



UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

MASTER FILE

12/31/02

DSSD A.C.E. REVISION II ESTIMATES MEMORANDUM SERIES # PP-55

PRED CENSUS AND SURVEY MEASUREMENT STAFF MEMORANDUM SERIES CSM-
A.C.E. REVISION II-R4R

MEMORANDUM FOR

Donna Kostanich
Chair, A.C.E. Revision II Planning Group
Decennial Statistical Studies Division

From:

David Whitford and Rita Petroni *signed 12/31/02*
Chairs, A.C.E. Revision II Measurement Group

Prepared by:

Elizabeth A. Krejsa and Tamara S. Adams

Subject:

Results of the A.C.E. Revision II Measurement Coding

This memorandum documents the results of the A.C.E. Revision II coding operations. It includes a comparison of enumeration and residence status to the original A.C.E. results as well as a quantification of the source of the data, unresolved cases, and the results of the review of conflicting cases.

cc:

A.C.E. Revision II Estimates Memorandum Series Distribution List
Ruth Ann Killion, PRED
David Hubble, PRED

1. BACKGROUND

1.1 A.C.E. Revision II Background

In the fall of 2001, the Evaluation Followup (EFU) coding results showed an increase in the number of erroneous enumerations and nonresidents as compared to the March 2001 A.C.E. An additional review (called the PFU¹/EFU Review) of a sample of E-sample cases (n=17,522) was conducted to verify the E-sample EFU results. The review shows that the March 2001 A.C.E.² underestimated the number of erroneous enumerations in the census but by a lesser amount than indicated by the EFU results.

These errors in the A.C.E. were corrected for the A.C.E. Revision II. For the A.C.E. Revision II, we wanted coding with the same level of quality as the PFU/EFU Review for a large enough sample in both the P-sample and E-sample to provide accurate subgroup estimates. Ideally we would recode the entire A.C.E. sample, but that was not possible because the EFU collected data in only 2,259 out of the 11,303 A.C.E. sample clusters. Even clerically recoding the approximately 70,000 E-sample cases and approximately 52,000 P-sample cases in the EFU sample was not feasible given the time constraints³.

1.2 Using the Keyed Data in A.C.E. Revision II

A new strategy was devised to provide the highest quality data in the time allotted by restricting the clerical review to the more difficult cases. In order to assign the highest quality codes while meeting scheduled dates, we used a computer algorithm with data keyed from both the PFU form and EFU form to augment clerical coding procedures. We then determined if the automated enumeration status coding was of high quality by assessing the level of agreement between the automated codes and the PFU/EFU Review codes, for cases that were coded by both procedures.

1.3 A.C.E. Revision II Clerical Review

Some of the computer coding had a low agreement rate with the code assigned in the PFU/EFU Review. Cases that were coded in this part of the algorithm were sent for clerical coding. Other cases were sent for clerical coding if write-in information, such as an address, was present. Even if a case was eligible for computer coding, if the code assigned by the computer did not agree with the code assigned during the original coding operations, the case was sent to A.C.E. Revision II clerical coding. This helped to protect us against any keying errors in the keyed data.

¹PFU is the Person Followup interview.

²Throughout this document we refer to the March 2001 A.C.E. estimates and any results from the March 2001 A.C.E. as "original A.C.E.".

³Notice that these numbers do not match those in tables 2 and 7. These tables include people who are in the A.C.E. Revision II Revision sample but who do not have forms to be reviewed.

Using keyed data to code cases and only sending the above mentioned types of cases reduced the clerical workload to 23,988 people, which could be completed in the allotted time, and ensured the largest sample possible for the A.C.E. Revision II estimates. Table 1 below details the operations in which the cases were coded. Note that matches are included in both the E- and P-sample counts. This means that the sum of the two A.C.E. Revision II Clerical columns, for example, yields more cases than the 23,988 workload in which a matched person is counted only once.

| Table 1. Final Coding of Cases in A.C.E. Revision II | | |
|--|-----------------|-----------------|
| | E-sample | P-sample |
| Cases not sent to Clerical | 39,509 | 31,528 |
| Cases sent to Clerical | | |
| PFU/EFU Review | 15,678 | 7,035 |
| A.C.E. Revision II Clerical | 14,131 | 14,108 |
| Cases without Forms to Review | | |
| In A.C.E. Revision II Sample (duplicates, insufficient information for matching and followup, cases without EFU, others) | 7,323 | 8,654 |
| Not in A.C.E. Revision II Sample | 90,477 | 106,422 |

As a result of the A.C.E. Revision II Clerical Review, some cases were coded conflicting when the PFU and EFU disagreed about the enumeration or residence status of a person given the same quality data. A special review of these conflicting cases was conducted.

The results of the A.C.E. Revision II Clerical Review coding and Conflicting Review are presented here.

2. E-SAMPLE A.C.E. REVISION II CODING RESULTS

- How does the A.C.E. Revision II coding compare to the original A.C.E. coding?

Table 2. Original A.C.E. vs. A.C.E. Revision II E-sample Coding Results - Unweighted

| Original A.C.E. Results | A.C.E. Revision II Results | | | | |
|-------------------------|----------------------------|-----------------------|------------|-------------|--------|
| | Correct Enumeration | Erroneous Enumeration | Unresolved | Conflicting | Total |
| Correct Enumeration | 57,150 | 1,155 | 1,037 | 265 | 59,607 |
| Erroneous Enumeration | 482 | 9,362 | 204 | 271 | 10,319 |
| Unresolved | 2,540 | 744 | 3,417 | 14 | 6,715 |
| Total | 60,172 | 11,261 | 4,658 | 550 | 76,641 |

Table 3. Original A.C.E. vs. A.C.E. Revision II E-sample Coding Results - Weighted (Standard Errors in Paranthesis)

| Original A.C.E. Results | A.C.E. Revision II Results | | | | |
|-------------------------|----------------------------|-------------------------|------------------------|----------------------|----------------------------|
| | Correct Enumeration | Erroneous Enumeration | Unresolved | Conflicting | Total |
| Correct Enumeration | 244,875,909 (6,314,994) | 2,612,745 (276,011) | 2,868,710 (367,577) | 239,389 (44,474) | 250,596,753 (6,399,632) |
| Erroneous Enumeration | 638,092 (92,949) | 11,265,915 (396,380) | 216,217 (31,810) | 480,265 (310,655) | 12,600,489 (518,811) |
| Unresolved | 2,412,669 (184,293) | 759,176 (69,963) | 3,291,007 (224,607) | 21,962 (9,569) | 6,484,814 (340,690) |
| Total | 247,926,669 (6,375,465) | 14,637,836 (506,818) | 6,375,934 (462,610) | 741,616 (314,230) | 269,682,055 (6,677,302) |

The weights used here use only the probability of selection and do not reflect additional weighting adjustments (e.g., TES weighting). Therefore, the results presented are not directly comparable to similar tables comparing original A.C.E. and Measurement Error Reinterview results (as in the ESCAP II reports numbered 3 and 24) nor are they directly comparable to the results of the PFU/EFU Review.

Additionally, the tables above include people who were not followed up in the EFU (i.e. duplicates, insufficient information for matching, etc.) These people were excluded from the previous reports.

- What is the net difference in erroneous enumerations according to the A.C.E. Revision II coding in comparison with those identified in original A.C.E.?

Correct to Erroneous – The estimated number of original A.C.E. correct enumerations coded as erroneous enumerations in A.C.E. Revision II is 2,612,745.

Erroneous to Correct – The estimated number of original A.C.E. erroneous enumerations coded as correct enumerations in A.C.E. Revision II is 638,092.

Net Difference in Erroneous Enumeration Coding – The net difference in the Correct Enumeration to Erroneous Enumeration and the Erroneous Enumeration to Correct Enumeration cells is 1,974,653. This number represents the erroneous enumerations not identified in the original A.C.E. as a result of clerical coding issues.

- How many cases are coded as unresolved or conflicting?

The estimated number of unresolved people in the A.C.E. Revision II is 6,375,934 (2.4 percent).

The estimated number of conflicting cases in the A.C.E. Revision II is 741,616 (0.3 percent).

The estimated number of unresolved people in the original A.C.E. is 6,484,814 (2.4 percent).

- What is the source of the A.C.E. Revision II codes?

Table 4. Coding Data Source of E-sample A.C.E. Revision II Cases

| | Weighted Percent (N=269,682,055) | Unweighted Percent (N=76,641) |
|------------------------------------|-------------------------------------|----------------------------------|
| PFU/EFU Review | 54.7 | 20.5 |
| Keyed Data Coding | 32.7 | 51.5 |
| A.C.E. Revision II Clerical Coding | 9.2 | 18.4 |
| Original A.C.E. | 3.4 | 9.6 |

In Table 4, we see that much of the E-sample coding was completed using keyed data (51.5%). However, the majority of the weighted E-sample (54.7%) was completed in the PFU/EFU Review. This is due to the sample design of the PFU/EFU Review. Cases completed in the original A.C.E. were those cases coded in before followup as a duplicate, insufficient information for matching and followup, and some possible matches.

- What are the results of the Conflicting Review?

Of the 741,616 weighted conflicting cases, all but 46,738 were resolved in a special review by analysts after the initial A.C.E. Revision II coding was complete. The cases were resolved as follows:

Table 5. Original A.C.E. vs. Conflicting Review Coding Results - weighted

| A.C.E. Revision II Results | | | | | |
|----------------------------|---------|-----------|------------|-------------|---------|
| Original A.C.E. Results | Correct | Erroneous | Unresolved | Conflicting | Total |
| Correct | 100,933 | 102,297 | 16,283 | 19,876 | 239,389 |
| Erroneous | 90,646 | 373,531 | 2,948 | 13,140 | 480,265 |
| Unresolved | 2,288 | 5,185 | 767 | 13,722 | 21,962 |
| Total | 193,867 | 481,013 | 19,998 | 46,738 | 741,616 |

After accounting for the special review of conflicting cases, the resulting coding changes becomes:

Table 6. Original A.C.E. vs. A.C.E. Revision II E-sample (with Conflicting Review)-
Weighted
(Standard Errors in Paranthesis)

| A.C.E. Revision II Results | | | | | |
|----------------------------|----------------------------|-------------------------|------------------------|--------------------|----------------------------|
| Original A.C.E. Results | Correct Enumeration | Erroneous Enumeration | Unresolved | Conflicting | Total |
| Correct Enumeration | 244,976,842 (6,316,475) | 2,715,042 (284,477) | 2,884,993 (367,665) | 19,876 (9,906) | 250,596,753 (6,399,632) |
| Erroneous Enumeration | 728,738 (104,256) | 11,639,446 (499,412) | 219,165 (31,856) | 13,140 (4,453) | 12,600,489 (51,811) |
| Unresolved | 2,414,957 (184,315) | 764,361 (70,090) | 3,291,774 (224,615) | 13,722 (8,181) | 6,484,814 (340,690) |
| Total | 248,120,536 (6,378,857) | 15,118,849 (594,900) | 6,395,931 (462,750) | 46,738 (13,003) | 269,682,055 (6,677,302) |

The resulting net difference in erroneous enumeration coding is 1,986,304. The number of conflicting cases has decreased from the PFU/EFU Review also.

3. P-SAMPLE A.C.E. REVISION II CODING RESULTS

- How does A.C.E. Revision II coding compare to the original A.C.E. coding?

Table 7. Original A.C.E. vs. A.C.E. Revision II P-sample Coding Results - Unweighted

| Original A.C.E. Results | A.C.E. Revision II Results | | | | | |
|-------------------------|----------------------------|-------------|------------|-------------|---------|--------|
| | Resident | Nonresident | Unresolved | Conflicting | Inmover | Total |
| Resident | 42,243 | 745 | 666 | 94 | 441 | 44,189 |
| Nonresident | 181 | 2,664 | 77 | 65 | 152 | 3,139 |
| Unresolved | 802 | 442 | 2,944 | 6 | 137 | 4,331 |
| Inmover | 1,089 | 246 | 43 | 0 | 8,279 | 9,657 |
| Total | 44,315 | 4,097 | 3,730 | 165 | 9,009 | 61,316 |

In the above table, both A.C.E. nonmovers and A.C.E. outmovers fall into the first three rows of the table; A.C.E. inmovers are in the next-to-last row of the table. Those persons who are A.C.E. nonmovers and A.C.E. outmovers can become inmovers, as shown in the first three rows of the inmover column. However, in the original A.C.E., a person who was discovered to be an inmover would have been treated as a nonresident. In addition, A.C.E. inmovers can become nonmovers and outmovers, as shown in the first three columns of the inmover row.

Table 8. Original A.C.E. vs. A.C.E. Revision II P-sample Coding Results - Weighted (Standard Errors in Paranthesis)

| Original A.C.E. Results | A.C.E. Revision II Results | | | | | |
|-------------------------|----------------------------|------------------------|------------------------|---------------------|-------------------------|----------------------------|
| | Resident | Nonresident | Unresolved | Conflicting | Inmover | Total |
| Resident | 246,935,082 (6,346,587) | 2,409,931 (277,384) | 2,601,443 (362,341) | 151,774 (34,474) | 1,018,276 (168,143) | 253,116,506 (6,447,245) |
| Nonresident | 251,139 (40,032) | 3,950,539 (226,995) | 83,911 (15,366) | 101,981 (18239) | 177,859 (33,607) | 4,565,429 (237,645) |
| Unresolved | 920,294 (70,867) | 509,949 (61,541) | 4,221,415 (275,433) | 14,467 (7667) | 161,179 (27,936) | 5,827,304 (309,143) |
| Inmover | 1,685,555 (107,030) | 366,148 (54,779) | 70,541 (16,000) | 0 (0) | 11,999,468 (515,816) | 14,121,712 (561,412) |
| Total | 249,792,071 (6,392,343) | 7,236,566 (379,412) | 6,977,310 (468,214) | 268,223 (39,500) | 13,356,782 (567,105) | 277,630,951 (6,879,364) |

The weights used here use only the probability of selection and do not reflect additional weighting adjustments (e.g., TES weighting and noninterview adjustment). Therefore, the results presented

are not directly comparable to similar tables comparing the original A.C.E. and MER results (as in the ESCAP II reports numbered 16 and 24) nor are they directly comparable to the results of the PFU/EFU Review.

Additionally, the tables above include people who were not followed up in the EFU (i.e. duplicates, insufficient information for matching, etc.) These people were excluded from the previous reports.

- What is the net difference in nonresidents according to the A.C.E. Revision II coding in comparison with those identified in the original A.C.E.?

Resident to Nonresident – The estimated number of original A.C.E. residents coded as nonresidents in A.C.E. Revision II is 2,409,931.

Nonresident to Resident – The estimated number of original A.C.E. nonresidents coded as residents in A.C.E. Revision II is 251,139.

Net Difference in Residence Coding – The net difference in the Resident to Nonresident and the Nonresident to Resident cells is 2,158,792. This number represents the nonresidents not identified in the original A.C.E. as a result of clerical coding issues.

- How many cases are coded as unresolved or conflicting?

The estimated number of unresolved people in the A.C.E. Revision II is 6,997,310 (2.5 percent).

The estimated number of conflicting cases in the A.C.E. Revision II is 268,223 (0.97 percent).

The estimated number of unresolved people in the original A.C.E. is 5,827,304 (2.6 percent).

- What is the source of the A.C.E. Revision II codes?

Table 9. Coding Data Source of P-sample A.C.E. Revision II Cases

| | Weighted Percent (N=277,630,951) | Unweighted Percent (N=61,316) |
|------------------------------------|-------------------------------------|----------------------------------|
| PFU/EFU Review | 49.0 | 11.5 |
| Keyed Data Coding | 35.2 | 51.4 |
| A.C.E. Revision II Clerical Coding | 11.2 | 23.0 |
| Original A.C.E. | 4.6 | 14.1 |

In Table 9, we see that much of the P-sample coding was completed using keyed data (51.4%). However, much of the weighted P-sample (49.0%) was completed in the PFU/EFU Review. This

is due to the sample design of the PFU/EFU Review. Cases completed in the original A.C.E. were those cases coded in before followup as a duplicate, insufficient information for matching and followup, and some possible matches.

- What are the results of the Conflicting Review?

Of the 268,223 conflicting cases, all but 63,457 were resolved in a special review by analysts after the original coding was complete. The cases were resolved as follows:

Table 10. Original A.C.E. vs. Conflicting Review Results - weighted

| | Resident | Nonresident | Unresolved | Conflicting | Total |
|-------------|----------|-------------|------------|-------------|---------|
| Resident | 42,523 | 56,969 | 17,132 | 35,151 | 151,775 |
| Nonresident | 17,916 | 59,899 | 215 | 23,951 | 101,981 |
| Unresolved | 129 | 5,173 | 4,810 | 4,355 | 14,467 |
| Total | 60,568 | 122,041 | 22,157 | 63,457 | 268,223 |

After accounting for the special review of conflicting cases, the resulting coding changes becomes:

Table 11. Original A.C.E. vs. A.C.E. Revision II P-sample (with Conflicting Review)-
Weighted
(Standard Errors in Parenthesis)

| Original A.C.E. Results | A.C.E. Revision II Results | | | | | Total |
|-------------------------|----------------------------|------------------------|------------------------|--------------------|-------------------------|----------------------------|
| | Resident | Nonresident | Unresolved | Conflicting | Inmover | |
| Resident | 246,977,604 (6,348,035) | 2,466,900 (277,789) | 2,618,575 (362,573) | 35,151 (15,756) | 1,018,276 (168,143) | 253,116,506 (6,447,245) |
| Nonresident | 269,055 (40,485) | 4,010,439 (228,282) | 84,125 (15,367) | 23,951 (8,179) | 177,859 (33,607) | 4,565,429 (237,645) |
| Unresolved | 920,423 (70,867) | 515,121 (61,701) | 4,226,225 (275,445) | 4,355 (3,768) | 161,179 (27,936) | 5,827,304 (309,143) |
| Inmover | 1,685,555 (107,030) | 366,148 (54,779) | 70,541 (16,000) | 0 | 11,999,468 (515,816) | 14,121,712 (561,412) |
| Total | 249,852,638 (6,393,997) | 7,358,608 (381,061) | 6,999,466 (468,649) | 63,457 (18,099) | 13,356,782 (567,105) | 277,630,951 (6,879,364) |

The resulting net difference in nonresidence coding is 2,197,845. Also, we see that there is a net decrease in the number of in-movers of 764,930.