not possible to separate these errors. The difference in the production code and the evaluation code was due to either coding the information on both follow up interviews differently or different information was collected in the field.

If we assume the interview obtained the same information in the production interview and in the evaluation interview, there was clerical coding error in assigning the DSE enumeration code. If we assume the clerical matchers accurately assigned the enumeration status code in the production and in the evaluation, the production and evaluation DSE follow up interview obtained different information.

Usually we assume the evaluation interview is correct and the errors are in the production interviewing, since the best interviewers are used for the evaluation. The interviewers are more thoroughly trained and the evaluation interviews are of the highest quality possible.

Because of the deadlines and workloads during production, there were interviews that were not of the

quality we would like for production work. The work may be late getting to the field, but still must be completed on time. A work force of high quality is not always available, particularly for a census test. For tests, the sample size is larger than in a decennial census for a city. The ICM sample was 10 percent of the test sites in the 1996 test. In a decennial census, the sample is closer to one percent, which makes recruiting more difficult for a test census. For an evaluation, we have a smaller workload and the best interviewers are selected to do the best interview possible. The unresolved rate for an evaluation should be near zero.

The field work for the evaluation interview for DSE follow up was conducted without additional training. Multiple evaluation projects were conducted at the same time, reducing the number of the "best" interviewers.

Enumeration Code	Production	Evaluation		
		Correct	Erroneous	Insufficient
Correct	2,000	4	2	83
Erroneous	30	92	0	51
Insufficient	1	0	155	2
Unresolved	46	15	0	120

Of the 173 E-sample people in the production that were coded as erroneously enumerated, 30 were coded as correctly enumerated, 92 were confirmed to be erroneously enumerated, and 51 were coded unresolved in the evaluation. If we assume the evaluation results are accurate, about one fourth of the resolved cases would be classified as errors. We can not assume the evaluation is accurate for this evaluation.

Errors identified by the evaluation interview in the E-sample for DSE were inconclusive because the percentage of E-sample people coded unresolved after the DSE follow up interview for the evaluation was 10 percent of the total E-sample cases and 30.9 percent of the cases followed up. This indicated that the evaluation interview may not have been of the quality expected. For the production DSE in the clusters sampled for the evaluation, 23.0 percent of the E-sample cases needing a DSE follow up interview were unresolved after the follow up interview.

3.2.3 CAPI Information to Reduce Follow up

Some of the data needed to code the census nonmatches as correctly or erroneously enumerated was obtained within the CAPI interview. Using this information to code the DSE reduced the production workload for DSE follow up. Use of these data was evaluated by sending these

cases to follow up during the evaluation. The final E-sample codes were compared to this evaluation.

Census nonmatches were coded as correctly or erroneously enumerated during production when the CAPI instrument collected the information indicating the person was a resident or not a resident of the housing unit on census day. This information reduced the follow up workload for DSE.

The people coded as correctly or erroneously enumerated without a follow up interview during production DSE were sent for an interview for this evaluation. There was no evidence to suggest that the data collected within the CAPI interview should not be used to code correct or erroneous enumeration and save a repeat visit to the housing unit.

3.2.4 Partial Household Nonmatches

The decision to exclude P-sample partial household nonmatches from the production follow-up was evaluated by sending them for a interview during the evaluation. Twelve of the 76 partial household nonmatches were in households that received a production follow up interview for another household member. Including these partial household nonmatches when another household member is followed up in the production interview would have only slightly increased the cost of the production follow up.

Of the 76 partial household nonmatches, 55 (72.4 percent) were confirmed by the follow up interview to have been residents of the housing unit on census day and were correctly coded as not matched to the census. The evaluation interview did not collect enough information to determine the census day residence of nine (11.8 percent) of the partial household nonmatches.

Eleven (14.5 percent) were removed from the P-sample, because of the results of the follow up interview for this evaluation. A person was removed from the P-sample when the follow up interview concluded the person was not a resident of the housing unit on census day. The P-sample people were identified from the ICM roster in the CAPI instrument as residents of the housing unit on census day. The follow up interview conducted for this evaluation concluded that they were not residents of the housing unit on census day. The partial household P-sample nonmatches should be sent for a follow up interview in the future.

3.2.5 Pristine Rosters

The P-sample in this evaluation contains 79 people from the production Pristine Roster in housing units converted to P-sample noninterviews because the interviewer compromised the independence of the P-sample. The people obtained before the interviewer saw the E-sample are in the Pristine Roster. The corresponding

census count in households that contained these 79 people in the Pristine Rosters was 146 census people.

Of these 79 people, 42 had names that were not real, such as "Household Head" and "Mary Not Sure" or were incomplete, such as "D. Jones". Of the 37 with complete names, 25 had no date of birth and were classified as insufficient information for matching. The definition of sufficient information for matching and follow up was revised for the evaluation to date of birth, because only age and date of birth were kept for the people in the Pristine Rosters. Of the 12 that contained sufficient information, one person was matched, one was confirmed to be missed in the census, 8 were unresolved during the follow up interview, and 2 were removed from the P-sample because they were not residents of the housing unit on census day.

The names classified as not real would have been converted to noninterviews in production when the entire household was not real. The decision to convert these housing units to P-sample noninterviews in production was a good decision.

3.3 Conclusions/Recommendations

A comparison of the production and the evaluation match codes for the P-sample cases indicates a high level of consistency between the two (96.3 percent). More importantly, for the cases that were resolved as matched or not matched, 99.5 percent were consistent. A comparison of the production and the evaluation enumeration status codes for the E-sample indicates a level of consistency of 91.0 percent between the two sets of codes. Ignoring the unresolved cases, the percentage with consistency is 98.4 percent, indicating consistency in the resolved cases for the E-sample.

Errors identified by the evaluation interview in the E-sample for DSE were inconclusive because the percentage of E-sample people coded unresolved after the DSE follow up interview for the evaluation was about 10 percent of the total E-sample cases and 30.9 percent of the cases followed up. This indicated that the evaluation interview may not have been of the quality expected.

We must increase the quality of the field interviews for future evaluations. The unresolved rate for any evaluations should be small.

4.0 References

Census Test Memorandum Series IP-MD-40, November 22, 1995, "Matching Design for the 1995 Integrated Coverage Measurement".

* This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.