Correcting Measurement Error in the Accuracy and Coverage Evaluation

Tamara S. Adams, Elizabeth A. Krejsa, and Douglas B. Olson
US Census Bureau, Washington DC

1. Introduction

As part of the Census 2000, the Census Bureau conducted the Accuracy and Coverage Evaluation (A.C.E.) and the A.C.E. Revision II. The A.C.E. measured the net coverage of the census using an area sample of block clusters. Following the A.C.E. production, the Census Bureau evaluated the quality of the A.C.E. operations using the Evaluation Followup (EFU). The Census Bureau then used the data from the March 2001 A.C.E. and the EFU to correct for measurement errors in the March 2001 A.C.E. in the A.C.E. Revision II (Kostanich 2003). In this paper, we examine the results of the correction of measurement error in the A.C.E. and errors resulting from that correction.

2. Background

The A.C.E. uses two independent samples, the P-sample and the E-sample, to measure net coverage of the census. The P-sample, or Population Sample, was an independently enumerated sample and was used to measure the census misses. The E-sample, or Enumeration Sample, consisted of the census people enumerated in the A.C.E. sample areas. The E-sample was used to measure the census erroneous enumerations and duplicates.

After the A.C.E. housing unit and the person interviewing operations were completed, the person followup matching process was conducted. There were several major steps to the March 2001 A.C.E. person matching and followup process:

1. Computer Match - The P-sample and the E-sample people were matched by computer. The results were used during the before followup clerical matching.
2. Before Followup Matching–The clerical matchers reviewed the P-sample and E-sample persons who were not matched, those who were possibly matched by the computer, and census cases with insufficient information for matching. The matchers also attempted to identify and code duplicated persons within both the P-sample and the E-sample.
3. Person Followup Interview (PFU) - Unresolved and selected unmatched persons were sent to a field interview. During the interview additional information was obtained to help assign a final match and/or residence status to each person. For the E-sample, nonmatches were sent for a follow-up interview to determine if they were correctly or erroneously enumerated in the block cluster. Certain whole household nonmatches in the P-sample were not sent for a person followup interview (Childers 2001). Possible matches were also sent for an interview to resolve their match status.
4. After Followup Coding –The information obtained in the PFU interview was used to code the match and/or residence or enumeration status of the persons in question. These statuses were assigned based on the Residence Rules for Census 2000 (Childers 2001).
5. Evaluation Followup – Following the PFU, the Evaluation Followup interview (EFU) was conducted to assess the quality of the A.C.E. and, specifically, to review the assignment of residence status in the P-sample and enumeration status in the E-sample. The EFU interview was an expanded PFU interview. The same people followed up in the PFU were interviewed using the EFU form. In addition, a sample of people (Krejsa 2000) not interviewed in the PFU were selected for the EFU. The EFU form was expanded to ask more detailed questions about other residences a person may have and about movers from a housing unit. Using this expanded information from the EFU, the Measurement Error Reinterview (MER) matching process (structured similarly to the after followup matching step) was conducted to determine residence and enumeration status of the persons in question.

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 PFU/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. underestimated the number of erroneous enumerations in the census (Adams and Krejsa, 2002).

These errors in the A.C.E. were corrected for the A.C.E. Revision II. For the A.C.E. Revision II, we

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

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 of net coverage. 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 constraints.

Fortunately, both the PFU and EFU questionnaires had been keyed and were available in electronic form for the A.C.E. Revision II process. Due to the complex nature of residence situations, the Census Bureau relies on clerical coding operations to take full advantage of interviewer notes and other information available on a paper form that are not available to use during automated coding. A new strategy evolved to combine automated coding and clerical coding to provide high quality data in the time allotted. The plan restricted the clerical review to the more difficult cases and automated the assignment of codes to the more straightforward cases. Initially an automated algorithm assigned an enumeration status code (or residence status code) and a why code which described the reason for the code assigned. A three-step process was followed to assign final codes to each case:

- Validation – Determine for each why code category if the automated enumeration status coding is 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.
- Targeting – Target only those why code categories that have automated enumeration status codes with low levels of agreement with the PFU/EFU Review data.
- Clerical Coding – Clerically recode only cases in the targeted why code categories. The clerical recoding took advantage of handwritten interviewer comments (Adams and Krejsa, 2002).

We targeted several types of cases for clerical coding. 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. We did this because the keyed data could have keying error, which could lead to an incorrect computer code. 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. The majority of cases in the both the P-sample and the E-sample did not need clerical review in the A.C.E. Revision II.

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 Not Needing Review		
In A.C.E. Revision II Sample	7,323	8,654
Not in A.C.E. Revision II Sample	90,477	106,422

2. Coding Methods

Unlike the production and evaluation coding operations, the Review was conducted by the most experienced staff, called analysts. Each analyst reviewed a workunit of sampled persons, coding the EFU form separately from the PFU form. Then, the analyst indicated which form contained the best code - both, EFU, PFU, or conflicting. We used the following rules to select the form with the best code:

- If either form was unresolved, we chose the other (resolved) form. An exception to this rule was when the unresolved form gave the analyst more information.
- If both forms were resolved and on one form the respondent was a proxy and on the other the

respondent was a household member, we chose the form with a household member respondent.
- If both forms were resolved and the same type of respondent answered both, we picked the form that gave more information.

In some cases a clear determination of the best code could not be made; these cases were termed "conflicting". We coded a case conflicting in the following circumstances
- Contradictory Information from the Same Respondent Type – A case was determined to be conflicting when both forms were completed by the same type of respondent - either both were household respondents or both were similar caliber proxies - who provided contradictory information that resulted in a different enumeration status for the followup person.
- Contradictory Geocoding[1] Information – In addition, if one form indicated through geocoding information that the housing unit was in one place and the other form indicated it was elsewhere, the case was coded as conflicting.

However, unlike the PFU/EFU Review, we used a strict set of guidelines in choosing the cases. After the PFU/EFU Review, we determined that some of the cases coded unresolved should have been coded as correct enumerations and we determined that there were too many conflicting cases. Since this was part of the uncertainty in the October 2001 Executive Steering Committee for A.C.E. Policy (ESCAP) II decision, we decided to standardize the coding as much as possible. We also allowed the analysts to combine the forms in order to code a case, rather than choosing one of the two forms. However, this was not done frequently. In order to minimize the number of conflicting cases, we performed a special review of all conflicting cases at the end of the A.C.E. Revision II clerical coding. In this review, we relaxed the coding rules and allowed the analysts to use their experience and judgement to determine the appropriate code for the case.

3. Limitations

Some data in this report were obtained from the EFU. The most significant limitation of the EFU is the 9 to 10 month time lag between Census Day (April 1, 2000) and when the EFU data were collected in January and February 2001. Given the time lag, people could forget or inaccurately report information. In addition, people may have moved during that time period so more proxy data were collected. The EFU questionnaire was developed, though, to attempt to minimize such problems by asking questions of the respondent that aid them in recalling the correct information. Another limitation is that the EFU did not have a full field quality assurance program as did the A.C.E. Person Interview and the PFU. For both the PFU and the EFU interviews, there is evidence that the questions were not always read as worded (Keeley 1999 and Krejsa 2001). This may have led to varying responses for questions. A third limitation is that standard errors presented in this report are simple jackknife estimates and do not fully capture all phases of the multiphase A.C.E. sampling.

4. Results

Table 2 presents the E-sample A.C.E. Revision II coding results.[2] In this table, we compare the March 2001 A.C.E. with the A.C.E. Revision II. We see the following results:

- Correct to Erroneous – The estimated number of original A.C.E. correct enumerations coded as erroneous enumerations in A.C.E. Revision II is 2,715,042.
- Erroneous to Correct – The estimated number of original A.C.E. erroneous enumerations coded as correct enumerations in A.C.E. Revision II 728,738.
- 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 1,986,304. This number represents the erroneous enumerations not identified in the original A.C.E. as a result of clerical coding issues.
- Unresolved Cases – The estimated number of unresolved people in the A.C.E. Revision II was 6,395,931 (2.4 percent). The estimated

[1] For certain types of cases, the Census Bureau checked the location of the housing unit with respect to census geography, also known as "geocoding".

[2] The weights used here use only the probability of selection and do not reflect additional weighting adjustments. 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.

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

- Conflicting Cases – The estimated number of conflicting cases before the special review in the A.C.E. Revision II was 741,616 (0.3 percent). After the conflicting review, there were 46,738 conflicting cases remaining; 193,867 of the conflicting cases were coded correct enumerations, 481,013 were coded erroneous enumerations, and 19,998 were coded as unresolved.

Table 2. March 2001 A.C.E. vs. A.C.E. Revision II E-sample (with Conflicting Review)- Weighted (Standard Errors in Parenthesis)

	A.C.E. Revision II Results				
March 2001 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)

In Table 3, we present results for the P-sample.

- Resident to Nonresident – The estimated number of original A.C.E. residents coded as nonresidents in A.C.E. Revision II is 2,466,900.
- Nonresident to Resident – The estimated number of original A.C.E. nonresidents coded as residents in A.C.E. Revision II is 269,055.
- Net Difference in Residence Coding – The net difference in the Resident to Nonresident and the Nonresident to Resident cells is 2,197,845. This number represents the nonresidents not identified in the original A.C.E. as a result of clerical coding issues.
- Unresolved Cases – The estimated number of unresolved people in the A.C.E. Revision II is 6,999,466 (2.5 percent). The estimated number of unresolved people in the original A.C.E. is 5,827,304 (2.6 percent).
- Conflicting Cases – The estimated number of conflicting cases before the special review in the A.C.E. Revision II is 268,223 (0.97 percent). Of these, 63,457 remained conflicting; 60,568 were coded as residents, 126,273 were coded as nonresidents, and 22,157 were coded as unresolved.

Table 3. March 2001 A.C.E. vs. A.C.E. Revision II P-sample (with Conflicting Review)- Weighted (Standard Errors in Parenthesis)

	A.C.E. Revision II Results					
March 2001 A.C.E. Results	**Resident**	**Nonresident**	**Unresolved**	**Conflicting**	**Inmover**	**Total**
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)

5. Evaluating the Coding

To assess the potential error in the dual system estimates (DSE) due to the at-risk[3] cases, we used the error rates observed in the PFU/EFU Review to derive estimated error factors for the at-risk cases. The underlying assumption for this approach is that the at-risk cases have the same error factor as the cases in their keyed code category[4] that were in the PFU/EFU Review. We used the following approach to calculate the potential error in the DSE:

- Create Donor Cells – These are cases in a given combined keyed code category[5] in the PFU/EFU Review.
- Calculate Error Factors – We calculated the error factor for each combined keyed code category in the PFU/EFU Review. The error factor is how much error we could incur by accepting that category without further review.
- Calculate Average Unresolved Probabilities – To calculate error factors for unresolved cases, we used the same correct enumeration probability[6] for all cases with that why code category.

Since the data from the A.C.E. Revision II was used only in the double-sample ratio (Kostanich, 2003), we examined the double-sampling ratios for each of the poststratum groups in both the E-sample and the P-sample. We see a summary of the results below in Table 4. Twelve of the 31 E-sample double-sampling ratios have significant differences between the A.C.E. Revision II and those with the at-risk adjustment. The differences in the double-sampling ratios range from –0.0025 (se=0.0016) to 0.0009 (se=0.00004). The largest increase in a poststratum due to the at-risk adjustment is an additional 8,743 (0.035%) correct enumerations; the largest decrease in a poststratum due to the at-risk adjustment is a decrease of 10,798 (0.246%) correct enumerations. In the P-sample, none of the double-sampling ratios with the at-risk adjustments were significantly different from the A.C.E. Revision II double-sampling adjustments. The differences ranged from –0.0009 (se=0.04) to 0.0006 (se=0.03). The largest increase in residents due to the at-risk adjustment is an increase of 7,988 (0.031%) additional residents in a poststratum; the largest decrease in residents due to the at-risk adjustment is a

[3] "At-risk" cases are those cases which received a code assigned by a computer algorithm as their final A.C.E. Revision II code.

[4] A "keyed code category" consists of all cases within a given why code and match code grouping (see Appendix A for details). For instance, all E-sample persons who are erroneous enumerations because they lived in a dormitory are within one "keyed code category".

[5] A "combined keyed code category" consists of all cases that have a keyed code category for PFU and a keyed code category for EFU. For instance, all E-sample persons who, in EFU, are erroneous enumerations because they lived in a dormitory, and who, in PFU, are correct enumerations because they had no other residences, moving, or group quarters are one combined keyed code category.

[6] Note that throughout this report we use the E-sample for discussion's sake. The P-sample is analogous, where enumeration status is equivalent to residence status.

decrease of 8,342 (0.093%) residents in a poststratum. Since these differences are small, we can conclude that the effects of using the keyed data to code cases and save time and money did not generate a large error in the dual system estimates for the A.C.E. Revision II.

Table 4 – Summary of the Effects of the At-Risk Cases on the A.C.E. Revision II Double Sampling Ratios

	A.C.E. Revision II Double-Sampling Ratio	At-Risk Double-Sampling Ratio	Difference
E-sample			
Lower Range	0.97670	0.97917	-0.0025
Upper Range	0.97450	0.97360	0.0009
P-sample			
Lower Range	0.99695	0.99788	-0.0009
Upper Range	0.99567	0.99584	0.0006

6. Conclusions

We see that there were significant changes in the enumeration status of E-sample people and the residence status of P-sample people in the A.C.E. Revision II as compared to the March 2001 A.C.E. Most notably, the net difference in the correct enumeration to erroneous enumeration and the erroneous enumeration to correct enumeration cells is 1,986,304. From previous studies (Adams and Krejsa, 2002), we can attribute much of this difference to people in non-traditional living situations such as group quarters; other residences for work, school, or vacation; children in joint custody situations, or moving in or out of the sample address.

We also note that using keyed data was an effective way of completing a high-quality coding operation with few negative effects on the final estimator. There were very small changes in the double-sampling ratio used in the A.C.E. Revision II estimator. However, in using the keyed data, we were able to use previous coding results and the PFU/EFU Review as a truth-deck. Consequently, the results presented here should not be taken as a carte blanche to use keyed or otherwise automated data in a production coding operation without first assessing the effects on the coding of residence and enumeration status.

7. References

Adams, T. and Krejsa, E. (August 2001), "Independent Review of a Sample of the Person Followup and Evaluation Followup Forms", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-PT-24 and Planning, Research and Evaluation Division TXE/2010 Memorandum Series:CM-MER-S-4

Adams, T. and Krejsa, E. (October 2002), "Results of the Person Followup/Evaluation Followup Review for the Census 2000 Accuracy and Coverage Evaluation", *ASA Proceedings of the Section on Survey Research Methods.*

Childers, D. (January 2001), "Accuracy and Coverage Evaluation: The Design Document", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-1.

Keeley, C. (1999), "Evaluation of the Integrated Coverage Measurement/Post Enumeration Survey Person Followup Questionnaire", Census 2000 Dress Rehearsal Evaluation Results Memorandum Series, Chapter C6.

Kostanich, D. (2003), "A.C.E. Revision II: Design and Methodology," DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-30, U.S. Bureau of the Census, Washington, DC.

Krejsa, E. (2000), "Measurement Error Reinterview Sample Selection", Planning, Research, and Evaluation Division TXE/2010 Memorandum Series: CM-MER-S-01.

Krejsa, E. (May 2001), "Observation of Accuracy and Coverage Evaluation - Evaluation Followup Interview and Office Edit January 22-24, 2001", Planning, Research, and Evaluation Division TXE/2010 Memorandum Series: CM-GEN-T-12.

Krejsa, E. and Raglin, D. (October 2001), "Evaluation Results for Changes in A.C.E. Enumeration Status", ESCAP II, #3.